AN INVESTIGATION OF THE NATURAL HISTORY OF EARLY CERVICAL HUMAN PAPILLOMAVIRUS INFECTION AND ITS RELATIONSHIP TO THE ACQUISITION OF EPITHELIAL ABNORMALITIES OF THE CERVIX

by

STUART IAN COLLINS

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Cancer Sciences
The University of Birmingham
September 2008

Abstract

Cervical human papillomavirus (HPV) infection is a very common sexually transmitted disease which is now considered to be a *necessary,* but not *sufficient,* cause of cervical cancer. It has been suggested that the association between HPV infection and cervical neoplasia can be exploited to improve the efficiency and effectiveness of primary- and secondary-prevention programmes for cervical cancer. However, whether this aspiration can be realized in practice requires a greater understanding of the natural history of early cervical HPV infection and its role in the acquisition of epithelial abnormalities of the cervix. In this thesis, a longitudinal study of young women who had recently embarked on sexual activity has provided sequential observations on the natural history of cervical HPV infection. This thesis addresses four aspects of this natural history: the association between HPV infection and the proximity of first sexual intercourse to menarche; the association between smoking, cervical HPV infection and high-grade cervical disease; the validation of a neutralising antibody assay and its use in defining the kinetics of the humoral immune response to cervical HPV16 and HPV18 infections; and the analysis of measurements of the viral load of HPV16 and HPV18, and their association with epithelial abnormalities of the cervix.

Dedication

This thesis is dedicated to the memory of my grandmother, Marjorie Brown, and my grandfather, Douglas Brown.

Acknowledgments

I am glad to be able to acknowledge the people who have made this thesis possible.

First and foremost, I would like to thank Professor Ciaran BJ Woodman for his supervision of this thesis, his constructive suggestions throughout the writing of it, his patience, and his entirely good natured cajoling to get me to finish it. Above all, however, I would like to thank him for his guidance, teaching, and constant support and encouragement over the last 16 years without which I would not have had the opportunity, the confidence, or the ability, to embark on an undertaking such as this.

I would also like to express my gratitude to:

Professor Lawrence Young for his commitment to me, his ongoing support, and above all his generosity.

Lucinda Billingham for her understanding and unwavering support, which has enabled me to clear the final hurdle.

Christothea Constandinou-Williams, Jane Steele and Professor Paul Murray for their patient instruction which has deepened my understanding of new concepts, and for their help and advice which has improved the quality of this thesis immensely.

The late Heather Winter for her mentorship, which kept me on the right track when required.

Kaisheng Wen, Sally Roberts and Gordon Ryan for their practical advice and support.

Professor Paola Dey and Alan Gibbs for their help in getting me started, and for technical advice.

My mum, Gillian Collins, and dad, Michael Collins, for their love and support.

And last, but not least, the women who participated in the study.

Above, I have gratefully acknowledged people who have been invaluable to me in completing this thesis. However, all opinions expressed in this thesis, and all errors, are mine.

Table of contents

This thesis is quite extensive. To ensure that it remained within the physical limits required by the binders when binding the final hardcopy version, I decided that it was necessary to print the hardcopy in three volumes. A white background indicates that the entry appears in volume 1 of the thesis; a green background indicates that the entry appears in volume 2 of the thesis; and a blue background indicates that the entry appears in volume 3 of the thesis.

List of illustrations

This thesis is quite extensive. To ensure that it remained within the physical limits required by the binders when binding the final hardcopy version, I decided that it was necessary to print the hardcopy in three volumes. A white background indicates that the entry appears in volume 1 of the thesis; a green background indicates that the entry appears in volume 2 of the thesis; and a blue background indicates that the entry appears in volume 3 of the thesis.

List of tables

This thesis is quite extensive. To ensure that it remained within the physical limits required by the binders when binding the final hardcopy version, I decided that it was necessary to print the hardcopy in three volumes. A white background indicates that the entry appears in volume 1 of the thesis; a green background indicates that the entry appears in volume 2 of the thesis; and a blue background indicates that the entry appears in volume 3 of the thesis.

List of abbreviations

Abbreviation	Description
ASCUS	Atypical squamous cells of unknown significance
BNA	Borderline nuclear abnormalities
BPV	Bovine papillomavirus
CDF	Cumulative distribution function
CI	Confidence interval
CIN	Cervical intraepithelial neoplasia (also grade 1, 2 and 3)
CRUK	Cancer Research UK
Ct	Cycle threshold (qPCR)
DNA	Deoxyribonucleic Acid
ELISA	Enzyme-linked immunoabsorbent assay
GAPDH	"Housekeeping" gene found in all human DNA (one copy per strand)
GLM	Generalized linear model
GP	General Practitioner
HARE	HAzard REgression: a computer programme
HC2	Hybrid capture II
HGCIN	High-grade cervical intraepithelial neoplasia (CIN2 or CIN3)
HPV	Human papillomavirus
HR	Hazards ratio
HSIL	High-grade squamous intraepithelial lesion
IC	Interval censored
ICC	Intraclass correlation coefficient
ICE	Name of a macro in SAS
ICESCC	International Collaboration of Epidemiological Studies of Cervical Cancer
IQR	Interquartile range
KM	Kaplan-Meier
LGCIN	Low-grade cervical intraepithelial neoplasia (CIN1 or a histological diagnosis of HPV)
LR	Likelihood ratio
LSIL	Low-grade squamous intraepithelial lesion
M	Name of a computer programming language
MAR	Missing at random
MCAR	Missing completely at random
MPLE	Maximum penalised likelihood estimator
NHS	National Health Service
NHSCSP	NHS cervical screening programme
NMAR	Not missing at random
NPMLE	Non-parametric maximum likelihood estimator
OR	Odds ratio
PCR	Polymerase chain reaction
PHMPL	Name of a computer program for estimating the MPLE
PJS	PJ Smith: acronym coined for a statistical method
PsV	Pseudovirion
QPCR	Quantitative polymerase chain reaction
RR	Relative risk
SCCC	Squamous cell carcinoma/cancer of the cervix
SD	Standard deviation
SEAP	Secreted (embryonic) alkaline phosphatise
SIL	Squamous intraepithelial lesion
UK	United Kingdom
VLP	Virus-like particle

Foreword

Cervical human papillomavirus (HPV) infection is a very common sexually transmitted disease which is now considered to be a necessary but not sufficient cause of cervical cancer.

It has been suggested that the association between cervical HPV infection and cervical neoplasia can be exploited to improve the efficiency and effectiveness of secondary-prevention programmes, if not to replace them outright with primary-prevention strategies. However, whether these aspirations can be realized in practice requires a greater understanding of the natural history of cervical HPV infection and its rôle in the acquisition of cervical abnormalities, than we currently possess. Such an understanding can best be derived from longitudinal observations on women who are free of both infection and disease at recruitment.

In this thesis a longitudinal study of young women who had recently embarked on sexual activity has provided sequential observations on the natural history of cervical HPV infection. This thesis addresses four distinct, but interrelated, aspects of the natural history of cervical HPV infection and its relationship to the subsequent development of epithelial abnormalities of the cervix. I address each of these four aspects in stand-alone groups of chapters which consist of background, methods, results and discussion. These groups of chapters are preceded by introductory chapters which describe: the epidemiology and aetiology of cervical cancer; cervical HPV infection; strategies for the prevention of cervical cancer; the design of the cohort study which has provided the data for this thesis; and some of the statistical and methodological issues which have arisen with this study design.

Chapter 1

THE DESCRIPTIVE EPIDEMIOLOGY OF CERVICAL CANCER

Orientation

In this chapter I describe the incidence of, mortality and survival, from cervical cancer, and how these vary with geographical location, calendar period, and age.

1.1 THE INCIDENCE OF CERVICAL CANCER

In 2002, 5,060,657 new cases of cancer, excluding skin cancer, were diagnosed in women, worldwide: 493,243 (9.7%) were cancers of the cervix, making cervical cancer the second most common cancer among women, after breast cancer and just ahead of colon and rectum cancer, and the seventh most common cancer overall. The incidence rate in 2002, standardised to the age distribution of the world population, was 16.2 per 100,000, and the cumulative risk of cervical cancer prior to the age of 64 years was 1.3%, compared with 9.5% for all cancers combined (Parkin 2005).

1.1.1 Geographical variations in the incidence of cervical cancer

In 2002, 83% of cervical cancers occurred in developing countries, where they accounted for 15% of all cancers occurring in females. The highest incidence rates

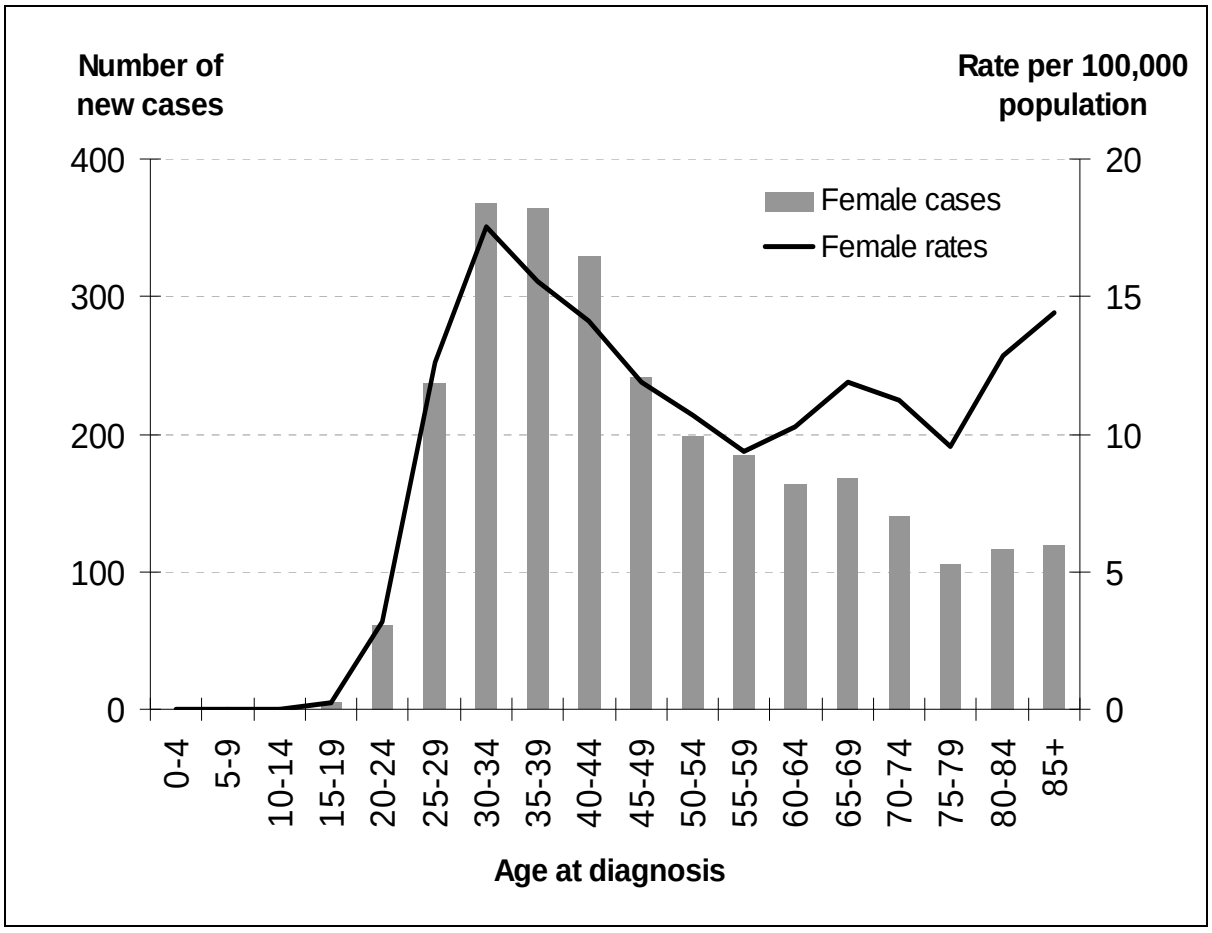
were in sub-Saharan Africa, Melanesia, Latin America, the Caribbean, south-central Asia, and southeast Asia. In comparison, incidence rates were low in developed countries, where cervical cancer accounted for 3.6% of cancers occurring in females, and also in China and Western Asia. This was not the case prior to the introduction of (organised) screening programmes from the 1960s onwards, when incidence rates in developed countries were comparable to those currently seen in developing countries (Parkin 2005). Age-specific incidence varies according to country, and typically depends upon the availability of a screening programme in that country: in the absence of such a programme, the incidence of cervical cancer tends to increase with age, as with the majority of the epithelial tumours. The greatest incidence occurs in women over the age of 50 years (Bosch 2003).

In 2005, 144,756 new cases of cancer, excluding non-melanoma skin cancer, were diagnosed in women in the UK: 2,803 (1.9%) were cancers of the cervix, making it the twelfth most common cancer in women (CRUK website, August 2008). The crude incidence rate was 9.1 per 100,000, with the incidence rate in England being 21% lower than that in Scotland.

In the UK, the incidence of cervical cancer varies with age, with the highest incidence rates seen in women aged 30 to 34 years, but with a second increase in women aged 85 years or over (figure 1.1.1). This phenomenon, i.e. two “peaks” in incidence, one early, and one late, in life, is not confined to the UK alone (Bosch 2003).

In the UK, the cumulative risk of cervical cancer by the age of 65 years is currently estimated to be 0.6%, with only a slight additional increase when calculated over a lifetime (CRUK website, August 2008).

Figure 1.1.1. Number of new cases of cervical cancer, and incidence rates of cervical cancer, by age group, for the UK in 2005 (CRUK website, August 2008).

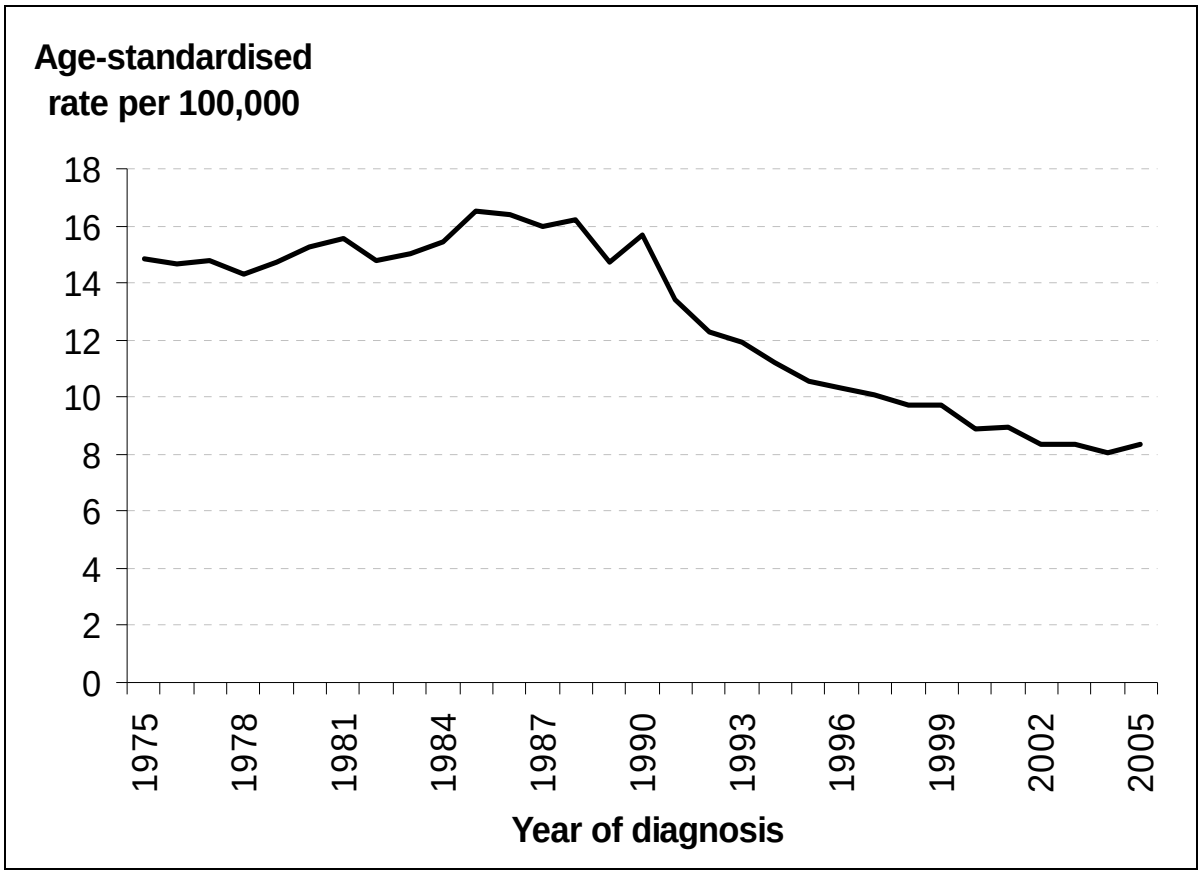


1.1.2 Temporal trends in the incidence of cervical cancer

Over time, there has been a substantial decline in the incidence of cervical cancer worldwide. This decline has not been confined just to those countries with well-established screening programmes, but has also been seen in some developing

countries, in particular in China (Parkin 2005). In Great Britain, the incidence rate of cervical cancer has fallen by nearly 50% since the mid-1980s (figure 1.1.2).

Figure 1.1.2. Trends in incidence rates of cervical cancer in Great Britain for the period 1975-2005, age-standardised to the European population (CRUK website, August 2008).



1.1.3 Squamous cell carcinoma of the cervix

Squamous cell carcinoma is the most frequently diagnosed histological type of cervical cancer, comprising around two-thirds of all cervical cancers diagnosed in England and Wales in 1997 (Quinn 2001), although it has also been reported to comprise up to 85% of cervical cancers (Vizcaino 1998, Souhami 2005). A study

based on cancer registry data, collected between 1962 and 1991 from 25 mostly developed countries, with either systematic screening programs, or high-levels of opportunistic screening, found a progressive decline over time in the incidence of invasive squamous cell carcinoma in a majority of countries, with essentially stable rates in the remainder (Vizcaino 2000). The exceptions was the UK, where, despite an overall decline, there was an increasing trend in both incidence and mortality among women under the age of 50 years in England and Scotland. The assertion that this had begun to decline by the early 1990s is not supported by a more recent study (Bray 2005a).

1.1.4 Adenocarcinoma of the cervix

Adenocarcinoma is the second most frequently diagnosed histological type of cervical cancer, comprising nearly 15% of all cervical cancers diagnosed in England and Wales in 1997 (Quinn 2001). The incidence of adenocarcinoma of the cervix is rising at the same time as that of squamous cell carcinoma of the cervix is falling (Bray 2005b). A study based on cancer registry data, collected between 1962 and 1991 from 25 mostly developed countries, with either systematic screening programs or high-levels of opportunistic screening, found an increase over time in the incidence of invasive adenocarcinoma and adenosquamous carcinoma of the cervix in 12 of the 25 countries, including the UK; no change in 11 countries; and a decline in three (Vizcaino 1998). The authors of the report on this study noted that most of the countries with increasing trends in the incidence of adenocarcinoma of the cervix were those which were devoting considerable effort toward screening, and that some

of the increase may have been due to increased screening efforts during the 1980s. This is unlikely to provide the full explanation however.

1.2 MORTALITY FROM CERVICAL CANCER

In 2002, there were 2,927,896 deaths from cancer among women worldwide, of which 273,505 (9.3%) were attributed to cancer of the cervix, making cervical cancer the third most common cause of cancer-death in women, behind breast and lung cancer, and the seventh most common cause of death overall. Mortality rates are lower than incidence rates: the mortality rate in 2002, standardised to the age distribution of the world population, was 13.2 per 100,000. The cumulative risk of death due to cervical cancer prior to the age of 64 years was 0.7%, compared with 4.9% for all cancers combined (Parkin 2005).

1.2.1 Geographical variation in mortality from cervical cancer

There is substantial variation in mortality rates from cervical cancer according to geographical location: in 2002, the age-standardised mortality rate ranged from 34.6 per 100,000 in Eastern Africa, to around 3 per 100,000 in developed countries; mortality rates in China and Western Asia were comparable to those seen in developed countries (Parkin 2005).

In 2006, there were 73,621 deaths due to cancer in women in the UK: 949 (0.6%) were deaths due to cancer of the cervix, making cervical cancer the sixteenth most

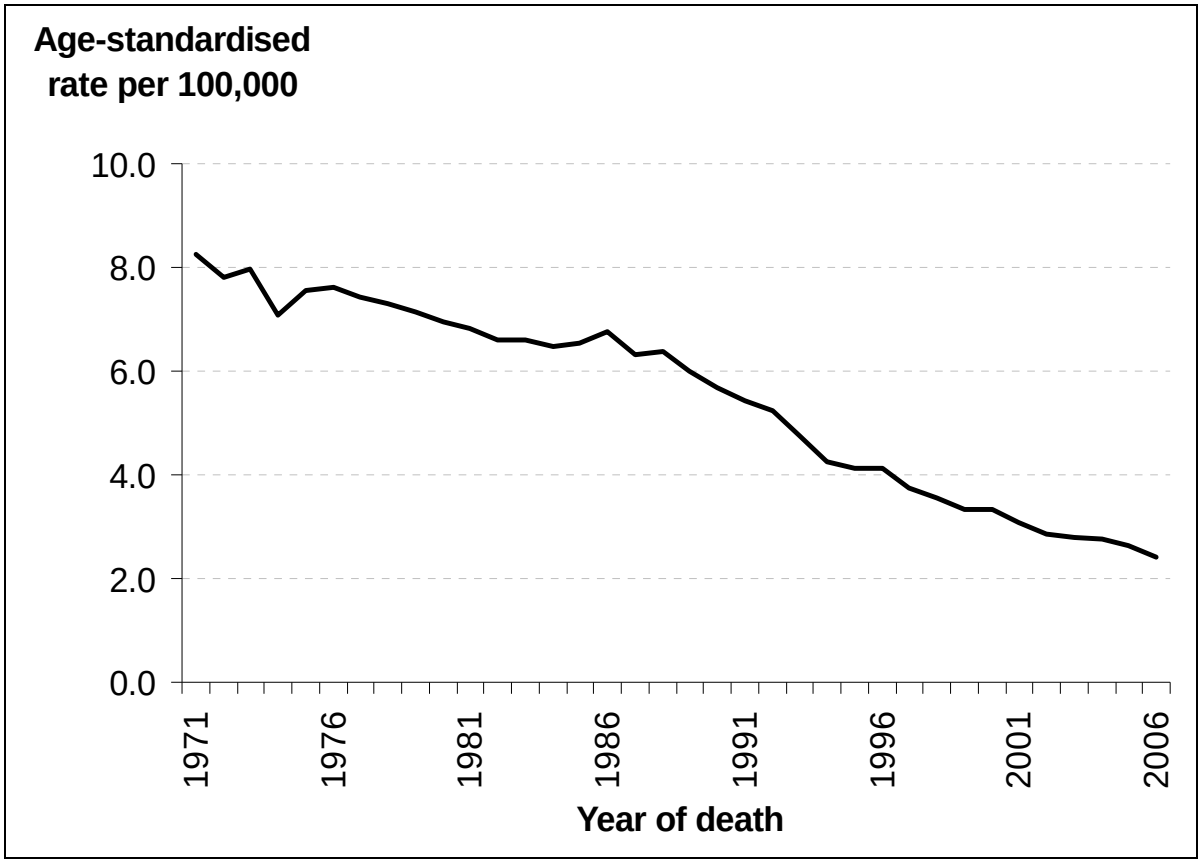
common cause of cancer-death in women in the UK (CRUK website, August 2008): the crude mortality rate was 3.1 per 100,000, with slight national variations among the “home” nations. Cervical cancer mortality rates increased steadily with increasing age.

1.2.2 Temporal trends in mortality from cervical cancer

Over time, there have been substantial declines in mortality from cervical cancer, in both developed countries with well-established screening programmes, and also in some developing countries without such programmes (Parkin 2005).

In the UK, mortality from cervical cancer has decreased dramatically since 1971 (CRUK website, August 2008). This decrease has occurred in all age groups, with the exception of women aged between 25 and 44 years, for whom incidence is high but mortality is low (figure 1.2.2).

Figure 1.2.2. Trends in mortality rates of cervical cancer in the United Kingdom for the period 1971-2006, age-standardised to the European population (CRUK website, August 2008).



1.3 SURVIVAL FROM CERVICAL CANCER

In 2002, the worldwide five-year age-adjusted survival rate for cancer of the cervix, estimated as the ratio of mortality to incidence, was 55% (Parkin 2005).

1.3.1 Geographical variations in survival from cervical cancer

In 2002, the five-year age-adjusted survival rate for cancer of the cervix was 61% (range 51% to 70%) among nations of the developed world, compared with 41% (range 21% to 58%) among developing nations, where cases are more likely to present at an advanced stage (Parkin 2005).

For cervical cancers diagnosed in 2000-2001 in the UK, one-year survival rates were 83%, age-standardised to the European population; this fell to 66% after five years, but decreased only slightly, to 62%, by ten years (CRUK website, August 2008). Five-year relative survival decreased with increasing age, with women diagnosed at age 80 years or over having a relative survival rate which was a quarter that of women diagnosed before the age of 40 years.

1.3.2 Temporal trends in survival from cervical cancer

In the UK, there has been a modest but steady increase in survival from cervical cancer since the 1970s (CRUK website, August 2008). For cancers diagnosed in 1971-75, the ten-year age-standardised survival rate was 46%, but this had improved to 62% by 2000-2001.

Summary

Cervical cancer is common throughout the world, but is becoming less so over time, with many countries experiencing a dramatic decline in incidence: the incidence rate of cervical cancer is highest in women over the age of 50 years in most countries. The fall in incidence is not seen in all histological types: the incidence of squamous cell cervical carcinoma is falling at the same time as that of adenocarcinoma and adenosquamous carcinoma is rising. Worldwide, cervical cancer is still a leading cause of cancer-death, although in developed countries with well-established screening programmes, such as the UK, this is no longer the case. Survival is poor in undeveloped countries, but has improved over time in the UK.

Chapter 2

THE AETIOLOGY AND PATHOGENESIS OF CERVICAL NEOPLASIA

Orientation

In this chapter I:

- describe the anatomy of the cervix
- describe the natural history of cervical neoplasia
- describe risk factors for the acquisition and progression of epithelial abnormalities of the cervic, and how these vary according to type of disease

2.1 ANATOMY OF THE CERVIX

The cervix is a continuation of the uterus, the upper two-thirds of which is a distinct site in terms of diagnostic, staging and treatment criteria (Souhami 2005). The cervix is the lower, narrower third of the uterus (figure 2.1). It is cylindrical, or conical, in shape, but size and shape vary widely with age, hormonal state, and parity. Under appropriate examination conditions, approximately half of its length is visible, typically 3 centimetres long by 2.5 centimetres wide, with the remainder lying above the vagina, beyond view (Singer 1989).

The upper limit of the cervix is an anatomically and histologically ill-defined junction; it is considered to be the *internal os*, the opening where the uterus narrows to meet

the cervix. At its opposite end, the cervix meets the vagina; the portion protruding into the vagina is called the *portio vaginalis*. The *external os* is the lowermost opening between the cervix and vagina: before childbearing, it is a small, circular opening at the center of the cervix. The portion of the cervix exterior to the *external os* is the *ectocervix*. The passageway between the *external os* and the endometrial cavity, with upper limit the *internal os*, is referred to as the *endocervical canal*.

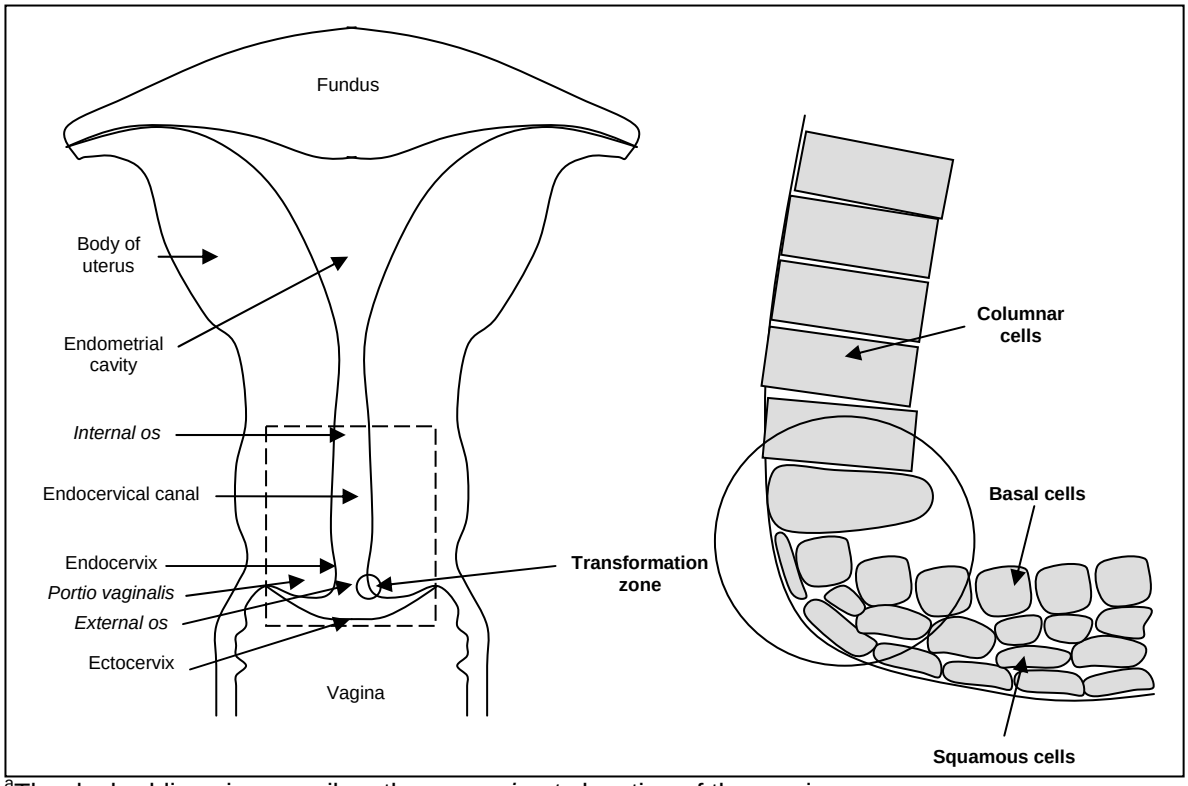
The nature of the epithelial lining of the cervix varies according to location, with both columnar and squamous epithelia present at different locations within the cervix. Most of the ectocervix consists of stratified squamous epithelium similar to that found in the vagina, i.e. layers of flattened cells which protect against the relatively harsh environment of the vagina. The parts of the cervix closer to, or contiguous with, the uterus are subject to a more mild environment: the epithelia here consists of simple columnar epithelium, i.e. a single layer of rectangular column-like cells, which secrete mucus. This “glandular” epithelium covers a varying portion of the *ectocervix*, as well as lining the *endocervical canal*. The epithelium in the *endocervical canal* has folds and invaginations that make up the *endocervical glands* (they are not true glands). These complex topographical features make the cytological screening, and the colposcopic examination of endocervical tissues technically more difficult, and less reliable, than for the smoother and more accessible squamous epithelium of the ectocervix (these folds are not illustrated in figure 2.1).

2.1.1 The transformation zone

The junction between the squamous epithelium and the glandular epithelium is called the “new” squamocolumnar junction (SCJ), which is marked by a line of metaplasia. The exact location of the SCJ varies, with age and hormonal status being the most important factors which influence its location. It is adjacent to the new SCJ that the process of squamous metaplasia occurs throughout the reproductive years: this is a normal process, during which columnar epithelium is replaced by squamous epithelium. The border between the metaplastic epithelium arising during the reproductive years, and the original squamous epithelium, is called the “original SCJ.” The transformation zone (TZ) is the area of metaplastic epithelium between the original and new SCJs.

It is within the TZ that over 90% of all cervical neoplasia arises. Metaplasia, which is particularly active during puberty and during a womans first pregnancy, increases the risk of abnormal changes, since there is a rapid turnover of new cells. The metaplastic epithelium adjacent to the new SCJ is the newest and the least mature squamous epithelium on the cervix. As new metaplastic epithelium arises, the older metaplastic epithelium is moved outward toward the original SCJ, leaving the newest and least mature metaplasia adjacent to the new SCJ (Singer 1989).

Figure 2.1. Anatomy of the cervix[a].



[a]The dashed-line circumscribes the *approximate* location of the cervix.

2.2 THE NATURAL HISTORY OF CERVICAL CANCER

The natural history of cervical cancer is thought to be characterized by a well-defined, and what was once believed to be a lengthy, pre-malignant phase. Preinvasive lesions are asymptomatic, although other coexisting conditions may produce incidental symptoms (Souhami 2005). During the preinvasive phase, the normal maturation of epithelial cells within the TZ becomes disrupted, resulting in dysplastic lesions (Franco 2002). Dysplasia refers to lesions which have morphological changes giving them the characteristics of malignancy (*atypia*), but without involvement of the full thickness of the epithelium by basal-type neoplastic cells (Miller 1999). The majority of low-grade dysplastic lesions will regress without

intervention, whereas others may become high-grade by expanding to cover the full thickness of the cervical epithelium. Thereafter, in a significant proportion of, but again not in all, women with high-grade dysplastic lesions, the lesion may progress through the microinvasive to the invasive stage by invading the basement membrane separating the epithelium from the underlying connective tissue. Once it has become invasive cancer, the tumour may subsequently reach blood and lymphatic vessels, resulting in metastatic tumours in local pelvic lymph nodes, and eventually in distant lymph nodes and other organs (Franco 2002).

2.2.1 Precursors of cervical cancer

The epithelial abnormalities of the cervix are assumed to form a continuous spectrum within one disease, an assumption recognised by the adoption of the term *cervical intraepithelial neoplasia* (CIN) to provide a unifying description of pre-malignant cervical abnormalities. According to this paradigm, cervical disease is a disease in which each degree of abnormality merges imperceptibly into the next. Differences between the cells and the structure of the epithelium in one part of the spectrum compared with another are thus quantitative, rather than qualitative, with no well-defined boundary between them (Singer 1989). The finding that cervical cancer and its precursors have similar risk factors, with the strength of association being lower for precursors than for cancer, lends credibility to this model of the natural history (Franco 2002). But clearly, progression along the spectrum is not inevitable.

The terminology used to describe premalignant lesions of the cervix has changed over time (see table 2.2.1). Adoption of terminology based on a rigid subdivision of an essentially continuous underlying disease process may appear counterintuitive, but is common practice in pathology (Singer 1989). Unfortunately, the manner in which the terminology has changed over time does not necessarily reflect an improvement in the understanding of the natural history of these lesions. There has been a tendency to combine categories in order to reduce interobserver variability (Miller 1999). The Bethesda system, for example, does not facilitate the study of natural history: the term “lesion” was adopted in preference to “neoplasia” since many of these entities are as likely to regress as they are to progress. Although there is some correspondence between the terminology used for cytological findings and histological diagnoses, these are imperfect. The Papanicolau categories, for example, have been shown not to have strict histological counterparts, and are no longer recommended for cytological reporting.

In the cohort study described in this thesis, cervical smears were reported using the WHO system, and the results of histological assessments were reported using the original CIN terminology. Using these systems, the implication is that a smear containing mild dyskaryosis, for example, indicates an underlying histological abnormality of at worst CIN1, moderate dyskaryosis indicates CIN2, etc. Note also that the use of the term “koilocytic atypia” in the original CIN terminology refers to *cytopathic* changes resulting from HPV infection, which can be recognised under a microscope: these are sometimes reported as a *histological* diagnosis of “HPV”, as they were in the study described in this thesis. There is a clearly a straightforward

conversion between the original and the modified CIN terminology (the modified CIN terminology is used throughout this thesis).

Table 2.2.1. Correspondence among reporting terminologies for cervical cytology and pathology reports[a].

System	Papanicolaou class system	Dysplasia terminology (WHO system)	SIL terminology (Bethesda system)	Original CIN terminology	Modified CIN terminology
Criteria	Cytology (originally)	Cytology	Cytology	Histology	Histology
Year of publication	1954	1953	1989	1968	1990
Categories	I	Normal	Within normal limits	Normal	Normal
	II	Atypia (multiple qualifiers)	Benign cellular changes (infection or repair)		
	II	Atypia (epithelial cell abnormalities)	ASCUS/AGCUS with qualifier[b]		
	II or III		LSIL	Koilocytotic atypia, flat condyloma, without epithelial changes	Low grade CIN
	III	Mild dysplasia/ dyskaryosis		CIN grade 1	
	III or IV	Moderate dysplasia/ dyskaryosis	HSIL	CIN grade 2	High-grade CIN
	IV	Severe dysplasia/ dyskaryosis		CIN grade 3	
	IV or V	Carcinoma in situ		CIN grade 3	
	V	Invasive carcinoma	Invasive carcinoma	Invasive carcinoma	Invasive carcinoma

[a]Miller 1999, Nanda 2000, Franco 2002; [b]whether a reactive, or a pre-malignant, or a malignant process is favoured.

2.2.1.1 Cervical intraepithelial neoplasia

Cervical intraepithelial neoplasia (CIN), regardless of grade, is asymptomatic and not visible, unaided, on examination (Hoffman 1996). According to the CIN histological classification system (original, or modified), lesions are graded primarily according to

the proportion of the epithelium which is occupied by basaloid, undifferentiated cells (figure 2.2.1.1) (Kiviat 1992):

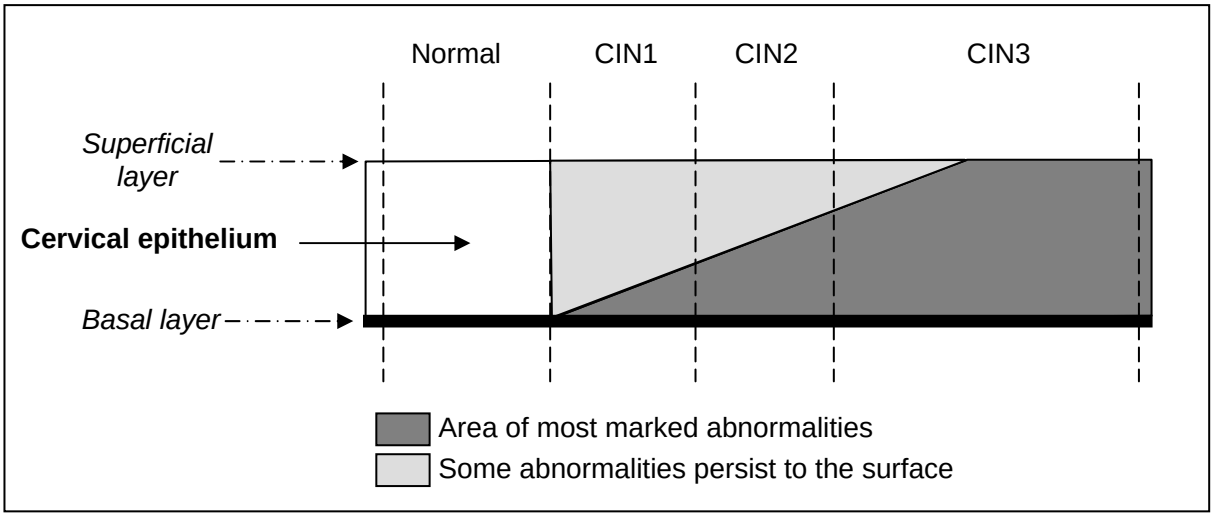
CIN1: CIN grade 1. The upper two thirds of the epithelium usually shows good differentiation, although some nuclear abnormalities persist up to the surface (if this were not the case, a corresponding cytological diagnosis would not be possible). Abnormalities of the nuclei are mild and are most marked in the basal third of the epithelium. Mitotic figures (*mitotic figure: the microscopic appearance of a cell undergoing mitosis; a cell of which the chromosomes are visible by the light microscope*) are present but not numerous; they are confined to the basal third of the epithelium, and abnormal forms are not expected.

CIN2: CIN grade 2. The upper half of the epithelium shows differentiation and maturation, again with nuclear abnormalities (*atypia*) persisting to the surface. Abnormalities of the nuclei are more marked than they are in CIN1, and more abnormal nuclei are found higher in the epithelium than in CIN1. Mitotic figures, which may be abnormal, are present in the basal two-thirds of the epithelium.

CIN3: CIN grade 3. Maturation is confined to the superficial third of the epithelium, or is completely absent. Nuclear abnormalities are marked throughout the whole thickness of the epithelium. Mitotic figures are present throughout the epithelium and may be numerous, with many abnormal configurations.

Determining whether CIN is present in a cervical biopsy is difficult since there are benign and physiological changes which can be mistaken for CIN. Having determined that CIN is present, deciding the degree of CIN is also difficult due to the number of features which have to be considered. Inevitably, there is a degree of subjectivity involved in the interpretation of these features, which vary in a quantitative rather than in a qualitative way (Singer 1989). Reproducibility is therefore a legitimate concern.

Figure 2.2.1.1 Schematic representation of the CIN histological classification system.



2.2.1.2 Cervical glandular intraepithelial neoplasia

The pre-invasive stage of adenocarcinoma of the cervix is believed to comprise a continuous spectrum of one disease, similar to that of squamous cell carcinoma. By analogy with CIN, the terminology adopted for this disease includes adenocarcinoma

in situ, cervical glandular atypia and cervical glandular intraepithelial neoplasia (CGIN) (Kumar 2000). These lesions arise from endocervical epithelium, but are much less common than CIN, and consequently much less is known about their natural history. There does, however, appear to be a definite association with the development of invasive adenocarcinoma (Hoffman 1996, Miller 1999).

In the cohort study described in this thesis, no cases of CGIN or worse were identified, and so this disease will not be considered further.

2.2.2 Acquisition of cervical abnormalities

Cervical cytological abnormalities are very common, in particular in young women (Winer 2005). In the cohort study described in this thesis, among 1,075 women who were HPV DNA negative and cytologically abnormal at study entry, the cumulative risk at three years of any cytological abnormality was 28% (95% CI 25% to 32%). And of the 246 of these women who had at least one cytological abnormality, 93 had at least one further abnormality after having an intervening normal smear: 80 women had two “episodes” of cytological abnormality, 12 had three, and one had four (Woodman 2001).

Risk factors for the acquisition of CIN are essentially identical to those for the acquisition of squamous cell carcinoma of the cervix, but with the strength of many associations being attenuated (Hoffman 1996). Risk factors for high-grade CIN (CIN2 and CIN3) may differ from those of low-grade CIN (CIN1 and lesser

histological abnormalities). For example, in a case-control study of women under the age of 40 years who were diagnosed with CIN, risk factors for CIN3 were found to be similar to those for invasive disease; CIN1 cases were similar to the controls; and women with CIN2 were intermediate between the two. Multiple sexual partners and age at first childbirth were strong risk factors for CIN2 and CIN3, as was smoking; however, smoking had only a weak association with CIN1 (Cuzick 1990).

The single most important risk factor for cervical neoplasia of all grades is infection with certain types of cervical human papillomavirus (HPV) infection. Cervical HPV is a very common sexually transmitted infection, which accounts for the prominence of sexual behaviour characteristics in the known risk-factors for cervical neoplasia. The epidemiological evidence which lends support to a role for cervical infection with HPV in the aetiology of cervical neoplasia is now indisputable. Initially, this was largely based on studies with a case-control design: these consistently revealed a strong association between cervical neoplasia and the detection of HPV DNA in samples of exfoliated cervical cells taken at, or subsequent to, diagnosis of disease (Munoz 1992). Evidence from these cross-sectional studies has now been superseded by that obtained from longitudinal studies, which have again revealed strong associations between exposure to infection with cervical HPV and the risk of the acquisition, and progression, of cervical disease (Woodman 2002). Cervical HPV infection is discussed in greater detail in chapter 3.

2.2.3 Progression and regression of epithelial abnormalities of the cervix

Determining an accurate estimate of rates of progression and regression for epithelial abnormalities of the cervix is not straightforward. In developed countries, in particular those with screening programmes for cervical cancer, all high-grade lesions which are detected are now treated, as are many low-grade lesions (see section 4.2.3). In general, therefore, the opportunity to observe the uninterrupted natural history of cervical disease is no longer possible. It is, however, clear that the vast majority of cervical lesions are transient, and are never destined to progress to invasive disease (Holowaty 1999). In the cohort study described in this thesis, the first episode of cervical cytological abnormality lasted 9 months (interquartile range 6 to 14 months) before regressing to normal; episodes of cytological abnormality lasted longer as the severity of the initial smear in the episode of cytological abnormality increased (Woodman 2001). High-grade lesions have a much greater probability of progressing to invasive disease than do low-grade lesions, although the majority of these lesions also almost certainly regress (Holowaty 1999). One cohort study including 17,000 women which, unlike many previous studies, used an appropriate time-to-event analysis to estimate risks of progression and regression (see chapter 6), found that both mild and moderate dysplasia were more likely to regress than to progress, and that the risk of progression increased with increasing severity of dysplasia. For example, within five years, 25% of women with moderate dysplasia, a cytological diagnosis corresponding to a presumed histological diagnosis of CIN2, *progressed* to severe dysplasia or worse, compared to 63% who *regressed* within the same time period (Holowaty 1999). This report also cites six

other studies, with follow-up ranging from 24 to 78 months, all of which report progression for women with moderate dysplasia/CIN2. The risk of disease progression in these studies ranged from 11% to 53%.

Time to progression from preclinical cervical disease to invasive cervical cancer has been estimated to be of the order of several years, if not decades: the age at which the maximum incidence of CIN occurs is around 25 to 34 years, at least two decades before the age at which the maximum incidence of cervical cancer occurs (Bosch 2003). Therefore, not only do the majority of preclinical lesions regress without treatment, even those which progress to invasive cancer only do so after an interval of many years (Miller 1999).

2.2.4 Invasive disease

The term cervical cancer includes a number of different diseases. The most common invasive lesion is squamous cell carcinoma, comprising up to 85% of cases; adenocarcinomas account for a further 5% to 10%; the remaining rare lesions include adenosquamous lesions, adenoacanthomas, small-cell cancers, and sarcomas (Souhami 2005).

Age at first sexual intercourse, number of sexual partners, use of oral contraceptives, high-parity, socioeconomic status, and smoking have all been implicated as risk factors for invasive cervical cancer (Castellsague 2003, International collaboration of epidemiological studies of cervical cancer (ICESCC)

2006a, 2006c and 2007a). However, the importance of these risk factors may vary according to age. In a two-to-one matched case-control study of 121 British women diagnosed with cervical cancer under the age of 40 years, only aspects of sexual behaviour were found to be important risk factors, and not other cofactors often identified in studies of older women. Age at first sexual intercourse and lifetime number of sexual partners were found to be the most important risk factors: weak associations with smoking and fruit consumption were “explained” by sexual behaviour; parity, educational attainment, and social class, were not significant risk factors (Cuzick 1996).

2.2.5 Squamous cell neoplasia

Squamous cell neoplasia includes the most common precancerous and cancerous conditions of the cervix, from the lowest grade of CIN, to advanced invasive squamous cell cervical cancer (SCCC). Risk factors for SCCC include: younger age at first sexual intercourse; increasing number of sexual partners; increasing parity; younger age at first full-term pregnancy; increasing duration of oral contraceptive use; a history of sexually transmitted diseases; low socioeconomic status; and smoking cigarettes (Miller 1999, Green 2003, Castellsague 2003, ICESCC 2006a and 2006b).

2.2.6 Adenocarcinoma of the cervix

The majority of adenocarcinomas of the cervix arise from the endocervical epithelium, frequently within the endocervical canal (Souhami 2005). Precursors of this cancer are therefore more likely to be missed by cervical cytology tests (see section 2.1). These cancers also tend to produce fewer early symptoms, and are therefore more likely to be diagnosed at a later stage, than SCCC. Risk factors for adenocarcinoma include: younger age at first sexual intercourse; increasing number of sexual partners; increasing parity; younger age at first full-term pregnancy; and oral contraceptive use; smoking does not appear to be a risk factor for adenocarcinoma (Green 2003, Castellsague 2006, ICESCC 2006a, 2006c and 2007a).

Chapter 3

CERVICAL HUMAN PAPILLOMAVIRUS INFECTION

Orientation

In this chapter I:

- describe the human papillomavirus
- describe the natural history of cervical human papillomavirus infections
- describe the role of cervical human papillomavirus in the acquisition of cervical abnormalities

3.1 CERVICAL HUMAN PAPILLOMAVIRUS

Cervical human papillomavirus (HPV) infection is a very common *sexually transmitted* infection. Infection with high-risk types of HPV is now considered to be a *necessary*, but not *sufficient*, cause of cervical cancer (Munoz 2000). HPV is a necessary cause since almost 100% of cervical cancers are positive for HPV DNA (Walboomers 1999, Parkin 2006). However, it is clearly not a sufficient cause because, while the majority of women appear to acquire cervical HPV infection soon after the onset of sexual activity (Collins 2002, Winer 2008), very few subsequently develop cervical cancer. Aetiological factors, additional to cervical HPV infection, are therefore required.

3.1.1 Viruses

Viruses are submicroscopic obligate intracellular parasites: they are parasites which can have no mode of existence other than moving from one host-cell to the next (Cann 2001). Viruses lack the genetic information which encodes the apparatus necessary for the generation of metabolic energy, or for protein synthesis. They are therefore absolutely dependent upon the host-cell for these functions. Viruses have no ulterior motive: they are “designed” merely for the effective transmission of a nucleic acid genome from one host-cell to another. Once inside a host-cell, they co-opt the cellular machinery in order to replicate themselves for transmission to a new host-cell.

Viruses vary enormously in size, shape, and the number, and nature, of the molecules from which they are constructed (Flint 2004). There is thus no “typical” virus, or virus structure. However, in general, viruses comprise an RNA- or DNA-genome enclosed in an outer protective “coat” of proteins, called the capsid. The capsid may itself consist of sub-units called capsomeres. On exit from a cell, some viruses also acquire an envelope surrounding the capsid, formed from structural components of the host-cell.

3.1.2 The human papillomavirus

Human papillomaviruses (HPV) are epitheliotropic viruses of the family Papillomaviridae (De Villiers 2004), with DNA genomes and non-enveloped

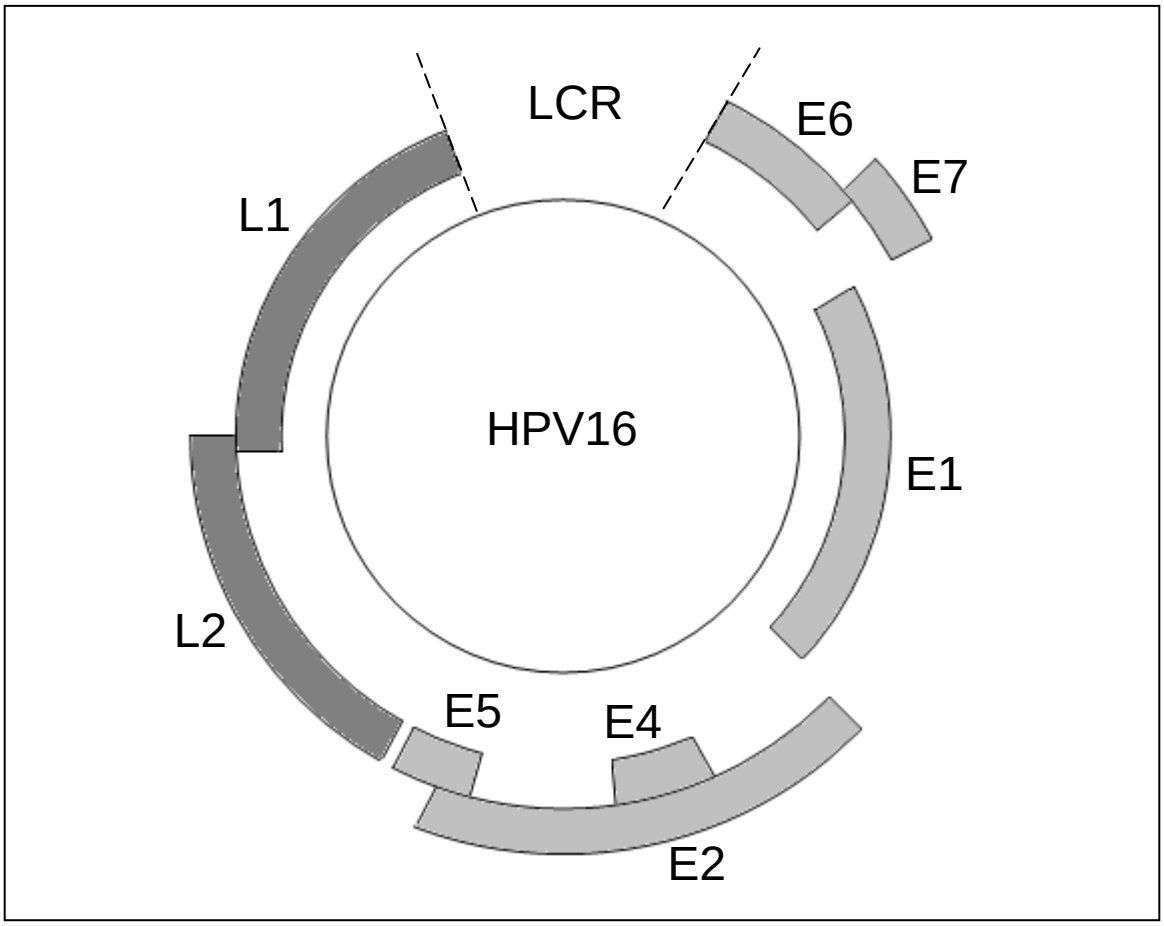
icosahedral capsids (Severson 2001, Doorbar 2006). Over 120 genotypes have now been identified, of which at least 118 have been fully sequenced, and approximately 40 types are associated with lesions of the anogenital tract (which includes, but is not restricted to, the cervix) (Peitsaro 2002, De Villiers 2004).

3.1.2.1 Organisation of the HPV genome

The HPV genome consists of double-stranded, circular, DNA, the size of which varies according to HPV type, but they are all typically around 8,000 base-pairs in length (Mahdavi 2005). For example, the total HPV16 genome is 7,904 basepairs in length (PubMed accession number NC_001526), compared with 7,857 for the HPV18 genome (PubMed accession number NC_001357). The HPV genome codes for eight proteins, E1, E2, E4, E5, E6, E7 (the *early* proteins), L1 and L2 (the *late* proteins) (Mahdavi 2005); these genes overlap in the genome. A non-coding region of about 1,000 base pairs, called the *upstream regulatory region* or the *long control region*, separates the early and late gene clusters (Severson 2001) (figure 3.1.2.1). The viral genome does not encode a unique DNA polymerase, which is required for DNA replication, and therefore the virus must depend upon the replication machinery of the host-cell for the generation of viral progeny (Alani 1998).

Comment: With respect to gene nomenclature, the apparently missing E3 gene is present only in a few papillomaviruses which infect animals; similarly there are E8, L3, and L4 genes, which are encoded by only a few papillomavirus types.

Figure 3.1.2.1. Schematic representation of the HPV16 genome. E1 to E7 are the early genes, L1 and L2 the late genes, and LCR the long control region (or the *upstream regulatory region*).



3.1.2.1.1 Late genes

The Late (L) genes encode the viral capsid proteins: the L1 and L2 genes code for the *major* and *minor* viral capsid proteins, respectively. The HPV capsid comprises 360 copies of the L1 protein, organized into 72 pentavalent capsomeres, with one copy of L2 at the centre of the pentavalent capsomeres, at the virion vertices (Doorbar 2006). The L1 and L2 genes are not expressed in precancerous or malignant cells, but these structural proteins are important in immunological settings,

such as vaccine development (Zur Hausen 2002), and in neutralizing antibody assays (see sections 16.6, 16.7 and chapter 17).

3.1.2.1.2 Early genes

The early (E) genes function primarily in episomal replication.

E1 is a regulatory protein, which plays a pivotal role in initiating viral-DNA replication and regulation of viral-gene expression, and is thus essential for permissive infection (Severson 2001).

E2 is a regulatory protein, which codes for a protein responsible for viral replication and transcription (Mahdavi 2005).

E4 appears to have a role in disrupting the cytoskeleton integrity of the host-cell, and thus in helping the assembly, maturation, and release, of the virus (Doorbar 2006). Despite its name, E4 is expressed *late* in the virus-replicating cycle (Severson 2001).

E5 has been implicated in cellular transformation (Severson 2001, Mahdavi 2005): it induces some transformed phenotypes in established cells, and can increase the proliferative capacity of human keratinocytes (Flint 2004).

E6 is an oncogene. The E6 gene of high-risk HPV types encodes a transforming protein which is capable of immortalizing epithelial cells and is thought to play a role

in the initiation of the oncogenic process (Flint 2004). The protein products of this gene interfere with the normal function of the p53 tumour suppressor gene: HPV E6 binds with p53, leading to its dysfunction, and thereby impairing its ability to block the cell-cycle when DNA errors occur. E6 also protects the cell from apoptosis (Mahdavi 2005).

E7 is an oncogene. The E7 gene of high-risk HPV types encodes a transforming protein capable of immortalizing epithelial cells, and is thought to play a role in the initiation of the oncogenic process. The protein-product of this gene binds to retinoblastoma protein (pRb), and activates genes which start the cell-cycle, leading to tissue proliferation (Mahdavi 2005). High-risk HPV E7 proteins bind to pRb with much greater efficiency than do low-risk HPV E7 proteins.

E1 and E2 are negative regulators of E6 and E7 expression: viral integration into the host-cell genome typically occurs within the viral E1 or E2 genomes, thereby disrupting these genes, and leading to the loss of repression of E6 and E7 transcription (Severson 2001). E6 and E7 are consistently found to be expressed in HPV-associated cervical cancers. Although continued expression of E6 and E7 is not sufficient for malignant conversion of squamous cells, continued expression of these genes appears to be necessary for maintaining the transformed phenotype. It has been shown that only the E6 and E7 genes of high-risk HPV types are able to immortalize human cells in tissue culture (Zur Hausen 2002).

3.1.2.2 Classification of HPV types

Genotyping of HPV, or the classification of HPV into types, is based on the DNA sequences of the L1, E6, and E7 genes (Mahdavi 2005), analysed by either hybridization techniques, or by direct sequence comparison: a difference of 10% or more in sequence with respect to previously established strains, is sufficient to define a new type (Mahdavi 2005); subtypes differ by between 2% and 10%; and variants differ by less than 2%. For example, HPV16 and HPV18 share only 50% homology at the nucleotide level.

HPV types differ according to their tissue affinity and pathogenicity. Therefore, they can also be classified into subtypes which infect stratified squamous or mucosal epithelia; or more typically, and of more direct relevance to this thesis, according to their observed association with cervical cancer in epidemiological studies. Oncogenic, or *high-risk*, types are those types most frequently detected at the time of diagnosis of cervical cancer, for example HPV16 and HPV18; non-oncogenic, or *low-risk*, types, for example HPV6 and HPV11, are less likely to be detected in association with cervical cancer (Mahdavi 2005). Intermediate-, or *medium-risk*, types are also sometimes separately distinguished.

3.1.2.3 The HPV life-cycle

All viruses have life-cycles which, for convenience of description, can be divided into distinct phases: attachment, penetration, uncoating, expression, assembly, maturation, and release (Cann 2001).

The papillomavirus life-cycle differs from that of all other virus families: infection requires the availability of epidermal or mucosal epithelial cells which are still able to proliferate, i.e. basal-layer cells (Zur Hausen 2002). All known HPVs are exclusively epitheliotropic; they do not infect or express their genes in the underlying dermis (Mahdavi 2005).

Basal cells in the cervical squamous epithelium rest on the basement membrane, which is supported by the dermis. HPV is thought to access the basal cells through micro-abrasions in the cervical epithelium (figure 2.1), since only undifferentiated, basal, replicating keratinocytes are infected by HPV (Woodman 2007). A viral receptor which mediates the attachment of HPV virions to the host-cell has yet to be definitively identified (Alani 1998). Penetration into a host-cell of attached virions is a slow process, with a half-life of hours, and occurs *via* endocytosis (Mahdavi 2005). Once inside the cell, uncoating occurs, and the viral genome is released from the virion, allowing viral DNA to be transported into the nucleus, where the cellular DNA replication machinery is located. All but one of the early HPV genes, E1, E2, E5, E6 and E7, are then expressed, and the viral DNA replicates from episomal DNA. It is thought that the virus maintains its genome as a low-copy-number episome in the

basal cells of the epithelium (typically quoted as 10 to 200 copies per cell, or 50 to 100); the contribution of E6 and E7 to basal cell proliferation during *in vivo* infection is currently uncertain, and it has been suggested that expression of E1 (and possibly also E2) may be sufficient for the maintenance of viral episomes in basal cells (Doorbar 2005).

Normal squamous human epithelia grow as stratified layers, where only cells in the basal-layers are able to actively divide (Hebner 2006). After division of infected basal-cells, some progeny remain in the basal-layer as infected cells; others detach and move away from the basement membrane towards the suprabasal-layers of the epithelium, and undergo cellular differentiation to become epithelial cells (Mahdavi 2005). In the upper layers of the epithelium (the midzone and the superficial zone) the viral genome is replicated further, increasing to several thousand copies per cell (Wilson 2005), and the late genes L1 and L2, and the "early" gene E4, are expressed (Woodman 2007). Progeny virions are assembled in the nucleus when L1 and L2 proteins encapsidate the viral genomes, which are then free to mature.

Virus replication and assembly is tightly linked to the differentiation of epithelial cells (Mahdavi 2005): infectious virions are produced only in the terminally-differentiated cell (i.e. cells which will not differentiate further). The virus-laden squamous cells are then shed, from which the virus is released to initiate a new infection.

3.1.3 The acquisition of cervical HPV infection

There is compelling evidence to suggest that a *cervical* HPV infection, as distinct from HPV infections at more superficial sites which may follow vertical *or* horizontal transmission, is only acquired as a result of penetrative vaginal sexual intercourse (Ley 1991, Andersson-Ellstrom 1996a, Koch 1997). The male sexual partner is thus the *vector* of HPV infection. The majority of women appear to acquire cervical HPV infection soon after the onset of sexual activity. In an analysis of 242 women from the cohort study described in this thesis who were recruited within six months of first sexual intercourse, and who were censored on acquiring a second sexual partner, 46% were found to be HPV-positive within three-years of first sexual intercourse (Collins 2002). In a repeat of this analysis in 130 university students aged 18 to 22 years recruited within three months of first sexual intercourse, 50% were found to be HPV-positive within three years of first sexual intercourse (Winer 2008).

Risk factors for the acquisition of cervical HPV infection are similar to those for the acquisition of cervical cancer and its precursors, as might be anticipated since HPV infection is in the causal pathway for these conditions. Reported risk factors have included various aspects of a woman's sexual behaviour, in particular: age at first sexual intercourse; lifetime number of sexual partners; recent number of sexual partners; history of other sexually transmitted diseases; and oral contraceptive use; other factors included the sexual behaviour of a woman's male sexual partner(s), and cigarette smoking, dietary factors, and human immunodeficiency virus status (Burk 1996, Moscicki 2001, Schiffman 2003, Baseman 2005, Vaccarella 2006a,

Vaccarella 2006b). There is evidence that risk factors differ according to oncogenic potential: the association between low-risk HPV types and sexual behaviour is not as pronounced as that with high-risk HPV types (Kjaer 1997, STD-2000-Richardson, Vaccarella 2006b). Reports vary, but use of condoms *may* be protective against the acquisition of HPV infection: however, it is probably only *consistent use* which provides protection (Winer 2006).

The prevalence of cervical HPV infection varies with age. For example, in 3,305 cytologically normal cervical smears taken from women aged between 15 and 69 years in the general population of Amsterdam, the Netherlands, prevalence was highest (20%) in women aged 25 to 29 years, declining rapidly thereafter (Jacobs 2000). Other studies have reported similar results (Burk 1996a, Castle 2006).

Longitudinal, cross-sectional, and case-control studies have shown that HPV16 is the most common type of HPV infection, followed by HPV18 (Woodman 2001, Winer 2003).

3.1.4 Persistence of HPV infection

HPV infections can be categorised into latent, subclinical, and clinical, infections. Latent HPV infections can only be detected by virological means, and cause no clinical signs. Subclinical HPV infections can be detected by pathology techniques, e.g. cytopathological changes seen (under the microscope) in cervical smears, or by colposcopically-directed biopsy and histological studies, but *not* by unaided

examination. Clinical HPV infections are defined as those which result in visible lesions (by the naked eye), or which display clinical symptoms, for example genital warts (De Sanjose 1992).

Most young women with cervical HPV infection appear to resolve their infection spontaneously in a very short time: in the cohort described in this thesis, median duration of an incident cervical HPV infection of any type was 13.7 months (inter-quartile range 8.0 to 25.4), with some differences among specific types (Woodman 2001). However, many viruses establish persistent infections which are characterized by continuous low- or high-levels of viral replication (for example the human immunodeficiency virus (HIV) and the hepatitis B virus) or by periodic reactivation of a latent infection following apparently disease-free intervals (for example the herpes simplex virus (HSV)) (Woodman 2007). Although it is now widely believed that a persistent infection with a high-risk HPV type is necessary for the development of high-grade cervical intraepithelial neoplasia (CIN) or invasive disease, the term “persistence” has often been loosely defined when testing this hypothesis. A fundamental problem relates to inferences drawn from observations made at indeterminate points during the natural history of the infection. In these circumstances, the distinction between a persistent and a transient infection is arbitrary to the extent that it is dependent on both the timing of the samples in relation to the natural history of the infection, and the time interval between the samples: the shorter the interval between tests, the more likely it is that an infection will be defined as persistent.

It is now clear that epithelial abnormalities of the cervix can be evident shortly after the first detection of HPV DNA in cervical samples (Woodman 2001). This is not to deny that HPV can establish a persistent infection, nor that a persistent cervical HPV infection is necessary for the development of invasive cervical disease. However, it remains to be determined whether persistent infections are characterized by the continuing detection of HPV, or by a state of latency during which the virus remains undetectable, only to reappear later. A mechanism for latency has not yet been established, nor is it clear whether the differences between a latent and an active cervical infection are *qualitative* or *quantitative*. However, a woman cannot be defined as having a persistent infection in any meaningful virological sense just because she tests positive for HPV on two occasions, some months apart, and therefore she should not, on the basis of this evidence, be considered to be at a high risk of cervical cancer. Nor can a woman who tests positive for a specific HPV type necessarily be considered to have cleared her infection when she first tests negative for that type. A clearer understanding of these issues is essential for the effective implementation of screening strategies which include HPV testing.

3.1.5 HPV-mediated progression from pre-cancerous lesions to cervical cancer

There is now overwhelming epidemiological and laboratory evidence linking cervical HPV infection with the acquisition of cervical cytological abnormalities, and of cervical cancer. Low-risk HPV types, such as types 6 and 11, cause benign genital warts; high-risk HPV types, such as types 16 and 18, are associated with the development of high-grade cervical abnormalities, and of cervical cancer. HPV16 is

the type most frequently detected at the time of diagnosis of squamous-cell cancer, followed by HPV18; HPV18 is the type most strongly associated with adenocarcinoma of the cervix (Woodman 2003). HPV16 is the type most commonly linked with cancer, since it is present in 50% of cervical cancers and high-grade CIN (and in 25% of low-grade CIN) (Koutsky 2002), and is estimated to account for approximately 60% of cervical cancers, with HPV18 adding another 10% (Harper 2004).

Not all viral genomes which are transported to the nucleus of the host-cell remain as episomal DNA. HPV can be found in cervical material in episomal forms, integrated forms (in which the viral genome becomes inserted into the host-cell genome), or in mixed forms which contain both. Although integration into the host-cell genome is effectively the end of the life-cycle for a virus, this change of *physical-state* from a "free" to a "captured" piece of DNA has implications for the malignant potential of the virus. Low-grade intraepithelial lesions support productive viral replication and shedding; only a fraction of these lesions will progress to high-grade cervical intraepithelial neoplasia. The integration of the HPV genome into the host-cell chromosomes is associated with the progression of untreated lesions. Viral integration into the host-cell genome often occurs downstream of the E6 and E7 genes, in the E1 or E2 region of the virus genome (both the host cell and viral genomes are affected by integration); disruption of these coding regions results in a loss of negative-feedback control of E6 and E7 oncogene expression by the viral regulatory E1 or E2 proteins, which can no longer be manufactured by the cell, since their DNA sequence has been disrupted. Oncogene (i.e. E6 and E7) expression is

thus free to proceed unchecked. The prevalence in exfoliated cervical cells, or in cervical tissue, of episomal or integrated forms of HPV, or both, varies with the severity of disease, the infecting HPV type, and the method used to determine the physical-state of the virus (Woodman 2007). Development of invasive cancer requires additional genetic events, facilitated by E6- and E7-mediated inactivation of the genome guardians p53 and pRb, genomic instability, and suppression of apoptosis (Mahdavi 2005).

Chapter 4

THE PREVENTION OF CERVICAL CANCER

Orientation

In this chapter I:

- define primary- and secondary-prevention
- describe strategies for the secondary-prevention of cervical cancer, and attempts to improve the efficiency of these strategies
- discuss primary-prevention strategies based on HPV vaccines

4.1 PRIMARY- AND SECONDARY-PREVENTION

Primary-prevention of a targeted disease refers to measures taken to prevent that disease occurring in susceptible individuals: for example vaccination against certain types of human papillomavirus in young women to prevent the occurrence of cervical cancers associated with those types. Successful primary prevention is potentially the most cost-effective form of health care, since it eliminates not only the disease itself but also the sequelae of that disease.

Secondary-prevention refers to measures taken to identify and treat *asymptomatic* subjects who already have preclinical disease, and thereby prevent the progression of disease in these already-diseased individuals. Secondary-prevention aims to

minimize the impact of disease. Successful secondary-prevention relies upon the targeted disease having a relatively long preclinical phase, and upon the availability of an effective treatment: there has to be an opportunity to identify the disease at an early stage, and the disease must be treatable, or the treatments must be capable of altering the natural history of the disease in an advantageous way.

4.2 SECONDARY-PREVENTION OF CERVICAL CANCER

A cervical cancer screening programme which uses cytological tests, or other methods, to identify women at high-risk of cervical cancer, is an example of a secondary-prevention initiative.

4.2.1 Screening tests for cervical cancer

In developing countries, where health facilities may be basic, *visual inspection* is, of necessity, one approach adopted for the detection of cervical cancer at an early stage. Direct visual inspection (or *cervicoscopy*) requires only a speculum, good light, and the use of dilute acetic acid. After swabbing the cervix with acetic acid, areas of dysplasia are stained white. *Aided* visual inspection, which is aimed at detecting cervical intraepithelial neoplasia grade 3 as well as early cancer, requires, in addition, a small magnifying instrument with which to view the acetic acid-stained cervix (Miller 1999, Wright 2003).

In developed countries, where health facilities are limited only by the allocation of resources, cervical cytological screening at regular intervals is the established method of screening for cervical cancer. The objective of screening in this case is primarily to detect precursors of cervical cancer, i.e. to identify disease which, if left untreated, would develop into cancer. The majority of cervical cancers, and therefore their precursors, arise in the transformation zone (see section 2.1.1). In cervical cytological screening, exfoliated cervical cells are collected using a device and sampling technique designed to sample cells from the entire transformation zone: a sample which does not contain endocervical cells, indicating that the entire transformation zone was not sampled, is considered inadequate. Sampling devices include the Ayre's spatula, as used in the cohort study described in this thesis, cervical brushes, cervical swabs, etc. According to the original practice, the cytological material which is collected is then "smeared" evenly on to a glass slide (hence cervical *smear*), and a fixative applied. The smear is then transferred to a laboratory where it is examined under a microscope ("read") by a cytopathologist for indications of cytological abnormalities. To aid interpretation of the smear result, an adequate "history" must accompany the smear e.g. details of the woman's menstrual cycle, current hormonal treatments etc.

There is now an alternative to the conventional smear, liquid-based cytology. The sample is taken, as described above, using a brush device, which is then either placed into liquid preservative in its entirety, or the cells on the brush are rinsed into the preservative. The liquid sample is then sent to the laboratory for preparation of a

cellular monolayer for examination. Examination of samples is currently (September 2008) manual, but it may be possible to automate this in future.

No screening test is perfect, and cervical smears are no exception. Cervical cytology screening involves taking smears in a standardised way, transferring the cytological material sampled to another medium, storage of the sample, and interpretation of the smear. Clearly this procedure is subject to error at a number of stages. The percentage of false-negative smears varies from 15% to 50%, while the percentage of false-positives is approximately 10%. This means that the sensitivity of cytology for cervical cancer is quite low (50% to 85%), although the specificity is quite high (approximately 90%) (Miller 1999, Franco 2002). The lack of sensitivity of an individual test is compensated for by repeating the tests after a relatively short time interval (Rozendaal 1996). And even if a smear is taken, and reported, without error or misinterpretation, cervical smears have *detection* rather than *diagnostic* accuracy: once a smear has been reported as positive, further tests are required for confirmation of the presence of disease. Nevertheless, taking cervical smears is minimally-intrusive, presents little or no risk to the woman, and has proven to be acceptable to the majority of women in developed countries.

4.2.2 Criteria for effective cervical cancer screening programmes

The aim of a cervical cancer screening programme is to reduce mortality from cervical cancer by identifying and treating women with cervical cancer precursors before they develop invasive cancer, and thereby prevent death due to cervical

cancer. The availability of a suitable screening test is clearly a crucial requirement, but this alone will not guarantee the success of a cervical screening programme. The essential crtiteria which must be met by an organised cervical screening programme are as follows (Miller 1999):

- Individual women in the target population are identifiable.
- Measures are available to guarantee high-coverage and high-attendance, such as personal letters of invitation.
- Adequate facilities exist for obtaining cervical smears, and adequate laboratory facilities are available to examine them
- An organised quality control programme for obtaining and interpreting cervical smears is in place
- Adequate facilities exist for the diagnosis and appropriate treatment of any confirmed neoplastic lesions identified
- A carefully designed (and agreed upon) referral system is in place for the management of any abnormality found, and for providing information about normal screening tests.
- The evaluation and monitoring of the complete programme is organised effectively.

4.2.3 The NHS cervical screening programme

An example of a population-based cervical screening programme is the NHS cervical screening programme (NHSCSP) in the UK. *Ad hoc* cervical screening

began in the UK in the mid-1960s. However, it was hoped that the introduction of the National Cervical Screening Programme in 1988 would not only increase population coverage, but also reduce the over-screening of women who are at low-risk of cervical cancer.

The NHS call and recall system is a national computerised system which automatically invites women who are registered with a general practitioner (GP) for cervical screening. Approximately four million women are now screened each year, with coverage approaching nearly 80% of the target population (NHSCSP website, September 2008). Originally, all women aged between 20 and 64 years were invited for screening at time intervals of three to five years; there were regional variations in screening intervals. Currently (September 2008), all women aged between 25 and 64 years are invited for screening, with those under 50 routinely invited every three years and those over 50 every five years; women aged 65 years or over are also invited if they have not been screened since the age of 50, or if they have had recent abnormal cervical smears; and opportunistic screening of women who attend their GP, or a family planning clinic, for other reasons, is also encouraged, by means of incentive payments. Some other programmes invite women to attend for the first time once they become sexually active.

The NHS call-recall system keeps track of smear results and any follow-up of abnormal smears, and recalls women for screening at the appropriate interval. Women with normal smears remain on routine recall. Women with cervical smears containing borderline nuclear abnormalities (BNA) or mild dyskaryosis, are recalled

for repeat smears at intervals of six months. Women who have three consecutive smears containing BNA, or two consecutive smears containing mild dyskaryosis, are referred for colposcopic assessment; women who have any smear containing moderate or severe dyskaryosis are immediately referred for colposcopic assessment (see section 5.2.3 for how this differs from the protocol used in the cohort study described in this thesis). Women who have smears suspected of containing invasive cancer, or glandular neoplasia, are (urgently) referred to a gynaecological oncologist: these women are not considered further here. Women only return to routine recall once they have had three consecutive normal smears, taken at six-month intervals.

Colposcopic assessment is the standard method for evaluating the cervix following the detection of cytological abnormalities. A colposcope is a device for viewing the cervix under magnification: the cervix is swabbed with acetic acid to highlight abnormal areas, and biopsies of these areas are taken for histological assessment. In the NHSCSP, either a punch biopsy is taken or an excision biopsy, usually using a diathermy loop. As with cytology, for a satisfactory colposcopic assessment, the entire transformation zone must be visualised (Hoffman 1996).

Women can be treated upon request at their first colposcopic assessment: an excision biopsy aims to remove the abnormality completely. Otherwise, only women who have histologically-confirmed disease are offered treatment. The NHS guidelines suggest that women with CIN1 can either be treated immediately, or kept under close colposcopic surveillance; in contrast, women with CIN2 or CIN3 should

be treated immediately. Treatment is usually given under general, rather than local, anaesthetic. Once a woman has been treated for CIN, she is followed-up, either in clinic or in primary care, with at least three smears being taken in the five-year period following treatment.

Smears are taken using an appropriate sampling device and technique; as of August 2008, liquid-based cytology, with manual interpretation, is in use. Quality assurance procedures are in place at all appropriate stages of the screening programme, and the programme is subject to regular audit.

4.2.4 Effectiveness of cervical cancer screening programmes

Screening for cervical cancer was introduced, in the UK and elsewhere, without awaiting evidence from prospective randomised controlled trials; such trials will not now be undertaken. In lieu of such trials, the decreasing mortality rate from cervical cancer is taken as evidence that screening is effective.

One study used routine statistics to assess the impact of screening on the incidence of, and mortality from, cervical cancer in England following the introduction of the NHS call-recall system (Quinn 1999). This study found that the overall incidence of invasive disease remained stable up to the end of the 1980s (figure 1.1.2 suggests similar results for Great Britain), but thereafter fell continuously. Mortality from cervical cancer in England has been decreasing since 1950 (figure 1.2.2 suggests similar results for the UK since 1970 at least), but the authors claim that this

accelerated at the end of the 1980s. Other studies using similar methods have reached similar conclusions regarding the efficacy of cervical screening, both in the UK (Peto 2004) and in young women in European countries (Levi 2000). The position of cervical screening programmes now seems unassailable.

4.2.5 Efficiency of cervical cancer screening programmes

Cervical cancer screening programmes may be effective in reducing mortality from cervical cancer, but they do so at the expense of substantial over-diagnosis and unnecessary treatment. Although the treatment of pre-malignant changes in the cervix is therapeutically efficacious, it is also procedurally inefficient. This situation has arisen because of uncertainties surrounding the natural history of CIN.

Cytological and histological examinations cannot reliably distinguish the few women with abnormal smears who will progress to invasive cancer, from the vast majority of those whose abnormalities will spontaneously regress (see section 2.2.3). Therefore it is necessary to treat all abnormalities which exceed a certain severity. For example, in Japan 35 million cervical smears are performed annually, with evidence of cervical cancer revealed in approximately 2,000 women (Tanaka 2004). In the US, for each new case of invasive cervical cancer found by using cytological smears, there are approximately 50 other cases of abnormal smears consistent with precursor lesions; to this must be added twice as many equivocal or borderline abnormalities (Franco 2002). Even in developing countries, where 83% of cases

occur and where cervical cancer accounts for 15% of all female cancers, the risk of cervical cancer before the age of 65 years is only 1.5% (Parkin 2005).

The identification and treatment of cytological abnormalities of uncertain malignant potential is an expensive undertaking. For example, in England the cost of cervical screening, including the cost of treating cervical abnormalities, is estimated to be approximately £157 million a year (NHSCSP website, September 2008). Presumably, in Japan, where ten times as many smears are taken annually, the cost of the programme is in excess of £1 billion. Improving the efficiency of their cytological screening programmes should therefore be a priority for any health system.

4.2.6 Improving the efficiency of cervical cancer screening programmes

Infection with high-risk types of the human papillomavirus (HPV) is the most important risk factor for cervical cancer (see chapter 3). Therefore, it has been suggested that adding testing for the presence of cervical infection with HPV to cervical screening programmes can be used to improve the efficiency of these programmes. If HPV testing can be used to identify women at low-risk of acquiring cervical cytological abnormalities, or to identify those women with abnormalities who are at low-risk of progression to high-grade CIN, or worse, then efficiency gains can be realised by screening these women less frequently.

Several randomised controlled trials have now reported on the effectiveness of either adding HPV testing to conventional cervical cytology screening, or replacing conventional screening by HPV testing, outright. A trial undertaken in Sweden has reported that in women aged 32 to 38 years, the addition of HPV testing, using a PCR-enzyme immunoassay detecting 14 high-risk types, to conventional cytology, screening significantly reduced the incidence of CIN2 and CIN3 compared to conventional cytology alone (Naucler 2007). A similar trial in the Netherlands has reported that in women aged 29 to 56 years, the addition of the same HPV test as used in the previous trial to conventional cytology screening significantly reduced the incidence of CIN3, or worse, compared to conventional cytology alone (Bulkmans 2007). A trial undertaken in Canada in women aged 30 to 69 years found that the hybrid capture 2 HPV test, used alone, had greater sensitivity for detecting CIN2 and CIN3 than conventional cytology (Mayrand 2007); a similar conclusion was reached by another trial of this HPV test undertaken in Italy in women in the same age-range (Ronco 2008). A trial investigating this issue is currently underway in the UK (the ARTISTIC trial), and is due to report in 2009 (NHSCSP website, September 2008). Of course none of these trials address the prevention of the primary end-point of cervical screening programmes, i.e. mortality from cervical cancer.

Other triage strategies relying upon HPV testing include identifying women with “persistent” infections with high-risk types; measuring HPV viral load; and the detection of integrated forms of HPV. The difficulties with the concept of persistent infections have been described earlier (section 3.1.4); the difficulties with using measurements of HPV viral load will be described subsequently (see section 23.7). It

has been proposed that the identification of integrated forms of HPV could be a useful biomarker for progressive disease. There are several problems with this proposal. First, the identification of small numbers of integrated forms of HPV in a background of mainly episomal forms will always be a technical challenge when only exfoliated cells are available for analysis. Second, if integrated genomes are often transcriptionally silent, or become so shortly after integration, then their detection might have limited prognostic usefulness (see section 3.1.5 for a discussion of the physical state of HPV).

4.3 PRIMARY-PREVENTION OF CERVICAL CANCER

We now appear to be entering the era of the primary-prevention of cervical cancer, beginning with the approximately 70% of cervical cancers associated with the two most common high-risk HPV types, type 16 and type 18.

Two prophylactic vaccines targeting HPV have now been developed. GARDASIL® is a quadrivalent vaccine against HPV16, HPV18, HPV6, and HPV11, marketed by Merck and Company; Cervarix® is a bivalent vaccine against HPV16 and HPV18, marketed by GlaxoSmithKline PLC. These vaccines have been evaluated in several multi-centre, randomised, double-blind, placebo-controlled, clinical trials (Villa 2005, Mao 2006, FUTURE II study group 2007a, Garland 2007, Paavonen 2007): data from the cohort reported in this thesis were used to inform the sample size calculation for some of these trials.

The vast majority of cervical HPV infections will not be followed by a diagnosis of cervical cancer, and for those which are, the time interval between the acquisition of the initial infection and the diagnosis of invasive disease is of the order of decades (see section 2.2.3). Therefore, the phase III trials evaluating these vaccines have used intermediate virological- and disease-outcomes as surrogates for estimating the impact of the vaccines on the incidence and mortality of invasive cervical cancer. One trial evaluating the quadrivalent vaccine used co-primary composite disease outcome measures: the incidence of genital warts, vulvar or vaginal intraepithelial neoplasia, or cancer; and the incidence of cervical intraepithelial neoplasia, adenocarcinoma *in situ*, or cancer associated with HPV types 6, 11, 16 and/or 18 (Garland 2007). Another used a single primary composite disease outcome measure: the incidence of CIN2, CIN3, adenocarcinoma *in situ*, or invasive carcinoma of the cervix, associated with the detection of HPV16 or HPV18 DNA. The trial evaluating the bivalent vaccine used the incidence of CIN2 or worse associated with HPV16 and/or HPV18 as its primary composite disease outcome measure (Paavonen 2007).

All three phase III trials have reported a significant reduction in the incidence of their primary outcome(s) in the vaccinated arm compared to the placebo arm, and both vaccines *appear* to offer almost 100% protection against high-grade cervical intraepithelial neoplasia (HGCIN), or worse, associated with HPV16 and HPV18.

These exciting results have persuaded many of the need for a population-based vaccination programme, including the UK Department of Health: as of September

2008, a national programme of HPV vaccination has been introduced in the UK (NHSCSP website, September 2008). Initially, the Cervarix® vaccine, given as three injections over six months, is being offered to girls aged 12 to 13 years, i.e. *prior* to the presumed onset of sexual activity. This will be followed by a catch-up programme which, by the year 2011, will result in vaccination having been offered to all girls up to 18 years of age. Women over the age of 18 years, i.e. *after* the presumed age of onset of sexual activity, will not be offered vaccination.

However, there are issues surrounding the vaccines which remain unresolved. A full cost-benefit analysis of a HPV vaccination programme must consider the impact of vaccination on the *overall* incidence of disease in the population, and not just that of disease associated with vaccine types. For example, a combined analysis of four randomised controlled trials (two phase II, and two phase III), reported a 44% reduction in the overall incidence of HGCIN following vaccination with the quadrivalent vaccine (Future II study group 2007b) (see table 4.3). This falls far short of the 100% protection popularly perceived to be offered by the vaccine, but this is not surprising, given that this analysis included women who were already seropositive to, or HPV DNA-positive for, vaccine types, or both, at study entry. Unfortunately, neither the combined analysis of the quadrivalent vaccine, nor the interim analysis of the phase III trial of the Cervarix® vaccine, reported the reduction in the overall incidence of HGCIN in women who were naïve for both types targeted by the vaccine, at study entry. However, one of the trials contributing to the combined analysis *does* separately report results for women who resemble those who are likely to be targeted for vaccination (Future II study group 2007a). This trial

found only a 27% reduction in the overall incidence of HGCIN among 9,396 women who tested negative for HPV16 and HPV18 at enrolment. This finding, which has received little attention, is critical to any cost-benefit analysis.

If the overall reduction in the incidence of HGCIN is less than that predicted, then the benefits of a vaccination programme will also be less than anticipated. The apparent shortfall in the reduction in the incidence of HGCIN may have arisen because of an overestimate of the contribution of vaccine types to the overall incidence of HGCIN: the attribution of causality to a specific HPV type will always be problematic when contemporaneous, or sequential, infection with more than one HPV type is common, and more than one HPV type precedes the diagnosis of the outcome (Woodman 2001, Gravitt 2007).

However, a plausible alternative explanation has been offered for the possible continuing increase in the overall incidence of HGCIN among vaccinated women (Sawaya 2007). The disappointing performance of the vaccine is consistent with the possibility that other oncogenic HPV types have “filled the biological niche” left after the elimination of HPV16 and HPV18. It is therefore now of considerable interest to know which HPV types continue to cause HGCIN in women vaccinated against HPV16 and HPV18. This information, which surely is readily available, will be critical when considering prospects for second-generation vaccines, but has yet to be published.

Table 4.3. Protective efficacy of Gardasil® against HPV types 6, 11, 16, and/or 18-associated high-grade CIN (HGCIN), and HGCIN associated with any HPV type, in "all comers", and in those naïve for one or more types targeted by the vaccine.

Population	Reduction in the incidence of HGCIN[1] associated with HPV types 6, 11, 16, and/or 18 (95% CI)	Reduction in incidence of all HGCIN[1] (95% CI)
Naïve	99% (93 to 100)[2]	27% (4 to 44)[3]
All comers	44% (31 to 55)[2]	18% (7 to 29)[2]

[1]High-grade CIN; [2]FUTURE II study group 2007b; [3]FUTURE II study group 2007a

There are still many issues to be resolved before a primary-prevention programme based on HPV vaccination can replace, outright, cervical cytology screening (Woodman 2007). Therefore screening will remain the main method of preventing cervical cancer for the foreseeable future (Miller 1999).

Chapter 5

STUDY DESIGN

Orientation

In this chapter I describe:

- the study design, study population, and study procedures, for the cohort study which forms the basis for this thesis
- the collection of socio-demographic and behavioural data from study subjects
- the collection and assessment of study samples
- data considerations

5.1 STUDY AIM

When it was designed, the aim of the study described in this thesis was to determine the natural history and aetiology of early cervical neoplasia, in young women.

5.2 STUDY DESIGN

This was a prospective cohort study, with follow-up until the diagnosis of high-grade cervical intraepithelial neoplasia (CIN).

5.2.1 Setting

A single family planning clinic (a Brook Advisory Centre) in Birmingham, UK.

5.2.2 Study population

Between March 1988 and September 1992, 2,011 women aged between 15 and 19 years, who were likely to be available for long-term follow-up because of their continuing need for contraceptive supplies, were recruited by research nurses.

5.2.3 Study protocol

The study protocol is illustrated in figure 5.2.3a, and a schematic representation of the study design is illustrated in figure 5.2.3b. All women found to have cervical cytological abnormality, irrespective of the degree of severity of that abnormality, were immediately referred to a dedicated research clinic for colposcopic assessment, where a sample of colposcopically abnormal epithelium was removed for histological examination. Colposcopic and cytological surveillance was maintained in these women, and treatment was postponed, until there was histological evidence of high-grade CIN (CIN2 or CIN3), at which point women were treated and then left the study (Woodman 2001).

5.2.3.1 Follow-up

Follow-up was six-monthly. Women asked to return to the study at intervals of six months, until the end of follow-up in August 1997. If they did not return for a follow-up visit, they were sent a letter reminding them.

5.2.3.1.1 Follow-up in the absence of cervical cytological abnormality

Women who were found to be free of cytological abnormality were asked to return six months later. Women found to have an inadequate smear were invited to return as soon as possible for a repeat smear.

5.2.3.1.2 Follow-up in the presence of cervical cytological abnormality

All women found to have cervical cytological abnormality, irrespective of degree of severity, were immediately referred to a dedicated research clinic for colposcopic examination. The first appointment at the colposcopy clinic was made as soon as possible after the first detection of abnormality, often within one month, and certainly within less than six months. Once a woman had acquired a cytological abnormality, she remained on colposcopic and cytological surveillance every six months until there was no evidence of abnormality, defined as three consecutive negative colposcopic assessments *and* three consecutive negative smears. The woman then returned to cytological surveillance every six months. If another cervical cytological abnormality was subsequently detected, the same procedure applied.

5.2.3.1.3 End of follow-up

Women could request treatment for histologically-confirmed abnormality at any time. However, the study protocol called for treatment to be postponed until there was histological evidence of CIN3, at which point women were recalled for treatment (there was of course a delay between taking the biopsy and receipt of the associated histological report), and thereafter left the study. During the initial period of the study, this was the procedure followed; however, subsequently, all cases of high-grade CIN were treated, i.e. CIN2 as well as CIN3. At the completion of the study, all women with evidence of histological abnormality were invited to return for treatment.

5.2.3.1.4 Treatment

Women were treated by either loop diathermy or laser cone biopsy. In loop diathermy, a thin, shaped, piece of wire (the loop) is heated and then used to cut through abnormal tissue, allowing it to be removed in one piece. Laser cone biopsy uses a laser to remove tissue.

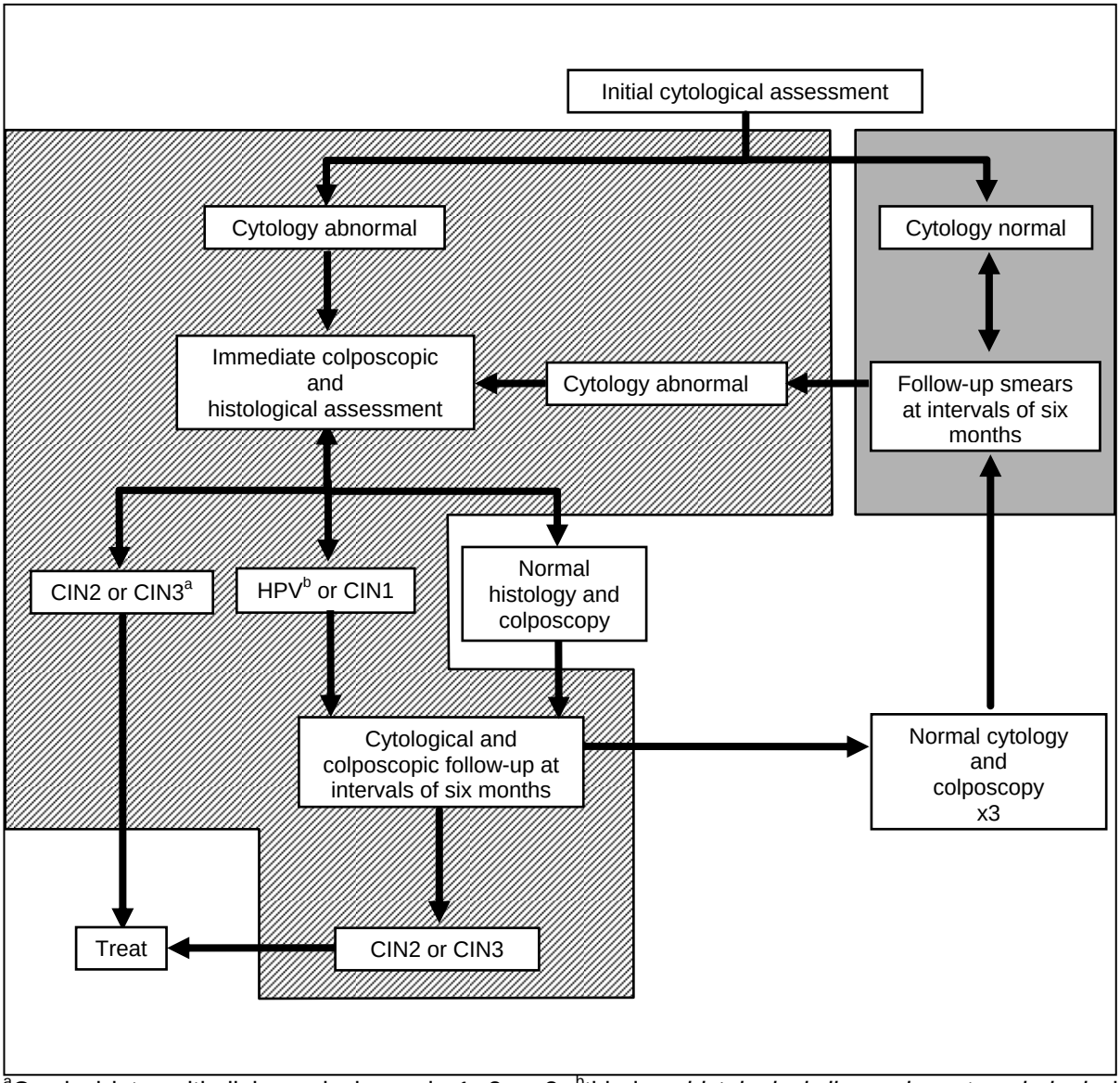
5.2.4 Ethical considerations

The study protocol was approved by the appropriate ethical committee.

5.2.4 Consent

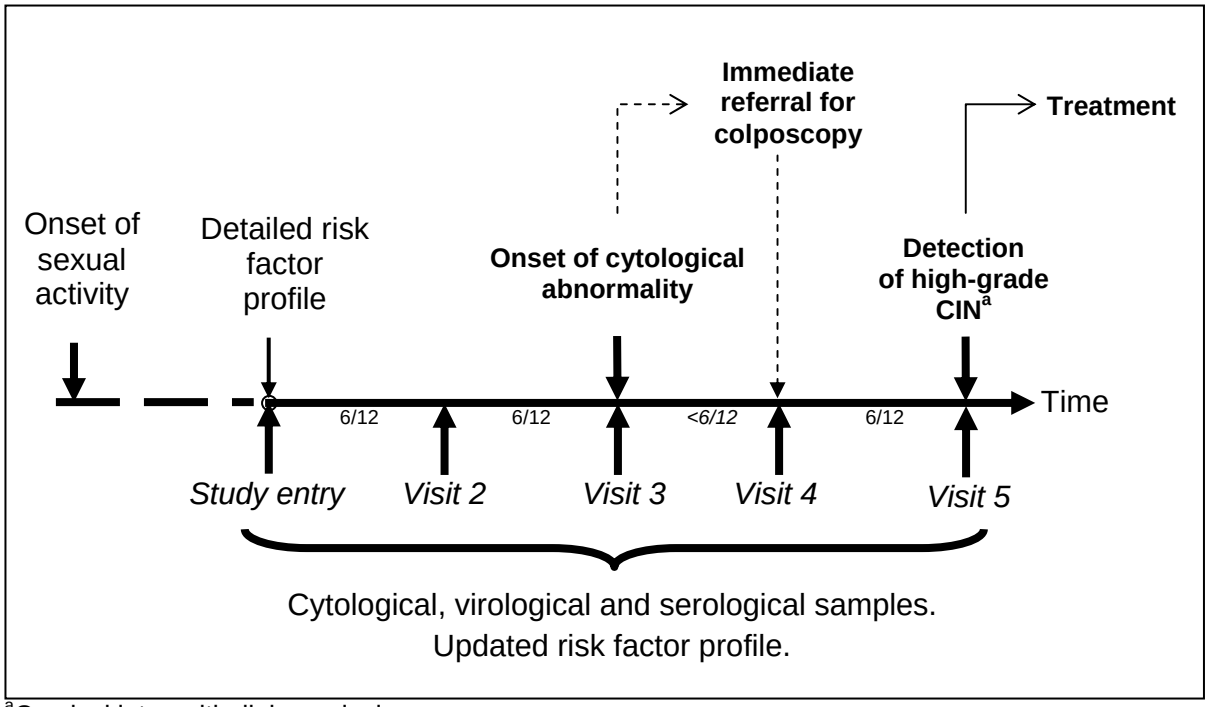
Informed oral consent was obtained from all women.

Figure 5.2.3a. The study protocol.



[a]Cervical intraepithelial neoplasia grade 1, 2 or 3; [b]this is a *histological diagnosis*, not a *virological* diagnosis, of cervical human papillomavirus (see section 2.2.1).

Figure 5.2.3b. Schematic representation of the study design.



[a]Cervical intraepithelial neoplasia

5.3 COLLECTION OF DEMOGRAPHIC AND BEHAVIOURAL DATA

On recruitment, all women were interviewed in-person: two medically-qualified and experienced female researchers conducted the majority of interviews during consecutive periods of the study (P Blomfield, H Winter), supplemented by two female research nurses: a structured questionnaire (appendix 1) was completed by the interviewer, and a detailed social, sexual, and behavioural risk factor profile at study entry was assembled (initially women who were 20 years old were also intended to be recruited: the wording on page five was unaltered when this intention was abandoned). At each subsequent visit to the study, women were re-interviewed, and changes in risk factors were collected using a second structured questionnaire

(appendix 2) (however, in practice, the interviewer didn't always take the first questionnaire to the interview).

On rare occasions, women were not interviewed for the first time contemporaneously with the taking of their first cervical smear; for example their visit to the Brook clinic may have coincided with menstruation, or they may not have been psychologically prepared to have a smear, etc. If a woman had a smear which was reported to be inadequate, she was invited back to the study for a repeat smear as soon as possible. Usually, this was within four months of the inadequate smear; when she returned to the study for a repeat smear she was not necessarily re-interviewed: this meant that the occurrence of an inadequate smear resulted in an extended period between data-collection interviews (the usual, or designed, six months, plus the time interval between the date of the inadequate smear and the date of the repeat smear).

5.4 SAMPLE COLLECTION

5.4.1 Cytological samples

At each visit, two cervical cytological samples were taken using the same wooden Ayre's spatula. The first sample was used to prepare a cervical smear for immediate cytological evaluation and reporting: a 360^{o} rotation of the transformation zone was performed, then the material collected was applied to a slide. The second sample was stored for future virological examination: a 360^{o} rotation of the transformation zone

was performed using the same spatula as was used to take the first sample, then the end of the spatula was broken off and placed into a sterile sample of phosphate-buffered saline. The solution was refrigerated both before and after use, and then transferred to the laboratory as soon after the sample was taken as feasible. During the final 12 months of the study there was some concern regarding reports that the use of a wooden spatula could affect either the nature, or the detection, of viral DNA, and so cotton-tipped orange sticks were used for taking the second (virological) sample. The sampling technique was identical to that used with the spatula; the first sample (the one used as a cervical smear for immediate reporting) was still obtained using the Ayre's spatula. In the laboratory, cells from the virological sample were pelleted and stored at -40^{o}C for future testing.

5.4.2 Blood samples

At each visit, women were requested to provide a blood sample. A sample of approximately 10ml of peripheral blood was taken by venipuncture. Serum was collected from the blood sample by centrifugation and stored at -40^{o}C for future analysis.

5.4.3 Histological samples

If there was colposcopic evidence of abnormal epithelium, a punch biopsy was removed for immediate histological examination; a second histological specimen was then taken from an immediately adjacent area of abnormal epithelium, and

stored for future histological studies. Whilst on colposcopic follow-up, further histological specimens were only taken when there was considered to be evidence of the progression of disease.

5.5 CLINICAL ASSESSMENT

Two researchers, in consecutive periods, were responsible for interviewing the majority of study subjects, taking cervical cytological samples, and performing colposcopic assessments (including the taking of histological samples) (P Blomfield, H Winter).

5.5.1 Cervical cytological assessments

All cervical cytological material was examined and reported by one cytologist (M Yates) using standard report forms for the NHS cervical screening programme, and employing the WHO dysplasia reporting system (see section 2.2.1).

5.5.2 Colposcopic assessments

Two researchers, in consecutive periods, performed all colposcopic assessments (P Blomfield, H Winter).

5.5.3 Histological assessments

All histological material was examined by one pathologist (TP Rollason) and reported using the original CIN terminology (see section 2.2.1).

5.5.4 Virological assessments

Two methods for detecting HPV DNA were used for testing cervical samples from the cohort study described in this thesis: one was used to test all available samples, the other to test only a specific subset.

5.5.4.1 HPV testing using GP5+/GP6+ general primer-mediated PCR

After *all* clinical follow-up had ended, all cervical (virological) samples were tested for the presence of HPV DNA using a polymerase chain reaction (PCR) (see section 21.1.1 for a description of PCR). Testing was carried out by a single laboratory scientist (J Selby) at the Department of Clinical Virology, Central Manchester Healthcare Trust, Manchester, during the period 1998 to 1999.

First, the presence of human DNA was determined by performing a PCR using primers for β-globin; this assay was discontinued when 97.8% of the first 2,997 samples tested were found to be β-globin positive (Woodman 2001). Thereafter, it was assumed that all samples contained human DNA and therefore that a HPV

DNA-negative result would only arise as a result of the sample in fact being HPV DNA-negative.

Cervical samples were tested for the presence of HPV DNA using general primer (GP5+/GP6+) mediated PCR. Amplified product was detected by electrophoresis through a 2% agarose gel, followed by ethidium bromide staining, and visualisation under ultraviolet illumination. All gels were then subject to Southern-blot transfer followed by hybridisation (under low-stringency conditions) with a digoxygenin labelled generic HPV probe. Samples found to be HPV-positive after ethidium bromide staining were then subject to further PCR tests using type-specific primers for “two” HPV types considered low-risk (6/11) [these types could not be further distinguished] and six HPV types considered high-risk (16, 18, 31, 33, 52 and 58) because of their association with cervical neoplasia.

A further presumptive typing of certain HPV-positive samples found to be non-reactive with the type-specific PCR tests was achieved, either by restriction-enzyme analysis of PCR product obtained after DNA amplification using the degenerate HPV primers (MY09/MY11), or by sequencing the GP5+/GP6+ PCR product. Repeat Southern-blot analysis using type-specific digoxygenin-labelled probes for HPV types 16, 18, and 33, was carried out on those samples found to be HPV-negative following ethidium bromide staining, but HPV-positive after hybridisation of the PCR product with a digoxygenin-labelled generic probe. The use of additional tests resulted in 60 samples from 40 women being assigned a numeric type (i.e. one of

the types specifically tested for: 6/11, 16, 18, 31, 33, 52, 58) not identified by type-specific PCR.

5.5.4.2 Measurement of HPV viral load

As described in the previous section, *all* cervical (virological) samples were tested for the presence of HPV DNA using a GP5+/GP6+ general primer-mediated PCR. However, a subset of samples was also tested for HPV16, or HPV18, viral load, or both. The method for measuring viral load is described in detail in section 21.2. The measurement of viral load was carried out by a single laboratory scientist (C Constandinou-Williams) at the Institute for Cancer Studies, University of Birmingham, Birmingham, during the period 2005 to 2006.

5.5.5 Immunological assessments

Sera samples from a subset of women were tested for their neutralizing antibody response to HPV16 and HPV18. The method for measuring the neutralizing antibody response is described in detail in chapter 17. Testing was carried out by a single laboratory scientist (K Wen) at the Institute for Cancer Studies, University of Birmingham, Birmingham, during 2006.

5.6 DATA CONSIDERATIONS

5.6.1 Database construction

The questionnaires (appendices 1 and 2) were designed to directly code information as it was collected. Questionnaires were entered onto two databases, one each for the retrospective and prospective questionnaires, with data entered by a limited (but unknown to the candidate) number of research staff. Construction and completion of these databases occurred during a period prior to the candidate's association with this study. However the candidate has been extensively involved in the editing of existing databases, and the design, collation, and management, of new databases.

The results from HPV testing using GP5+/GP6+ general primer-mediated PCR were entered into a Microsoft Access database, constructed and maintained by the laboratory technician who undertook the testing, with additional entries made by her supervisor (A Bailey).

The raw results from the testing of HPV16 and HPV18 viral load were entered into several Microsoft Excel spreadsheets, one for each assay, by the laboratory scientist who undertook the assay. These were then processed and combined into a single dataset in Microsoft Excel, by the candidate.

The raw results from testing for the neutralizing antibody response were output into a Microsoft Excel spreadsheet, one for each virus type in each assay, by the

laboratory technician who undertook the assay. These were then processed by the candidate and combined into a single dataset, also in Microsoft Excel.

The final study database combining these five separate data sources is maintained in the software system *Cache* by the candidate, and interrogated using the programming language *M* (*Cache ObjectScript*).

5.6.2 Data editing

Given the extensive nature of the data collected by the questionnaires, and the various types of samples collected, a considerable amount of data editing was required for the various databases ultimately constructed. With the exception of the data cleaning which occurred during the initial quality checks undertaken for the questionnaire databases, for the most part data editing was conducted either by, in conjunction with, or under the supervision of, the candidate. In particular, all variables and outcomes used in analyses presented in this thesis were scrutinised for internal consistency by the candidate, as well as consistency with the hardcopy questionnaires.

5.7 PUBLICATIONS

Data from this cohort study have now been published in ten reports, including four based on the work described in this thesis (Blomfield 1998, Woodman 2001, Collins 2002, Woodman 2003, Collins 2005, Woodman 2005, Steele 2008, Collins 2009,

Constandinou-Williams 2009, Collins 2010). These reports are described in table 5.7. The candidate was a co-author on all of these papers, except the first.

Table 5.7. Publications based on the cohort study described in this thesis.

Publication	Largest Population Described	Description
Blomfield 1998	1,219	Analysis to determine whether women at risk of cervical abnormality can be identified. Opportunistic analysis.
Woodman 2001	2,011	Analysis of the natural history of cervical HPV[a] infection in young women. Main report on the cohort study described in this thesis.
Collins 2002	242	Describes the incidence of cervical HPV infection in women during their first sexual relationship. Opportunistic analysis.
Woodman 2003	1,075	Investigation of the possibility that HPV18 is associated with rapidly progressing CIN[b]. Opportunistic analysis.
Collins 2005	474	Analysis of the association between the proximity of first sexual intercourse to menarche and the risk of cervical HPV infection. Opportunistic analysis. Opportunistic analysis. Based on work described in this thesis.
Woodman 2005	1,023	Investigation into the role of sexual behaviour in the acquisition of asymptomatic Epstein-Barr virus infection. Opportunistic analysis.
Steele 2008	42	Describes the measurement of the humoral immune response to HPV16 and HPV18 infections following an incident cervical HPV16 or HPV18 infection, using a pseudovirion-based neutralizing antibody assay. Subject of a new grant application based on the original cohort. Based on work described in this thesis.
Collins 2009	66	Analysis of the integration of cervical HPV16 and HPV18 infections. Subject of a new grant application based on the original cohort.
Constandinou-Williams 2009	60	Describes the measurement of HPV16 and HPV18 viral load following an incident cervical HPV16 or HPV18 infection using quantitative (real-time) PCR. Subject of a new grant application based on the original cohort. Based on work described in this thesis.
Collins 2010	1,485	Analysis of the association between cigarette smoking and the incidence of cervical HPV infection, the duration of those infections, and the incidence of high-grade CIN. Addresses a secondary objective of the original cohort study. Based on work described in this thesis.

[a]Human papillomavirus; [b]cervical intraepithelial neoplasia.

Summary

When it was designed, the aim of the study described in this thesis was to determine the natural history and aetiology of early cervical neoplasia, in young women. This was a prospective cohort study. 2,011 women aged 15 to 19 years were recruited from a single Brook Advisory Centre (a family planning clinic) in Birmingham, United Kingdom, between 1988 and 1992, and asked to re-attend at intervals of six months: follow-up ended on 31st August 1997. At recruitment**,** a standardised interview questionnaire was used to construct a detailed social, sexual, and behavioural risk-factor profile, including smoking. Risk factor profiles were updated at each follow-up visit using a second standardised interview questionnaire. Women were asked to provide a blood sample at each visit. At each visit two cervical cytological samples were taken using the same Ayres spatula: the first was used to prepare a smear for immediate cytological evaluation; the second was placed into 10 ml of phosphate-buffered saline and stored at -80°C for subsequent virological examination. All women in whom a cytological abnormality was identified were immediately referred to a dedicated research clinic for histological examination, irrespective of the severity of that abnormality. Also at each visit, women were asked to provide a serum sample. Colposcopic and cytological surveillance were maintained in these women and treatment was postponed until there was histological evidence of high-grade cervical intraepithelial neoplasia (CIN: CIN2 or CIN3), at which point women left the study. The study was approved by the appropriate research ethics committee, and informed oral consent was obtained from all women. After all clinical follow-up had ended, cervical samples were tested for the presence of HPV DNA using a general primer (GP5+/GP6+) mediated polymerase chain reaction (PCR) and further PCR tests were done with type-specific primers on samples which were HPV-positive. Data from this cohort study have now been published in 10 reports.

Chapter 6

STATISTICAL AND METHODOLOGICAL CONSIDERATIONS

Orientation

In this chapter I:

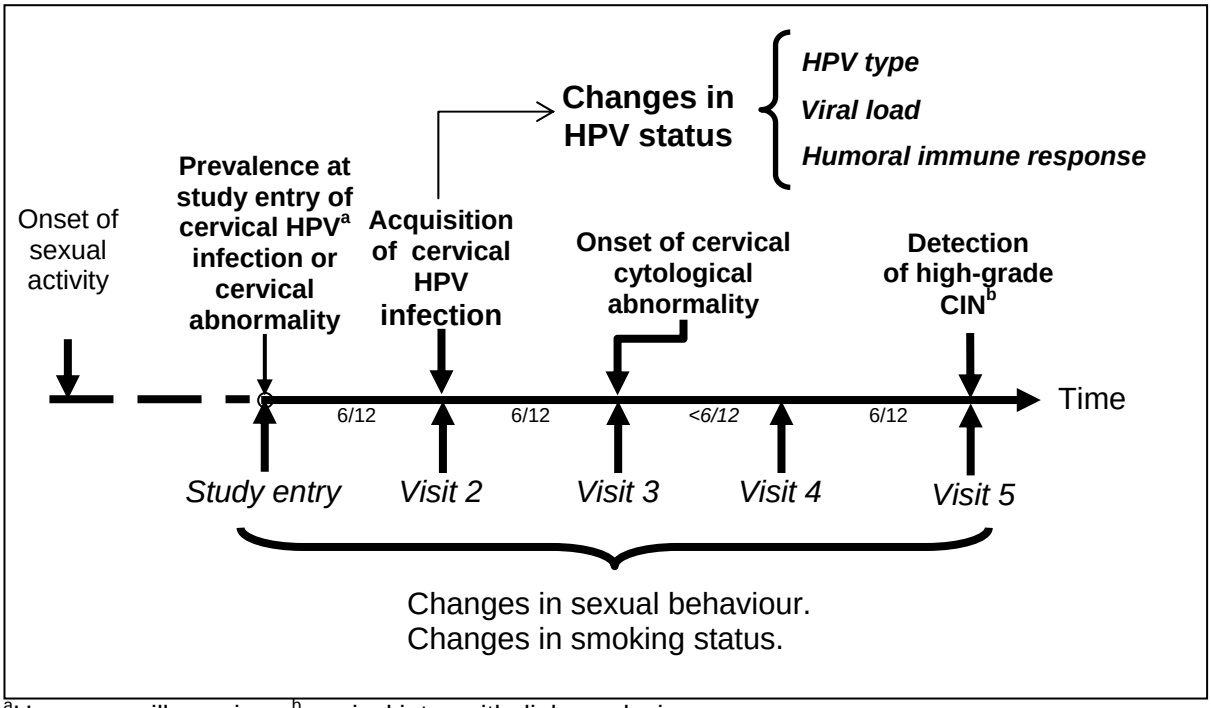
- describe the statistical nature of the data collected in the study described in this thesis
- describe time-to-event data, and interval-censoring
- discuss the process used to select methods for the analysis of the study described in this thesis
- describe the methods used to analyse interval-censored time-to-event data in this thesis
- discuss issues relating to time-varying covariates which are common to many analyses in this thesis
- discuss issues relating to missing data which are common to many analyses in this thesis

6.1 STATISTICAL CHALLENGES OF THE ANALYSIS OF THE STUDY DESCRIBED IN THIS THESIS

When it was designed, the aim of the cohort study described in this thesis was to determine the natural history and aetiology of early cervical neoplasia, in young women. Repeated measurements on cytological, histological, virological, and

serological, status were made, as well as on several behavioural risk factors (figure 6.1).

Figure 6.1. Statistical challenges of the analysis of the study described in this thesis.



[a]Human papillomavirus; [b]cervical intraepithelial neoplasia

The data from the cohort study described in this thesis present the following challenges:

- Several possible outcomes

There were several possible outcomes for the cohort study described in this thesis: illustrative examples are given in table 6.1a. The prevalence at study entry of cervical HPV infection, and of cervical cytological abnormality, are just two examples

of cross-sectional outcomes, and clearly there are several possible others; but most of the outcomes in table 6.1a involve an element of follow-up. Some of the outcomes in table 6.1a are explanatory, or intermediate, for others. For example, cervical HPV infection may generate an immune response to that infection, which, in turn, may "explain" the subsequent clearance of that infection, or lead to changes in HPV viral load.

Table 6.1a. Examples of potential outcomes for the cohort study described in this thesis.

Outcome
Prevalence at study entry of cervical HPV infection, of any type, or of types 6/11, 16, 18, 31, 33, 52 or 58
Acquisition during follow-up of an incident cervical HPV infection, of any type or of types 6/11, 16, 18, 31, 33, 52 or 58
Acquisition during follow-up of an incident cervical HPV infection of a type not present at study entry
Time to clearance of an incident cervical HPV infection, of any type, or of a specific type
Acquisition and clearance of a second, or subsequent, cervical HPV infection, of any type, or of a specific type
Prevalence at study entry of cervical cytological abnormality
Acquisition of an incident cervical cytological abnormality, of any severity, or of a specific severity
Time to clearance of an incident cervical cytological abnormality, of any severity, or of a specific severity
Acquisition and clearance of a second or subsequent cervical cytological abnormality, of any severity, or of a specific severity
Acquisition of incident histologically-confirmed cervical abnormalities of any severity, or of a specific severity
Time to acquisition of cervical cytological abnormality following the first detection of HPV DNA of any type, or of a specific type.
Time to progression of cervical cytological abnormality
HPV16 and HPV18 viral load, and how this changes during follow-up
The humoral immune response to cervical HPV infection, of any type, or of a specific type, and how this changes during follow-up
Time from acquisition of an incident cervical HPV16 or HPV18 infection until seroconversion to that type
Time from seroconversion to cervical HPV16 or HPV18 infection until clearance of infection with that type

- Outcome data of different types

The cohort study described in this thesis generated data of several different types. The two which concern the analyses presented in this thesis are time-to-event data, and longitudinal data. Time-to-event data was interval-censored. Such data require special techniques.

- Repeated observations

Measurements were made on the same woman on several occasions over time. Data were not *balanced*: different numbers of observations were made on each subject, with a different time interval between consecutive observations for each pair of observations *and* for each subject (see section 6.3.3.1). The time interval between the first cervical cytological abnormality and the first colposcopy (which followed "immediately") was systematically shorter than any other interval between visits (see section 5.2.3.1.2).

- Time-varying covariates

Explanatory, and exposure, covariates were updated at each visit, yielding a mixture of time-fixed and time-varying covariates, for both explanatory variables and exposure status.

• Multiple exposures

There were several possible exposure variables: examples are given in table 6.1b. Some exposures were potentially dependent, e.g. interaction between exposure to cervical HPV infection and smoking, or between exposure to cervical HPV16 infection and to cervical infections with other HPV types.

Table 6.1b. Examples of exposure variables for the cohort study described in this thesis.

Exposure variable
Changes in HPV viral load during follow-up
Age at starting smoking
Smoking status at study entry
Changes in smoking status during follow-up
Age at first sexual intercourse, or at menarche
Acquisition of a new sexual partner during follow-up
Changes in sexual behaviour during follow-up
Characteristics of sexual partners
HPV status at study entry, with respect to any type or an individual type
Changes in cervical HPV status during follow-up, with respect to any type or an individual type
Cervical cytological status at study entry
Changes in cervical HPV viral load during follow-up

• Measurements on different types of scale

Cervical HPV DNA, cytological, and histological, status, and measures of the immune response to HPV infection, can all be considered as both binary and categorical measurements, but with an assumed underlying continuous process. In the case of HPV status, measurements on the continuous process were also available for two HPV types in a limited number of women (see chapter 21 on viral load).

- Measurements subject to misclassification

Explanatory, exposure, and outcome, variables were all potentially subject to misclassification. Many explanatory variables were self-reported. Cytological and histological assessment were based on objective criteria; nevertheless tests could give false-positive or false-negative results, due to sampling, or observer, error. The colposcopist only took a biopsy for histological evaluation when there was evidence of disease progression, a subjective decision on their part.

- Missing observations

There were missing explanatory variables, “exposure” variables, and measurements of outcomes. There were difficulties in distinguishing which observations were missing because women were requested to return to the study every six months, but were not given specific appointment schedule: a woman who returned to the study after nine months may or may not generate a “missing” observation at six months in the true sense.

Not all of the challenges described above are of concern in this thesis, and some will only be discussed in the relevant chapters. Other challenges apply to many of the analyses presented in this thesis, and will be discussed here in general terms, with more specific details given in the appropriate chapters.

6.2 DEFINITION OF TIME-TO-EVENT DATA

Time-to-event data (also referred to as failure-time data, or survival data) consists of measurements of the time to the occurrence of an event of interest in a sample of study subjects. In this thesis, there are many examples of time-to-event data.

The essential elements of a time-to-event are a well-defined origin from which time is measured, typically an event such as birth or diagnosis of disease (the origin event); the time of occurrence of an unambiguous event or outcome, such as death (the outcome event); and a measurement scale for time, such as calendar time or age. A time-to-event can be measured on a continuous or on a discrete scale, and it is not necessary for the event to be considered a “failure” in any sense: the event could be the development of a disease, a positive response to a treatment, relapse of disease, or death, in which case the time-to-event is more commonly known as the survival time. A typical example of a time-to-event measured on a continuous scale is the lifetime of an individual, the interval from date of birth to date of death, measured in calendar time. As an example of a discrete time-to-event, suppose that a subject is tested for the presence of antibodies to a virus every hour on the hour: the time-to-event in this case could then be taken as the number of hours which elapse until antibodies are first detected, measured on the scale 0, 1, 2, 3 hours, etc.

Time-to-event studies are follow-up studies in which individuals are followed until the occurrence of an event. Technically this distinguishes them from longitudinal studies, in which study subjects are measured on more than one occasion, with exposure

and outcome data obtained at each occasion. However, clearly a longitudinal study can also generate time-to-event data, as in the longitudinal cohort study described in this thesis. In the analyses presented in this thesis, it is assumed that the event of interest, which varies according to the objectives of the analysis, is irreversible, and that once a subject has experienced the event they are removed from further consideration; in general, this is not always a requirement.

6.3 CENSORING

What distinguishes time-to-event data from other types of data is the occurrence of censoring. Censoring occurs when the time of occurrence of the outcome event is either not known, or is known only approximately. To deal appropriately with censoring in the analysis of time-to-event data, the study design which generated that data must be considered. There are several types of censoring which must be carefully distinguished, but they fall into three broad categories.

6.3.1 Left-censoring

Left-censoring arises when the event of interest occurs before the subject is first observed in the study. For example, in a study of the time from the date of the detection of the first cervical HPV infection (assumed known) to the date of the first acquisition of cervical cytological abnormality, some women may already have cervical abnormality when they are first recruited. Therefore their exact time-to-event will be unknown: but what is known is that their time-to-event is less than that

observed, in this case the time interval between the acquisition of cervical HPV infection and recruitment.

6.3.2 Right-censoring

Right-censoring occurs when subjects do not experience the event of interest before the end of the observation period: all that *is* known is that their time-to-event is greater than the duration of the observation period.

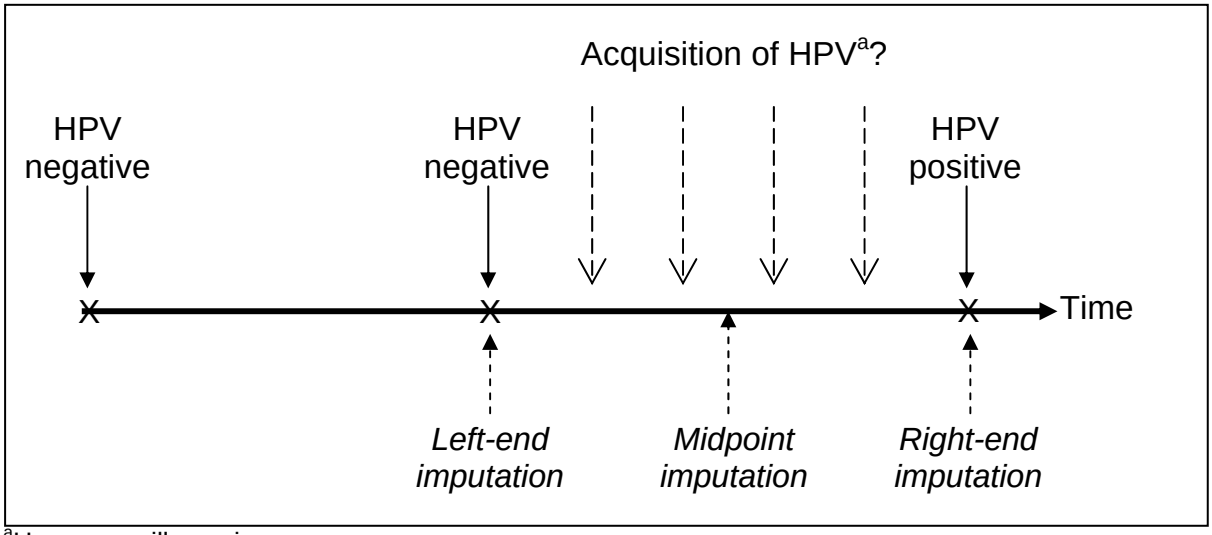
For example, in an analysis of the time from recruitment until the acquisition of high-grade CIN, most women in the cohort study described in this thesis had not had a diagnosis of high-grade CIN by the end of the follow-up period. Therefore their exact time-to-event was unknown: all that was known was that their time-to-event was greater than that observed, in this case the time interval between recruitment and the end of follow-up.

6.3.3 Interval-censoring

Interval-censoring occurs when an event is known only to have occurred within an interval of time, but the exact time of occurrence of that event is unknown: the associated time-to-event is said to be *interval-censored*. In this case the observed time-to-event is actually an interval of times rather than a single time i.e. the time-to-event is known only approximately.

For example, in this thesis, I describe measurements made on cervical HPV infection, cervical cytological abnormalities, and cervical histological abnormalities, all three of which are *asymptomatic*. For example, since cervical HPV infection is an asymptomatic condition, its onset can only be detected by taking cervical samples from a woman periodically, and then testing these samples for the presence of HPV DNA (figure 6.3.3). In this case, the time from the date of study entry to the date of the first detection of cervical HPV infection will not be known exactly. However, it *will* be known that the time at which cervical HPV infection was first acquired lies in the interval between: the date of the cervical sample which first tested positive for HPV DNA; and the date of the sample taken at the immediately preceding visit, which was the last sample taken that was known to be HPV DNA-negative. Time from date of study entry to date of first detection of cervical HPV infection is therefore interval-censored.

Figure 6.3.3. Example of interval-censored observations on the time to acquisition of cervical HPV infection.



[a]Human papillomavirus

Interval-censoring is a generalisation of left- and right-censoring; or alternatively, left- and right-censoring are special cases of interval-censoring. In practice, time-to-event data can be subject to any combination of left-, right-, or interval-censoring, and the time of occurrence of the origin event can also be interval-censored.

6.3.3.1 Types of interval-censoring

There are several types of interval-censoring, which arise from different study designs. The applicability of different methods of analysis to data from these different study designs depends on the assumptions made about the observation scheme.

Current status data. Under the simplest observation scheme, only one observation is made on each subject. This observation scheme gives rise to perhaps the most extreme form of interval-censored time-to-event data, known as current status data.

These data arise when every subject is measured (observed) on a single occasion, although the time (with respect to an origin event) at which the measurement is made may vary for each subject. In this case all times-to-event are interval-censored. With this censoring scheme, the time-to-event is *never* known exactly: for each subject, either the outcome is known to have occurred before the measurement time, or the outcome is not observed. Within the context of this thesis, methods for the analysis of current status data are of no relevance, and will not be considered further.

Grouped interval-censored data. When repeated observations are made on every subject under study at a common and fixed set of time points (e.g. every hour on the hour), then the time-to-event data generated are called grouped interval-censored data. The same number of observations can be made on all subjects, with the intervals between observations being of equal length (an X indicates an observation):

1. X - - - - - X - - - - - X - - - - - X - - - - - X - - - - - X
2. X - - - - - X - - - - - X - - - - - X - - - - - X - - - - - X
3. X - - - - - X - - - - - X - - - - - X - - - - - X - - - - - X

Or, the same number of observations can be made on all subjects, with the time intervals between observations being of different lengths, with given intervals being of the same length for each subject:

1. X - - - - - - - - - X - - - - - X - - - X - - - - - - X - - - - - - - - - - - - - - - X
2. X - - - - - - - - - X - - - - - X - - - X - - - - - - X - - - - - - - - - - - - - - - X
3. X - - - - - - - - - X - - - - - X - - - X - - - - - - X - - - - - - - - - - - - - - - X

The two designs above give rise to balanced data, and what has been called parallel data: the number of observations is fixed by design, and this may be allowed to vary for individuals, so long as this variation is pre-specified (Hougaard 2000). In contrast, longitudinal data arise when the number of observations is allowed to be random. A different number of observations can be made on each subject, with the time intervals between the observations being of equal lengths:

1. X - - - - - X - - - - - X - - - - - X - - - - - X
2. X - - - - - X - - - - - X - - - - - X - - - - - X - - - - - X
3. X - - - - - X - - - - - X - - - - - X

Or a different number of observations can be made on each subject, with the time intervals between observations being of different lengths, but with given time intervals being of the same length for each subject:

1. X - - - - - - - - - X - - - - - X - - - X - - - - - - X
2. X - - - - - - - - - X - - - - - X - - - X - - - - - - X - - - - - - - - - - - - - - - X
3. X - - - - - - - - - X - - - - - X - - - X

In theory many studies would be expected to generate grouped interval-censored data *by design*, in particular the cohort study described in this thesis. However, the ideal of a common and fixed set of time points at which observations are made is rarely achieved in practice: observations tend to be asynchronous, with time intervals between observations varying both within a specific subject, and between different subjects. In the study described in this thesis, for example, women could not be compelled to return at intervals of six months, and could not be compelled to remain in the study until a set amount of follow-up had been completed. This meant

that women returned to the study at intervals of between four and eight months after the previous visit, assuming that they didn't miss a visit and return after 12 months, for example; the time interval between the date of the first visit at which a cervical cytological abnormality was detected and the date of the subsequent smear (which was taken in a colposcopy clinic) was systematically shorter than other time intervals between study visits; an inadequate smear meant a rapid return to the study for a repeat smear; and some women made more visits than others, before they stopped returning to the study completely.

Asynchronous interval-censored data. For studies generating this type of data, either the same number of observations are made on all subjects, with the time intervals between observations being of different length, and with given time intervals being of different lengths for each subject:

1. X - - - - - - - - - X - - - - - X - - - X - - - - - - X - - - - - - - - - - - - - - - - X
2. X - - - - - X - - - - X- - - - - - - - - X - - - X - - - - - - - - - - X
3. X - - - - - - - - - - - - - - - - X - - - - X - - - X - - - X - - - - - - - - X

Or, more commonly, a different number of observations are made on each subject, with the time intervals between observations being of different lengths, and with given time intervals being of different lengths for each subject:

1. X - - - - - - - - - X - - - - - X - - - X - - - - - - X - - - - - - - - - - - - - - - - X
2. X - - - - - X - - - - X- - - - - - - - - X - - - - - - - - - - X
3. X - - - X - - - - X

This is the situation which arose in the cohort study described in this thesis. In each of the above illustrations, observations may or may not continue after the event of interest has occurred, depending on the event of interest. For example, in cohort study described in this thesis, samples were only tested for HPV DNA after *all* visits had been made and *all* samples had been collected. Therefore, for an analysis of time to first cervical infection with HPV, observations would *unknowingly* have continued beyond the event of interest, i.e. the first detection of cervical HPV infection. In contrast, for an analysis of the time to the first detection of cervical cytological abnormality, status of the previous observation, i.e. the previous smear result, would have been known *prior to* the next observation being made; observations on this outcome would therefore *knowingly* have continued beyond the event of interest. For an analysis of the time to the diagnosis of high-grade CIN, once the outcome was reached, further observations ceased, by design.

In some circumstances it may be possible to ignore variation in time interval lengths, in which case a *grouped interval-censored analysis* will provide satisfactory results. The analysis of grouped interval-censored data is easily handled by standard methods (Carstensen 1996). However, this observation scheme is not directly relevant to the data considered in this thesis and so methods for the analysis of this type of interval-censored time-to-event data will not be discussed further. For the analysis of the data the study described in this thesis, more general methods are needed.

Doubly censored data. Usage of the term “doubly censored” is not consistent. It has been used to mean: the time of occurrence of the origin event is interval-censored, with the outcome event either being observed exactly, or being right-censored (e.g. Law 1992); or the data consist of left-censored, right-censored and exact observations (e.g. Samuelsen 1994); or the origin event and outcome event are both interval-censored (e.g. Kim 1993). The latter definition is the most useful definition in the context of this thesis.

6.4 THE STATISTICAL ANALYSIS OF INTERVAL-CENSORED TIME-TO-EVENT DATA

The aim of an analysis of time-to-event data is to describe how the risk of the occurrence of an event of interest changes over time. Except in the simplest of situations, the analysis of time-to-event data requires specialised statistical methods, for two reasons. First, time-to-event data tends to be non-normally distributed, and in particular positively-skewed. This means that the vast number of statistical techniques available which are based on the assumptions of normally distributed data are not directly applicable to time-to-event data. Second, in practice, in most studies which generate time-to-event data, those data are subject to censoring. Censoring requires appropriate handling to ensure that the data are analysed efficiently using a method which produces unbiased estimates of statistics of interest.

Many powerful methods for the analysis of right-censored time-to-event data have now been developed: the basic methods are available in all of the major statistical packages (e.g. SPSS, SAS, Stata), with more advanced techniques available in specialist statistical packages (e.g. S-plus, R). In contrast, methods for the analysis of interval-censored time-to-event data, whilst now becoming more common, have developed relatively slowly over the past 40 years. One popular approach to the analysis of interval-censored time-to-event data is to ignore the interval-censoring: in particular, recognition of the requirement for specific methods for analysing such data has not yet become widespread in the medical literature. This may reflect genuine ignorance on the part of most investigators, or it may have a more pragmatic explanation: many of the methods for the analysis of interval-censored time-to-event data are quite basic at present, with counterparts of the rich array of techniques which are available for right-censored time-to-event data yet to become widely available; software for analysing these data has only recently started to become accessible to the non-specialist.

6.4.1 Estimation of survival functions with interval-censored time-to-event data

For time-to-event data, the *"survival" function* provides a summary of how the risk of the occurrence of an event of interest changes over time. Theoretically, from the survival function can be derived all of the other functions which can be used to describe time-to-event data, in particular the hazard function, the density function and the cumulative distribution function (CDF). For right-censored time-to-event data, the Kaplan-Meier (KM) method is usually used to estimate the survival

function. For interval-censored time-to-event data, an analogue of the KM estimator is the nonparametric (generalized) maximum likelihood estimator (NPMLE). This was first discussed by Peto (Peto 1973) and Turnbull (Turnbull 1976), whose papers can be taken as the starting point(s) for the development of the statistical literature on the analysis of interval-censored time-to-event data. Both authors proposed algorithms for deriving the NPMLE of the CDF for continuous interval-censored time-to-event data, but Turnbull's estimator is applicable in more general situations i.e. the assumptions made by the Turnbull estimator are less restrictive than those made by the Peto estimator.

6.4.1.1 The non-parametric maximum likelihood estimator

Unlike for the KM estimator, there is no closed formula for the NPMLE with interval-censored time-to-event data, which must therefore be derived *via* an iterative algorithm. This is done using methods for multinomially-distributed data, after having formulated the problem in terms of a discrete survival analysis (Lawless 2003). Several algorithms may be used to estimate the NPMLE. Turnbull used a self-consistency algorithm. With this approach, the observed data are used to construct disjoint time intervals which can gain possible probability masses; outside these intervals, the estimator is undefined (the original construction of the intervals receiving positive probability mass was flawed, and was subsequently corrected (Allioum 1996)). The estimator is therefore very unusual: it is a decreasing and very "rough" function, consisting of a series of horizontal lines and/or points, with gaps between the lines and/or points on the time axis. This is difficult to interpret visually,

and often the lines and/or points in the graph of the estimator are extended and joined to form a continuous step-function, similar in appearance to the KM estimator. However, this is aesthetically pleasing, rather than technically correct (but the same is true of the KM estimator), and since the extension of the lines can be performed in several ways, the visual representation is not necessarily unique (one book gives an algorithm which does not suffer from this visually difficult interpretation, but it is not clear why this algorithm is fundamentally different (Klein 1997)). In addition, methods based on the analysis of discrete time-to-event data can be numerically unstable because discrete hazards are probabilities which need to be constrained between zero and one (Goetghebeur 2000). Therefore, different approaches to estimating the NPMLE, or altering convergence criteria, may produce visually, if not statistically, or numerically, different representations of the estimator. Further, neither standard maximum likelihood estimation, nor martingale theory, is directly applicable to interval-censored time-to-event data measured on a continuous scale, and the method has not been shown to provide a consistent estimate of the variance of survival estimates for continuous time-to-event data in a general setting (Petroni 1994).

The Turnbull, or other, non-parametric maximum likelihood estimator has yet to attain the wide acceptance of the KM estimator. Given the unusual nature of the estimator, and the difficulties with its estimation and interpretation, this seems unlikely to change. So why use this type of estimator? The disadvantages are outweighed by the fact that no restrictive assumptions are made about the underlying distribution from which the observed data are generated. The estimator

can therefore be used to suggest a possible parametric model, which will rarely be known *a priori*, or, as with any nonparametric method, to help assess the robustness of conclusions drawn from a parametric analysis (Davison 1997).

6.4.1.2 "Smoothing" the non-parametric maximum likelihood estimator

The NPMLE is a "rough" estimator. In reality, we would expect the true survival function (and its CDF) to be a "smooth" function. There are methods available for smoothing estimators of certain of the survival functions. One approach is to use splines to estimate the survival function, or one of its associated functions, directly (Kooperberg 1995, Kooperberg 1997) (see later). Another is to "penalize" the likelihood function to force it to be smooth.

The penalized maximum likelihood method (for which a FORTRAN programme is available, called PHMPL), defines the NPMLE as the function which maximizes the penalized likelihood (Joly 1998, Joly 1999). A "smoothing parameter" is introduced into the likelihood, which takes large values for rough functions, thereby "penalizing" the likelihood. The likelihood is then maximised over the class of all functions of the appropriate form, to yield the maximum penalised likelihood estimator (MPLE). However, the MPLE cannot be calculated directly, it must be approximated; in this case, splines are used. *Splines* are piecewise polynomial functions which can be combined, in a linear fashion, to approximate a function over an interval. A spline function is completely defined by a sequence of increasing "knots" on the time-axis ($t_1,..., t_l$), and the coefficients $\theta = (\theta_1,...,\theta_l)^T$ of the splines. PHMPL uses splines of

order four (called cubic splines), and allows a minimum of five, and a maximum of 25, knots: in theory, the more knots that are used, the better the approximation. The recommended method is to start with a small number of knots and increase this number until the graph (*sic*) of the hazard function remains unchanged, which seems a little subjective. The smoothing parameter can be specified empirically, or it can be determined automatically. Explanatory variables can be included in a proportional hazards model, but PHMPL will not allow for the incorporation of time-varying covariates.

6.4.1.3 Comparison of estimates derived from an analysis of interval-censored time-to-event data with those derived from the Kaplan-Meier analysis of right-censored time-to-event data: in the cohort study described in this thesis

The method used to analyse data should correspond to the underlying process which generated that data. However, as stated earlier, many investigators have failed to observe this general rule. Investigators who ignore the true nature of their data tend to use imputation of some type to convert their interval-censored data into “exact” (or more likely at worst right-censored) time-to-event data. If the interval-censoring is recognised, midpoint imputation (stating that the event occurred in the middle of the interval in which the event is known to have occurred) is likely to be used (as indeed the candidate did in Woodman 2001 and Woodman 2003); if not, right-imputation is most likely to be used, for example using the date of diagnosis as the time at which the event occurred for an asymptomatic condition measured at intervals (see figure 6.3.3). Standard methods for use with right-censored time-to-

event data are then used to analyse the resulting data. But what impact does this have on statistical inferences?

To compare the Turnbull NPMLE, MPLE, and KM estimator, I undertook three analyses: one with a large number of events in a small population, one with a large number of events in a large population, and one with a small number of events in a large population. These were, respectively: the analysis of the incidence of cervical HPV infection in women recruited within six months of first sexual intercourse and before they acquired their second sexual partner (242 women, 78 incident cervical HPV infections) (Collins 2002); the analysis of the incidence of cervical cytological abnormality in women who were cytologically normal and negative for cervical HPV DNA at recruitment, and who had further follow-up (1,075 women, 246 incident cervical cytological abnormality events) (Woodman 2001); and the analysis of high-grade CIN in the same women as used in the previous analysis (1,075 women, 28 incident high-grade CIN events) (Woodman 2001). The NPMLE was calculated using the ICE macro in the statistical package SAS. The MPLE was obtained using the FORTRAN programme PHMPL, with the maximum number of knots used (25), and with linear interpolation *of the estimator* to obtain estimates at specific points in time. Three versions of the KM estimator were calculated, one each assuming the event of interest occurred at the left-hand end of the interval, at the right-hand end, or at the midpoint. The results of the comparison with the Turnbull NPMLE are presented in table 6.4.1.3 (note that cumulative risk is the complement of survival i.e. one minus survival).

Estimates based on the assumption that an event occurs at the left or right limits of the interval will, in theory, produce the most extreme estimates of the survival function (and therefore the cumulative risk function); these estimates should then form an interval within which estimates produced by making any other interpolation assumption will lie. In particular, this interval is expected to contain the *true* survival estimate, and the estimate based on the midpoint assumption, is considered to be the “best” estimate, in some (unspecified) sense.

In the analysis of the incidence of cervical HPV infection, the NPMLE estimate does tend to lie between the left and right KM estimates, except at 24 months, where it lies below both limits. At other time points, the midpoint estimate provides a reasonable, although not always the best (among the three estimates investigated), approximation. The MPLE always lies between the left and right limits. In the analysis of the incidence of cervical cytological abnormalities, all three versions of the KM estimator underestimate the risk of acquiring cervical cytological abnormality, as estimated by the NPMLE, from 24 months onwards, with a substantial discrepancy at 60 months (this may be due to small numbers at risk by this stage). The MPLE estimate always lies between the left and right KM limits. In the analysis of the incidence of high-grade CIN, all three versions of the KM estimator underestimate the risk of acquiring high-grade CIN, as estimated by both the NPMLE and MPLE, up to 48 months, but overestimate the risk at 60 months.

In all cases, the confidence interval associated with the NPMLE is substantially wider than the confidence intervals around any of the three versions of the KM estimator.

However, the confidence intervals for the MPLE are comparable in width to those for the three versions of the KM estimator. The paper presenting the PHMPL method states that it generates confidence *bands* for the penalized likelihood estimators, and thus for the survival estimates; the instructions for the PHMPL computer programme calls these statistics *confidence limits* (Joly 1999). If they are in fact confidence bands, this makes the narrow confidence intervals difficult to understand.

Table 6.4.1.3. Comparison of estimates of cumulative risk obtained using methods appropriate to interval-censored time-to-event data, with those obtained using interpolated data, and the Kaplan-Meier estimator. The cumulative risk (CR) of: cervical HPV infection following first sexual intercourse in women recruited within six months of first sexual intercourse whilst they had only one sexual partner; cervical cytological abnormality (CA) in women who were cytological normal and HPV DNA negative at recruitment, and who had further follow-up; and the analysis of high-grade CIN (HGCIN) in the same women. Estimates obtained using the Turnbull NPMLE (T), the MPLE (P), and the Kaplan-Meier estimator, assuming that events occur at the left-hand end of the interval (L), at the midpoint (M), or at the right-hand end (R). Number of women (n) and number of events (e) are also presented. n=number of subjects, e=number of events.

Time (months)	**Method**	**HPV (n=242, e=78)**		**CA (n=1075, e=246)**		**HGCIN (n=1075, e=28)**	
		CR (%)	*95% CI*	*CR (%)*	*95% CI*	*CR (%)*	*95% CI*
0	T	0.0	–	0.0	–	0.0	–
	P	0.0	–	0.0	–	0.0	–
	L	15.3	10.8 to 19.8	8.4	6.7 to 10.0	0.0	–
	M	0.0	–	0.0	–	0.0	–
	R	0.0	–	0.0	–	0.0	–
12	T	24.6	3.9 to 45.3	13.4	0.6 to 26.2	1.2	0.0 to 2.9
	P	25.5	18.9 to 31.6	11.4	9.4 to 13.4	1.4	0.8 to 2.1
	L	27.1	21.0 to 33.2	15.3	13.1 to 17.5	0.2	0.3 to 1.5
	M	26.7	20.6 to 32.8	11.5	9.5 to 13.4	0.8	0.2 to 1.4
	R	23.5	17.6 to 29.4	9.0	7.2 to 10.8	0.4	0.0 to 0.8
24	T	34.2	18.5 to 49.8	22.0	10.6 to 33.3	2.4	0.0 to 5.1
	P	38.2	30.0 to 45.4	20.8	17.9 to 23.5	2.5	1.5 to 3.5
	L	39.3	31.5 to 47.1	21.8	19.1 to 24.6	1.3	0.5 to 2.1
	M	38.3	30.5 to 46.1	21.3	18.5 to 24.1	1.4	0.6 to 2.2
	R	36.9	29.1 to 44.7	20.2	17.4 to 22.9	1.5	0.7 to 2.3
36	T	45.4	27.1 to 63.7	31.2	11.2 to 51.3	3.3	0.1 to 6.5
	P	45.1	34.8 to 53.9	27.7	24.3 to 30.9	3.1	2.0 to 4.3
	L	42.5	33.9 to 51.1	27.3	24.1 to 30.6	2.5	1.3 to 3.7
	M	45.6	36.8 to 54.4	28.4	25.0 to 31.8	1.8	0.8 to 2.8
	R	51.0	41.4 to 60.6	28.5	25.0 to 32.0	1.7	0.7 to 2.7
48	T	49.8	30.3 to 69.3	35.6	23.5 to 47.7	3.6	0.0 to 8.3
	P	48.4	37.2 to 57.6	32.2	28.1 to 36.0	3.4	2.1 to 4.8
	L	44.8	35.4 to 54.2	31.3	27.4 to 35.1	3.2	1.6 to 4.8
	M	50.0	40.0 to 60.0	33.2	29.2 to 37.2	2.8	1.4 to 4.2
	R	53.0	43.0 to 63.0	34.3	30.2 to 38.4	3.0	1.6 to 4.4
60	T	58.9	34.3 to 83.5	55.5	39.7 to 71.3	3.6	0.0 to 8.3
	P	54.2	40.7 to 64.7	37.2	31.5 to 42.3	3.5	1.5 to 5.5
	L	51.5	39.5 to 63.5	34.2	29.7 to 38.7	4.5	2.3 to 6.7
	M	57.2	44.5 to 69.9	37.9	32.8 to 43.1	4.5	2.3 to 6.7
	R	59.5	47.3 to 71.7	40.6	35.4 to 45.8	3.8	2.0 to 5.6

6.4.1.4 Comparison of estimates derived from an analysis of interval-censored time-to-event data with those derived from the Kaplan-Meier analysis of right-censored time-to-event data: in other studies

Others have also examined the impact of imputation of time-to-event data on statistical inferences, in the context of univariate, multivariate, and doubly-interval-censored, analyses (Law 1992).

Samulesen *et al.* (Samuelsen 1994) found that the NPMLE was substantially less than the KM estimate, but NPMLE estimates were closer to KM estimates using midpoint- rather than endpoint-imputation; in their multivariate analysis, there was very little difference between the estimates produced by the three methods, or their standard errors, but there were differences for certain covariates.

In the analysis of a randomised controlled trial of 262 patients treated for gastric cancer, Nakazato *et al.* (Nakazato 1994) used methods suitable for right-censored time-to-event data, with right-endpoint imputation at the date of first detection of gastric cancer, to analyse interval-censored disease-free survival (DFS) times. In this study, follow-up was at every two weeks during the first month, then once a month to three months, every three months for one year, and thereafter every six months to five years. Subsequently, this data was reanalysed using methods appropriate to interval-censored time-to-event data (Sakamoto 1997). First the authors found that the biggest discrepancy in DFS rates between two analyses using the KM method, one with left-hand, the other right-hand, imputation, was only 1.2%.

Clearly, this will depend on the event rate, as seen in table 6.4.1.3. Similarly, although estimated regression coefficients and their standard errors were all slightly higher in the analysis using interval-censored time-to-event data, there were no substantial differences between this analysis and the original analysis. However, in their repeat analysis a parametric Weibull regression model was used, possibly because this was the only way the authors could proceed with an interval-censored analysis, given the stage of development of available software at that time. If this parametric assumption was not justified, it is difficult to interpret their conclusions. In the context of a Weibull model, Odell *et al.* (Odell 1992) found that when the hazard rate is flat (i.e. when the risk of an event occurring is essentially constant over time), or when the proportion of interval-censored data in a dataset is small (with most time-to-event data being exact, or right-censored), an analysis based on midpoints is adequate.

Goggins *et al.* (Goggins 1998) reached a different conclusion for the semiparametric Cox model with interval-censored time-to-event data. Using simulation, they found that midpoint- or right-imputation produced bias in parameter estimates from the proportional hazards model, and underestimated standard errors. Kim *et al.* (Kim 2002) reached similar conclusions in the context of multivariate interval-censored time-to-event data.

The problem with single imputation, i.e. imputing only one time-to-event for each subject, such as midpoint, left- and right-endpoint imputation, is that it treats the imputed values as if they were known exactly. Inevitably, this will underestimate the

true variability in time-to-event data. Multiple imputation, on the other hand, attempts to overcome this problem by replacing each of the original interval observations with two or more “exact” values, thereby yielding multiple data sets. These data sets are then analysed separately using a standard method for right-censored time-to-event data, and the results combined in such a way as to take proper account of the variability in the original data set (Pan 2000). Sun (Sun 2006) used an approach involving multiple imputation to analyse two different datasets; Sun then compared the results against results obtained using the KM method with midpoint imputation, and the Cox regression model, and found very little difference overall. Nevertheless, Sun’s discussion is concluded by reiterating that an imputation approach can produce biased estimates, the precision of estimates may be overstated, and the effects of covariates may be measured incorrectly. Whether any or all of these adverse factors have arisen in an analysis of a specific dataset cannot be determined with any certainty.

6.4.2 What is the most appropriate way to analyse interval-censored time-to-event data in the cohort study described in this thesis?

So, can an imputation method, in particular one using midpoint-imputation, be used to make inferences based on the interval-censored time-to-event data generated by the cohort study described in this thesis? Unfortunately, the answer seems to be “it depends”. An imputation approach *appears* to yield satisfactory results, provided the time intervals between measurements are short compared to the mean time-to-event, i.e. the closer the actual observation scheme is to observations made in

continuous time, the better are the results of imputation within the interval (Collett 2003). Based on the limited examples presented in table 6.4.1.3, the MPLE and KM estimate are quite close, even when the KM estimate differs from the NPMLE; the CI around the estimates also seem very narrow, since one would anticipate some concession to the fact that interval-censored observations must always be less precise than exact observations. Similarly, the imputation procedure described by Sun, while more complicated than midpoint-imputation, produces results which do not differ substantially from those produced by the midpoint analysis described. Given the added complexity of implementing these approaches, and the suspicion (on the candidates part, at least) that they yield spuriously precise estimates, they appear to offer little or no advantage compared to performing an analysis suitable to right-censored time-to-event data with midpoint imputation. It therefore appears that the best approach to dealing with interval-censored time-to-event data for the cohort study described in this thesis is to perform a nonparametric analysis appropriate to interval-censored time-to-event data; and then to check the results for consistency by using the more easily implemented methods which are available for right-censored time-to-event data, obtained *via* midpoint-imputation. Any large discrepancies between the results of these two analyses can then be investigated in an attempt to determine the reason (that being said, it is the results of the interval-censored analysis which should be accepted). However, it remains to identify suitable methods for the analysis of interval-censored time-to-event data when covariates are available.

6.5 IDENTIFICATION OF METHODS FOR THE ANALYSIS OF INTERVAL-CENSORED TIME-TO-EVENT DATA FROM THE COHORT STUDY DESCRIBED IN THIS THESIS

A search of the literature was undertaken with the aim of identifying methods for the analysis of interval-censored time-to-event data with: time-varying covariates; varying times between observations, both within and between individuals; varying numbers of observations for different individuals; and for which it would be possible for the candidate to implement in practice. Suitable methods were those for which there were existing computer programmes, or those which did not require extensive computer programming.

All available published papers relating to methods used in the analysis of interval-censored time-to-event data were identified *via* a search of the PubMed database, and electronic journals databases at the Universities of Manchester and Birmingham; the term “interval-censored” was used in searches. The reference lists of the papers identified were then also searched. Two hundred and ten papers which were potentially relevant were identified: 54 were subsequently considered to be irrelevant for the purposes of this thesis (e.g. only applicable to current status data); 65 were considered to be essentially of only theoretical interest, or to require excessive computer programme development time; and 17 were of indirect interest (e.g. review papers, or examples of applications of methods for interval-censored time-to-event data to real-life data). This left 74 to be examined in detail. Thirty-one were of potentially practical use, i.e. with computer programmes available, or

requiring easily-programmed techniques, and allowed for variation in the number and timing of measurements on each subject. Ten of these 31 papers presented methods for estimating time-to-event functions, including two which allowed for the inclusion of time-fixed covariates only; six discussed the comparison of survival curves, with time-fixed covariates; and 15 presented methods for the regression analysis of interval-censored time-to-event data. Of the 15 papers dealing with regression analysis: three were based on methods making parametric assumptions (one allowed for the inclusion of time-fixed covariates only, two for time-varying covariates); and 12 presented nonparametric, or semiparametric, methods. Of these 12, eight allowed for the inclusion in the analysis of time-fixed covariates only, and four allowed for the inclusion of time-varying covariates, of which however, two assumed that discrete time-to-event data were of interest. The remaining two methods were considered for use with the data described in this thesis; Cox regression with midpoint imputation provided a benchmark against which results generated by the two methods identified were compared.

I now briefly describe the two methods identified.

6.5.1 Regression methods for analysing interval-censored time-to-event data: HARE

HARE (the acronym is derived from HAzard Regression) is an adaptive regression method for analysing interval-censored time-to-event data in the presence of covariates (Kooperberg 1995, Kooperberg 1997). In HARE, linear or cubic splines

(see section 6.4.1.2) and tensor products are used to estimate the logarithm of the conditional hazard function (conditional upon covariates). Unknown parameters in a model are estimated using maximum likelihood; selection of the final model is either fully automatic, incorporating the stepwise addition and deletion of covariates using the Bayesian Information Criterion; or specific models can be investigated by forcing certain covariates into a model. Proportional hazards models assume that the conditional hazard function is a multiplicative function of time and covariates; with HARE, this assumption is not made, although proportional hazards models can be fitted as a special case. HARE places no restriction on the number, or frequency, of observations made on subjects, and allows both of these to vary between subjects. HARE can be used to analyse time-varying covariates, provided, as with other regression models with this capability, covariate values are known for each at-risk subject at each event time. However, HARE models the *complete* conditional hazard function, so time-varying covariates present a great numerical challenge than time-varying covariates usually do. A computer programme in the statistical package S-Plus is available for implementing HARE.

HARE *appeared* to be an ideal method for analysing the data generated by the cohort study described in this thesis. In practice, implementing HARE proved problematic, with difficulty repeating apparently identical analyses on different occasions (sometime a result would be obtained, and sometimes it would not; however, whenever a result *was* obtained, it was always identical for identical analyses). It is unclear why this was the case. It is most likely to have been due to the performance of the computer(s) used to run the analyses, but may also have

been due to the statistical package used, flaws in the HARE programme itself, or the competence of the user. HARE was therefore only used in one set of analyses in this thesis.

6.5.2 Regression methods for analysing interval-censored time-to-event data: PJS

This method (the acronym PJS is mine, and is taken from the initials of the author) is a semi-parametric proportional hazards regression model for analysing interval-censored time-to-event data in a generalized linear model (GLM) framework (Smith 1997). This is a major advantage of the method, since it means that it falls within the mainstream of statistics, unlike HARE for example, which is highly specialized in nature. Covariates are incorporated parametrically (c.f. Cox regression), but the hazard function is either modeled empirically (a semi-parametric model), or is specified parametrically (a parametric model). This method makes no restrictive assumptions with respect to the observation scheme: subjects may be seen at irregular time intervals and/or on a varying number of occasions. The unit of analysis for this method is the interval between consecutive observation times. Each subject contributes a set of binary variables which indicate whether a failure occurred during each of the intervals of observation, prior to the known occurrence of the failure event, or to the end of the last time interval. These binary indicators are then modeled in a GLM. The logarithm of the baseline hazard function is modeled empirically (i.e. based on the observed data) as a smooth function. The suggested method is to model the logarithm of the baseline hazard function as an appropriate

function of the midpoints of the intervals of observation. For example, the logarithm of the baseline hazard function can be approximated by a quadratic polynomial in the interval midpoints. Alternatively, the baseline hazard function can be specified parametrically. However, in this case, it is not clear why the PJS method offers any advantage over the standard parametric methods for interval-censored time-to-event data. The PJS method allows for the incorporation of time-varying covariates defined on each interval of observation: therefore, a decision has to be made about how to code a covariate which varies throughout an interval (see section 6.6.1).

6.5.3 Regression methods for analysing interval-censored time-to-event data: Other methods

The PJS method is not the only method which has been proposed for the analysis of interval-censored time-to-event data which is based on GLM. For example Carstensen (Carstensen 1996) and Farrington (Farrington 1996) both present semi-parametric methods for modeling interval-censored time-to-event data as a GLM, when patients are seen at irregular time intervals and/or on a varying number of occasions. As with the PJS method, models in the Carstensen and Farrington methods are based on an empirical estimate of the hazard function, with covariates incorporated parametrically. Both Carstensen and Farrington describe approaches for incorporating time-fixed covariates, but also state that time-varying covariates may be incorporated, with differing degrees of difficulty. For both the Carstensen and Farrington methods, models are fitted as GLM, with the contribution to the likelihood from a single individual equivalent to the likelihood from at most two independent

Bernoulli trials. This differs from the PJS method, since even when only fixed baseline covariates are included in the analysis, all intervals for each subject are used so that both the baseline hazard function can be estimated, and time-varying covariates can be included. Both the Carstensen and the Farrington method assume a piecewise-constant hazard function: the time axis is divided into a fixed number of intervals of varying width, independent of the time intervals yielded by the interval-censored observations, and within these fixed intervals the hazard function is assumed to be a non-negative constant. The potential advantage of a piecewise-constant hazards function is that, as the number of intervals used to model the hazard function increases, the model becomes more non-parametric in nature (in other words, this is the next-best-thing to a non-parametric method), which allows for flexibility in modeling, while retaining the advantages of parametric methods for hypothesis testing and estimation (Lindsey 1998). In the Farrington method, as the number of observations on subjects increases, the number of parameters in the discrete-time model increases, but for the additive and multiplicative models, the number of parameters remains constant; in the Carstensen method, the number of parameters used to estimate the hazard function remains under the analysts control. This last point is an important consideration when analysing a large cohort, such as the cohort of 2,011 women described in this thesis; with the PJS method, the number of parameters also remains constant as the number of observations increases.

The Farrington method was rejected for use in the analysis of the cohort study described in this thesis on grounds of the requirement for substantial computer

programming to implement the method, as well as the extensive amount of work perceived to be required to include time-varying covariates in a model. In contrast, the Carstensen method *was* considered suitable, initially, since it appeared to be relatively straightforward to implement, and the incorporation of time-varying covariates seemed straightforward. An attempt was therefore made to use this method. However, when the data was analysed as a set of at most two records per subject, a mixture of positive- and negative-intensities (or estimates of the hazard function) was obtained, without including any covariates. Given that the hazard function is strictly non-negative by definition, this is clearly not satisfactory. When the data was coded as a set of binary outcomes with only one non-negative covariate per record (as suggested by the author to overcome this problem), all parameters were negative. This was the case whether analysing time to diagnosis of high-grade CIN (28 outcomes in 1,075 women) or HPV (407 outcomes in 1,075 women). Seaman *et al.* (Seaman 2001) also attempted to use this method and seem to have had the same problem. The authors state: “however, the generalized linear model computer algorithm may converge to an impossible solution, because it does not take into account the parameter constraints that exist when the survival model is formulated in this way as a generalized linear model [however, this is why Carstensen suggested coding the data with only one non-negative covariate per record]. In fact, I attempted to use Carstensen’s method, but found that negative estimates of the hazard function were produced for some time periods, even before covariates were introduced. I decided not to pursue this method any further.

The PJS method was found to be entirely satisfactory in terms of both ease of implementation, and consistency of results with the Cox model using midpoint-interpolated right-censored time-to-event data. This method has been used in preference to HARE in all relevant analyses in this thesis, save the first undertaken (see chapters 8, 9, 10 and 11).

6.6 TIME-VARYING COVARIATES IN TIME-TO-EVENT ANALYSES

Covariates whose values change over time are called time-dependent, or time-varying, covariates, for example temperature during the day, or smoking status during follow-up in the cohort study described in this thesis; covariates whose values do not change with time are called time-independent, or time-fixed, covariates, for example sex, or cytological status at recruitment into the cohort study described in this thesis. Time-varying covariates may be more plausible causal factors and better explain observed outcomes than time-fixed covariates (Bacchetti 2002). However, time-varying covariates present practical challenges in terms of: assigning them a value; the ability to include them in an analysis; their description; and their interpretation.

6.6.1 Assigning a value to a time-varying covariate

When including a time-varying covariate in an analysis, a decision must usually be made as to how to assign a value to that covariate. In particular, with a longitudinal repeated-measurements design, time-varying covariates are updated only at distinct

points in time, although the underlying process which they measure is likely to be a “smooth” continuously varying process. Therefore, the most representative, or relevant, value for that covariate during the interval between which the measurements are made must be decided upon. This decision will vary according to the covariate, and the study hypothesis, but tends to be based on a limited range of options. For example: previous measured value; current value; mean value; cumulative value; maximum value attained to date, etc.

The value(s) of some time-varying covariates may be missing: subjects with missing values could be excluded from an analysis, but this is likely to be very inefficient, especially if values are missing only intermittently (see section 6.7.1). If the subjects are to be retained in the analysis, their missing values must be assigned a value in some way. Possible approaches are: to impute the missing value using the most recent measurement available, for categorical covariates; to interpolate between adjacent measurements, for continuous covariates; or to use several adjacent measurements to assign a value, for continuous covariates.

These are general issues with assigning a value to a time-varying covariate. How time-varying covariates were assigned values, and how missing values were dealt with, in specific analyses of the cohort study described in this thesis, are discussed in the relevant sections of the thesis.

6.6.2 Including and describing time-varying covariates in an analysis

Not all statistical methods for the analysis of time-to-event data will allow for the inclusion of time-varying covariates. For example, the nonparametric maximum likelihood estimators for both right-censored and interval-censored time-to-event data, the KM and Turnbull estimators, respectively, can only be estimated for time-fixed covariates. And there is no equivalent graphical, or tabular, representation which will capture the effects of a time-varying covariate “at a glance” in the same way as a graph of the KM or Turnbull estimator will.

Only methods capable of including time-varying covariates were considered for use in relevant analyses in this thesis.

6.6.3 Interpretation of time-varying covariates

Time-varying covariates can be categorised into groups according to the ability of study subjects, or investigators, to influence their values, either directly or indirectly: the greater the ability, the more challenging the question of interpretation.

An *external* time-varying covariate is one which can be measured without regard to the status of the study subject (Kalbfleisch 2002). External covariates cannot be influenced by study subjects, or manipulated by investigators, and generally present few problems with respect to statistical inferences. There are different types of external time-dependent covariate (Singer 2003). A *defined* covariate is one for

which values are predetermined for every subject under study. Most defined covariates are functions of time e.g. time itself, or calendar age. An *ancillary* covariate is a random covariate whose values are determined by a stochastic (i.e. random) process independent of any particular study subject. For example a time-varying environmental exposure, such as temperature: knowing the temperature on a given date at a certain time does not convey any information about the status with respect to any given outcome of a study subject at that time.

An *internal* time-varying covariate is one for which the status of the study subject determines whether the covariate can be measured (Kalbfleisch 2002). For example, number of cigarettes smoked per day by a person is an internal time-varying covariate: this covariate changes over time in a way which cannot be determined in advance, and which can be changed by the subject themselves, either as a result of external influences, or merely on a whim. Knowing that a person smoked a certain number of cigarettes on a certain date implies that the person was alive on that date; knowledge of the variable therefore conveys knowledge of survival status.

Internal time-varying covariates present difficulties of interpretation. One of the strengths of a longitudinal study design is that it establishes *temporality*, i.e. that the cause, as measured by the covariate, precedes the effect, as measured by the outcome variable. In contrast, a cross-sectional study, where the predictor (i.e. the covariate of interest) and outcome are measured contemporaneously, can never establish temporality with certainty. If a longitudinal analysis links outcomes with contemporaneous values of time-varying covariates, this effectively converts a

longitudinal analysis into a cross-sectional analysis. Therefore, even though a longitudinal analysis may demonstrate an association between a time-varying covariate and an outcome, this would not necessarily establish temporality. There are two types of time-varying covariate for which temporality may be an issue, state- and rate-dependent covariates (Singer 2003).

A time-varying covariate is *state dependent* if its values at a given time are affected by whether the subject has experienced the event of interest at that time. HPV viral load, and the first detection of cervical cytological abnormality, may be an example of this: the risk of acquiring cytological abnormality may increase with increasing viral load; but cytological abnormality may also facilitate an increase in HPV replication, and therefore an increase in HPV viral load. This is further complicated by the relative sensitivities of the two methods of measurement. The detection of cytological abnormality is relatively insensitive, whereas qPCR is very sensitive at detecting HPV. Therefore, a time-varying covariate measuring HPV viral load may demonstrate changes resulting from the acquisition of the outcome well before it is possible to detect the outcome itself.

A time-varying covariate is *rate-dependent* if its values at time t are affected by an individual's value of hazard at time t. Consider the relationship between the stress experienced by an aircraft pilot and the risk of an aircraft accident. Stress increases with workload and with knowledge of risk. A person under stress is more likely to make a mistake; thus the risk of an accident (and therefore the underlying hazard of an accident) increases as stress increases. The workload of a pilot, and the risk of

an accident, are greatest at take off and at landing, but are negligible in between. A pilot is well aware of this fact, and so the pilot's stress levels follow the same pattern, hence pilot-stress is a rate-dependent covariate.

To avoid difficulties with interpretation, covariate processes should be predictable. A *predictable* process is a stochastic (random) process whose value at time t is known infinitesimally before t, if not sooner. A time-fixed covariate measured at baseline is predictable, since once it is measured it is known at all subsequent times. Measuring a covariate simultaneously with the event of interest is analogous to placing a bet on a race just as the winner crosses the line (Therneau 2002). Therefore, ideally time-varying covariates should take values from earlier chronological times, i.e. prior status on a potential predictor of the risk of a given outcome should be linked with current status on that outcome (Singer 2003).

In practice, the best way to deal with time-varying covariates is to adopt the usual approach to complicated issues in statistical analysis: see how sensitive the conclusions are to different assumptions and definitions, and attempt to reach a consensus amongst the various possible conclusions. How time-varying covariates were interpreted in specific analyses of the cohort study described in this thesis is discussed in the relevant sections of the thesis.

6.7 MISSING VALUES

In any longitudinal analysis involving repeated measurements, the assumption is

made that each subject's observed data are a random sample from the underlying process of interest within that subject. However, in practice, most longitudinal studies will include some subjects for whom data is missing for one or more of the repeated measurements.

6.7.1 General description

Missing values arise whenever one or more of the sequences of measurements on subjects within the study population are *incomplete*, in the sense that intended measurements are not taken, are lost, or are otherwise unavailable (Diggle 2000). Missing values result in *unbalanced* data. However, although missing data requires special handling from a statistical point of view, the existence of missing data does not, in itself, necessarily present a problem in an analysis. However, the nature of the missing values must also be considered: it must be determined whether the fact that values are missing bears any relationship to the study hypothesis. If the missing values have arisen as a result of systematic sources, this may invalidate statistical inferences (Singer 2003).

Little describes procedures for the statistical analysis of data which are subject to missing values, and provides a now widely-accepted hierarchy of missing value mechanisms (Little 2002). The crucial distinction in likelihood-based analyses is between *informative* and *non-informative*, or random, missing values; in the latter case the missing value mechanism is referred to, suggestively, as *ignorable*.

The first mechanism, *missing completely at random* (MCAR), means that the observed values can be considered to be a random sample of those values which would have been observed had there been no missing data. MCAR data may be analysed by any method which can handle unbalanced data. Therefore, if missing data are unavoidable, this is the ideal type to have. Longitudinal data are MCAR if the probability of data being missing on any given occasion is independent of: (1) the time at which the measurement is made; (2) the values of covariates; (3) the value of the (missing) outcome (Singer 2003). To establish that data are MCAR it must be shown that the probability of data being missing at a given point in time is unrelated to the contemporaneous (but missing) value of the associated outcome. Because the contemporaneous value of the associated outcome is unobserved, this can only be established by a theoretical argument. Therefore, *any* reasonable potential association invalidates the MCAR assumption.

The second mechanism, *missing at random* (MAR), is less ideal, but also less restrictive in its assumptions, than MCAR, and still allows valid statistical inferences to be made. When data are MAR the probability of data being missing can depend on *any* observed data, for either the covariates or any outcome values; it cannot, however, depend on any unobserved values of either covariates or outcome.

If data is *not missing at random* (NMAR), the third mechanism, standard statistical methods based on likelihood principles cannot be used without accounting for the missing data mechanism.

Whether missing values occur intermittently or as “dropouts” is also an important consideration. Missing values occur as dropouts if, whenever a measurement at time t is missing, so are all measurements which should have been made after time t; otherwise the missing values are said to be “intermittent”. It may be reasonable to assume that intermittent missing values arise as a result of reasons unrelated to the measurement process, and are therefore MCAR. Dropouts arise as a result of subjects being withdrawn from a study prematurely. This can often be assumed to be related to the measurement process, in which case the missing data mechanism is said to be informative (Diggle 2000, Singer 2003).

6.7.2 Missing values in the study described in this thesis

The assumption has been made throughout all analyses based on the cohort study described in this thesis, that dropouts are non-informative, and that missingness is ignorable. Ignorable missingness permits valid statistical inferences; the burden of establishing that missingness is ignorable lies with the researcher (Singer 2003).

6.7.2.1 Observation scheme

In the cohort study described in this thesis, it is difficult to say for certain which missing values are dropouts in the sense described above. The study design did not specify the number of visits which each study subject was expected to make, only the desired frequency of those visits; and women could not be, and would not have been, *compelled* to return to the study. Some women only ever made one visit to the

study, whereas others remained on follow-up for several years; some women returned regularly at intervals of six months, others returned regularly but at intervals of their own choosing, and others returned irregularly. Dropout, due to inconvenience or to loss of interest in the study, *may* have occurred. There certainly does not appear to be an association between the number of visits made and the results of any study measurements of which the woman would have been aware, but this can never be established with certainty.

6.7.2.2 Missing HPV observations

Cervical HPV infection status can be considered to be both an exposure variable and an outcome, depending on the analysis: the method for dealing with missing values when HPV is used as an exposure variable is described in the appropriate sections of this thesis. In this study cytological samples were only tested for HPV DNA after all visits had been made and all clinical follow-up had ended. Neither a woman, nor any of the clinicians involved in her care, would have been aware of the woman's HPV status at any point during the follow-up phase of the study, i.e. on or after recruitment to the study. Therefore, any missing observations are unlikely to have been directly related to cervical HPV status.

Missing observations on cervical HPV status were mostly intermittent, given that a woman's observation status is assumed to be known completely (dropouts are difficult to identify). Reasons an observation was missing include the following: a sample may not always have been taken at a visit; a sample for HPV DNA analysis

may not have been taken when an inadequate smear was repeated; the sample may have been physically lost or misplaced, for example some samples which were missing in the GP5+/GP6+ analysis were subsequently found and were included in the HPV viral load analysis; others included in the GP5+/GP6+ analysis could not be traced for inclusion in the viral load analysis. The sample may have been lost *via* failures in storage procedures: for example, some samples arrived in the laboratory with no material, or had clearly been stored inadequately, and so could not be tested.

It would appear to be safe to conclude that missing measurements of cervical HPV status are at the very least MAR, if not MCAR.

6.7.2.3 Missing cytological and histological status measurements

Cervical cytological abnormality is essentially asymptomatic, but cervical cytological status was known to both the woman and the clinicians responsible for her care, after only a slight delay. Therefore, prior cytological status could, in theory, have influenced whether or not a subsequent measurement was missing. Histological measurements were only made following the detection of cytological abnormality, and therefore the issue of missingness of these measurements are closely related.

Some missing observations on cervical cytological and histological status were clearly intermittent (see section 6.7.1), with women missing one scheduled visit and either attending shortly after, or returning at the next scheduled visit. Similarly, a visit

with an inadequate smear may nevertheless have provided a sample which was evaluable for HPV status, thus generating a “missing” sample for cytological status, but not for HPV status. Ignorable missingness appears to be a safe assumption in these cases.

Non-informative dropout is likely to have occurred following increasing familiarity with study procedures, but for some cervical cytological outcomes, e.g. having a severe cervical cytological abnormality following an abnormality of a lower severity, informative dropout is a distinct possibility. All women were made fully aware of the design and objectives of the study when they agreed to participate. Therefore they would have known that any cytological abnormality they experienced would have been followed-up in the first instance without treatment, but nevertheless they would have been kept under close clinical surveillance; and that the frequency of study visits would not change according to their cervical cytological status, only the nature of the surveillance. This *should* have meant that the detection of a cervical cytological abnormality would not have affected the likelihood that a woman would return to the study. However, whereas cytological smears were taken in the Brook Advisory Centre (the family planning clinic), to which women would likely be returning for contraceptive supplies anyway, colposcopic assessments were performed in a different, and possibly less convenient, location. And agreement to participate in the study when cytological abnormality was a theoretical possibility is of course very different from being faced with the reality. The colposcopic assessments and histological diagnoses which followed the detection of a cytological abnormality were surely more intrusive and alarming in nature than having a

cytological smear was. Declining to participate in the study on the grounds of inconvenience certainly seems possible. However, refusing to return to the study because of concern over a diagnosis of cytological or histological abnormality seems counterintuitive, since women could address their concerns by requesting treatment at any point. Very few women exercised this right. Therefore, informative missingness, whilst theoretically unlikely, can never be ruled out entirely.

Summary

The study described in this thesis provides repeat observations on both outcome status and covariates, with, for each woman: correlation between observations; a variable number of observations; and a variable time interval between each observation. Outcomes are measured on both continuous and binary scales, can be either single events in time or multiple events measured at several points in time. For time-to-event data, outcomes are interval-censored, which require the use of special analytical methods. Covariates include both time-fixed and time-varying types, which are also known only to change within an interval. The study described in this thesis is also subject to “missing data”, the nature of which must be dealt with appropriately to ensure that valid statistical inferences are made.

Chapter 7

THESIS AIM AND OBJECTIVES

7.1 AIM

To describe the natural history of cervical human papillomavirus (HPV) infection and its relationship to the acquisition of epithelial abnormalities of the cervix in a cohort of young women.

7.2 OBJECTIVES

- To investigate if the adolescent cervix is inherently more susceptible to HPV infection.

    - To measure the association between the time interval between date of menarche and date of first sexual intercourse and the risk of acquiring an incident cervical infection with HPV DNA of any type.

- To investigate the association between cigarette smoking and the risk of cervical HPV infection, the natural history of this infection, and the subsequent development of epithelial abnormalities of the cervix.

- To determine the association between exposure to cigarette smoking and the risk of acquiring an incident cervical infection with HPV DNA of any type, HPV16 and HPV18.

- To determine the association between exposure to cigarette smoking and the duration of an incident cervical infection with HPV DNA of any type, HPV16, and HPV18.

- To determine the association between exposure to cigarette smoking and the risk of acquiring incident high-grade cervical intraepithelial neoplasia

- To investigate how the natural history of cervical HPV infection varies with the humoral immune response.

    - To evaluate the performance of an assay for the measurement of the neutralizing antibody response to cervical infection with HPV16 and HPV18.

    - To describe the kinetics of the neutralizing antibody response in a cohort of young women who were recruited soon after first sexual intercourse, and who first tested positive for HPV16 DNA, or HPV18 DNA, or both, in cervical samples during follow-up.

- To investigate how the natural history of cervical HPV infection varies with viral load, and to determine the association between HPV viral load and the acquisition of epithelial abnormalities of the cervix.

    - To evaluate the performance of an assay for the measurement of HPV16 and HPV18 viral load

    - To describe how the distribution of HPV viral load varies with the infecting HPV type in cervical samples.

    - To describe how viral load changes over time for women with incident cervical HPV16, or HPV18, infections

    - To describe the relationship between HPV16 and HPV18 viral load in cervical samples and the detection of epithelial abnormalities of the cervix.

Chapter 8

PROXIMITY OF MENARCHE TO FIRST SEXUAL INTERCOURSE AND THE RISK OF CERVICAL HPV INFECTION

INTRODUCTION

The process of squamous metaplasia, during which columnar epithelium in the cervix is replaced by squamous epithelium, begins during puberty and is ongoing throughout the reproductive years (see section 2.1.1). One mechanism by which this process is thought to occur is by reserve-cell hyperplasia. Reserve cells, i.e. undifferentiated epithelial cells, proliferate around the exposed endocervical glands and ultimately replace the columnar epithelium with squamous epithelium. It has been suggested that reserve cells in adolescent and young women are especially vulnerable to human papillomavirus (HPV) infection (Singer 1989).

Cross-sectional studies have suggested that, compared with women who delay the start of their sexual career, those who first have sexual intercourse soon after menarche are more susceptible to cervical HPV infection and thus have a greater risk of cervical neoplasia (Shew 1994, Kahn 2002a). However, these studies are limited in so far as outcomes are first defined some years after the woman's first sexual experience, and therefore after the first possible *exposure*; at this point in time, it is not possible to distinguish the effects of early-, late- and cumulative-exposures to HPV. The longer the time interval between date of first sexual intercourse and date of first sampling to test for the presence of HPV DNA in cervical samples, the more likely it is that a woman will have acquired, and then cleared, at

least one cervical HPV infection before first sampling, or possibly have established a latent infection, following which HPV may be only transiently detectable in cervical samples (Wallin 1999). Clearly, the occurrence of outcomes of interest in the interval immediately following the onset of sexual activity, and those factors which bear on it, can only be described when observations are made during this interval.

I investigate, using longitudinal data, the risk of acquiring cervical HPV infection associated with the proximity of first sexual intercourse to menarche, during the first sexual relationship, in women who first had a cervical sample taken within twelve months of the start of their sexual career: both outcome and exposure are thus measured during the most relevant period.

Chapter 9

PROXIMITY OF MENARCHE TO FIRST SEXUAL INTERCOURSE AND THE RISK OF CERVICAL HPV INFECTION

METHODS

Orientation

In this chapter I:

- define the outcome used in the analysis of the proximity of menarche to first sexual intercourse and the risk of cervical HPV infection
- define the study population for the analysis
- describe the statistical methods used in the analysis
- define covariates used to measure sexual behaviour

9.1 RESTATEMENT OF OBJECTIVES

To investigate if the adolescent cervix is inherently more susceptible to HPV infection. Specifically, to measure the association between the time interval from date of menarche to date of first sexual intercourse and the risk of acquiring an incident cervical infection with HPV of any type.

9.2 DEFINITION OF OUTCOME

The outcome for this analysis is the acquisition of an incident cervical HPV infection of any type, subsequent to the date of first sexual intercourse, and prior to the acquisition of a second sexual partner. In this analysis, the method used to detect HPV DNA was the GP5+/GP6+ general primer-mediated PCR system (see section 5.5.4.1).

9.3 STUDY POPULATION

The study population for this analysis is a subset of the full cohort of 2,011 women recruited to the main study, comprising all women who were recruited within twelve months of first having had sexual intercourse and who had yet to acquire a second sexual partner. The study population comprises 474 women.

9.4 STATISTICAL METHODS

Methods appropriate to the analysis of interval censored time-to-event data were used in analyses of time-to-event data.

Time-to-event was measured from the date of first sexual intercourse until the date of the first acquisition of a cervical infection with HPV DNA of any type. Given that cervical HPV infection is a *sexually transmitted* infection, women were assumed to be HPV DNA-negative on their date of first sexual intercourse. Subsequent

observations on HPV status were interval censored: time of acquisition of a cervical infection with HPV DNA of any type was known only to lie in the interval between the date of the first test which yielded a HPV-positive result and the date of the immediately preceding test which yielded a HPV-negative result, or the date of first sexual intercourse if the woman was HPV-positive in her first evaluable sample. Time-to-event was right-ensored on the earliest of: date of acquisition of a second sexual partner; date of diagnosis of high-grade cervical intraepithelial neoplasia; or the date of the last visit. Time-to-event was thus restricted to the period when a woman had only had one sexual partner.

Turnbull's generalisation of the product-limit estimator for interval censored data was used to estimate the cumulative risk of an incident cervical HPV infection of any type following first sexual intercourse, and to construct confidence intervals based on the logarithm of the hazard function (Turnbull 1976). Multivariate analyses were undertaken using HARE (HAzard REgression), an adaptive regression method for analysing interval-censored time-to-event data in the presence of covariates (Kooperberg 1997). Interactions were tested for, and although not so constrained, all models constructed using HARE were proportional hazards models. See section 6.5.1 for a more detailed discussion of these statistical methods.

9.5 DEFINITION OF SEXUAL BEHAVIOUR-RELATED STUDY VARIABLES

A detailed sexual-behaviour history was collected at study entry for all women using an interviewer-completed questionnaire (see section 5.3); these details were then

updated at each visit using another interviewer-completed questionnaire recording changes since the previous visit (see appendices 1 and 2 for these two questionnaires).

Covariates available for analysis were: the time interval between the date at menarche and the date of first sexual intercourse; age at first sexual intercourse; age at menarche; age of sexual partner; sexual experience of partner; and use of barrier contraception.

9.5.1 Age at first sexual intercourse

This is the calendar age at which a woman first experienced penetrative vaginal sexual intercourse.

All women were sexually active prior to the end of follow-up. Age at first sexual intercourse was recorded on the questionnaire collected at the first interview for the vast majority of women (appendix 1). However, among the full cohort of all 2,011 women recruited, 21 had yet to have sexual intercourse at their first interview, and so their age at first sexual intercourse was recorded prospectively.

At their first interview, women were asked two questions relevant to their age at first sexual intercourse. One question, “How old were you when you first had vaginal intercourse”, elicited an explicit age in years. “Loss of virginity” is typically a memorable event, so *age* at first sexual intercourse is likely to have been accurately

recalled. However, the other question, “When did you first have intercourse?”, was recorded as a date, accurate to nearest month and year. Although this date is probably less memorable than age at first sexual intercourse, it could be used as a means of verifying the recorded age. During follow-up, women were asked the date on which they had acquired any new sexual partners since their previous interview, recorded as a full date. Given this date, age at first sexual intercourse could be calculated.

9.5.2 Age of first sexual partner

This is the calendar age of the woman’s first male sexual partner on the date at which the woman first experienced penetrative vaginal intercourse.

Recorded ages were accepted at face value. At their first interview and during follow-up, women were always asked the age of their sexual partner. However, their responses are subject to difficulties with recall, and possibly with the partner lying about their age. One additional potential source of confusion is that women occasionally *appeared* to give the current age of their current partner, rather than the age of that partner on the date they first had sexual intercourse with that partner.

9.5.3 Age at menarche

This is the calendar age of the woman when she first menstruated.

Age at menarche was recorded on the questionnaire collected at the woman's first interview, as an age in years (appendix 1). Getting their first "period" is typically a memorable event for a woman, so age at menarche is likely to have been accurately recalled.

9.5.4 Use of barrier contraception

This is a binary variable recording whether a woman ever used barrier contraception with her partner during sex.

For each of their sexual relationships, women were asked whether they had used barrier contraception during that relationship, and, if so, for how long, i.e. the length of the period during the relationship when barrier contraception was used. Note that, as defined here, this is not strictly speaking a *baseline* variable since it is defined as use *at any time prior to the first evaluable sample*.

9.5.5 Sexual experience of partner

This is a categorical variable, with categories "experienced", "inexperienced" and "unknown", recording whether or not the woman thought that her sexual partner was already sexually experienced prior to first sexual intercourse with her.

At her first interview, for each of her sexual partners prior to that date, women were asked whether they thought that this partner was a "virgin" at the start of their sexual

relationship. During the prospective phase of the study, women were *not* asked about the sexual experience of their partners.

9.5.6 Gynaecological age

In general, this is a term coined to denote the elapsed time since date at menarche. This analysis focuses specifically on gynaecological age *at first sexual intercourse*, the time interval between the date at menarche and the date of first sexual intercourse.

The year and month when first sexual intercourse occurred was known, or at least recorded, for all women; the assumption was made that first sexual intercourse occurred at the midpoint of the month, and therefore date of first sexual intercourse was accurate to within 15 days. Age at menarche (AM) was known, or at least recorded, but the exact date on which menarche occurred is known rarely, if ever, in studies of this nature, and therefore the assumption was made that menarche occurred on the day before her $(AM+1)^{st}$ birthday. Age at menarche is therefore the minimum possible age, but note that if menarche had instead been located on the day of the AM^{th} birthday, this would have yielded identical risk estimates of hazards ratios, given the method used to estimate these quantities. Age at menarche and age at first sexual intercourse were identical for one woman, whose gynaecological age was set to zero.

Chapter 10

PROXIMITY OF MENARCHE TO FIRST SEXUAL INTERCOURSE AND THE RISK OF CERVICAL HPV INFECTION

RESULTS

Orientation

In this chapter I:

- describe the sexual behaviour characteristics of the study population and the association among these variables
- investigate, in univariate analyses, association between sexual behaviour characteristics and the risk of acquiring a cervical HPV infection of any type
- investigate, in univariate analyses, the association between the proximity of first sexual intercourse to menarche and the risk of acquiring a cervical HPV infection of any type
- investigate, in multivariate analyses, the association between the proximity of first sexual intercourse to menarche and the risk of acquiring a cervical HPV infection of any type

10.1 SEXUAL BEHAVIOUR CHARACTERISTICS OF THE STUDY POPULATION

The frequency distribution of the sexual behaviour characteristics of the 474 women comprising the study population used in the analysis described in these chapters of the thesis are shown in table 10.1a. The median calendar age of a woman's first sexual partner was 19 years (range 15 to 42); 14 (3.0%) women had a sexual partner aged 25 years or older. The median calendar age of a woman at first sexual

intercourse was 17 years (range 13 to 19). The median calendar age of a woman at menarche was 13 years (range 8 to 16). Most women used barrier contraception at some point during their first sexual relationship. Just over half of the women believed that their partner was already sexually experienced prior to first sexual intercourse with them, but a substantial proportion of women were unsure.

Table 10.1a. Frequency distribution of the sexual behaviour characteristics of the study population.

Characteristic	Number of women (n=474)	%
Calendar age of first sexual partner (years)		
≤17	109	23.0
18	88	18.6
19	76	16.0
20	50	10.5
21	48	10.1
22	35	7.4
≥23	43	9.1
Calendar age at first sexual intercourse (years)		
≤15	65	13.7
16	155	32.7
17	138	29.1
18	82	17.3
19	34	7.2
Calendar age at menarche (years)		
≤11	70	14.8
12	114	24.1
13	160	33.8
≥14	130	27.4
Use of barrier contraception		
No	126	26.6
Yes	348	73.4
Sexual experience of partner		
Inexperienced	136	28.7
Experienced	245	51.7
Unknown	93	19.6

The frequency distribution of gynaecological age at first sexual intercourse for the study population used in this analysis is shown in table 10.1b. The median gynaecological age at first sexual intercourse of the women was 3.3 years (range 0 to 9.3).

Table 10.1b. Frequency distribution of gynaecological age at first sexual intercourse.

Gynaecological age at first sexual intercourse (years)[a]	Number of women (n=474)	%
[0 , 1)	21	4.4
[1 , 2)	67	14.1
[2 , 3)	104	21.9
[3 , 4)	127	26.8
[4 , 5)	84	17.7
[5 , 6)	46	9.7
[6 , 7)	18	3.8
≥7	7	1.5

[a][x, y) is a mathematical convention meaning values ≥x and <y.

Summary

The median age of a woman's first sexual partner was 19 years (range 15 to 42); 14 (3.0%) women had a sexual partner aged 25 years or older; median age at first sexual intercourse was 17 years (13 to 19); median age at menarche was 13 years (8 to 16); most women used barrier contraception at some point during their first sexual relationship; just over half of the women believed that their partner was already sexually experienced prior to first sexual intercourse with them, but a substantial proportion of women were unsure; and median gynaecological age at first sexual intercourse was 3.3 years (0 to 9.3).

10.2 THE ASSOCIATION BETWEEN GYNAECOLOGICAL AGE AT FIRST INTERCOURSE AND THE SEXUAL BEHAVIOUR CHARACTERISTICS OF THE STUDY POPULATION

10.2.1 The association between gynaecological age at first sexual intercourse and calendar age at first sexual intercourse

The association between gynaecological age and calendar age at first sexual intercourse is shown in table 10.2.1. These variables are positively-correlated: as gynaecological age at first sexual intercourse increases, calendar age at first sexual intercourse increases (correlation coefficients: Pearson's=0.51; Spearman's=0.52; Kendall's Tau=0.40; $p_{two\text{-}sided}<0.01$ in each case). This is as expected, given the definition of gynaecological age at first sexual intercourse and the narrow calendar age range within which women were eligible for recruitment into the study.

Table 10.2.1. The association between gynaecological age at first sexual intercourse and calendar age at first sexual intercourse.

Calendar age at first sexual intercourse (years)	Gynaecological age at first sexual intercourse (years)[a]					
	[0 , 2)[b]	[2 , 3)	[3 , 4)	[4 , 5)	≥ 5	**Total**
≤15	29 (33)	21 (20)	12 (9)	3 (4)	0 (0)	65 (14)
16	39 (44)	50 (48)	41 (32)	15 (18)	10 (14)	155 (33)
17	15 (17)	25 (24)	54 (43)	29 (35)	15 (21)	138 (29)
18	5 (6)	5 (5)	18 (14)	27 (32)	27 (38)	82 (17)
≥19	0 (0)	3 (3)	2 (2)	10 (12)	19 (27)	34 (7)
Total	88 (19)	104 (22)	127 (27)	84 (18)	71 (15)	474 (100)

[a]Frequency counts, i.e. number of women, with percentages in brackets. Percentages are calculated based on column totals, with the exception of those in the final row, which are percentages of the row total; [b][x, y) means ≥x and <y.

10.2.2 The association between gynaecological age at first sexual intercourse and age of first sexual partner

The association between gynaecological age at first sexual intercourse and the age of the woman's first (and for this analysis, only) sexual partner, is shown in table 10.2.2. These variables are only slightly positively correlated (correlation coefficients: Pearson's=0.19; Spearman's=0.17; Kendall's Tau=0.12; $p_{two\text{-}sided}<0.01$ in each case), with no consistent trend noticeable.

Table 10.2.2. The association between gynaecological age at first sexual intercourse and calendar age of first sexual partner.

Age of first sexual partner (years)	Gynaecological age at first sexual intercourse (years)[a]					
	[0 , 2)[b]	[2 , 3)	[3 , 4)	[4 , 5)	≥ 5	Total
≤16	9 (10)	11 (11)	6 (5)	4 (5)	3 (4)	33 (7)
17	18 (20)	22 (21)	18 (14)	13 (15)	5 (7)	76 (16)
18	11 (13)	22 (21)	24 (19)	18 (21)	13 (18)	88 (19)
19	20 (23)	17 (16)	16 (13)	8 (10)	15 (21)	76 (16)
20	5 (6)	8 (8)	16 (13)	10 (12)	11 (15)	50 (11)
21	10 (11)	6 (6)	18 (14)	8 (10)	6 (8)	48 (10)
22	7 (8)	5 (5)	12 (9)	4 (5)	7 (10)	35 (7)
≥23	8 (9)	13 (13)	17 (13)	19 (23)	11 (15)	68 (14)
Total	88 (19)	104 (22)	127 (27)	84 (18)	71 (15)	474 (100)

[a]Frequency count (number of women) with percentage in brackets. Percentages are calculated based on column totals, with the exception of those in the final row, which are percentages of the row total; [b][x, y) means ≥x and <y.

10.2.3 The association between gynaecological age at first sexual intercourse and calendar age at menarche

The association between gynaecological age at first sexual intercourse and calendar age at menarche is shown in table 10.2.3. These variables are negatively-correlated:

as gynaecological age at first sexual intercourse increases, calendar age at menarche decreases (correlation coefficients: Pearson's=–0.70; Spearman's=–0.66; Kendall's Tau=–0.52; $p_{two\text{-}sided}<0.01$ in each case). This is as expected, given the definition of gynaecological age at first sexual intercourse and the narrow calendar age range within which women were eligible for recruitment into the study.

Table 10.2.3. The association between gynaecological age at first sexual intercourse and calendar age at menarche.

Age at menarche (years)	Gynaecological age at first sexual intercourse (years)[a]					
	[0 , 2)[b]	[2 , 3)	[3 , 4)	[4 , 5)	≥ 5	Total
≤11	0 (0)	1 (1)	12 (9)	18 (21)	39 (55)	70 (15)
12	4 (5)	20 (19)	41 (32)	29 (35)	20 (28)	114 (24)
13	17 (19)	50 (48)	54 (43)	27 (32)	12 (17)	160 (34)
≥14	67 (76)	33 (32)	20 (16)	10 (12)	0 (0)	130 (27)
Total	88 (19)	104 (22)	127 (27)	84 (18)	71 (15)	474 (100)

[a]Frequency counts (number of women) with percentages in brackets. Percentages are calculated based on column totals, with the exception of those in the final row, which are percentages of the row total; [b][x, y) means ≥x and <y.

10.2.4 The association between gynaecological age at first sexual intercourse and the use of barrier contraception

The association between gynaecological age at first sexual intercourse and the use of barrier contraception with the first sexual partner is shown in table 10.2.4. The association between these two variables is not statistically significant.

Table 10.2.4. The association between gynaecological age at first sexual intercourse and the use of barrier contraception.

Use of barrier contraception	Gynaecological age at first sexual intercourse (years)[a]					
	[0 , 2)[b]	[2 , 3)	[3 , 4)	[4 , 5)	≥ 5	Total
No	23 (26)	21 (20)	35 (28)	24 (29)	23 (32)	126 (27)
Yes	65 (74)	83 (80)	92 (72)	60 (71)	48 (68)	348 (73)
Total	88 (19)	104 (22)	127 (27)	84 (18)	71 (15)	474 (100)

[a]Frequency counts (number of women) with percentages in brackets. Percentages are calculated based on column totals, with the exception of those in the final row, which are percentages of the row total; [b][x, y) means ≥x and <y.

10.2.5 The association between gynaecological age at first sexual intercourse and sexual experience of partner

The association between gynaecological age at first sexual intercourse and sexual experience of the first sexual partner is shown in table 10.2.5. The association between these two variables is not statistically significant.

Table 10.2.5. The association between gynaecological age at first sexual intercourse and sexual experience of the first sexual partner.

Sexual experience of the first partner	Gynaecological age at first sexual intercourse (years)[a]					
	[0 , 2)[b]	[2 , 3)	[3 , 4)	[4 , 5)	≥ 5	Total
Experienced	26 (30)	38 (37)	34 (27)	21 (25)	17 (24)	136 (29)
Inexperienced	44 (50)	50 (48)	67 (53)	46 (55)	38 (54)	245 (52)
Unknown	18 (21)	16 (15)	26 (21)	17 (20)	16 (23)	93 (20)
Total	88 (19)	104 (22)	127 (27)	84 (18)	71 (15)	474 (100)

[a]Frequency counts (number of women) with percentages in brackets. Percentages are calculated based on column totals, with the exception of those in the final row, which are percentages of the row total; [b][x, y) means ≥x and <y.

10.2.7 The association between age of sexual partner and sexual experience of that partner

A woman with an older sexual partner was more likely to state that her partner was sexually experienced than a woman with a younger sexual partner; a woman's knowledge, or perception, of her partners' sexual experience decreased slightly as the age of her sexual partner increased.

Summary

Gynaecological age and calendar age at first sexual intercourse were positively-correlated (Pearson's correlation coefficient=0.51). Gynaecological age at first sexual intercourse and age of first sexual partner were only slightly positively-correlated (Pearson's correlation coefficient=0.19). Gynaecological age at first sexual intercourse and calendar age at menarche were negatively-correlated (Pearson's correlation coefficient=–0.70). The associations between gynaecological age at first sexual intercourse and both the use of barrier contraception with the first sexual partner, and the sexual experience of that partner, were not statistically significant. A woman with an older sexual partner was more likely to state that her partner was sexually experienced than a woman with a younger sexual partner.

10.3 UNIVARIATE ANALYSIS

With the exception of gynaecological age at first sexual intercourse, all "continuous" variables in the following analyses are nevertheless highly discrete, taking only a very small number of integer values: continuous is thus somewhat of a misnomer.

10.3.1 Incident cervical HPV infections

The median duration of follow-up was 22 months, and the median number of visits was three. During follow-up, 145 (31%) of the 474 women became HPV-positive; the cumulative risk of cervical HPV infection three years after first sexual intercourse was 45.0% (95% CI 37.9% to 51.2%).

10.3.1.1 HPV types detected in the first HPV DNA-positive cervical sample

The frequency distribution of the HPV types detected in the first HPV DNA-positive cervical sample is shown in table 10.3.1.1. Eighty six (59%) of the first HPV-positive samples were non-informative with respect to HPV type: just over half of all cervical HPV infections were uncharacterized types (unknown types which, however, were known to not be among the specific types tested for), with a further 7% being “positive” (HPV-positive samples which were not, or could not be, tested for specific types). Among the 59 women for whom at least one numerical HPV type could be detected in their first HPV-positive cervical sample, the most frequently detected numerical type was HPV16, present in the samples of 28 (19%) women, followed by HPV18, in 16 (11%). HPV6/11, a low-risk type(s), was the only numerical type present for seven women.

Table 10.3.1.1. HPV types detected in the first HPV DNA-positive cervical sample

Types detected in the first HPV-positive cervical sample	Number of women (%) (n=145)	Risk category[a]
58	2 (1)	H
52	2 (1)	H
33	6 (4)	H
31	3 (2)	H
18	10 (7)	H
18;Type X[b]	1 (1)	H
16	16 (11)	H
16;58	1 (1)	H
16;31	1 (1)	H
16;31;Type X[b]	1 (1)	H
16;18	3 (2)	H
16;18;33	1 (1)	H
16;18;31	1 (1)	H
16; 6/11	3 (2)	H
16;52;6/11	1 (1)	H
6/11	7 (5)	L
Uncharacterized types	76 (52)	U
Positive	10 (7)	U

[a]H=high, L=low, U=unknown: conventionally accepted categories based on previous reports of association with cervical cancer; [b]Uncharacterized types: unknown types which, however, were known to *not* be among the specific types tested for.

10.3.2 Calendar age at first sexual intercourse

The results of the univariate regression analysis of the association between calendar age at first sexual intercourse and the risk of an incident cervical HPV infection of any type in women who were first sampled within twelve months of the start of their sexual career is shown in table 10.3.2. When analysed as a continuous variable, there was a significant association between calendar age at first sexual intercourse

and the risk of an incident cervical HPV infection, with an increase in the hazards ratio of 21% (1.21-1.00) for every year of increase in age at first sexual intercourse.

There did not appear to be a non-linear effect with this variable. When calendar age at first sexual intercourse was treated as a categorical variable with five categories (an arbitrary categorisation based upon “reasonable” numbers of women in each category), it remained a significant predictor.

Table 10.3.2. Univariate regression analysis of the association between calendar age at first sexual intercourse and the risk of an incident cervical HPV infection of any type.

Calendar age at first sexual intercourse (years)	Number of women	%	Number HPV-positive	Hazards ratio (95% CI)[a]
≤15	65	13.7	14	1.00 (Reference)
16	155	32.7	39	1.16 (0.63 to 2.14)
17	138	29.1	52	1.89 (1.05 to 3.40)
18	82	17.3	29	1.49 (0.79 to 2.82)
19	34	7.2	11	2.47 (1.12 to 5.44)
Continuous (per year)[b]	474	100.0	145	1.21 (1.05 to 1.40)

[a]Derived from a fixed univariate HARE analysis; [b]variable analysed as a centred continuous variable.

10.3.3 Age of first sexual partner

The results of the univariate regression analysis of the association between the age of the first sexual partner and the risk of an incident cervical HPV infection of any type, in women who were first sampled within twelve months of the start of their

sexual career, is shown in table 10.3.3. When analysed as a continuous variable, there was a significant association between the age of the first sexual partner and the risk of an incident cervical HPV infection of any type, with an increase in the hazards ratio of 8% (1.08-1.00) for every year of increase in age of sexual partner.

There was some evidence of non-linear effects for age of sexual partner: when a transformed variable was constructed by subtracting the mean age of the sexual partner for all women in the study population (19 years) from the age of the sexual partner observed for the woman, the square of this variable was of borderline significance (p=0.062 (Wald test)) when added to a model containing the age of the sexual partner (transformed) alone; however the coefficient of this variable was very small. When age of partner was treated as a categorical variable with seven categories (an arbitrary categorisation based upon “reasonable” numbers of women and events in each category), it remained a significant predictor of the risk of an incident cervical HPV infection of any type; but it is doubtful whether an 8.4% increase for each year of age of partner is a true description of the relationship, possibly reflecting non-linear effects for this variable. It appears that a dichotomous categorisation with ≥22 compared to <22 may be the best categorisation for this variable.

Table 10.3.3. Univariate regression analysis of the association between the age of the first sexual partner and the risk of an incident cervical HPV infection of any type.

Age of first sexual partner (years)	Number of women	%	Number HPV-positive	Hazards ratio (95% CI)[a]
≤17	109	23.0	26	1.00 (Reference)
18	88	18.6	23	1.08 (0.61 to 1.89)
19	76	16.0	14	0.87 (0.45 to 1.66)
20	50	10.5	17	1.21 (0.66 to 2.24)
21	48	10.1	16	1.35 (0.72 to 2.51)
22	35	7.4	20	3.03 (1.69 to 5.44)
≥23	43	9.1	29	2.52 (1.48 to 4.29)
Continuous[b]	474	100.0	145	1.08 (1.05 to 1.13)

[a]Derived from a fixed univariate HARE analysis; [b]variable analysed as a centred continuous variable.

10.3.4 Age at menarche

The results of the univariate regression analysis of the association between calendar age at menarche and the risk of an incident cervical HPV infection of any type, in women who were first sampled within twelve months of the start of their sexual career, is shown in table 10.3.4. When analysed as a continuous variable, age at menarche was not significantly associated with the risk of an incident cervical HPV infection of any type; the point estimate of the hazards ratio suggests a slight decrease in risk as age at menarche increases.

There was some evidence of non-linear effects for age at menarche; when a transformed variable was constructed by subtracting the mean age at menarche for all women in the study population (13 years) from the woman's age at menarche, the square of this variable was of borderline statistical significance (Wald-test p-

value=0.053) when added to a model containing this variable alone; however this model was not itself statistically significant. Analysing age at menarche as a categorical variable with four categories (an arbitrary categorisation based on "reasonable" numbers of women and numbers of events in each category) suggested some fluctuation in the hazards ratio, although it did not yield a statistically significant predictor, either overall or on a category-specific basis.

Table 10.3.4. Univariate regression analysis of the association between calendar age at menarche and the risk of an incident cervical HPV infection of any type.

Age at menarche (years)	Number of women	%	Number HPV-positive	Hazards ratio (95% CI)[a]
≤11	70	14.8	22	1.00 (Reference)
12	114	24.1	29	0.78 (0.44 to 1.38)
13	160	33.8	55	1.11 (0.67 to 1.86)
≥14	130	27.4	39	0.87 (0.51 to 1.49)
Continuous[b]	474	100.0	145	0.97 (0.86 to 1.10)

[a]Derived from a fixed univariate HARE analysis; [b]variable analysed as a centred continuous variable.

10.3.5 Use of barrier contraception

The results of the univariate regression analysis of the association between use of barrier contraception during the first sexual relationship and the risk of an incident cervical HPV infection of any type in women who were first sampled within twelve months of the start of their sexual career, are shown in table 10.3.5. Use of barrier contraception at any time prior to the date of the first evaluable sample was not significantly associated with the risk of an incident cervical HPV infection; the point

estimate of the hazards ratio suggests that the use of barrier contraception results in a decrease in risk (i.e. barrier contraception is protective against cervical HPV infection, as might be expected).

Table 10.3.5. Univariate regression analysis of the association between the use of barrier contraception during the first sexual relationship and the risk of an incident cervical HPV infection of any type.

Use of barrier contraception	Number of women	%	Number HPV-positive	Hazards ratio (95% CI)[a]
Never	126	26.6	39	1.00 (Reference)
Ever	348	73.4	106	0.92 (0.64 to 1.33)

[a]Derived from a fixed univariate HARE analysis; [b]variable analysed as a centred continuous variable.

10.3.6 Sexual experience of partner

The results of the univariate regression analysis of the association between the sexual experience of the first sexual partner and the risk of an incident cervical HPV infection of any type, in women who were first sampled within twelve months of the start of their sexual career, is shown in table 10.3.6. There was a significant association between the sexual experience of the woman's first sexual partner and the risk of an incident cervical HPV infection of any type. Women with sexually experienced partners were almost three times as likely to acquire an incident cervical HPV infection of any type as those with inexperienced partners; women who did not know whether their partner was sexually experience were at twice the risk.

Twenty six (19%) of the 136 women with supposedly sexually inexperienced partners acquired an incident cervical HPV infection of any type. Assuming that, in common with women, men only acquire HPV *via* sexual transmission, this result casts some doubt on the ability of women to distinguish sexually experienced from sexually inexperienced partners. Of course some of the male sexual partners may have had other sexual partners after the start of their sexual relationship with this woman (for those acquiring cervical HPV infections, the median interval (months) between date of first sexual intercourse and date of first detection of cervical HPV infection was 20.40, 2.92 and 9.99 for women with inexperienced, experienced and unknown experience partners, respectively). The unknown category is clearly a mixture of experienced and inexperienced partners, and so it is probably unwise to try and combine this group with either the experienced or unexperienced groups, alone. Note that 67% of women who were first sampled within six months of the start of their sexual career (n=242) reported that they believed that their partner was sexually experienced, compared with 52% in this study population (n=474), with the difference being in the unknown rather than the sexually inexperienced category.

Table 10.3.6. Univariate regression analysis of the association between the sexual experience of the first sexual partner and the risk of an incident cervical HPV infection of any type.

Sexual experience of partner	Number of women	%	Number HPV-positive	Hazards ratio (95% CI)[a]
Inexperienced	136	28.7	26	1.00 (Reference)
Experienced	245	51.7	89	2.79 (1.80 to 4.33)
Unknown	93	19.6	30	1.98 (1.17 to 3.36)

[a]Derived from a fixed univariate HARE analysis; [b]variable analysed as a centred continuous variable.

10.3.7 Gynaecological age at first sexual intercourse

Gynaecological age at first sexual intercourse is measured on a continuous scale, and initially it was analysed as a continuous variable. It was then further analysed as both a categorical variable, and as a series of binary variables, with different gynaecological ages at first sexual intercourse used for dichotomisation. The results of these univariate analyses of the association between gynaecological age at first sexual intercourse and the risk of an incident cervical HPV infection of any type, in women who were first sampled within twelve months of the start of their sexual career, are shown in table 10.3.7.

When analysed as a continuous variable, there was a significant association between gynaecological age at first sexual intercourse and the risk of an incident cervical HPV infection, with an increase in the hazards ratio of 12.9% (1.129≈1.13-1.00) for every year of increase in gynaecological age at first sexual intercourse. When a transformed variable (GAC) was constructed by subtracting the mean gynaecological age at first sexual intercourse for all women in the study population (3.3 years) from the observed gynaecological age at first sexual intercourse, the square of this variable was significant (p=0.03) when added to a model containing this variable alone, indicating a departure from linearity. However, in practice the effect of departures from linearity appear to be slight: model GAC gives an increase in the hazards ratio of 12.9% for each year of increase in the interval, compared to 13.2% for model GAC+GAC^2. This is also consistent with the fact that a model containing GAC alone is selected using the automated model-building option with

HARE. On the grounds of parsimony, results from the linear model are therefore accepted.

Three of the five variables which were constructed to analyse gynaecological age at first sexual intercourse as a categorical variable yielded significant predictors of the risk of an incident cervical HPV infection of any type, with an increased risk associated with greater gynaecological age at first sexual intercourse. The exceptions were the dichotomous variables based on comparing ≥4 or ≥5 with a complementary reference group: the direction of the effect was the same, with women having gynaecological age ≥5 years being at no greater risk than women with gynaecological age <5; similarly for ≥4 versus <4.

Table 10.3.7. Univariate regression analysis of the association between gynaecological age at first sexual intercourse and the risk of an incident cervical HPV infection of any type.

Gynaecological age at first sexual intercourse (years)	Number of women	%	Number HPV positive	Hazards ratio (95% CI)[a]
Continuous (per year)[b]	474	100.0	145	1.13 (1.02 to 1.25)
<2	88	18.6	19	1.00 (Reference)
2-3	104	21.9	28	1.16 (0.65 to 2.07)
3-4	127	26.8	47	1.89 (1.11 to 3.22)
4-5	84	17.7	28	2.10 (1.17 to 3.76)
≥5	71	15.0	23	1.63 (0.89 to 2.98)
<2	88	18.6	19	1.00 (Reference)
≥2	386	81.4	126	1.66 (1.01 to 2.71)
<3	192	40.5	47	1.00 (Reference)
≥3	282	59.5	98	1.72 (1.22 to 2.44)
<4	319	67.3	94	1.00 (Reference)
≥4	155	32.7	51	1.34 (0.96 to 1.89)
<5	403	85.0	122	1.00 (Reference)
≥5	71	15.0	23	1.09 (0.70 to 1.70)

[a]Derived from a fixed univariate HARE analysis; [b]variable analysed as a centred continuous variable.

Summary

In univariate analyses, the hazards ratio (HR) of cervical HPV infection increased significantly with age at first sexual intercourse (HR=1.21 (95% CI 1.05 to 1.40) per year), partner age (HR=1.08 (95% CI 1.05 to 1.13) per year), and when women reported a sexually experienced partner (HR=2.79 for an experienced compared to an inexperienced sexual partner; 95% CI 1.80 to 4.33); the interval between menarche and first sexual intercourse (gynaecological age at first sexual intercourse) was a significant predictor of the risk of infection, with an increase in the HR of 12.9% for every year of increase in this interval (95% CI 2.1% to 24.9%).

10.4 MULTIVARIATE ANALYSIS

Forward and backward stepwise multivariate regression analyses identified as significant independent predictors of incident cervical HPV infection of any type: gynaecologic age at first sexual intercourse; age of sexual partner; sexual experience of partner; and use of barrier contraception. Calendar age at first sexual intercourse and calendar age at menarche were not significant predictors. In this section, I consider how each of these variables influences the association between gynaecological age at first sexual intercourse and the risk of an incident cervical HPV infection of any type.

10.4.1 Gynaecological age at first sexual intercourse and the use of barrier contraception

In a multivariate regression analysis, use of barrier contraception was identified as an independent predictor of the risk of incident cervical HPV infection of any type, but the magnitude of its effects were small and its deletion from the final model had little impact on the estimates of effects of the remaining variables. Barrier contraception will therefore not be considered further.

10.4.2 Gynaecological age at first sexual intercourse, calendar age of sexual partner and sexual experience of partner

The results of a multivariate regression analysis of the association between gynaecological age at first sexual intercourse and the risk of an incident cervical HPV infection of any type, controlling for age and sexual experience of partner, in women who were first sampled within twelve months of the start of their sexual career, are shown in table 10.4.2.

Although gynaecological age at first sexual intercourse was found to be an independent predictor of the risk of incident cervical HPV infection of any type, when analysed as a continuous variable, it was of borderline statistical significance, and when analysed as a categorical variable the confidence intervals for each category included the null value of one, although the increasing risk with increasing gynaecological age at first sexual intercourse is still apparent. The same set of independent predictors were identified when gynaecologic age at first sexual intercourse was dichotomised into $\geq$3 years versus <3 years, but in this case gynaecological age was clearly significant (p=0.009); however the variable comparing $\geq$2 years versus <2 years was not. In all models, age of sexual partner was selected as a significant predictor.

In a model containing gynaecological age at first sexual intercourse and calendar age of sexual partner, both treated as continuous variables, gynaecological age at first sexual intercourse was not statistically significant, but there was some evidence

of an interaction between the two variables. When gynaecological age at first sexual intercourse was dichotomised into ≥3 years versus <3 years, age of sexual partner "explained" some of the effect of gynaecological age at first sexual intercourse, but gynaecological age at first sexual intercourse remained statistically significant; age of sexual partner also remained statistically significant, with a slightly reduced hazards ratio, and there was no evidence of an interaction. When non-linear effects of age of sexual partner were included, the hazards ratio for this dichotomisation of gynaecological age at first sexual intercourse was reduced only slightly, with the non-linear effect of age of sexual partner remaining statistically significant. In this model, the hazards ratio for gynaecological age at first sexual intercourse was closer to that from the model containing only independent effects, than to that from the model containing the interaction, which is surprising since one would expect the categorical variable to 'capture' in some sense the non-linear effect of age of sexual partner: there was no interaction between the variables when both were analysed as categorical. However, the hazards ratios for gynaecological age at first sexual intercourse categorised as ≥3 years versus <3 years, were very similar in all three models.

Although age of sexual partner and sexual experience of partner may be expected to be positively-correlated, they were both still significant predictors of the risk of acquiring an incident cervical HPV infection of any type when included in a model together, and there was no evidence of an interaction. In a model including gynaecological age at first sexual intercourse, either as a continuous variable alone or with non-linear effects, and sexual experience of partner, the significance of

gynaecological age at first sexual intercourse was not explained by sexual experience of partner. In a model including gynaecological age at first sexual intercourse as a continuous variable with non-linear effects, experience and age of sexual partner as a categorical variable, the variables measuring gynaecological age at first sexual intercourse remained significant.

When gynaecological age at first sexual intercourse was categorised into three categories [0,2), [2,3), [3,∞), (the non-dichotomous categorisation of this variable with the greatest statistical significance in a univariate regression analysis), the significance of this variable dropped just below the 5% level, i.e. gynaecological age at first sexual intercourse became non-significant, after controlling for age of sexual partner, either as a continuous or as a categorical variable. When age of sexual partner was categorised into a dichotomous variable with categories ≥22 years and <22 years, gynaecological age at first sexual intercourse with the three categories became significant once more. Again, gynaecological age at first sexual intercourse with this categorisation remained significant when only experience of sexual partner was included.

Table 10.4.2. Multivariate regression analysis of the association between gynaecological age at first sexual intercourse and the risk of an incident cervical HPV infection of any type.

Gynaecological age at first sexual intercourse (years)	Number of women	%	Number HPV positive	Hazards ratio (95% CI)[a]
Continuous (per year)[b]	474	100.0	145	1.10 (0.99 to 1.22)
<2	88	18.6	19	1.00 (Reference)
2-3	104	21.9	28	1.07 (0.61 to 1.90)
3-4	127	26.8	47	1.62 (0.96 to 2.72)
4-5	84	17.7	28	1.74 (0.98 to 3.09)
≥5	71	15.0	23	1.42 (0.78 to 2.58)
<2	88	18.6	19	1.00 (Reference)
≥2	386	81.4	126	1.55 (0.95 to 2.51)
<3	192	40.5	47	1.00 (Reference)
≥3	282	59.5	98	1.58 (1.11 to 2.25)
<4	319	67.3	94	1.00 (Reference)
≥4	155	32.7	51	1.27 (0.90 to 1.78)
<5	403	85.0	122	1.00 (Reference)
≥5	71	15.0	23	1.06 (0.68 to 1.67)

[a]derived from a fixed multivariate HARE analysis, controlling for age of partner as a centred continuous variable, and sexual experience of partner; [b]variable analysed as a centred continuous variable.

10.4.3 Gynaecological age at first sexual intercourse and calendar age at first sexual intercourse

Given the positive correlation between gynaecological age at first sexual intercourse (G) and calendar age at first sexual intercourse (A) (section 10.2.1) and also the association between calendar age at first sexual intercourse and the risk of an incident cervical HPV infection of any type (K) (section 10.3.2), it is possible that the

association observed between gynaecological age at first sexual intercourse and the risk of an incident cervical HPV infection of any type is merely a (weaker) reflection of the former association, i.e. that the association between K and the variable G-A is merely a consequence of the association between K and G, with the subtraction of A being of no effect, except possibly to attenuate the observed association.

The results of a multivariate regression analysis of the association between gynaecological age at first sexual intercourse and the risk of an incident cervical HPV infection of any type, controlling for calendar age at first sexual intercourse, age of sexual partner and sexual experience of partner, is presented in table 10.4.3.

Table 10.4.3. Multivariate regression analysis of the association between gynaecological age at first sexual intercourse and the risk of an incident cervical HPV infection of any type, controlling for calendar age at first sexual intercourse, and for calendar age at first sexual intercourse, age of partner and sexual experience of partner.

Gynaecological age at first sexual intercourse (years)	Number of women	%	Number HPV positive	Hazards ratio (95% CI)		
				Uncontrolled	Controlled for AFI[a]	Controlled for AFI+PA[b]+PX[c]
Continuous (per year)	474	100	145	1.13 (1.02 to 1.25)	1.17 (0.88 to 1.54)	1.21 (1.12 to 1.31)
<2	88	19	19	1.00 (Reference)	1.00 (Reference)	1.00 (Reference)
≥2	386	81	126	1.66 (1.01 to 2.71)	1.15 (0.58 to 2.27)	1.20 (1.11 to 1.30)
<3	192	41	47	1.00 (Reference)	1.00 (Reference)	1.00 (Reference)
≥3	282	60	98	1.72 (1.22 to 2.44)	1.08 (0.66 to 1.77)	1.12 (1.04 to 1.21)
<4	319	67	94	1.00 (Reference)	1.00 (Reference)	1.00 (Reference)
≥4	155	33	51	1.34 (0.96 to 1.89)	1.22 (0.87 to 1.69)	1.26 (1.17 to 1.37)
<5	403	85	122	1.00 (Reference)	1.00 (Reference)	1.00 (Reference)
≥5	71	15	23	1.09 (0.70 to 1.70)	1.12 (0.83 to 1.52)	1.17 (1.08 to 1.26)

[a]Calendar age at first sexual intercourse; [b]age of sexual partner; [c]sexual experience of the woman's sexual partner (as stated by the woman).

In univariate analyses of the association between gynaecological age at first sexual intercourse and the risk of an incident cervical HPV infection of any type, gynaecological age at first sexual intercourse was a significant predictor of risk when it was included in a model as either a centred continuous variable or as a categorical variable dichotomised into ≥2 years versus <2 years or ≥3 years versus <3 years. However, all of these variables were no longer statistically significant once the model had been controlled for calendar age at first sexual intercourse. Compared to it's magnitude in the univariate (uncontrolled) analysis, the magnitude of the hazards ratio for gynaecological age at first sexual intercourse decreased when this variable was measured as a categorical variable dichotomised into ≥2 years versus <2 years or ≥3 years versus <3 years. This is what one might expect if the observed association between gynaecological age at first sexual intercourse and the risk of an incident cervical HPV infection of any type was due to the association between calendar age at first sexual intercourse and the risk of cervical HPV infection (although there are other explanations). However, the magnitude of the hazards ratio for gynaecological age at first sexual intercourse *increased* compared to it's magnitude in the univariate analysis when this variable was measured as a either a centred continuous variable, or as a categorical variable dichotomised into ≥5 years versus <5 years; when this variable was measured as a categorical variable dichotomised into ≥4 years versus <4 years, the hazards ratio was attenuated only slightly, essentially remaining the same. However, when the analysis was controlled for age of sexual partner and sexual experience of partner, in addition to calendar age at first sexual intercourse, all hazards ratios increased slightly compared to their values in the analysis controlling for calendar age at first sexual intercourse alone,

and every variable measuring gynaecological age at first sexual intercourse became statistically significant

The association between gynaecological age at first sexual intercourse and the risk of an incident cervical HPV infection of any type does not appear to be due to the association between calendar age at first sexual intercourse and the risk of cervical HPV infection. No matter how gynaecological age at first sexual intercourse is measured in this analysis, there appears to be an increase, not a decrease, in the risk of an incident cervical HPV infection of any type with increasing gynaecological age at first sexual intercourse

10.4.4 Gynaecological age at first sexual intercourse and calendar age at menarche

Analysing these two variables simultaneously presents difficulties, because the study population is restricted to women who first had intercourse within a short interval before recruitment, at which time they were between the ages of 15 and 19 years, but menarche occurred between the ages of eight and 16 years; so a cross-tabulation of these two variables reveals many zero-cells, and cells with no, or few, events.

In a multivariate regression analysis in which gynaecological age at first sexual intercourse and calendar age at menarche were treated as continuous variables, age at menarche was not statistically significant. There was an interaction between

these variables, but given that both have non-linear effects, it may only be appropriate to look at models where both variables are categorical. Age at menarche, when analysed as a categorical variable, was not significant when gynaecological age at first sexual intercourse was treated as a continuous variable, and there was no evidence of an interaction. Age at menarche, when analysed as a categorical variable, was not significant when gynaecological age at first sexual intercourse was categorised into [0,2), [2,3), [3,∞); the model failed to converge when looking for an interaction in this analysis.

10.4.5 Age of partner and calendar age at first sexual intercourse

Women whose first coital episode was at a later age were more likely to form relationships with older men, and the association between incident cervical HPV infection of any type and calendar age at first sexual intercourse was no longer significant when this variable was controlled for in the multivariate analysis.

Summary

In a multivariate analysis, compared with women who first had sexual intercourse within three years of menarche, those who postponed first sexual intercourse beyond this time had a greater risk of cervical infection with HPV of any type (HR=1.58; 95% CI 1.11 to 2.25), after controlling for age and sexual experience of partner.

Women who first have intercourse soon after menarche are not at greater risk of cervical HPV infection than women who delay the onset of sexual activity. The adolescent cervix does not appear to be inherently more susceptible to HPV infection.

Chapter 11

PROXIMITY OF MENARCHE TO FIRST SEXUAL INTERCOURSE AND THE RISK OF CERVICAL HPV INFECTION

DISCUSSION

I have described, using longitudinal observations, how the risk of acquiring cervical HPV infection of any type varies with the time interval between menarche and first sexual intercourse in 474 women aged 15 to 19 years recruited within twelve months of first sexual intercourse and before the acquisition of a second sexual partner. One hundred and forty-five women became HPV-positive; the cumulative risk of acquiring a cervical HPV infection of any type three years after first sexual intercourse was 45% (95% CI 38% to 51%). In univariate analyses, the hazards ratio (HR) of acquiring an incident cervical HPV infection of any type increased significantly with increasing age at first sexual intercourse (per year of increase in age, HR=1.21; 95% CI 1.05 to 1.40), partner age (per year of increase in age, HR=1.08; 95% CI 1.05 to 1.13), and when women reported a sexually experienced partner (for women with an experienced compared to an inexperienced partner, HR=2.79; 95% CI 1.80 to 4.33); the time interval between date of menarche and date of first sexual intercourse was a significant predictor of infection, with an increase in the HR of 13% for every year of increase in this time interval (95% CI 2% to 25%). In a multivariate analysis, compared with women who first had intercourse within three years of menarche, those who postponed first sexual intercourse beyond this time had a greater risk of infection (HR=1.58; 95% CI 1.11 to 2.25), after controlling for age and sexual experience of partner.

The results described in this thesis demonstrate that the risk of an incident cervical HPV infection of any type increases as the time interval between menarche and first sexual intercourse increases, after controlling for other determinants of infection. This finding is biologically plausible, given that the life-cycle of HPV is dependent on host-cell replication and differentiation, which are characteristic of the squamous metaplasia occurring in the cervical transformation zone (see sections 2.1.1 and 3.1.2.3); basal cells, the target cells for HPV infection, are most accessible in the transformation zone which, in young women, increases in size with time from menarche, independent of sexual and other reproductive factors (Jacobson 1999, Howley 2001).

Findings from this analysis contradict those from studies which suggest that a short interval between menarche and first sexual intercourse increases the risk of HPV infection (Shew 1994, Kahn 2002a). In these studies, women were first sampled some years after first sexual intercourse and, as the authors of these reports concede, there is no evidence that infection was acquired during the hypothesised period of increased vulnerability. In contrast, the study population used in the analysis described in this thesis comprises women first sampled within a year of first sexual intercourse and, unlike the studies cited, was also restricted to women who had had only one sexual partner. By restricting the analysis to the period prior to the start of the second sexual relationship, confounding due to multiple sexual partners is eliminated. This additional restriction allowed a more precise examination of the role of the “male factor” in the acquisition of cervical HPV infection of any type (Kjaer 2001, Kahn 2002b).

The study population for this analysis comprises all women who were recruited within twelve months of first having sexual intercourse, and who had yet to acquire a second sexual partner. To be eligible, women were not required to be in a sexual relationship during the entire period they were followed-up, nor were women who had only a brief sexual relationship with their first partner distinguished from those with longer relationships. The period 12 months was chosen because we had previously found that the median time to acquisition of the (presumed) first cervical HPV infection was over three years, and the median duration of cervical HPV infection was 13.7 months (range 8.0 to 25.4) (Woodman 2001). It was therefore unlikely that a woman recruited so soon after first sexual intercourse would have both acquired and cleared her first cervical HPV infection before she was first tested for the presence of HPV DNA. Estimates of the cumulative risk of the incidence of cervical HPV infection of any type were stable when the study population was based on either a six-month interval between first sexual intercourse and study entry, as used in a previous analysis (Collins 2002), or a twelve-month interval, as used in the analysis described in this thesis. This provides reassurance that the larger sample size provided by the “twelve-month population” did not compromise the greater proximity of first sexual intercourse and study entry in the “six-month population”.

Times-to event were measured from date of first sexual intercourse until the date of the first detection of a cervical HPV infection of any type. Women were assumed to be negative for cervical infection with HPV DNA of any type on their date of first sexual intercourse: *cervical* HPV infection is a *sexually transmitted* infection acquired as a result of penetrative vaginal intercourse. Date of first sexual intercourse is thus

a woman's first opportunity to acquire this *sexually transmitted* disease, and so this assumption appears reasonable.

Older sexual partners were more likely to be reported as sexually experienced and, among women with sexually experienced partners, the risk of acquiring a cervical HPV infection of any type increased with increasing age of sexual partner. The "sexual experience of partner" variable was based on information collected from the woman: she was asked whether she knew, or thought, her partner had previous sexual experience prior to the start of her sexual relationship with that partner. However, the nature of that sexual experience was not further ascertained, e.g. penetrative sexual intercourse was not distinguished from other types of sexual activity. Responses had to be accepted at face value, with no opportunity to verify the accuracy of this variable with the partners themselves: there are no data on the partner's number of sexual partners.

It was necessary to define use of barrier contraception as use at any time prior to study entry, since first use was not reported consistently. The assumption is made that use of barrier contraception is a marker for low-risk sexual behaviour, and *consistent* condom use has been reported to be protective against the acquisition of cervical HPV infection (Winer 2006). In the analysis described in this thesis, use of barrier contraception (which primarily means condoms, but not exclusively) was slightly protective, but the association was neither strong nor statistically significant. Given that many women were also using oral contraceptives, barrier contraception may only have been used intermittently: data was not collected with sufficient

resolution to determine whether this was the case. Thus, only limited inferences are possible with this variable.

The work in this chapter of the thesis was published in 2005 (Collins 2005: see appendix 3). Subsequently, a similar study citing this work (for ease of reference, the "repeat study") was undertaken in 130 university women aged 18 to 24 years who reported first sexual intercourse within three months of enrolment into the study, or during follow-up, and who had further follow-up after first sexual intercourse (Winer 2008). Biological samples were taken at time intervals of four months, with sexual behaviour updated at intervals of two-weeks *via* a web-based diary. HPV DNA was detected using PCR.

In a univariate analysis, the repeat study found an association with gynaecological age at first sexual intercourse in the opposite direction to that found in the study described in this thesis: women reporting menarche two to five years before first sexual intercourse were more likely to acquire HPV than women reporting menarche eight to 14 years before first sexual intercourse (HR 2.57; 95% CI 1.13 to 5.82); women with a gynaecological age at first sexual intercourse of six years were at a substantially reduced risk of acquiring an incident cervical HPV infection of any type, although this category was not significantly different from the null value. When the woman's sexual partner's number of previous sexual partners, gynaecological age at first sexual intercourse, and sexual partner's age were entered into a multivariate model, gynaecological age at first sexual intercourse was not statistically significant. As in the study described in this thesis, in univariate analyses women with sexually

experienced partners and older sexual partners were more likely to acquire *genital* HPV infection, but age at first sexual intercourse was not a significant risk factor *for* the acquisition of genital HPV infection.

The authors of the repeat study did not undertake an analysis appropriate to interval-censored time-to-event data. The results of the interval censored analyses used in the analysis described in this thesis were tested for “plausibility” using Cox proportional hazards regression, assuming that events occurred at the midpoint of an interval in which the first detection of HPV was known to have occurred (see section 6.4); the authors of the repeat study imputed the time of event at the right-hand end of the interval. There were no material differences between the results of the interval censored and the midpoint analyses in the analysis described in this thesis. And given that the repeat study used a shorter time interval between study visits than that used in the study described in this thesis (four compared to six months), the conclusions reached by the authors of the repeat study are also likely to have remained materially unaltered had they undertaken an analysis appropriate to interval-censored time-to-event data, provided that late-occurring events were not overly influential (see section 6.4.1.3). That being said, it is not clear why the authors of the repeat study did not undertake the appropriate analysis, with the example of the publication based on the analysis described in this thesis before them.

The cumulative incidence of *genital* HPV infection 36 months after reporting a first sexual partner was 49% (95% CI 36% to 64%), similar to that found in the study described in this thesis for *cervical* HPV infections of any type, specifically.

Interestingly, five women were reported to have tested positive for HPV before first sexual intercourse, and were excluded from the repeat study; four were positive for a single type (types 6, 51, 53, and 66), and one for multiple types (types 40, 61, and 66): the authors of the repeat study state that there are plausible explanations for this, although they don't provide them. Presumably, these infections were at superficial sites, rather than cervical HPV infections. Indeed, the repeat study only describes *genital* HPV infections, obtained by combining results from self-collected vaginal swabs and clinician-collected cervical and vulvovaginal swabs; cervical HPV infections are not separately distinguished. Therefore the repeat study was unable to address the hypothesis of interest directly.

The authors of the repeat study state that "previous studies have suggested that the interval between menarche and first intercourse is related to the risk of HPV infection, but the directions of the observed associations have been inconsistent. Given that 95% of participants in the repeat study were at least 5 years past menarche at the time of first intercourse, we were unable to assess whether short intervals between menarche and first intercourse increase the risk of infection from a first partner".

This appears to be an attempt by the authors of the repeat study to overstate the novelty of their findings. The previous studies cited are that based on the work described in this thesis, and the two studies which the work described in this thesis has now superseded, for reasons stated both above and in the published paper. In fact, the criticism which the authors correctly level at their own work is precisely the

weakness which the work in this thesis is uniquely able to address. The mean gynaecological age at first sexual intercourse in the study population on which this thesis was based was 3.4 years (range 0 to 9), compared to 6.9 (range 2 to 14) in the repeat study. The study described in this thesis, and reported in the published paper, *was* thus able to assess whether short intervals between menarche and first sexual intercourse increased the risk of acquiring cervical HPV infection from a first sexual partner.

The observations described in this thesis suggest that the adolescent cervix is *not* at greater risk of cervical HPV infection *per se*. Postponement of first sexual intercourse will not reduce a woman's risk of acquiring a cervical HPV infection, and the risk is high even among women engaged in low-risk sexual behaviour.

11.1 PUBLICATION

As stated above, a paper based on the work described in chapters 8, 9, 10 and 11 of this thesis was published in 2005 (Collins 2005). This paper is reproduced in appendix 3.

Chapter 12

SMOKING, CERVICAL HPV INFECTION, AND THE RISK OF CERVICAL NEOPLASIA

INTRODUCTION

Orientation

In this chapter I discuss:

- the mechanisms by which smoking may have an impact on the risk of acquiring cervical HPV infection and epithelial abnormalities of the cervix
- the evidence of the association between smoking and epithelial abnormalities of the cervix
- the evidence of the association between sexual behaviour and epithelial abnormalities of the cervix
- the evidence of the association between smoking and cervical HPV infection
- the limitations of previous studies which have investigated the association between smoking and epithelial abnormalities of the cervix, controlling for cervical HPV infection

12.1 BACKGROUND

Cervical human papillomavirus (HPV) infection is considered a necessary, but not sufficient, cause of cervical cancer, and attention is now focused on identifying cofactors which modulate the progression of cervical HPV infection to high-grade cervical intraepithelial neoplasia (CIN) and invasive disease (Castellsague 2003).

Cigarette smoking is one potential cofactor. Cigarette smoking may act through any of several biological mechanisms, either alone or in combination with other factors, to increase the risk of acquiring cervical HPV infection and cervical neoplasia. Smoking results in systemic exposure to compounds which damage DNA and increase the risk of cancer in various organs; in particular, these compounds are known to cause genomic damage in squamous epithelial cells (Castle 2002b). The components of tobacco smoke reach cervical cells *via* the bloodstream and by diffusion through tissue; high contents of the chemical components of cigarette smoke, e.g. nicotine, and its major metabolite cotinine, as well as other tobacco-specific carcinogens, have been detected in the cervical mucus of smokers. These components may act as carcinogens (i.e. they may initiate the development of malignant tumours) or as mutagens (i.e. they may cause an increase in the rate of mutation), with high concentrations of these components leading to damage to the cervical epithelium, thus making it more susceptible to HPV infection. The *order* of exposures may therefore also be important (Olsen 1998). *In vitro* studies show that benzo[a]pyrene, a carcinogen found in cigarette smoke, can stimulate genome amplification and virion synthesis in cell lines productively infected with HPV (Alam 2008). Or smoking may influence the course (persistence, viral load, extent) of a HPV infection *via* its impact on the immune response. Smoking has been associated with a “generalized suppression of the immune system”, which includes: decreases in the number of natural killer (NK) cells, and in NK cell activity; decreases in circulating levels of IgG and IgA; and reductions in numbers of cervical Langherans' cells (Giuliano 2002).

12.2 THE ASSOCIATION BETWEEN SMOKING, SEXUAL BEHAVIOUR, CERVICAL HPV INFECTION AND EPITHELIAL ABNORMALITIES OF THE CERVIX

In this section I provide an overview of the evidence regarding the association between smoking and epithelial abnormalities of the cervix, and potential confounders of that association.

12.2.1 Association between smoking and epithelial abnormalities of the cervix

While it should be noted that not all studies have reported a statistically significant association between smoking and epithelial abnormalities of the cervix (Koutsky 1992, Olsen 1995, Sasagawa 1997, Schiff 2000), the weight of evidence suggests that there *is* a positive association, at least for squamous disease.

A meta-analysis has previously been undertaken of 72 epidemiological studies of cervical disease and cigarette smoking (Haverkos 2003). Thirty-nine of the studies had cervical cancer as an outcome, 23 non-malignant cervical disease, and 10 included patients with both cancer and non-malignant disease. Of the 71 studies tabulated in the report of the meta-analysis (one study with an Asian population appears to be missing), 50 are case-control studies, 10 cross-sectional, and 11 longitudinal studies.

Results of the meta-analysis of the association between cervical disease and cigarette smoking are presented in table 12.2.1. Overall, a significant association

was found between cigarette smoking and both malignant and non-malignant disease. This was the case whether exposure status was measured as current smoking status, or as cumulative history of smoking, although the strongest associations were those between disease and current smoking status. Odds ratio estimates were all approximately equal to two, indicating only a small increase in risk. However, given the large number of smokers in the population, and the high rates of smoking amongst young women in particular, this increased risk has serious public health implications.

Table 12.2.1. Results of a previously reported meta-analysis of 72 epidemiological studies of cervical disease and cigarette smoking.

Outcomes included	**Odds ratio; 95% CI; Number of *studies***	
	Ever smokers	***Current smokers***
Cervical cancer	1.72 (1.63 to 1.82); n=31	2.25 (1.91 to 2.19); n=23
Non-malignant disease	1.87 (1.70 to 2.05); n=18	2.19 (1.99 to 2.41); n=20
Mixed disease	2.16 (1.89 to 2.46); n=9	2.45 (2.08 to 2.89); n=8
Total	1.80 (1.72 to 1.88); n=58	2.13 (2.02 to 2.25); n=51

A similar “pooled” analysis of 23 epidemiological studies based on approximately 14,000 women was undertaken more recently, with a substantial overlap between the studies involved in this pooled analysis and the earlier meta-analysis (International Collaboration of Epidemiological Studies of Cervical Cancer 2006a). However, the pooled analysis separately distinguished between squamous cell carcinoma of the cervix (SCCC) and adenocarcinoma, and reported associations for smoking exposures at a greater resolution.

Compared to women who had never smoked, women who were current smokers were at a significantly increased risk of both SCCC (RR=1.46; 95% CI 1.32 to 1.61) and CIN3 (RR=1.83; 95% CI 1.61 to 2.08). The risks for ex-smokers were lower for both disease outcomes, but there was no trend with time since stopping smoking. In analyses combining SCCC and CIN, in current smokers the risk of squamous disease increased with increasing number of cigarettes smoked per day, and with decreasing age at starting smoking; there was no such evidence for ex-smokers. There was no association between the risk of squamous disease and the duration of smoking. There was no association between smoking and the risk of adenocarcinoma, or adenocarcinoma *in situ*, irrespective of how smoking exposure was measured, with smoking appearing to have a slightly protective effect.

Smoking may be a more important risk factor for CIN2 and CIN3 than for CIN1, suggesting that this exposure may act at a later stage during the pre-invasive phase of the natural history of cervical neoplasia (Cuzick 1990, Parazzini 1992, Schiffman 1993, Brisson 1994, Ho 1998c, Franco 2002).

12.2.2 The association between sexual behaviour and cervical disease

Various aspects of sexual behaviour are known to be strong risk factors for cervical disease (see chapter 2), and sexual behaviour is also associated with smoking behaviour (International Collaboration of Epidemiological Studies of Cervical Cancer 2006a). Therefore, we would anticipate that the association between smoking and cervical disease would be confounded by sexual behaviour.

Several studies which have attempted to control for confounding of the smoking-cervical disease association by sexual behaviour have reported that smoking is an independent risk factor for CIN (Lyon 1983, La Vecchia 1986, Brock 1989, Cuzick 1990, Gram 1992, Parazzini 1992, Munoz 1993, Becker 1994, Brisson 1994, De Vet 1994, Nobbenhuis 1999, Mosckiki 2001, Franco 2002); some have also reported an increasing trend in the risk of CIN with increasing number of cigarettes smoked, and with increasing duration of smoking (Lyon 1983, La Vecchia 1986, Brock 1989, Becker 1994, Brisson 1994, De Vet 1994). In the pooled analysis, which includes some of these studies, four potential confounding factors were included in analyses: lifetime number of sexual partners, age at first sexual intercourse, parity, and oral contraceptive use. Only stratification by lifetime number of sexual partners altered the estimates of relative risk more than marginally.

The association between sexual behaviour and the risk of cervical disease is almost certainly entirely due to the association between sexual behaviour and cervical HPV infection. Therefore the most appropriate control of confounding of the association between smoking and cervical disease by sexual behaviour is obtained by controlling for exposure to cervical HPV infection directly, rather than by using surrogate measures.

12.2.3 The association between cervical HPV infection and smoking

Is cervical HPV infection a potential confounder of the association between smoking and cervical disease? To determine this it must be shown that cervical HPV infection is associated with both smoking and cervical disease. Cervical HPV infection is the

pre-eminent risk factor for cervical disease in both smokers and non-smokers. The reported associations between smoking and the acquisition of HPV infections, or the duration of those infections, present a conflicting picture. One longitudinal study found no association between smoking and the prevalence of HPV at study entry (Kotloff 1998), whilst one did (Minkoff 2004). Four longitudinal studies have reported no association between the risk of an incident HPV infection and smoking, including one recruiting women aged 13 to 21 years from two family planning clinics (Ho 1998b, Kotloff 1998, Mosckiki 2001, Sellors 2003). Two longitudinal studies have reported a statistically significant association, including one in HIV-seropositive and at-risk HIV-seronegative women, and one studying genital-tract HPV infections (Minkoff 2004, Winer 2003). Smoking has been reported to: have a significantly protective effect against persistent HPV infection in two studies (Hildesheim 1994, Ho 1998b); have no association in four studies (Kotloff 1998, Sellors 2003, Winer 2003, Minkoff 2004); and to be associated with an increased duration of HPV infection, with a dose response with duration of smoking, in another (Giuliano 2002).

Any reported association implies that cervical HPV infection has the potential to confound the measurement of the association between smoking and cervical disease, and so cervical HPV infection must be adequately controlled for in an analysis.

12.2.4 The association between smoking and cervical disease controlling for cervical HPV infection

Case-control studies restricted to women who test positive for HPV DNA have consistently reported an increased risk of cervical neoplasia in smokers, with some reporting a dose-response effect (Olsen 1998, Kjaer 1998, Deacon 2000, Hildesheim 2001, Plummer 2003, Harris 2004, McIntyre-Seltman 2005). A longitudinal study, restricted to women who tested positive for HPV DNA, has also reported an increased risk of CIN3 or worse in those who were smokers at enrolment (Castle 2002b). Another longitudinal study, similarly restricted, has reported an increased risk of low-grade squamous intraepithelial lesions associated with daily smoking, but not with pack-years of smoking (Moscicki 2001). However, these studies are limited in so far as women were only tested for HPV on a single occasion, and therefore the duration of infection prior to the onset of disease is unknown. Residual confounding by time from first HPV exposure could still explain the excess risk associated with smoking, if women who smoke are more likely to have had an early exposure to HPV, and if early exposure is also a risk factor for disease.

In the cohort study described in this thesis, it was possible to avoid residual confounding by time from first exposure to cervical HPV infection by making repeated measurements of smoking and HPV status in a cohort of young women, who were free of disease and HPV DNA-negative at study entry. Thus the design of the cohort study described in this thesis allowed not only the measurement of the risk of high-grade CIN in relation to changes in smoking and HPV status, but also the

exploration of the impact of smoking on the acquisition and duration of incident cervical HPV infections.

Chapter 13

SMOKING, CERVICAL HPV INFECTION, AND THE RISK OF CERVICAL NEOPLASIA

METHODS

Orientation

In this chapter I:

- define the outcomes used in the analysis of smoking, cervical HPV infection, and the risk of epithelial abnormalities of the cervix
- describe the study populations used in these analyses
- describe the statistical methods used in these analyses
- describe the collection of smoking data
- describe the method used to interpolate dates of smoking “events”
- describe the construction of variables used to measure exposure to smoking
- describe the variables used to measure sexual behaviour
- describe the construction of variables used to measure exposure to cervical HPV infection

13.1 RESTATEMENT OF OBJECTIVES

To investigate the association between cigarette smoking and: the risk of incident cervical HPV infection; the natural history of incident cervical HPV infection; and the subsequent development of epithelial abnormalities of the cervix.

- To determine the association between exposure to cigarette smoking and the risk of acquiring an incident cervical infection with HPV DNA of any type, HPV16, and HPV18.

- To determine the association between exposure to cigarette smoking and the duration of an incident cervical infection with HPV DNA of any type, HPV16, or HPV18.

- To determine the association between exposure to cigarette smoking and the risk of acquiring incident high-grade cervical intraepithelial neoplasia

13.2 DEFINITION OF OUTCOMES

The time to the acquisition of an incident cervical HPV infection of any type measured from the date of study entry until the date of the first detection of HPV DNA in a cervical sample. Equivalent definitions were used in the analyses of incident cervical HPV16 and HPV18 infections.

Duration of an incident cervical HPV infection measured from the date of the cervical sample in which HPV DNA was first detected until the date of the first subsequent sample which tested negative for HPV DNA. Equivalent definitions were used for the analysis of the duration of incident cervical HPV16 and HPV18 infections.

Time to the diagnosis of incident high-grade cervical intraepithelial neoplasia (CIN) was measured from the date of study entry until the date of the first diagnosis of high-grade CIN (a histological diagnosis of CIN2 or CIN3).

In these analyses, the method used to detect HPV DNA was the GP5+/GP6+ general primer-mediated PCR system (see section 5.5.4.1).

13.3 STUDY POPULATION

The study population for a given analysis varied according to the outcome analysed.

13.3.1 The study population for the analysis of incident cervical HPV infections

In analyses of incident cervical HPV infections, of any type, or of types 16 or 18, the study population was restricted to women who were HPV DNA-negative in their first evaluable cervical sample, cytologically normal in their first evaluable smear, and who had further follow-up after the earliest of these two events.

13.3.2 The study population for the analysis of the duration of incident cervical HPV infections

13.3.2.1 The study population for the analysis of the duration of incident cervical HPV infections of any type

Women who had an incident cervical HPV infection of any type, and who had further follow-up after the first detection of HPV DNA.

13.3.2.2 The study population for the analysis of the duration of incident cervical HPV16 infections

Women who had an incident cervical HPV16 infection, and who had further follow-up after the first detection of HPV16 DNA.

13.3.2.3 The study population for the analysis of the duration of incident cervical HPV18 infections

Women who had an incident cervical HPV18 infection, and who had further follow-up after the first detection of HPV18 DNA.

13.3.3 The study population for the analysis of incident high-grade CIN

In analyses of incident high-grade CIN, the study population was restricted to women who were HPV DNA-negative in their first evaluable cervical sample, cytologically normal in their first evaluable smear, and who had further follow-up after this time.

13.4 STATISTICAL METHODS

Methods appropriate to the analysis of interval censored time-to-event data were used.

Observations on the acquisition of cervical HPV infection of any type, or of types 16 or 18, and the acquisition of high-grade CIN were interval censored; time of onset was known only to lie in the interval between the date of the visit at which they were first detected and the date of the immediately preceding visit. Similarly, time to clearance of HPV infection was known only to lie in the interval between the date of the visit at which HPV was last detected and the date of the immediately following visit. All outcomes were thus interval censored time-to-event data.

Time to incident cervical HPV infection of any type was measured from the date of study entry until the interval containing the date of the first acquisition of HPV DNA of any type in a cervical sample, with censoring at the date of the end of follow-up. Equivalent definitions were used in the analysis of incident cervical HPV16 and HPV18 infections.

The duration of an incident cervical HPV infection of any type was measured from the date of the first detection of HPV DNA of any type in a cervical sample until the interval containing the date of the first subsequent sample which tested HPV DNA-negative, with censoring at the end of follow-up. Equivalent definitions were used for the analysis of the duration of incident cervical HPV16 and HPV18 infections.

The time to the diagnosis of incident high-grade CIN was measured from the date of study entry until the interval containing the date of the first diagnosis of high-grade CIN, with censoring at the date of the end of follow-up.

The date of the end of follow-up was the earliest of: the date of diagnosis of high-grade CIN; the date of treatment; the date of the last clinical follow-up (the latest of: the date of the last smear; the date of the last colposcopy visit).

Estimates of the cumulative risk of an event of a given type were obtained using a non-parametric maximum likelihood estimator (Turnbull 1976) (see section 6.4.1). Estimates of hazards ratios were obtained in univariate and multivariate analyses using a semi-parametric method for modelling interval-censored time-to-event data with time-varying covariates, as a generalized linear model (see section 6.5.2 for further details of this method).

Time-varying covariates were assigned their current values at each study visit. Construction of time-varying smoking exposure variables is described in following sections. When testing the statistical significance of linear trends of smoking

variables, an indicator variable for ever having smoked, or current smoking status, as appropriate, was also included: this is equivalent to restricting the analysis to ever, or current, smokers, respectively (Leffondre 2002). Sexual behaviour variables available for analysis (see section 13.8 below) were only retained in final models if they were significant in multivariate analyses which included smoking.

Tests of statistical significance were conducted at the 5% two-sided significance level using likelihood ratio tests, and 95% confidence intervals (CI) were constructed as appropriate.

13.5 THE COLLECTION OF SMOKING DATA

Smoking data was collected *via* in-person interview, using one questionnaire at study entry to collect retrospective data, and another during follow-up to collect "prospective" data (see section 5.3 and appendices 1 and 2). Smoking status was self-reported, with no further objective verification.

13.5.1 Retrospective data collection

At recruitment, women were asked the following questions regarding their smoking history (see appendix 1):

1) Have you ever regularly smoked cigarettes?

2) How old were you when you started?

3) Do you still smoke?

4) If you have given up, how old were you?

5) How many cigarettes per day?

6) Are the cigarettes filter, non-filter or handmade?

The possible answers were: for questions 1 and 3, Y (yes) or N (no); for questions 2 and 4, an age in years; for question 5, a number corresponding to a categorised range (1:1-9 2:10-19 3:20-29 4:30-39 5:40+); and for question 6, a number corresponding to one of the three categories (filter, non-filter or handmade).

The use of the word “regularly” in question 1 was crucial: this should have eliminated the women who tried smoking once or twice, but then stopped.

The recording of answers to question 6 was abandoned early during the study (hence this question does not appear on the questionnaire in appendix 1) since very few of these young women smoked anything other than filter-cigarettes. For example, within the study population for the incidence of cervical HPV infection: for 951 women this field was left blank on the database, with 119, one and four in the filter, non-filter and handmade categories, respectively.

In the coding rules for the entry of data onto the computerized database, it states that smoking (i.e. the answers to questions 1-6) is recorded in a field of 8 characters and gives the following example: a person who started smoking at age 16 and gave up at age 21, and who had smoked 15 filter cigarettes a day, would be coded

Y16N2121. This implies that even people who had given up smoking prior to study entry should have a value recorded for the amount they smoked prior to giving up (note that the restriction in age at recruitment was changed after these instructions were drawn up): this rule was not applied consistently. Of 91 women within the cohort of 1,075 who said that they were non-smokers at their first interview, but had been regular smokers in the past, 37 (41%) had no recorded value for the number they smoked. The vast majority of these women were among the last third recruited to the study, i.e. study numbers after 1241, or women recruited on or after 9th June 1990.

13.5.2 Prospective data collection

At follow-up (prospective) interviews, women were asked the following questions regarding their smoking history (see appendix 2):

7) Are you [a] cigarette smoker?
8) Have you changed the number you smoke since [the] last visit?
9) If smoking, how many per day?

The answers possible were: for question 7, Y (yes) or N (no); for question 8, a number corresponding to a categorised change (1:stopped, 2:same, 3:started, 4:cut down and 5:increased); and for question 9, a number corresponding to a categorised range (1:1-9, 2:10-19, 3:20-29, 4:30-39 and 5:40+).

Note that the word “regular” no longer appears, and that there is no question regarding the type of cigarettes smoked.

Question 8 was subsequently considered to be unreliable for various reasons. For example, it was felt that, when talking to health professionals, women were naturally inclined to say that they were smoking less than they were at the previous visit, even when there was no apparent change in the quantity smoked (i.e. they would give “socially desirable answers”). This variable was certainly internally inconsistent, with at least one woman saying she had cut down since the last visit when, according to the recorded information, she had been a non-smoker at the previous visit. Data collected in response to this question was not analysed.

Comment: This questionnaire was supposed to be used to record changes in status since the previous visit: the questionnaire has ‘DETAILS-change only’ printed at the top of the first sheet. This policy was not adhered to, with sequences of identical results recorded for various questions; also, apparently, the questionnaire was not always taken to the interview with the woman, so it is not clear how this policy could have been adhered to.

13.6 INTERPOLATION OF “SMOKING EVENT” DATES

A “smoking event” is defined as an occurrence of a change in smoking status e.g. starting or stopping smoking, or changing the number of cigarettes smoked per day.

The smoking data has limited resolution for detecting smoking events. Distinguishing the number of times a woman started and stopped smoking was limited by the number of interviews she had, and the proximity of interviews to smoking events. For example, if a woman who was smoking at her first interview stopped for a period and then started smoking again before her second interview, this pattern of behaviour would have gone unrecorded. If, however, she had been interviewed during the brief period of time during which she had stopped smoking, the fact that she had stopped *would* have been recorded.

The date on which a woman started smoking was interval censored: it was not known exactly. If she began smoking prior to her first interview, the date on which she started smoking was known to lie within a given year of age. Similarly, if once having started smoking, she subsequently stopped prior to her first interview, the date of this event was only known to lie within a given year of age for the woman. In some cases, women who reported that they had started and stopped smoking prior to recruitment, again reported that they were smoking at recruitment.

Interpolation was used to determine an “exact” date for smoking events. The most complex interpolations occured in the retrospective data, with interpolation

depending on the proximity of the ages at which the woman started smoking, stopped smoking, and was recruited. The possible patterns of smoking and recruitment are shown in figure 13.6a: note that for ease of illustration only, events are shown as occurring in consecutive years of age. The dates of smoking events were interpolated by equally distributing the dates of events within the year of age (i.e. between the relevant consecutive birthdays) in which they were known to occur, with additional restrictions based on recruitment date and smoking status at that time.

For example, in figure 13.6a *panel a*, there are three dates of smoking events to be interpolated between date of birthday Bd_b and date of recruitment R: the date of first starting smoking, the date of stopping smoking and the date of starting smoking for the second time, since this woman was a smoker at recruitment. These three dates are determined by calculating the time interval in days (d) between the date of birthday Bd_b and the date of recruitment, both of which are known exactly. This interval is then divided by the number of events plus one, in this case four. The nth event (n=1,2,3) thus occurs at the date given by $Bd_b+[n \times (d/4)]$ (figure 13.6b *panel b*). Similarly, figure 13.6b *panels b* and *c* illustrate the interpolation of two and one smoking events, respectively.

After recruitment, smoking events can only correspond to a single change in smoking status between interviews; consequently, all dates of smoking events are interpolated at the midpoint between the interview dates at which smoking status was determined.

Figure 13.6a. The ten possible patterns of smoking events prior to recruitment. S=Started smoking; P=Stopped smoking; R=Recruitment, with smoking status at recruitment (+: smoking, –: not smoking); Bd_i=Birthday, age i.

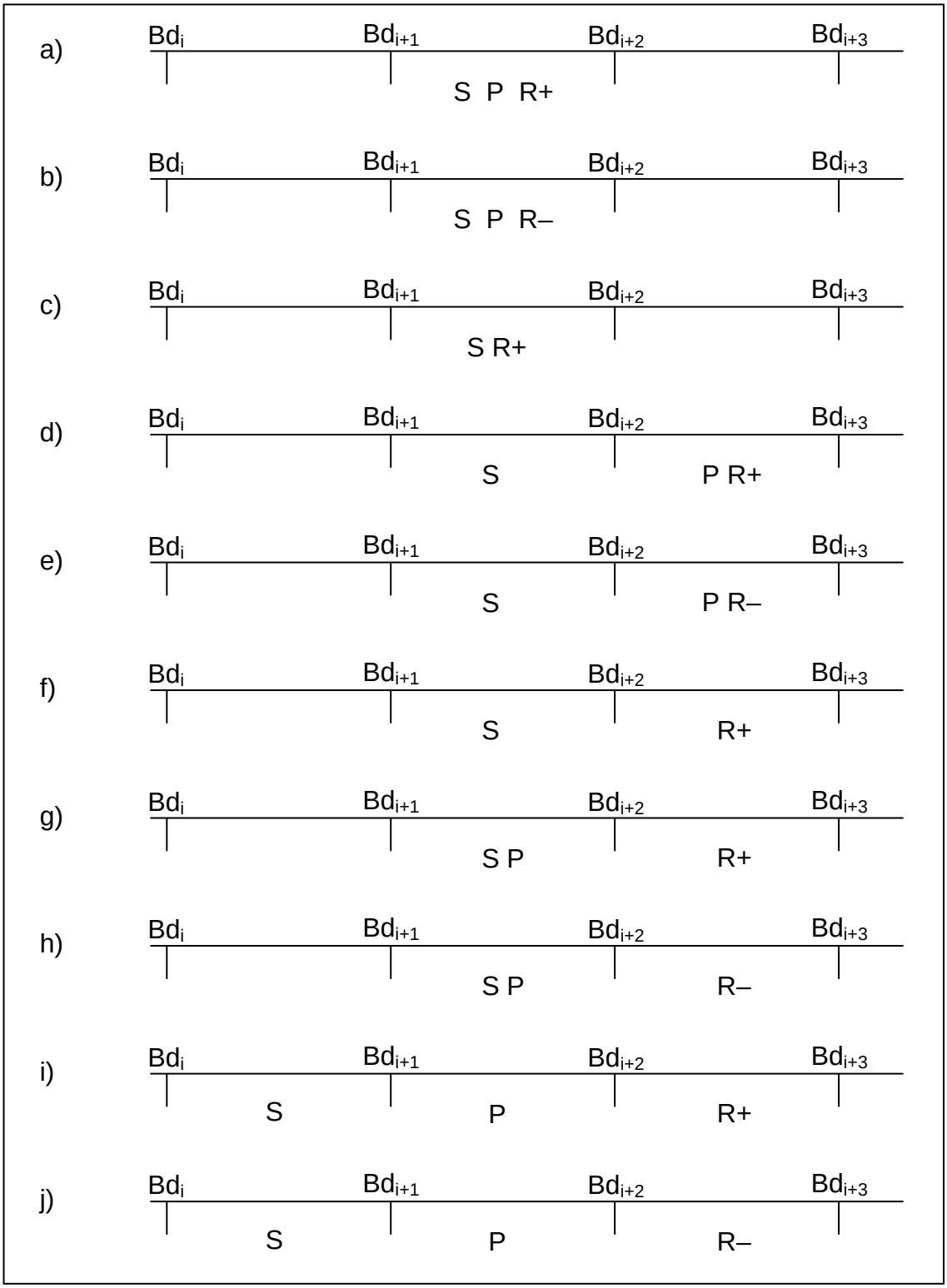
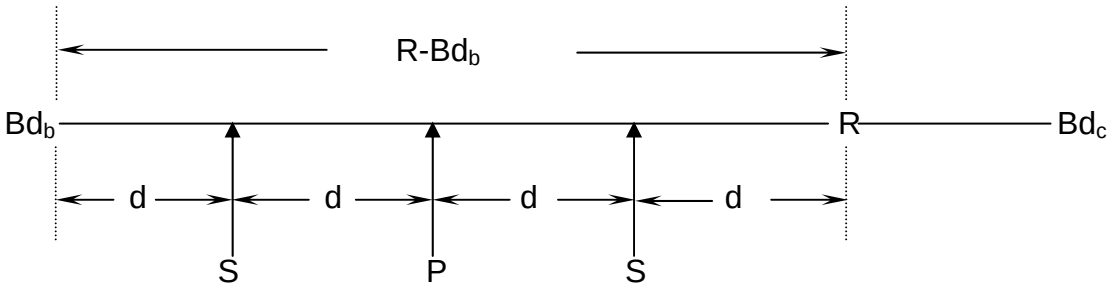


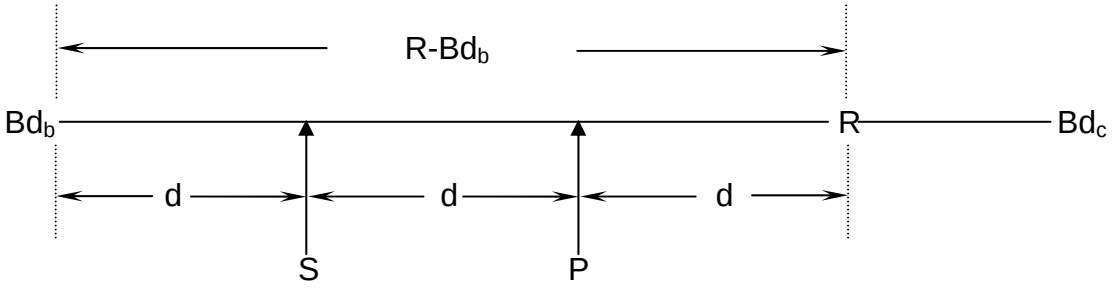
Figure 13.6b. Interpolation of smoking events prior to recruitment. S=Started smoking; P=Stopped smoking; R=Recruitment, with smoking status at recruitment (+: smoking, –: not smoking); Bd_i=Birthday, age i.

a) Three events in the same year of age.



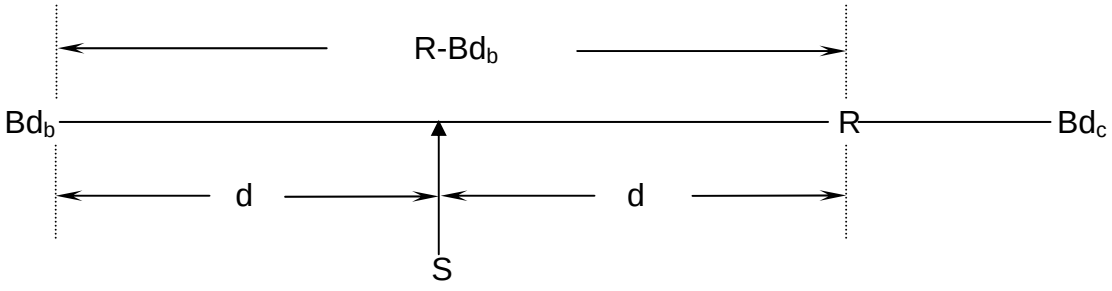
where $d = (R\text{-}Bd_b)/4$.

b) Two events in the same year of age.



where $d = (R\text{-}Bd_b)/3$, and d is measured in days.

c) One event in a given year of age.



where $d = (R\text{-}Bd_b)/2$, and d is measured in days.

13.7. DEFINITION OF SMOKING-RELATED STUDY VARIABLES

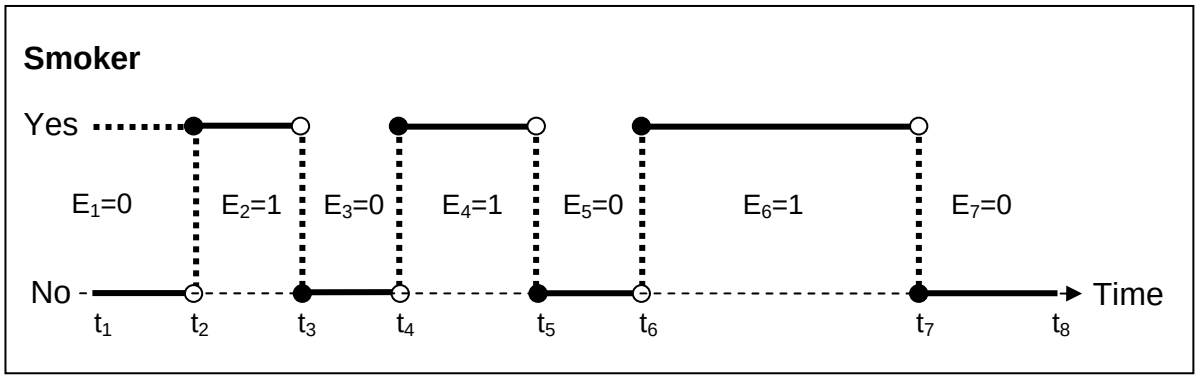
Interpolated smoking dates, as described in the previous section, were used to determine the values for all smoking exposure variables. Suppose a woman's smoking status is known to change at the m times t_i, i=1,2,...,m. For ease of notation, a change in smoking status is said to have occurred when smoking status is determined for the first time (i.e. at t_1) and at end of follow-up (i.e. at t_m). Define indicator variables E_j, j=1,2,...,m-1 for a woman's smoking status in each of the intervals $[t_i, t_{i+1})$ (if T is in the period [a,b) then $a \leq T < b$), i.e.:

$E_j = 1$ if the woman is a smoker during the period $[t_i , t_{i+1})$

0 if the woman is not a smoker during the period $[t_i , t_{i+1})$

for j=1,2,...m-1. Figure 13.7 is an illustration of how the smoking status of a woman may vary with time (Note: a solid circle means that the point is included; an open circle means that the point is excluded).

Figure 13.7. An illustrative example of how the smoking status of a woman may vary with time. t_i, i=1,2,...,8, are the times at which the womans smoking status is known to change; E_j, j=1,2,...,7, are indicators of current smoking status during the interval $[t_j, t_{j+1})$ (E_j=1 if the woman is a smoker, and 0 otherwise).



13.7.1 Smoking status at study entry

Two definitions of “smoking status at study entry” were used in analyses. The first definition distinguished ex-smokers, i.e. women who had once smoked but had since stopped, from women who had *never* smoked. With this definition, smoking status at study entry is a time-independent, i.e. time-fixed, categorical covariate with three levels, “never smoker”, “ex-smoker” and “current smoker”. At study entry: women who had never been regular smokers were categorised as “never smoker”; women who had been regular smokers, but who had stopped and were no longer regular smokers, were categorised as “ex-smokers”; women who were still regular smokers were categorised as “current smokers”. Women who were regular smokers at study entry, but who also stated that they had stopped smoking at some point prior to study entry, were still categorized as current smokers.

Some of the outcome variables analysed in this section of the thesis had a small total number of associated events; this meant that the ex-smoker category often had no events, so it was not possible to estimate risks for this category. To enable analysis of this variable for all outcomes, and to enable comparisons of risk estimates to be made across these outcomes, a second definition of “smoking status at study entry” was also used in analyses. The second definition combined women who had quit smoking prior to study entry with those who had never smoked, into a single “non-smoker” category. With this definition, smoking status at study entry is a time-fixed *binary* covariate with categories “non-smoker” and “current smoker”.

Note that the second definition is sub-optimal in that the reference group (non-smokers) combines women with some degree of exposure (current smokers and ex-smokers) with those who have none (never-smokers); in contrast, the first definition has a clear-cut non-exposed reference group.

13.7.2 Smoking history: ever versus never

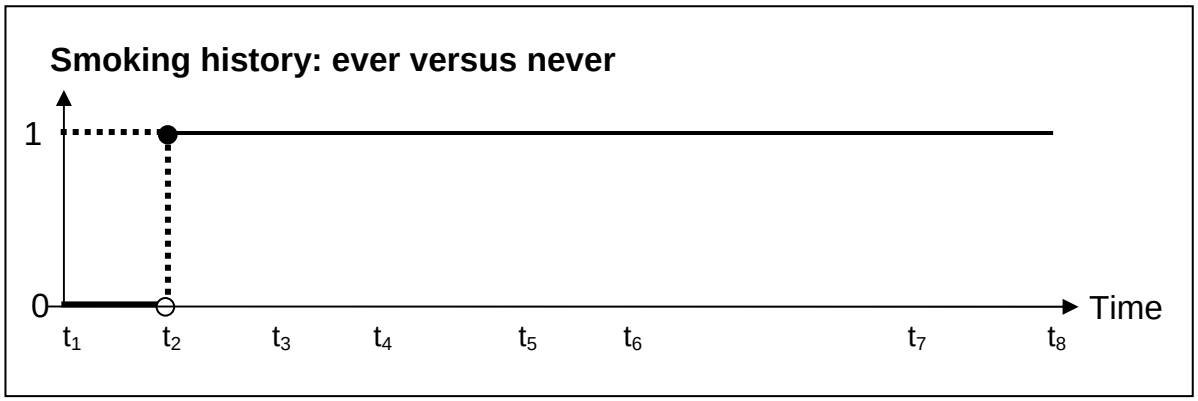
Smoking history (ever versus never) is a time-varying binary covariate. For the purposes of computation, it was defined as taking the value zero prior to the date of first starting smoking, and the value one thereafter. The value of this variable at time t is given by the indicator function $S_1(t)$:

$$S_1(t)=I[\, t \geq T_S \,]$$

where t is time and T_S is the time at which the woman first started smoking.

Figure 13.7.2 presents the graph of S_1, calculated as the indicator function $I[t \geq t_2]$, as a function of time for the woman illustrated in figure 13.7.

Figure 13.7.2. The time-varying binary covariate "smoking history (ever versus never)" ($S_1(t)=I(t \geq t_2)$) for the woman illustrated in figure 13.7 as a function of time.



13.7.3 Current smoking status

Current smoking status is a time-varying categorical covariate. Two definitions of this variable were used in analyses. The first definition combined women who had quit smoking (ex-smokers) with those who had never smoked (never-smokers), into a single "non-smoker category". This definition yields a binary covariate taking the value zero for a non-smoker, and one for a current smoker, constructed using the following method. Define the time-varying indicator function $V_i(t)$:

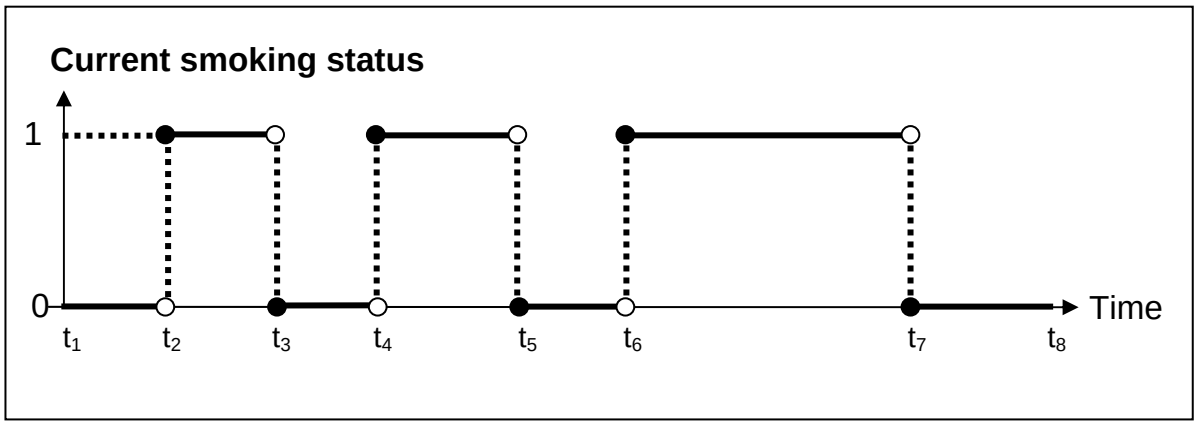
$$V_i(t) = I(t_i \leq T < t_{i+1}) \qquad i=1,2,...m-1$$

Then the current smoking status (binary) of the woman at time t, $S_2(t)$, is given by:

$$S_2(t) = \sum_{i=1}^{m-1} E_i V_i(t)$$

S_2 takes the value zero for a non-smoker and one for a current smoker. The graph of S_2 as a function of time for the woman illustrated in figure 13.7 is presented in figure 13.7.3; note that this figure effectively reproduces figure 13.7.

Figure 13.7.3a. The time-varying binary covariate "current smoking status" ($S_2(t)$) for the woman illustrated in figure 13.7 as a function of time: **two categories** (0=non-smoker, 1=current smoker).



The second definition distinguishes women who had smoked at some point in time but who were currently non-smokers (ex-smokers), from women who had never smoked (non-smokers). The second variable, a three-level categorical variable taking the value zero for a non-smoker, one for an ex-smoker, and two for a current smoker, was derived from the first variable using the following method. Define the time-varying indicator function T(t):

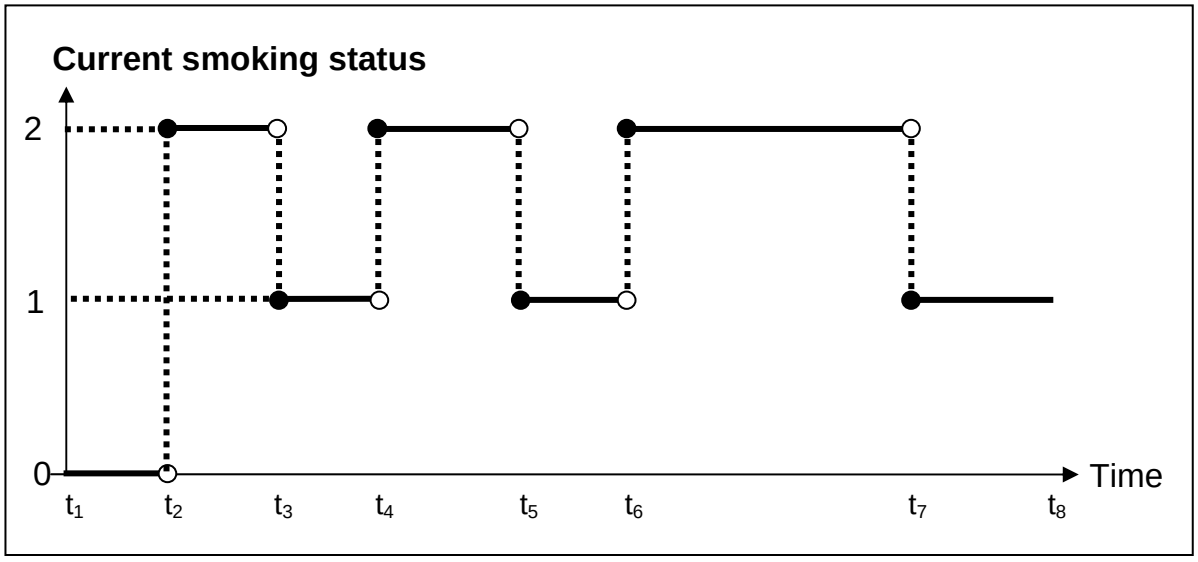
$$T(t) = I(S_2(t)>0)$$

Then the current smoking status (three-categories) of the woman at time t, $S_3(t)$, is given by:

$$S_3(t) = 2S_2(t)T(t)+S_1(t)(1–T(t))$$

The graph of S_3 as a function of time for the woman illustrated in figure 13.7 is presented in figure 13.7.3b.

Figure 13.7.3b. The time-varying categorical covariate "current smoking status" ($S_3(t)$) for the woman illustrated in figure 13.7 as a function of time: three categories (0=never-smoker, 1=ex-smoker, 2=current smoker).



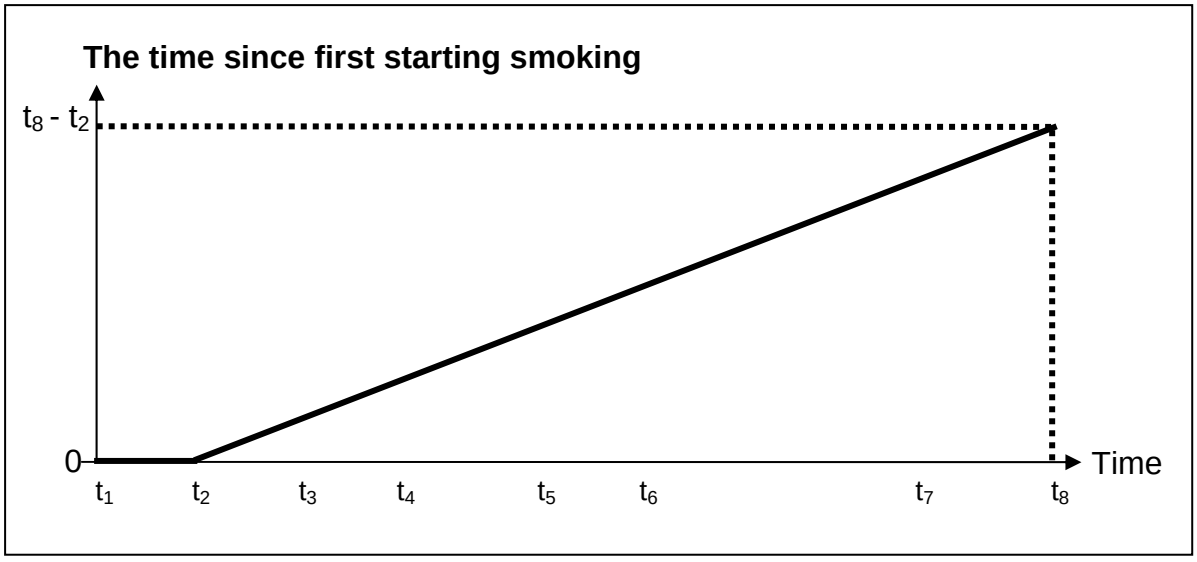
13.7.4 The time since first starting smoking

The time since first starting smoking is a time-varying covariate measuring the total amount of time since a woman first became a smoker; the value of this variable thus increases even during the periods of time when the woman was a non-smoker. The time since first starting smoking at time t ($S_4(t)$) is defined as follows:

$$S_4(t)=(t-T_S)S_1(t)$$

where t is current time and T_S is the time at which the woman first started smoking. The graph of S_4 as a function of time for the woman illustrated in figure 13.7 is presented in figure 13.7.4.

Figure 13.7.4. The time-varying continuous function "the time since first starting smoking" ($S_4(t)$) for the woman illustrated in figure 13.7 as a function of time.



Time since first starting smoking is only well-defined for women who have started smoking at some point; for women who have never smoked it is not defined (a value

of zero would be incorrect). This variable was therefore analyzed as a continuous covariate in a model which also included a time-varying covariate indicating whether a woman had ever smoked, effectively restricting the analysis to ever-smokers for this variable, but allowing maximum information to be used for other variables in the analysis.

This variable was also analysed as a categorical covariate with categories chosen because they were “logical”, iin this case multiples of whole years, and had “sufficient” numbers of events in each category; these categories were chosen prior to the estimation of risks. When analyzed as a categorical covariate, the reference category was never-smokers: this category is labelled “non-smokers”, which is strictly correct since all women are non-smokers until they start smoking, and only women who never start smoking remain in this category.

13.7.5 The cumulative duration of smoking

Smoking status can change over time, i.e. not all women who start smoking remain smokers thereafter; and some women who quit smoking subsequently re-start. Time since first starting smoking, as analyzed in the previous section, assumes that risk “accumulates” even while a woman is a non-smoker. In contrast, the cumulative duration of smoking distinguishes smoking and non-smoking time periods. The cumulative duration of smoking is a time-varying covariate measuring the total amount of time during which a woman has been an active smoker; the value of this variable remains constant during the periods of time when the woman was a non-

smoker. This variable is constructed using the following method. Define the time-varying indicator function U(t):

$$U_i(t) = I(t \geq t_i)$$

Define:

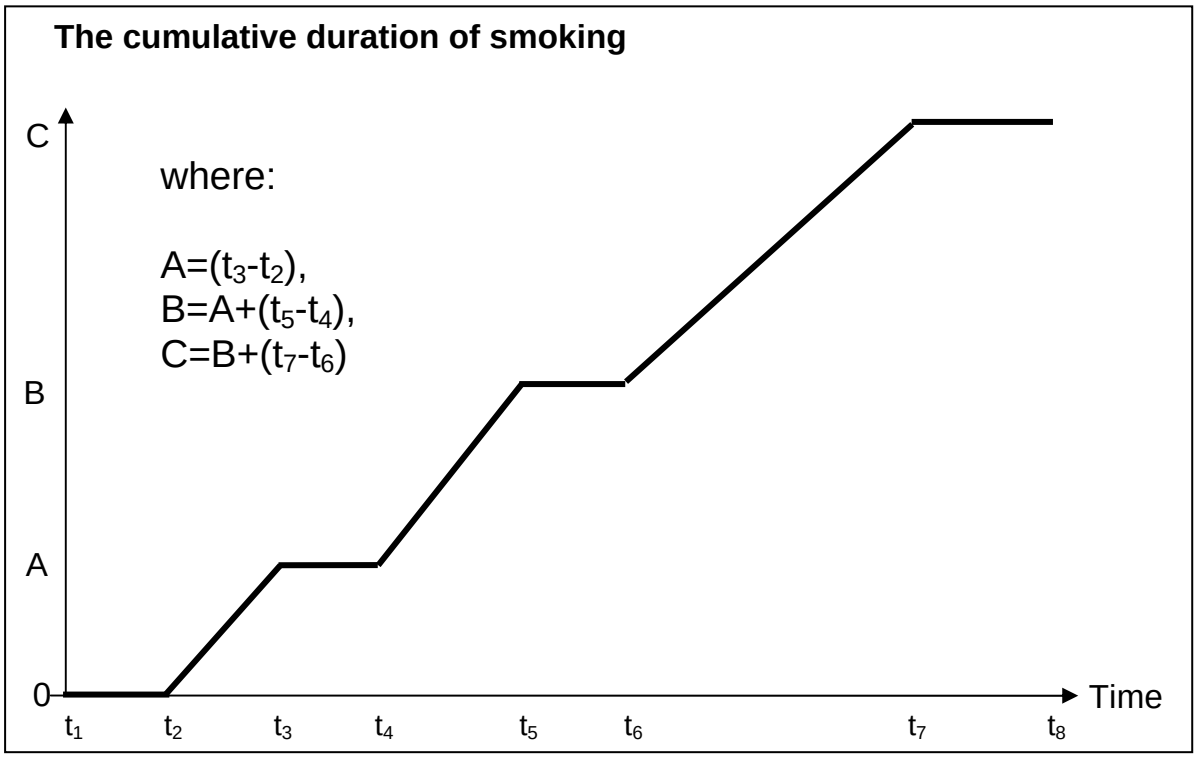
$$V_i(t) = tU_i(t)(1–U_{i+1}(t)) + t_{i+1}U_{i+1}(t)$$

Then, the cumulative duration of smoking at time t, $S_5(t)$, is given by:

$$S_5(t) = \sum_{i=1}^{m-1} (V_i(t) – t_i)U_i(t)E_i$$

The graph of S_5 as a function of time for the woman illustrated in figure 13.7 is presented in figure 13.7.5.

Figure 13.7.5. The time-varying covariate "the cumulative duration of smoking" ($S_5(t)$) for the woman illustrated in figure 13.7 as a function of time.



This variable is well-defined for all women: women who have never smoked at a given point in time have a true cumulative duration of smoking of zero.

The assumption that an increase from zero to one carries the same risk as an increase from four to five, for example, would need to be verified. In this case, the numerical change in the variable is the same, i.e. an increase of one, but a change from zero to one represents a change from never having smoked to starting smoking (a qualitative change), which may be expected to convey a greater increase in risk than merely continuing smoking, but at a higher rate. Therefore, this variable was also analyzed as a categorical covariate with categories chosen because they were "logical", in this case multiples of whole years, and had "sufficient" numbers of events in each category; these categories were chosen prior to the estimation of risks. Never-smokers were used as the reference category.

13.7.6 The duration of the current smoking episode

The duration of the current smoking episode is a time-varying covariate measuring the total amount of time a woman has been smoking *since she last started smoking*. For current smokers, the current smoking episode was considered to have begun at the midpoint between the most recent date at which a woman reported having started to smoke, and the date of the preceding visit. So, if at time T a woman has made one or more unsuccessful attempts to quit smoking, but is a smoker at time T, the duration of the current smoking episode at time T is the interval between the date on which she started smoking for the second time, and time T. This variable is only

well-defined for women who are current smokers. This variable was therefore initially analysed as a continuous covariate together with a time-varying covariate indicating whether a woman was a current smoker, thereby effectively restricting the analysis to current smokers.

This variable was also analysed as a categorical covariate with categories chosen because they were “logical”, in this case multiples of whole years, and had “sufficient” numbers of events in each category: these categories were chosen prior to the estimation of risks. Two versions of this variable were analyzed: one had non-smokers as the reference category (a category which combined never- and ex-smokers); the other had never-smokers alone as the reference category, with ex-smokers in a separate category.

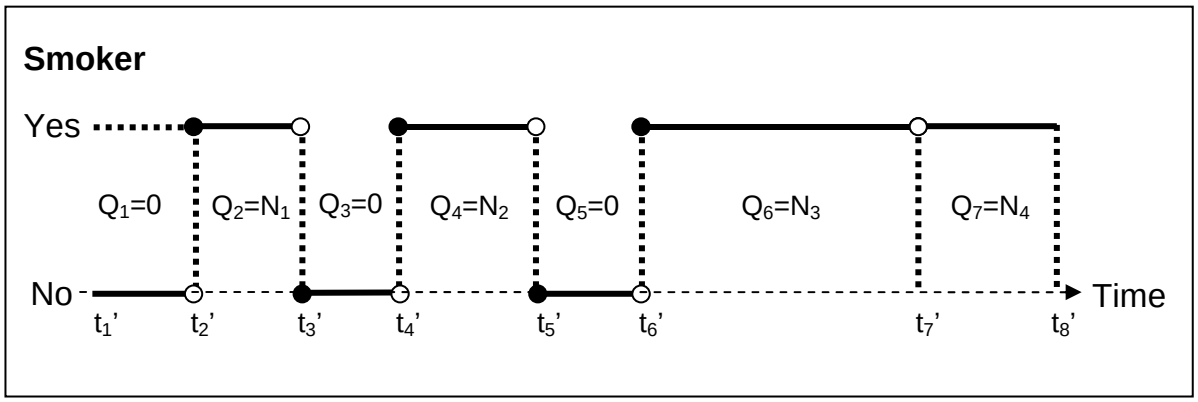
13.7.7 Pack-years smoked

Pack-years are a conventional measure of cumulative exposure to smoking over the lifetime of a woman, combining intensity and duration of smoking into one measurement, giving each *equal weight*. Pack-years smoked is a time-varying covariate: they are obtained by multiplying quantity smoked, which may vary over time, by the time for which that quantity was smoked, accumulating the resulting product over the lifetime of the woman, and expressing the result in suitable units.

Figure 13.7 illustrates how a woman’s smoking status may change over time with respect to both status (smoker or non-smoker) and quantity smoked (smoking

intensity). However, the situation is likely to be more complex than this illustration, since the amount a woman smokes, as well as her smoking status, may vary with time. Suppose the number of cigarettes a woman smokes per unit time is known to change at times t'_i, i=1,2,...,m' (once again, for ease of notation, a "change" is said to have occurred when the number smoked per unit time is determined for the first time, and at the end of follow-up) and let Q_k, k=1,2,...m'-1 be variables measuring the number of cigarettes smoked per unit time during each of the intervals $[t_i, t_{i+1})$. Figure 13.7.7a illustrates this situation, which is slightly different from that illustrated in figure 13.7 since *this* woman does not stop smoking at time t_7, but instead changes (increases, say) the quantity she smokes.

Figure 13.7.7a. An illustration of how the smoking status of a woman may vary with time. t'_i, i=1,2,...,8 are the times at which the amount a woman smokes is known to change; Q_k, k=1,2,...,7 are numbers of cigarettes smoked per unit time.



The number of cigarettes smoked per day was not recorded exactly, but was categorised into one of six categories (including a quantity of zero for non-smokers); the assumption was therefore made that the actual number of cigarettes smoked per day was equal to the midpoint of each category, i.e. 5, 15, 25, and 35 cigarettes per-

day, respectively, for the first four categories, with the number smoked in the final (open-ended) category arbitrarily set to 45. This number was then assumed to apply from the date of the interview at which that quantity-category was collected, until the date of the day preceding the subsequent interview, at which point the quantity was updated. For women who, prior to study entry, had started smoking, stopped, and then started again, the assumption was made that the number of cigarettes smoked per day during the first smoking episode was the same as the number smoked when she started smoking for the second time.

The cumulative quantity of cigarettes smoked is a time-varying covariate measuring the total number of cigarettes smoked by a woman; the value of this variable remains constant during the periods of time when the woman was a non-smoker. This variable was constructed using the following method. Define the time-varying indicator function U(t):

$$U_i(t) = I(t \geq t_i)$$

Define:

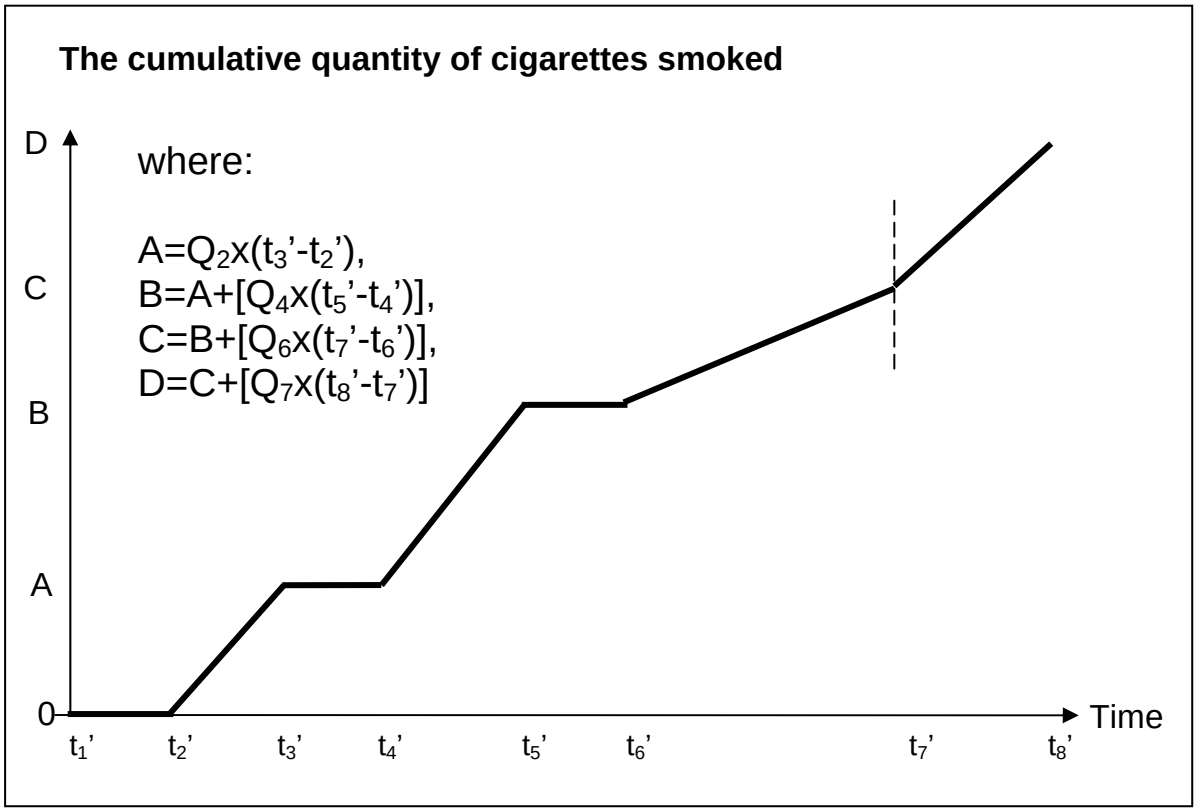
$$V_i(t) = tU_i(t)[1–U_{i+1}(t)] + t_{i+1}U_{i+1}(t)$$

Then, the cumulative quantity of cigarettes smoked at time t is given by:

$$S_6(t) = \sum_{i=1}^{m' -1} (V_i(t) – t_i)U_i(t)Q_i$$

The graph of S_6 as a function of time for the woman illustrated in figure 13.7.7a is presented in figure 13.7.7b.

Figure 13.7.7b. The time-varying covariate "the cumulative quantity of cigarettes smoked" ($S_6(t)$) for the woman illustrated in figure 13.7.7a as a function of time.



The cumulative quantity of cigarettes was then categorised using categories chosen to correspond to "pack-years", where one pack-year corresponds to 20*365.25=7,305 cigarettes. Note that a woman who smokes one pack of 20 cigarettes every day for one year, has an exposure of one pack-year; as does a woman who smokes half as many cigarettes for twice as long. This variable is well-defined for all women: women who have never smoked at a given point in time have

a number of pack-years of zero; there is thus no need to distinguish never-, ex-, and current-smokers.

13.7.8 Current smoking intensity

Current smoking intensity is a time-varying categorical covariate measuring the number of cigarettes smoked per day. Categories recorded for this variable were: 0, 1 to 9, 10 to 19, 20 to 29, 30 to 39, and 40 or more. Due to the small number of events for some of the outcome variables, and to enable comparisons to be made across these outcomes, this variable was analyzed as a four-category variable, with never-smokers as reference category, ex-smokers in a separate category with a smoking intensity of 0, and categories 1 to 9 and 10 or more for current smokers.

13.8 DEFINITION OF SEXUAL BEHAVIOUR-RELATED VARIABLES

Variables measuring some aspect of sexual behaviour which were available for analysis were age at first sexual intercourse, age of oldest sexual partner (two versions: in total, and in the previous six months), lifetime number of sexual partners, all measured on a continuous scale; and the acquisition of a new sexual partner in the previous six months, measured as a categorical variable. The first three of these variables are self-explanatory: age at first sexual intercourse was a time-fixed covariate; the other two variables were time-varying, with values updated at each study visit. When values were missing, these variables took their last known value.

Acquisition of a new sexual partner in the previous six months was a time-varying binary covariate indicating whether a woman had acquired one or more new sexual partners in the previous six months. A return to a previous sexual partner was also counted as acquiring a new sexual partner, whether or not the woman had acquired another new (distinct) sexual partner in the interim. Sometimes, a woman would state that she had started a sexual relationship with one partner, had ended that relationship, but that subsequently she had "returned" to this original sexual partner. If she had acquired a new (distinct) sexual partner in the interim, she would be categorised as having acquired *two* sexual partners; if not, she would be categorised as having acquired one sexual partner. This is because the assumption was made that any current sexual relationship was monogamous for both partners; but that the male partner was sexually active, and therefore exposed to cervical HPV infections carried by additional sexual partners, whilst he was not in a relationship with the study woman.

13.9 DEFINITION OF HPV EXPOSURE VARIABLES

In some analyses in this section, cervical HPV infection is an *outcome*, but for the analysis of incident high-grade CIN, cervical HPV infection is an *exposure variable*: in this analysis, cervical HPV infection status was treated as an irreversible time-varying indicator variable. The risk of cervical disease associated with exposure to cervical HPV infection is known to vary according to type; therefore controlling for HPV exposure using a variable measuring only exposure to any HPV type, may be suboptimal and provide insufficient control of this variable. Cervical HPV infection

status was therefore controlled by including in each analysis three separate time-varying binary variables constructed to measure whether a woman had ever been exposed to cervical infection with HPV16, HPV18, or other HPV types; HPV16 and HPV18, the two most common high-risk HPV types, were detected in sufficient numbers to make this analysis feasible.

For example, suppose a woman who was HPV DNA-negative in a cervical sample at study entry (T_0), acquired HPV16 at time T_1, acquired HPV18 at time T_2 ($T_2>T_1$), acquired HPV of a type other than type 16 or type 18, at time T_3 ($T_3>T_2$), was positive for both HPV16 and HPV18 (but no other types) at time T_4 ($T_4>T_3$), and was HPV DNA-negative at times T_5 and T_6 ($T_5>T_4$, $T_6>T_5$). Table 13.9 illustrates the coding of the three HPV exposure variables as they change over time for this woman. As soon as HPV of the relevant type is detected, the status for that variable changes from 0 (unexposed) to 1 (exposed) and thereafter remains at 1, whatever additional changes in HPV status occur thereafter. This variable assumes that once a woman has been exposed to HPV, her risk status for high-grade CIN is permanently altered, even if she subsequently becomes HPV DNA-negative.

Table 13.9. Example of the coding of the set of three time-varying variables measuring HPV exposure as they were used in the analysis of incident high-grade CIN. This example relates to a woman who was HPV DNA-negative at study entry (T_0), acquired HPV16 at time T_1, HPV18 at time T_2, HPV of another type at time T_3, was positive for HP16 and HPV18 (but no other types) at time T_4, and was HPV DNA-negative at times T_5 and T_6.

Time	HPV16	HPV18	Other HPV types
T_0	0	0	0
T_1	1	0	0
T_2	1	1	0
T_3	1	1	1
T_4	1	1	1
T_5	1	1	1
T_6	1	1	1

Chapter 14

SMOKING, CERVICAL HPV INFECTION, AND THE RISK OF CERVICAL NEOPLASIA

RESULTS

Orientation

In this chapter I:

- describe the association between exposure to cigarette smoking and the risk of acquiring an incident cervical infection with HPV DNA of any type
- describe the association between exposure to cigarette smoking and the risk of acquiring an incident cervical infection with HPV16 or HPV18
- describe the association between exposure to cigarette smoking and the duration of an incident cervical infection with HPV of any type, HPV16, or HPV18
- describe the association between exposure to cigarette smoking and the risk of acquiring incident high-grade cervical intraepithelial neoplasia

14.1 THE INCIDENCE OF CERVICAL HPV INFECTION IN RELATION TO SMOKING HISTORY

The study population for this analysis comprises 1,075 women who were HPV DNA-negative in their first evaluable sample, cytologically normal in their first evaluable cervical smear, and who had further follow-up after this time. The study population is identical for the analysis of the incidence of cervical HPV infections of any type, HPV16 and HPV18; only the outcome changes for these analyses.

Note: When presenting analyses, to describe the distribution of a variable, or to convey some measure of the reliability of estimates, it is conventional to report numbers of subjects and the number of events they experience for each of the categories of that variable, or overall for continuous variables. For time-varying covariates, however, it is the history of how these quantities change over time which are relevant; these require a more extensive and complex presentation than space allows. Therefore, in this section, women were categorised at their observed event time, and it is these frequencies and number of events which are reported in tables.

14.1.1 Univariate analysis

This section presents the results of univariate analyses describing the association between each of the variables used to measure exposure to cigarette smoking and the risk of: an incident cervical HPV infection of any type; an incident cervical HPV16 infection; and an incident cervical HPV18 infection.

HPV16 and HPV18 are the two HPV types which are most frequently detected at the time of diagnosis of cervical cancer, and are therefore of particular interest. Within the cohort on which this thesis is based, these two high-risk types were also the most commonly detected types (at any time during follow-up), and were the only two types detected with sufficient frequency to enable their inclusion as outcomes in separate analyses.

14.1.1.1 Cervical HPV infection of any type

Four hundred and seven women first acquired a cervical infection with HPV DNA of any type during follow-up.

14.1.1.1.1 The univariate analysis of the association between smoking status at study entry and the risk of an incident cervical HPV infection of any type

Three hundred and fifty six (33%) women were smokers at study entry, of whom one had stopped smoking at some point prior to study entry, but had since re-started. Seven hundred and nineteen (67%) women were non-smokers at study entry; 626 (87%) had never been smokers compared with 93 (13%) who reported that they had been smokers prior to study entry, but had since stopped. Of the 626 women who had never smoked prior to study entry, 228 (36%) became HPV-positive during follow-up: HPV was first detected after first exposure to smoking in 19 (8%) of these women; coincident with first exposure to smoking in 17 (7%); and before first exposure to smoking in 192 (84%) (15 (8%) of whom were known to have begun smoking before the end of follow-up). Of the 449 women who had first become smokers at some point prior to study entry, 179 (40%) became HPV-positive in cervical samples during follow-up.

The univariate analysis of the association between smoking status at study entry and the risk of an incident cervical HPV infection of any type is described in table 14.1.1.1.1. The association between smoking status at study entry and the risk of an

incident cervical HPV infection of any type was not statistically significant using the definition of the exposure variable which explicitly accounts for changes in smoking status prior to study entry. When ex-smokers were combined with never-smokers, smoking status at study entry remains non-significant; however, when they were combined with current smokers (forming an “ever-smoker” category), smoking status at study entry became statistically significant. Current smokers at baseline had an approximately 20% increase in risk in all three analyses; in the analysis using the first definition, current and ex-smokers were at a similar risk.

Table 14.1.1.1.1. The univariate analysis of the association between smoking status at study entry and the risk of an incident cervical infection with HPV DNA of any type.

Smoking Status At Study Entry	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Never-smoker	626 (58.2)	228	1.00 (Reference)
Ex-smoker	93 (8.7)	38	1.26 (0.89 to 1.79)
Current smoker	356 (33.1)	141	1.23 (0.99 to 1.52)
			χ^2=4.40; 2df; p=0.11[a]
Non-smoker	719 (65.0)	266	1.00 (Reference)
Current smoker	356 (35.0)	141	1.19 (0.97 to 1.46)
			χ^2=2.74; 1df; p=0.10[a]
Non-smoker	626 (58.2)	228	1.00 (Reference)
Ever smoker	449 (41.8)	179	1.23 (1.01 to 1.51)
			χ^2=4.38; 1df; p=0.04[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

In women who had never smoked prior to study entry, the cumulative risk of an incident cervical HPV DNA infection of any type at one, two and three years after study entry were 19.5%, 35.5% and 43.4% respectively.

14.1.1.1.2 The univariate analysis of the association between age at starting smoking and the risk of an incident cervical HPV infection of any type

Age at first starting smoking was analysed as both a time-fixed variable, fixed at study entry, and as a time-varying covariate. Four hundred and forty nine (41.8%) women had regularly smoked cigarettes at some point prior to study entry; the median age at starting smoking for these women was 15 years (range 9 to 19).

The univariate analysis of the association between age at starting smoking, as recorded at study entry, and the risk of an incident cervical HPV infection of any type, is described in table 14.1.1.1.2. Among women who had ever smoked at study entry, there was a decrease in the risk of an incident cervical HPV infection of any type as age at starting smoking increased: the hazard ratios decreased by 6% (1.00-0.94) for every year of increase in age at starting smoking; however, the association was not statistically significant. There was a statistically significant association between age at starting smoking at study entry and the risk of an incident cervical HPV infection of any type, when age at starting smoking was analysed as a categorical variable. However, there was no consistent pattern: statistical significance appears to be due to a single category, with no apparent trend in risk associated with age at starting smoking.

Table 14.1.1.1.2. The univariate analysis of the association between age at starting smoking as measured at study entry and the risk of an incident cervical infection with HPV DNA of any type.

Age At Starting Smoking (at study entry)	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per year)[a]	1075 (100)	407	0.94 (0.86 to 1.02)
			χ^2=2.04; 1df; p=0.15[b]
Never-smoker	626 (58.2)	228	1.00 (Reference)
≤13	100 (9.3)	41	1.38 (0.98 to 1.94)
14-15	182 (16.9)	83	1.38 (1.07 to 1.79)
≥16	167 (15.5)	55	1.00 (0.74 to 1.34)
			χ^2=8.59; 3df; p=0.04[b]

[a]Includes a variable indicating "ever smoker" status at study entry; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.1.1.1.3 The univariate analysis of the association between a history of ever having smoked and the risk of an incident cervical HPV infection of any type

Nearly 50% of women had smoked cigarettes at some point prior to the first detection of HPV DNA of any type during follow-up.

The univariate analysis of the association between a history of ever having smoked and the risk of an incident cervical HPV infection of any type is described in table 14.1.1.1.3. Compared to women who had never smoked, women who had ever smoked had a small (39%=1.39-1.00), but statistically significant, increase in the risk of an incident cervical HPV infection of any type.

Table 14.1.1.1.3. The univariate analysis of the association between a history of ever having smoked and the risk of an incident cervical infection with HPV DNA of any type.

Have you ever been a cigarette smoker?	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
No	550 (51.2)	192	1.00 (Reference)
Yes	525 (48.8)	215	1.39 (1.14 to 1.69)
			χ^2=10.79; 1df; p=0.001[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.1.1.1.4 The univariate analysis of the association between the time since first starting smoking and the risk of an incident cervical HPV infection of any type

At the earliest of the date of the first detection of HPV DNA of any type in a cervical sample, or the date of the end of follow-up, the median time since first starting smoking was 43.1 months (range 2.2 to 157.0) in women who had ever smoked.

The “univariate” association (the model actually contains more than one variable, but the focus of the analysis is on only one of those variables) between time since starting smoking and the risk of an incident cervical HPV infection of any type is described in table 14.1.1.1.4. There was a very slight and non-significant decrease in the risk of an incident cervical HPV infection of any type with increasing time since starting smoking, when this variable was analysed as a continuous variable, controlling for ever having smoked. The risk of an incident cervical HPV infection of any type decreased by 3% (1.00-0.97) for every increase of one year in time since starting smoking. When the variable was analysed as a categorical covariate, with

never-smokers as reference group, the time since first starting smoking was significantly associated with the risk of an incident cervical HPV infection of any type. However, there were only very slight differences between the risks estimated for each of the non-zero categories, suggesting that the association observed was a consequence of ever having smoked, rather than duration of smoking *per se*.

Table 14.1.1.1.4. The univariate analysis of the association between the time since first starting smoking and the risk of an incident cervical infection with HPV DNA of any type.

Time Since First Starting Smoking	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per year)[a]	1075 (100)	407	0.97 (0.91 to 1.03)
			χ^2=1.05; 1df; p=0.31[b]
Never-smoker	550 (51.2)	192	1.00 (Reference)
(0, 36 months)[c]	181 (16.8)	85	1.44 (1.11 to 1.87)
[36, 72)[c]	242 (22.5)	91	1.31 (1.02 to 1.69)
[72, +)[c]	102 (9.5)	39	1.46 (1.01 to 2.09)
			χ^2=11.26; 3df; p=0.01[b]
Never-smoker	550 (51.2)	192	1.00 (Reference)
(0, 36 months)[c]	181 (16.8)	85	1.44 (1.11 to 1.87)
[36, 72)[c]	242 (22.5)	91	1.31 (1.02 to 1.69)
[72, 84)[c]	40 (3.7)	18	1.44 (0.87 to 2.38)
[84, +)[c]	62 (5.8)	21	1.47 (0.91 to 2.36)
			χ^2=11.26; 4df; p=0.02[b]

[a]Model includes a time-varying indicator variable indicating “ever smoker” status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [c][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.1.1.1.5 The univariate analysis of the association between the cumulative duration of smoking and the risk of an incident cervical HPV infection of any type

At the earliest of the date of the first detection of HPV DNA of any type in a cervical sample, or the date of the end of follow-up, the median cumulative duration of smoking was 35.6 months (range 2.2 to 157.0) among women who had ever smoked.

The univariate analysis of the association between the cumulative duration of smoking and the risk of an incident cervical HPV infection of any type is described in table 14.1.1.1.5. There was a slight non-significant decrease in the risk of an incident cervical HPV infection of any type with increasing cumulative duration of smoking, when this variable was analysed as a continuous variable, controlling for ever having smoked. The risk of an incident cervical infection decreased by 4% (1.00-0.96) for every year of increase in the cumulative duration of smoking. When the variable was analysed as a categorical covariate, with never-smokers as reference group, cumulative duration of smoking was significantly associated with the risk of an incident cervical HPV infection of any type. However, there were only very slight differences between the risks estimated for each of the categories, with no trend apparent. Note that this analysis differs from that in section 14.1.1.1.4 because not all women smoked continuously once they had started to smoke. As expected, cumulative duration of smoking is highly correlated with time since starting smoking:

the risks estimates associated with the latter are slightly less extreme, but the pattern of risk estimates is similar.

Table 14.1.1.1.5. The univariate analysis of the association between the cumulative duration of smoking and the risk of an incident cervical infection with HPV DNA of any type.

Cumulative Duration Of Smoking	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per year)	1075 (100.0)	407	0.96 (0.90 to 1.02)
			χ^2=1.64; 1df; p=0.20[b]
Never-smoker	550 (51.2)	192	1.00 (Reference)
(0, 36 months)[c]	251 (23.3)	113	1.50 (1.18 to 1.90)
[36, 72)[c]	205 (19.1)	75	1.20 (0.92 to 1.57)
[72, +)[c]	69 (6.4)	27	1.57 (1.03 to 2.40)
			χ^2=13.45; 3df; p=0.004[b]

[a]Model includes a time-varying indicator variable indicating "ever smoker" status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [c][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.1.1.1.6 The univariate analysis of the association between the duration of the current smoking episode and the risk of an incident cervical HPV infection of any type

The duration of a smoking episode was measured from the date a woman reported smoking until the first date subsequent to this on which she reported stopping smoking. At the earliest of the date of the first detection of HPV DNA of any type in a cervical sample, or the date of the end of follow-up, the median duration of the current smoking episode was 34.8 months (range 1.2 to 157.0), among women who were current smokers at that time.

The univariate analysis of the association between the duration of the current smoking episode and the risk of an incident cervical HPV infection of any type is described in table 14.1.1.1.6. There was a very slight non-significant decrease in the risk of an incident cervical HPV infection of any type with increasing duration of current smoking episode, when this variable was analysed as a continuous variable, and the analysis was controlled for current smoking status. The risk of an incident cervical HPV infection of any type decreased by 3% (1.00-0.97) for every increase of year in time since the start of the current smoking episode. When the variable was analysed as a categorical covariate with non-smokers as reference group, there were very slight differences between the risks estimated for each of the non-zero categories. When the variable was analysed as a categorical covariate with *never*-smokers as reference group, there were very slight differences between the risks estimated for each of the non-zero categories and the risk for ex-smokers: inclusion of ex-smokers in the non-smokers category in the previous analysis “diluted” the risk estimates, but only slightly.

Table 14.1.1.1.6. The univariate analysis of the association between the duration of the current smoking episode and the risk of an incident cervical infection with HPV DNA of any type.

Duration Of Current Smoking Episode	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per year)[a]	1075 (100)	407	0.97 (0.90 to 1.04)
			χ^2=0.75; 1df; p=0.39[b]
Nonsmoker	679 (63.2)	243	1.00 (Reference)
(0, 24 months)[c]	132 (12.3)	60	1.41 (1.06 to 1.88)
[24, 48)[c]	120 (11.2)	52	1.26 (0.93 to 1.70)
[48, +)[c]	144 (13.4)	52	1.22 (0.90 to 1.66)
			χ^2=6.96; 3df; p=0.07[b]
Never-smoker	550 (51.2)	192	1.00 (Reference)
Ex-smoker	129 (12.0)	51	1.41 (1.03 to 1.93)
(0, 24 months)[c]	132 (12.3)	60	1.50 (1.12 to 2.01)
[24, 48)[c]	120 (11.2)	52	1.34 (0.98 to 1.82)
[48, +)[c]	144 (13.4)	52	1.30 (0.95 to 1.78)
			χ^2=11.44; 4df; p=0.02[b]

[a]Model includes a time-varying indicator variable indicating "current smoking" status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [c][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.1.1.1.7 The univariate analysis of the association between pack-years smoked and the risk of an incident cervical HPV infection of any type

This analysis was restricted to the 1,038 (out of 1,075) women for whom the quantity smoked was known throughout follow-up. At the earliest of the date of the first detection of HPV DNA of any type in a cervical cytological sample, or the date of the end of follow-up, the median pack-years smoked was 1 pack-year (range 0.1 to 8) in women who had ever smoked prior to that time.

The univariate analysis of the association between pack-years smoked and the risk of an incident cervical HPV infection of any type is described in table 14.1.1.1.7. When analysed as a continuous variable, the hazards ratio decreased by 8% (1.00-0.92) for every increase of one pack year, but the association was not statistically significant. There was no evidence of a departure from linearity for this variable. However, when analysed as a categorical variable, the risk estimates showed no consistent pattern across categories: women who had smoked less than one pack-year, or between two and three pack-years, were at a significantly increased risk of an incident cervical infection with HPV DNA of any type (the confidence intervals around the hazards ratios exclude 1), but women in the intervening category were not at a significantly increased risk; women who had smoked more than three pack-years had no detectable increase in risk compared with women who had never smoked.

Table 14.1.1.1.7. The univariate analysis of the association between pack-years smoked and the risk of an incident cervical infection with HPV DNA of any type.

Pack-years Smoked	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per pack-year)[a]	1038 (100)	394	0.92 (0.80 to 1.05)
			χ^2=1.66; 1df; p=0.20[b]
Never-smoker	550 (51.2)	192	1.00 (Reference)
(0 , 1)[c]	232 (22.4)	110	1.59 (1.25 to 2.01)
[1 , 2)[c]	143 (13.8)	50	1.16 (0.84 to 1.58)
[2 , 3)[c]	73 (7.0)	30	1.50 (1.01 to 2.21)
[3 , +)[c]	40 (3.9)	12	1.03 (0.56 to 1.90)
			χ^2=15.84; 4df; p=0.003[b]

[a]Model includes a time-varying indicator variable indicating "ever smoker" status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [c][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.1.1.1.8 The univariate analysis of the association between current smoking status and the risk of an incident cervical HPV infection of any type

The univariate analysis of the association between current smoking status and the risk of an incident cervical HPV infection of any type is described in table 14.1.1.1.8. Current smoking status was significantly associated with the risk of an incident cervical HPV infection of any type. Current smokers and current ex-smokers had a small (38%=1.38-1.00) but significantly increased risk of an incident cervical HPV infection of any type compared to women who had never smoked. As expected, the magnitude of this risk was similar to that seen for the "smoking history (ever versus never)" variable: these two variables are highly correlated, since for any woman who never smokes, or whose smoking status changes only once (the majority of women),

the variables are identical. When never-smokers and ex-smokers were combined into the reference group, this variable remained statistically significant.

Table 14.1.1.1.8. The univariate analysis of the association between current smoking status and the risk of an incident cervical infection with HPV DNA of any type.

Current Smoking Status	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Never-smoker	550 (51.2)	192	1.00 (Reference)
Ex-smoker	129 (12.0)	51	1.41 (1.04 to 1.93)
Current smoker	396 (36.8)	164	1.38 (1.12 to 1.70)
			χ^2=10.82; 2df; p=0.004[a]
Non-smoker	679 (63.2)	243	1.00 (Reference)
Current smoker	396 (36.8)	164	1.29 (1.06 to 1.58)
			χ^2=6.32; 1df; p=0.01[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.1.1.1.9 The univariate analysis of the association between current smoking intensity and the risk of an incident cervical HPV infection of any type

This analysis was restricted to the 1,038 (out of 1,075) women for whom quantity smoked was known throughout follow-up.

At the earliest of the date of the first detection of HPV DNA of any type in a cervical cytological sample, or the date of the end of follow-up, most women who were current smokers were smoking between one and nine cigarettes per day.

The univariate analysis of the association between current smoking intensity and the risk of an incident cervical infection with HPV DNA of any type is shown in table

14.1.1.1.9. When analysed as a categorical variable, current smoking intensity was significantly associated with the risk of an incident cervical infection with HPV DNA of any type. However, this seems to reflect the risk associated with being a current smoker rather than smoking intensity *per se*: among current smokers, there was no significant association between current smoking intensity and the risk of an incident cervical HPV infection of any type.

Table 14.1.1.1.9. The univariate analysis of the association between current smoking intensity and the risk of an incident cervical infection with HPV DNA of any type.

Current Smoking Intensity	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per category)[a]	1038 (100)	394	0.97 (0.82 to 1.14)
			χ^2=0.16; 1df; p=0.69[b]
Never-smoker	550 (53.0)	192	1.00 (Reference)
0 per day	104 (10.0)	40	1.41 (1.04 to 1.93)
1 to 9 per day	196 (18.9)	92	1.48 (1.15 to 1.90)
10 or more per day	188 (18.1)	70	1.26 (0.96 to 1.67)
			χ^2=11.81; 3df; p=0.008[b]

[a]The increase in risk for every increase of one in category of smoking intensity; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

Summary

There are several possible ways of measuring exposure to smoking in the study described in this thesis. In univariate analyses, most of the alternative exposure variables measuring exposure to smoking were significantly associated with the risk of an incident cervical HPV infection of any type, although the magnitude of the risk estimates were typically small.

14.1.1.2 HPV16 infection

One hundred and ten women acquired an incident cervical infection with HPV16 during follow-up.

Note: In the tables, frequencies and numbers of events are reported at the time of occurrence of an event. Therefore, since "HPV of a specific type" events are less common than "HPV of any type" events (110 and 64 for HPV16 and HPV18, respectively, compared with 407 for HPV of any type), women in the type-specific analyses are often followed-up for longer before experiencing an event. Therefore they have greater opportunity to change smoking status. This accounts for the apparent "discrepancy" between the number of women in specific categories of the smoking variable in the type-specific analyses and the "all-types-combined" analysis.

14.1.1.2.1 The univariate analysis of the association between smoking status at study entry and the risk of an incident cervical HPV16 infection

The univariate analysis of the association between smoking status at study entry and the risk of an incident cervical HPV16 infection is described in table 14.1.1.2.1. The association between smoking status at study entry and the risk of an incident cervical HPV16 infection was not statistically significant using any of the three possible definitions of this exposure variable.

Table 14.1.1.2.1. The univariate analysis of the association between smoking status at study entry and the risk of an incident cervical infection with HPV16.

Smoking Status At Study Entry	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Never-smoker	626 (58.2)	62	1.00 (Reference)
Ex-smoker	93 (8.7)	10	1.22 (0.62 to 2.38)
Current smoker	356 (33.1)	38	1.08 (0.72 to 1.63)
			χ^2=0.40; 2df; p=0.82[a]
Non-smoker	719 (65.0)	72	1.00 (Reference)
Current smoker	356 (35.0)	38	1.06 (0.71 to 1.57)
			χ^2=0.08; 1df; p=0.78[a]
Non-smoker	626 (58.2)	62	1.00 (Reference)
Ever smoker	449 (41.8)	48	1.11 (0.76 to 1.62)
			χ^2=0.29; 1df; p=0.59[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.1.1.2.2 The univariate analysis of the association between age at starting smoking and the risk of an incident cervical HPV16 infection

The univariate analysis of the association between age at starting smoking, as measured at study entry (i.e. this was a time-fixed variable whose value was measured at study entry), and the risk of an incident cervical infection with HPV16 is shown in table 14.1.1.2.2. Age at first starting smoking was analysed as both a time-fixed variable, fixed at study entry, and as a time-varying covariate, updated during follow-up. In neither analysis was there a statistically significant association between age at starting smoking and the risk of an incident cervical HPV16 infection.

Table 14.1.1.2.2. The univariate analysis of the association between age at starting smoking as measured at study entry and the risk of an incident cervical infection with HPV16.

Age At Starting Smoking (at study entry)	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per year)[a]	1075 (100)	110	1.02 (0.87 to 1.20)
			χ^2=0.06; 1df; p=0.80[b]
Never-smoker	626 (58.2)	62	1.00 (Reference)
≤13	100 (9.3)	9	0.95 (0.47 to 1.91)
14-15	182 (16.9)	22	1.21 (0.74 to 1.97)
≥16	167 (15.5)	17	1.09 (0.64 to 1.87)
			χ^2=0.69; 1df; p=0.88[b]

[a]Includes a variable indicating "ever smoker" status at study entry; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.1.1.2.3 The univariate analysis of the association between a history of ever having smoked and the risk of an incident cervical HPV16 infection

The univariate analysis of the association between a history of ever having smoked and the risk of an incident cervical infection with HPV16 is shown in table 14.1.1.2.3. Women who had ever smoked had a small (40%=1.40-1.00) but non-significant increase in the risk of an incident cervical HPV16 infection compared to women who had never smoked.

Table 14.1.1.2.3. The univariate analysis of the association between a history of ever having smoked and the risk of an incident cervical infection with HPV16.

Have you ever been a cigarette smoker?	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
No	539 (50.1)	48	1.00 (Reference)
Yes	536 (49.9)	62	1.40 (0.96 to 2.05)
			χ^2=3.14; 1df; p=0.08[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.1.1.2.4 The univariate analysis of the association between the time since first starting smoking and the risk of an incident cervical HPV16 infection

At the earliest of the date of the first detection of HPV16 in a cervical cytological sample, or the date of the end of follow-up, the median time since first starting smoking was 49.1 months (range 2.2 to 188.0) in women who had ever smoked.

The univariate analysis of the association between the time since first starting smoking and the risk of an incident cervical infection with HPV DNA of any type is shown in table 14.1.1.2.4. There was a small but statistically significant decrease in the risk of an incident cervical HPV16 infection with increasing time since starting smoking, when this variable was analysed as a continuous variable, controlling for ever having smoked. The risk of an incident cervical HPV16 infection decreased by 12% (1.00-0.88) for every increase of one year in the time since starting smoking. When the variable was analysed as a categorical covariate, with never-smokers as reference group, time since first starting smoking was not significantly associated with the risk of an incident cervical HPV16 infection. Note that this is a reversal in the

situation which occurred in the analysis of the incidence of a cervical infection with HPV DNA of any type.

Table 14.1.1.2.4. The univariate analysis of the association between the time since first starting smoking and the risk of an incident cervical infection with HPV16.

Time Since First Starting Smoking	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per year)[a]	1075 (100)	407	0.88 (0.79 to 0.99)
			χ^2=4.45; 1df; p=0.03[b]
Never-smoker	539 (50.1)	48	1.00 (Reference)
(0, 36 months)[c]	146 (13.6)	22	1.52 (0.91 to 2.54)
[36, 72)[c]	237 (22.0)	30	1.57 (0.99 to 2.49)
[72, +)[c]	153 (14.2)	10	0.92 (0.45 to 1.89)
			χ^2=5.33; 3df; p=0.15[b]
Never-smoker	539 (50.1)	48	1.00 (Reference)
(0, 36 months)[c]	146 (13.6)	22	1.52 (0.91 to 2.54)
[36, 72)[c]	237 (22.0)	30	1.58 (1.00 to 2.49)
[72, 84)[c]	47 (4.4)	5	1.22 (0.48 to 3.11)
[84, +)[c]	106 (9.9)	5	0.72 (0.27 to 1.92)
			χ^2=6.01; 4df; p=0.20[b]

[a]Model includes a time-varying indicator variable indicating “ever smoker” status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [c][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.1.1.2.5 The univariate analysis of the association between the cumulative duration of smoking and the risk of an incident cervical HPV16 infection

At the earliest of the date of the first detection of HPV16 in a cervical sample, or the date of the end of follow-up, the median cumulative duration of smoking was 41.0 months (range 2.2 to 178.7) among women who had ever smoked.

The univariate analysis of the association between the cumulative duration of smoking and the risk of an incident cervical infection with HPV16 is shown in table 14.1.1.2.5. There was a significant decrease in the risk of an incident cervical HPV16 infection with increasing cumulative duration of smoking when this variable was analysed as a continuous variable, controlling for ever having smoked. The risk of an incident cervical infection with HPV16 decreased by 12% (1.00-0.88) for every year of increase in the cumulative duration of smoking. When the variable was analysed as a categorical covariate, with never-smokers as reference group, cumulative duration of smoking was not significantly associated with the risk of an incident HPV16 overall; however, women who had smoked for a total of at most 36 months were at a significantly greater risk of acquiring an incident cervical HPV16 infection than women who had never smoked. A trend in risk with increasing smoking duration was apparent, although numbers of events were small in the last category in particular.

Table 14.1.1.2.5. The univariate analysis of the association between the cumulative duration of smoking and the risk of an incident cervical infection with HPV16.

Cumulative Duration Of Smoking	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per year)	1075 (100.0)	110	0.88 (0.78 to 0.99)
			χ^2=5.06; 1df; p=0.02[b]
Never-smoker	539 (50.1)	48	1.00 (Reference)
(0, 36 months)[c]	227 (13.6)	34	1.76 (1.13 to 2.74)
[36, 72)[c]	200 (22.0)	21	1.20 (0.72 to 2.01)
[72, +)[c]	109 (4.4)	7	0.92 (0.40 to 2.10)
			χ^2=6.43; 3df; p=0.09[b]

[a]Model includes a time-varying indicator variable indicating "ever smoker" status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [c][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.1.1.2.6 The univariate analysis of the association between the duration of the current smoking episode and the risk of an incident cervical HPV16 infection

At the earliest of the date of the first detection of HPV16 in a cervical sample, or the date of the end of follow-up, median duration of the current smoking episode was 38.8 months (range 1.2 to 157.0) among women who were current smokers at this time.

The univariate analysis of the association between the duration of the current smoking episode and the risk of an incident cervical infection with HPV16 is shown in table 14.1.1.2.6. There was a significant decrease in the risk of an incident cervical HPV16 infection with increasing duration of the current smoking episode, when this variable was analysed as a continuous variable controlling for current

smoking status. The risk of an incident cervical infection decreased by 13% (1.00-0.87) for every increase of one year in time since the start of the current smoking episode. When this variable was analysed as a categorical covariate with either non-smokers or never-smokers as reference group, it was also significant; however, a clear-cut trend with increasing duration was not apparent.

Table 14.1.1.2.6. The univariate analysis of the association between the duration of the current smoking episode and the risk of an incident cervical infection with HPV16.

Duration Of Current Smoking Episode	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per year)[a]	1075 (100)	110	0.87 (0.76 to 1.00[b])
			χ^2=4.05; 1df; p=0.04[c]
Non-smoker	677 (63.0)	66	1.00 (Reference)
(0, 24 months)[d]	123 (11.4)	22	1.81 (1.11 to 2.95)
[24, 48)[d]	107 (10.0)	7	0.61 (0.28 to 1.33)
[48, +)[d]	168 (15.6)	15	1.01 (0.57 to 1.77)
			χ^2=8.20; 3df; p=0.04[c]
Never-smoker	539 (50.1)	48	1.00 (Reference)
Ex-smoker	138 (12.8)	18	1.77 (1.03 to 3.05)
(0, 24 months)[d]	123 (11.4)	22	2.05 (1.23 to 3.41)
[24, 48)[d]	107 (10.0)	7	0.69 (0.31 to 1.52)
[48, +)[d]	168 (15.6)	15	1.15 (0.64 to 2.06)
			χ^2=12.07; 4df; p=0.02[c]

[a]Model includes a time-varying indicator variable indicating “current smoking” status; [b]To two decimal places, but <1; [c]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [d][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.1.1.2.7 The univariate analysis of the association between pack-years smoked and the risk of an incident cervical HPV16 infection

This analysis was restricted to the 1,038 (out of 1,075) women for whom the quantity smoked was known throughout follow-up.

At the earliest of the date of the first detection of HPV16 in a cervical sample, or the date of the end of follow-up, the median pack-years smoked was 1.2 pack-years (range 0.05 to 7.93) in women who had ever smoked prior to that time.

The univariate analysis of the association between pack-years smoked and the risk of an incident cervical infection with HPV16 is shown in table 14.1.1.2.7. There was a statistically significant association between pack-years of smoking and the risk of an incident cervical HPV16 infection. When analysed as a continuous variable, controlling for ever having smoked, the hazards ratio decreased by 26% (1.00-0.74) for every increase of one pack year. However, when analysed as a categorical variable, although statistically significant, the risk estimates showed no consistent pattern across categories: women who had smoked less than one pack-year, or between two and three pack-years, were at an increased risk of HPV16, but women in the remaining categories were at A decreased risk, based on the point estimates.

Table 14.1.1.2.7. The univariate analysis of the association between pack-years smoked and the risk of an incident cervical infection with HPV16.

Pack-years smoked	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per pack-year)[a]	1038 (100)	108	0.74 (0.57 to 0.96)
			χ^2=5.97; 1df; p=0.01[b]
Never-smoker	539 (52.0)	48	1.00 (Reference)
(0 , 1)[c]	210 (21.0)	36	1.96 (1.27 to 3.03)
[1 , 2)[c]	149 (14.0)	11	0.90 (0.46 to 1.73)
[2 , 3)[c]	78 (7.3)	11	1.82 (0.94 to 3.55)
[3 , +)[c]	62 (5.8)	2	0.39 (0.09 to 1.62)
			χ^2=15.34; 4df; p=0.004[b]

[a]Model includes a time-varying indicator variable indicating 'ever smoker' status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [c][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.1.1.2.8 The univariate analysis of the association between current smoking status and the risk of an incident cervical HPV16 infection

The univariate analysis of the association between current smoking status and the risk of an incident cervical infection with HPV16 is shown in table 14.1.1.2.8. The association was not statistically significant overall. However, when the variable was analysed as a three-category variable, current ex-smokers had a significantly increased risk of HPV16 compared to women who had never smoked; current smokers did not, although confidence intervals around the risk estimates overlapped.

Table 14.1.1.2.8. The univariate analysis of the association between current smoking status and the risk of an incident cervical infection with HPV16.

Current Smoking Status	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Never-smoker	539 (50.1)	48	1.00 (Reference)
Ex-smoker	138 (12.8)	18	1.77 (1.03 to 3.06)
Current smoker	398 (37.0)	44	1.29 (0.86 to 1.95)
			χ^2=4.34; 2df; p=0.11[a]
Non-smoker	677 (63.0)	66	1.00 (Reference)
Current smoker	398 (37.0)	44	1.14 (0.78 to 1.67)
			χ^2=0.45; 1df; p=0.50[a]
Never-smoker	539 (50.1)	48	1.00 (Reference)
Ever-smoker	536 (49.9)	62	1.77 (1.03 to 3.06)
			χ^2=3.14; 1df; p=0.08[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.1.1.2.9 The univariate analysis of the association between the current smoking intensity and the risk of an incident cervical HPV16 infection

This analysis was restricted to the 1,038 (out of 1,075) women for whom the quantity smoked was known throughout follow-up.

At the earliest of the date of the first detection of HPV16 in a cervical sample, or the date of the end of follow-up, most women who were current smokers were smoking between one and nine cigarettes per day, but the median quantity smoked was between 10 and 19 cigarettes per day.

The univariate analysis of the association between current smoking intensity and the risk of an incident cervical infection with HPV16 is shown in table 14.1.1.2.9. The

association was not statistically significant whether the variable was measured on a categorical or on a continuous scale.

Table 14.1.1.2.9. The univariate analysis of the association between the current smoking intensity and the risk of an incident cervical infection with HPV16.

Current Smoking Intensity	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per category)[a,b]	1038 (100)	108	0.86 (0.64 to 1.15)
			χ^2=1.11; 1df; p=0.29[c]
Never-smoker	539 (50.1)	48	1.00 (Reference)
0 per day	114 (11.0)	16	1.77 (1.03 to 3.06)
1 to 9 per day	188 (18.1)	23	1.28 (0.77 to 2.10)
10 or more per day	197 (19.0)	21	1.32 (0.79 to 2.20)
			χ^2=4.35; 3df; p=0.23[c]

[a]Model includes a time-varying indicator variable indicating "current smoking" status; [b]the increase in risk for every increase of one in category of current smoking intensity; [c]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.1.1.3 HPV18 infection

In univariate analyses, there were no statistically significant associations between smoking exposure and the risk of an incident cervical HPV18 infection. Hazards ratios tended to be closer to the null value of one than those seen in the analysis of HPV16 infection presented above.

Summary

Time since starting smoking, the cumulative duration of smoking, the duration of the current smoking episode, and pack-years of smoking, were all statistically significant predictors of the risk of an incident cervical HPV16 infection; hazards ratios were typically small. There were no significant associations between smoking exposure and the risk of an incident cervical HPV18 infection, and measures of association tended to be closer to the null value than in the analysis of HPV16.

14.1.2 Multivariate analysis

This subsection presents the results of multivariate analyses describing the association between each of the variables used to measure exposure to cigarette smoking and the risk of incident cervical HPV infections of any type, and incident HPV16 and HPV18 infections, after controlling for variables measuring some aspect of sexual behaviour.

Various aspects of sexual behaviour are known risk factors for cervical HPV infection (see section 3.1.3). Since high-risk sexual behaviour is also associated with smoking, sexual behaviour may confound the association between smoking and the risk of incident cervical HPV infection. A univariate analysis was therefore undertaken to identify potential confounders amongst the sexual behaviour variables. Variables measuring some aspect of sexual behaviour which were included in the analysis were: calendar age at first sexual intercourse; lifetime number of sexual partners; whether a woman had acquired a new sexual partner in the previous six months; the age of the woman's oldest sexual partner; and the age

of the woman's oldest sexual partner in the previous six months. The age of the (only) sexual partner was considered more reliable as a predictor of the sexual experience of that partner than the variable recording sexual experience of partner itself, which was therefore not included in analyses (see chapter 11).

I first describe the distribution of the available sexual behaviour variables.

14.1.2.1 Description of the sexual behaviour variables

14.1.2.1.1 Age at first sexual intercourse

The frequency distribution of calendar age at first sexual intercourse is shown in table 14.1.2.1.1. Strictly speaking, this variable is time-varying, and the ages in this table are those which were measured at the earliest of the date of the first detection of a cervical HPV infection of any type in a cervical sample, or the date of the end of follow-up. But note that there were only 21 virgins at study entry in the entire cohort of 2,011 women recruited into the study described in this thesis.

The median age at first sexual intercourse was 16 years (range 5 to 20).

Table 14.1.2.1.1. Frequency distribution of the age at first sexual intercourse.

Age at first sexual intercourse (years)[a]	Frequency[b] (n=1075)	%
5	1	0.1%
11	1	0.1%
12	7	0.7%
13	34	3.2%
14	119	11.1%
15	227	21.1%
16	357	33.2%
17	211	19.6%
18	98	9.1%
19	17	1.6%
20	3	0.3%

[a]Measured at the date of the first detection of HPV DNA of any type in a cervical sample; [b]Number of women.

14.1.2.1.2 Number of changes of sexual partner

The frequency distribution of the number of changes of sexual partner during follow-up is shown in table 14.1.2.1.2. This variable is time-varying, and the values in this table are those which were measured at the earliest of the date of the first detection of cervical HPV infection of any type in a cervical sample, or the date of the end of follow-up.

Most women did not change sexual partner, or did not acquire new sexual partners, during follow-up. Among women who did change sexual partner, the median number of changes was 2 (range 1 to 17).

Table 14.1.2.1.2. Frequency distribution of the number of changes of sexual partner during follow-up.

Number of changes of sexual partner during follow-up[a]	Frequency[b] (n=1075)	%
0	574	53.4%
1	243	22.6%
2	125	11.6%
3	67	6.2%
4	23	2.1%
5	13	1.2%
6	12	1.1%
7	5	0.5%
8	4	0.4%
9	1	0.1%
10	3	0.3%
15	1	0.1%
16	2	0.2%
17	2	0.2%

[a]Measured at the date of the first detection of HPV DNA of any type in a cervical sample; [b]Number of women.

14.1.2.1.3 Age of a woman's oldest sexual partner

The frequency distribution of the age of a woman's oldest sexual partner is shown in table 14.1.2.1.3. This variable is time-varying, and the ages in this table are those which were measured at the earliest of the date of the first detection of HPV infection of any type in a cervical sample, or the date of the end of follow-up.

The median age of a woman's oldest sexual partner was 21 years (range 15 to 50).

Table 14.1.2.1.3. Frequency distribution of the age of a woman's oldest sexual partner.

Age of oldest sexual partner (years)[a]	Frequency[b] (n=1075)	%
15	6	0.6%
16	22	2.0%
17	46	4.3%
18	99	9.2%
19	133	12.4%
20	132	12.3%
21	103	9.6%
22	104	9.7%
23	93	8.7%
24	72	6.7%
25	60	5.6%
26	46	4.3%
27	36	3.3%
28	32	3.0%
29	20	1.9%
30	22	2.0%
31	12	1.1%
32	8	0.7%
33	4	0.4%
34	5	0.5%
35	4	0.4%
36	1	0.1%
37	3	0.3%
38	1	0.1%
39	1	0.1%
40	2	0.2%
42	4	0.4%
43	1	0.1%
44	1	0.1%
45	1	0.1%
50	1	0.1%

[a]Measured at the date of the first detection of HPV DNA of any type in a cervical sample; [b]Number of women.

14.1.2.2 Identification of sexual behaviour variables which may act as confounders of the association between smoking and the risk of an incident cervical HPV infection of any type

The results of the univariate analysis to identify potential confounders of the association between smoking and the risk of an incident cervical HPV infection of any type are shown in table 14.1.2.2. In this table, all models are compared to the model called “Quadratic”, the model containing quadratic midpoint-modelling of the hazard function only (see section 6.5.2: this is the appropriate “null” model in this case). Differences in deviance between two nested models were referred to a chi-squared distribution with degrees of freedom equal to the difference in degrees of freedom between the two respective models. Estimates of association are not presented: it is the existence of an association which is relevant in this analysis.

By inspection, it can be seen that all variables were significantly associated with the risk of an incident cervical HPV infection of any type, except for age at first sexual intercourse; this was true whether this variable was analysed as a categorical or as a continuous covariate. In multivariate analyses including smoking, only those sexual behaviour variables which remained statistically significant were retained in the final model.

Table 14.1.2.2. Model deviance and degrees of freedom obtained from models analysing the association between the sexual behaviour variables and the risk of an incident cervical HPV infection of any type.

Model[a]	Deviance	Degrees of freedom	Significant
Null	2534.208	3240	-
Quadratic	2517.961	3238	-
Lifetime number of sexual partners (C)	2464.043	3233	Yes
Lifetime number of sexual partners (V)	2484.147	3237	Yes
Age at first sexual intercourse (C)	2514.9	3234	No
Age at first sexual intercourse (V)	2516.568	3237	No
Age of oldest sexual partner (C)	2475.495	3233	Yes
Age of oldest sexual partner (V)	2489.347	3237	Yes
New sexual partner in the previous 6 months (C)	2499.664	3237	Yes
Age of oldest partner in previous 6 months (C)	2495.879	3233	Yes
Age of oldest partner in previous 6 months (V)	2499.672	3237	Yes

[a]C=Variable analysed as categorical covariate, V=Variable analysed as continuous covariate.

14.1.2.3 HPV infection of any type

For each of the variables used to measure exposure to smoking, only lifetime number of sexual partners and age of the oldest sexual partner were statistically significant in multivariate models examining the association between smoking and the risk of an incident cervical HPV infection of any type. Therefore, in this section the phrase “controlling for sexual behaviour” is used as shorthand to indicate that associations are controlled for these two sexual behaviour variables only.

14.1.2.3.1 The multivariate analysis of the association between smoking status at study entry and the risk of an incident cervical HPV infection of any type, after controlling for sexual behaviour

The association between smoking status at study entry and the risk of an incident cervical HPV infection of any type was not statistically significant, after controlling for sexual behaviour, using any of the definitions of this exposure variable (table 14.1.2.3.1). The slightly elevated risks seen in the uncontrolled analysis are essentially reduced to null in the controlled analysis.

Table 14.1.2.3.1. The multivariate analysis of the association between smoking status at study entry and the risk of an incident cervical infection with HPV DNA of any type, after controlling for sexual behaviour.

Smoking Status At Study Entry	Number of Women (%)	Number of Events	Hazards Ratio[a] (95% CI)
Never-smoker	626 (58.2)	228	1.00 (Reference)
Ex-smoker	93 (8.7)	38	1.07 (0.75 to 1.52)
Current smoker	356 (33.1)	141	1.01 (0.81 to 1.26)
			χ^2=0.15; 2df; p=0.96[b]
Non-smoker	719 (65.0)	266	1.00 (Reference)
Current smoker	356 (35.0)	141	1.00 (0.81 to 1.24)
			χ^2=0.00; 1df; p=0.99[b]
Non-smoker	626 (58.2)	228	1.00 (Reference)
Ever smoker	449 (41.8)	179	1.02 (0.83 to 1.26)
			χ^2=0.05; 1df; p=0.83[b]

[a]Controlling for lifetime number of sexual partners and age of oldest sexual partner; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.1.2.3.2 The multivariate analysis of the association between the age at starting smoking (as recorded at study entry) and the risk of an incident cervical HPV infection of any type, after controlling for sexual behaviour

The multivariate analysis of the association between age at starting smoking (as recorded at study entry) and the risk of an incident cervical HPV infection of any type, after controlling for sexual behaviour, is described in table 14.1.2.3.2. The association was not statistically significant. Compared to the uncontrolled analysis, in the analysis treating age at starting smoking as a continuous variable, the hazards ratio and p-value decreased slightly; the hazards ratio in the 16-or-over age category (≥16) also decreased.

Table 14.1.2.3.2. The multivariate analysis of the association between the age at starting smoking (as recorded at study entry) and the risk of an incident cervical infection with HPV DNA of any type, after controlling for sexual behaviour.

Age At Starting Smoking (at study entry)	Number of Women (%)	Number of Events	Hazards Ratio[a] (95% CI)
Continuous (per year)[b]	1075 (100)	407	0.93 (0.86 to 1.02)
			χ^2=2.47; 1df; p=0.12[c]
Never-smoker	626 (58.2)	228	1.00 (Reference)
≤13	100 (9.3)	41	1.13 (0.80 to 1.60)
14-15	182 (16.9)	83	1.15 (0.88 to 1.50)
≥16	167 (15.5)	55	0.83 (0.62 to 1.13)
			χ^2=3.87; 3df; p=0.28[c]

[a]Controlling for lifetime number of sexual partners and age of oldest sexual partner; [b]Includes a variable indicating "ever smoker" status at study entry; [c]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.1.2.3.3 The multivariate analysis of the association between a history of ever having smoked and the risk of an incident cervical HPV infection of any type, after controlling for sexual behaviour

The multivariate analysis of the association between a history of ever having smoked and the risk of an incident cervical HPV infection of any type, after controlling for sexual behaviour, is described in table 14.1.2.3.3. The association was not statistically significant. The small (39%), but statistically significant, increased risk in women who had ever smoked compared to those who had never smoked seen in the univariate analysis, was reduced to a 15% (1.15-1.00) non-significant increase after controlling for sexual behaviour.

Table 14.1.2.3.3. The multivariate analysis of the association between a history of ever having smoked and the risk of an incident cervical infection with HPV DNA of any type, after controlling for sexual behaviour.

Have you ever been a cigarette smoker?	Number of Women (%)	Number of Events	Hazards Ratio[a] (95% CI)
No	550 (51.2)	192	1.00 (Reference)
Yes	525 (48.8)	215	1.15 (0.93-1.41)
			χ^2=1.73; 1df; p=0.19[b]

[a]Controlling for lifetime number of sexual partners and age of oldest sexual partner; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.1.2.3.4 The multivariate analysis of the association between the time since first starting smoking and the risk of an incident cervical HPV infection of any type, after controlling for sexual behaviour

The multivariate analysis of the association between the time since first starting smoking and the risk of an incident cervical infection with HPV DNA of any type, after controlling for sexual behaviour, is described in table 14.1.2.3.4. The association was not statistically significant whether the variable was analysed as a continuous, or as a categorical, covariate. Compared with the univariate analysis, all hazards ratios were attenuated after controlling for sexual behaviour.

Table 14.1.2.3.4. The multivariate analysis of the association between the time since first starting smoking and the risk of an incident cervical infection with HPV DNA of any type, after controlling for sexual behaviour.

Time Since First Starting Smoking	Number of Women (%)	Number of Events	Hazards Ratio[a] (95% CI)
Continuous (per year)[b]	1075 (100)	407	0.95 (0.89 to 1.01)
			χ^2=2.55; 1df; p=0.11[c]
Never-smoker	550 (51.2)	192	1.00 (Reference)
(0, 36 months)[d]	181 (16.8)	85	1.27 (0.97 to 1.65)
[36, 72)[d]	242 (22.5)	91	1.06 (0.81 to 1.37)
[72, +)[d]	102 (9.5)	39	1.13 (0.78 to 1.65)
			χ^2=3.10; 3df; p=0.38[c]
Never-smoker	550 (51.2)	192	1.00 (Reference)
(0, 36 months)[d]	181 (16.8)	85	1.27 (0.97 to 1.66)
[36, 72)[d]	242 (22.5)	91	1.06 (0.81 to 1.37)
[72, 84)[d]	40 (3.7)	18	1.18 (0.71 to 1.95)
[84, +)[d]	62 (5.8)	21	1.10 (0.68 to 1.79)
			χ^2=3.14; 4df; p=0.53[c]

[a]Controlling for lifetime number of sexual partners and age of oldest sexual partner; [b]Model includes a time-varying indicator variable indicating "ever smoker" status; [c]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [d][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.1.2.3.5 The multivariate analysis of the association between the cumulative duration of smoking and the risk of an incident cervical HPV infection of any type, after controlling for sexual behaviour

The multivariate analysis of the association between the cumulative duration of smoking and the risk of an incident cervical HPV infection of any type, after controlling for sexual behaviour, is shown in table 14.1.2.3.5. There was a small, but statistically significant, decrease in the risk of an incident cervical HPV infection of any type with increasing cumulative duration of smoking when this variable was analysed as a continuous variable. The association was not significant when this variable was analysed as a categorical covariate, and no clear-cut trend was apparent. In the univariate analysis, the reverse was true i.e. the categorical variable was a significant predictor of risk, but the continuous variable was not.

Table 14.1.2.3.5. The multivariate analysis of the association between the cumulative duration of smoking and the risk of an incident cervical infection with HPV DNA of any type, after controlling for sexual behaviour.

Cumulative Duration Of Smoking	Number of Women (%)	Number of Events	Hazards Ratio[a] (95% CI)
Continuous (per year)[b]	1075 (100.0)	407	0.93 (0.87 to 1.00)
			χ^2=4.43; 1df; p=0.04[c]
Never-smoker	550 (51.2)	192	1.00 (Reference)
(0, 36 months)[d]	251 (23.3)	113	1.31 (1.03 to 1.67)
[36, 72)[d]	205 (19.1)	75	0.95 (0.72 to 1.27)
[72, +)[d]	69 (6.4)	27	1.14 (0.74 to 1.76)
			χ^2=6.06; 3df; p=0.11[c]

[a]Controlling for lifetime number of sexual partners and age of oldest sexual partner; [b]Model includes a time-varying indicator variable indicating "ever smoker" status; [c]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [d][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.1.2.3.6 The multivariate analysis of the association between the duration of the current smoking episode and the risk of an incident cervical HPV infection of any type, after controlling for sexual behaviour

The multivariate analysis of the association between the duration of the current smoking episode and the risk of an incident cervical infection with HPV DNA of any type, after controlling for sexual behaviour, is described in table 14.1.2.3.6. The association was not statistically significant, whether this exposure variable was analysed as either a categorical or as a continuous variable. In the univariate analysis, this variable was not significant when analysed as a continuous variable, or as a categorical variable with non-smokers as reference group, but *was* significant when ex-smokers were separately distinguished from never-smokers. The p-values for both categorical variables *decreased* after controlling for sexual behaviour, whereas that of the continuous variable *increased*.

Table 14.1.2.3.6. The multivariate analysis of the association between the duration of the current smoking episode and the risk of an incident cervical infection with HPV DNA of any type, after controlling for sexual behaviour.

Duration Of Current Smoking Episode	Number of Women (%)	Number of Events	Hazards Ratio[a] (95% CI)
Continuous (per year)[b]	1075 (100)	407	0.95 (0.88 to 1.02)
			χ^2=2.10; 1df; p=0.15[c]
Non-smoker	679 (63.2)	243	1.00 (Reference)
(0, 24 months)[d]	132 (12.3)	60	1.23 (0.92 to 1.65)
[24, 48)[d]	120 (11.2)	52	1.06 (0.77 to 1.44)
[48, +)[d]	144 (13.4)	52	0.96 (0.70 to 1.31)
			χ^2=2.19; 3df; p=0.53[c]
Never-smoker	550 (51.2)	192	1.00 (Reference)
Ex-smoker	129 (12.0)	51	1.22 (0.88 to 1.67)
(0, 24 months)[d]	132 (12.3)	60	1.28 (0.95 to 1.73)
[24, 48)[d]	120 (11.2)	52	1.10 (0.80 to 1.52)
[48, +)[d]	144 (13.4)	52	1.00 (0.72 to 1.39)
			χ^2=3.61; 4df; p=0.46[c]

[a]Controlling for lifetime number of sexual partners and age of oldest sexual partner; [b]Model includes a time-varying indicator variable indicating "current smoker" status; [c]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [d][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.1.2.3.7 The multivariate analysis of the association between pack-years smoked and the risk of an incident cervical HPV infection of any type, after controlling for sexual behaviour

This analysis was restricted to the 1,038 (out of 1,075) women for whom quantity smoked was known throughout follow-up.

The multivariate analysis of the association between pack-years smoked and the risk of an incident cervical infection with HPV DNA of any type, after controlling for sexual behaviour, is described in table 14.1.2.3.7. When analysed as a continuous

variable, and after controlling for sexual behaviour, there was a statistically significant association, with a decrease in the risk of an incident cervical HPV infection of any type of 13% (1.00-0.87) for every increase of one in pack-years smoked; in the univariate analysis, the 8% decrease observed was not significant.

In contrast, in the univariate analysis the association was highly significant when the variable was analysed as a categorical variable, but in the multivariate analysis the association was no longer significant. However, there was no evidence of a departure from linearity for the pack-years variable in this analysis. Therefore, once lifetime number of sexual partner and age of oldest sexual partner had been controlled for, there was a significant decrease in the risk of an incident cervical HPV infection of any type as pack-years smoked increased.

Table 14.1.2.3.7. The multivariate analysis of the association between pack-years smoked and the risk of an incident cervical infection with HPV DNA of any type, after controlling for sexual behaviour.

Pack-years Smoked	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)[a]
Continuous (per pack-year)[b]	1038 (100)	394	0.87 (0.76 to 1.00)
			χ^2=4.36; 1df; p=0.04[c]
Never-smoker	550 (51.2)	192	1.00 (Reference)
(0 , 1)[d]	232 (22.4)	110	1.36 (1.06 to 1.74)
[1 , 2)[d]	143 (13.8)	50	0.99 (0.71 to 1.37)
[2 , 3)[d]	73 (7.0)	30	1.15 (0.77 to 1.71)
[3 , +)[d]	40 (3.9)	12	0.73 (0.39 to 1.37)
			χ^2=8.50; 4df; p=0.07[c]

[a]Controlling for lifetime number of sexual partners and age of oldest sexual partner; [b]Model includes a time-varying indicator variable indicating "current smoker" status; [c]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [d][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.1.2.3.8 The multivariate analysis of the association between current smoking status and the risk of an incident cervical HPV infection of any type, after controlling for sexual behaviour

The multivariate analysis of the association between current smoking status and the risk of an incident cervical infection with HPV DNA of any type, after controlling for sexual behaviour, is shown in table 14.1.2.3.8. After controlling for sexual behaviour, the association between current smoking status and the risk of an incident cervical HPV infection of any type was not statistically significant; this is in contrast to the univariate analysis, where all three of the variables used to measure current smoking status were significant.

Table 14.1.2.3.8. The multivariate analysis of the association between current smoking status and the risk of an incident cervical infection with HPV DNA of any type, after controlling for sexual behaviour.

Current Smoking Status	Number of Women (%)	Number of Events	Hazards Ratio[a] (95% CI)
Never-smoker	550 (51.2)	192	1.00 (Reference)
Ex-smoker	129 (12.0)	51	1.22 (0.89 to 1.68)
Current smoker	396 (36.8)	164	1.13 (0.90 to 1.41)
			χ^2=1.97; 2df; p=0.37[b]
Non-smoker	679 (63.2)	243	1.00 (Reference)
Current smoker	396 (36.8)	164	1.08 (0.87 to 1.33)
			χ^2=0.51; 1df; p=0.48[b]
Never-smoker	550 (51.2)	192	1.00 (Reference)
Ever-smoker	525 (48.8)	215	1.15 (0.93-1.41)
			χ^2=1.73; 1df; p=0.19[b]

[a]Controlling for lifetime number of sexual partners and age of oldest sexual partner; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.1.2.3.9 The multivariate analysis of the association between current smoking intensity and the risk of an incident cervical HPV infection of any type, after controlling for sexual behaviour

This analysis was restricted to the 1,038 (out of 1,075) women for whom quantity smoked was known throughout follow-up.

The multivariate analysis of the association between current smoking intensity and the risk of an incident cervical infection with HPV DNA of any type, after controlling for sexual behaviour, is described in table 14.1.2.3.9. The association was not statistically significant: the significant association between the categorical variable and the risk of an incident cervical HPV infection of any type seen in the univariate analysis was “explained” by controlling for lifetime number of sexual partners and age of oldest sexual partner (sexual behaviour).

Table 14.1.2.3.9. The multivariate analysis of the association between current smoking intensity and the risk of an incident cervical infection with HPV DNA of any type, after controlling for sexual behaviour.

Current Smoking Intensity	Number of Women (%)	Number of Events	Hazards Ratio[a] (95% CI)
Continuous (per category)[b,c]	1038 (100)	394	0.92 (0.78 to 1.08)
			χ^2=1.04; 1df; p=0.34[d]
Never-smoker	550 (53.0)	192	1.00 (Reference)
0 per day	104 (10.0)	40	1.22 (0.89 to 1.68)
1 to 9 per day	196 (18.9)	92	1.25 (0.97 to 1.62)
10 or more per day	188 (18.1)	70	0.99 (0.74 to 1.32)
			χ^2=4.16; 3df; p=0.24[d]

[a]Controlling for lifetime number of, and age of oldest, sexual partner; [b]Model includes a time-varying indicator variable indicating "current smoker" status; [c]the increase in risk for every increase of one in category of smoking intensity; [d]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.1.2.4 HPV16 infection

Only lifetime number of sexual partners was found to be associated with the incidence of cervical HPV16 infection in a univariate analysis. In analyses of the association between exposure to smoking and the risk of an incident cervical HPV16 infection, after controlling for sexual behaviour, the duration of the current smoking episode and pack-years smoked were both found to be statistically significant predictors of risk. However, the associated estimates of effect (hazards ratios) were small in magnitude.

14.1.2.5 HPV18 infection

There were no significant associations between sexual behaviour and the incidence of HPV18 infection in univariate analyses. Therefore analyses controlling for sexual behaviour were not necessary in this case.

Summary

In analyses of the association between exposure to smoking and the risk of an incident cervical HPV infection of any type, after controlling for sexual behaviour, the cumulative duration of smoking and pack-years smoked were found to be significant predictors of risk. Only lifetime number of sexual partners was found to be associated with the incidence of HPV16 infection in a univariate analysis, but after controlling for sexual behaviour, duration of current smoking episode and pack-years were both found to be significant predictors of risk: estimates of effect were all small in magnitude. There were no significant associations between sexual behaviour and the incidence of cervical HPV18 infection in univariate analyses, therefore analyses controlling for sexual behaviour were not performed.

14.2 THE DURATION OF INCIDENT CERVICAL HPV INFECTION IN RELATION TO SMOKING HISTORY

14.2.1 Univariate analysis

In the analysis of the duration of incident cervical HPV infection, of any type or of types 16 and 18, the event of interest is *clearing* an HPV infection. Therefore, since a hazards ratio (HR) greater than one means the event of interest (clearance) is

more likely to occur, exposures with HR>1 are associated with infections of shorter duration, i.e. the event is more likely to occur, so the infection is more likely to be cleared, so the infection is more likely to be of short duration. In contrast, a HR<1 indicates that the exposure is protective, and therefore the event of interest is *less* likely to occur, i.e. the relevant exposure is associated with infections of longer duration.

14.2.1.1 HPV infection of any type

The study population for this analysis comprises 328 women who had an incident cervical HPV infection of any type, and who had further follow-up after the first detection of HPV DNA, i.e. women had to have an incident infection to be able to clear that infection.

14.2.1.1.1 The univariate analysis of the association between a history of ever having smoked and the duration of an incident cervical HPV infection of any type

The univariate analysis of the association between a history of ever having smoked and the risk of clearing an incident cervical HPV infection of any type is described in table 14.2.1.1.1. The association was not statistically significant, with a hazards ratio very close to the null value of 1.

Table 14.2.1.1.1. The univariate analysis of the association between a history of ever having smoked and the risk of clearing an incident cervical infection with HPV DNA of any type.

Have you ever been a cigarette smoker?	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
No	149 (45.4)	120	1.00 (Reference)
Yes	179 (54.6)	149	1.05 (0.82 to 1.35)
			χ^2=0.17; 1df; p=0.68[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.2.1.1.2 The univariate analysis of the association between the time since first starting smoking and the duration of an incident cervical HPV infection of any type

The univariate analysis of the association between the time since first starting smoking and the risk of clearing an incident cervical HPV infection of any type is described in table 14.2.1.1.2. The association was not statistically significant.

Table 14.2.1.1.2. The univariate analysis of the association between the time since first starting smoking and the risk of clearing an incident cervical infection with HPV DNA of any type.

Time Since First Starting Smoking	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per year)[a]	328 (100.0)	269	0.97 (0.90 to 1.04)
			χ^2=1.05; 1df; p=0.31[b]
Never-smoker	149 (45.4)	120	1.00 (Reference)
(0, 36 months)[c]	50 (15.2)	41	1.21 (0.84 to 1.76)
[36, 72)[c]	75 (22.9)	63	0.96 (0.70 to 1.32)
[72, +)[c]	54 (16.5)	45	1.06 (0.73 to 1.54)
			χ^2=1.42; 3df; p=0.70[b]

[a]Model includes a time-varying indicator variable indicating “ever smoker” status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [c][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.2.1.1.3 The univariate analysis of the association between the cumulative duration of smoking and the duration of an incident cervical HPV infection of any type

The univariate analysis of the association between the cumulative duration of smoking and the risk of clearing an incident cervical HPV infection of any type is described in table 14.2.1.1.3. There was a slight non-significant decrease in the risk of clearing an incident cervical HPV infection of any type with increasing cumulative duration of smoking, when this variable was analysed as a continuous variable, controlling for ever having smoked. When the variable was analysed as a categorical variable, the model failed to converge. Alternative categorisations of the variable also led to failure in convergence of the model.

Table 14.2.1.1.3. The univariate analysis of the association between the cumulative duration of smoking and the risk of clearing an incident cervical infection with HPV DNA of any type.

Cumulative Duration Of Smoking	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous[a] (per year)	328 (100.0)	269	0.96 (0.90 to 1.02)
			χ^2=1.64; 1df; p=0.20[b]

[a]Model includes a time-varying indicator variable indicating 'ever smoker' status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.2.1.1.4 The univariate analysis of the association between the duration of the current smoking episode and the duration of an incident cervical HPV infection of any type

The univariate analysis of the association between the duration of the current smoking episode and the risk of clearing an incident cervical HPV infection of any type is described in table 14.2.1.1.4. The association was not statistically significant, and all hazards ratios were close to the null value of one, whether the exposure variable was measured on a continuous, or on a categorical, scale.

Table 14.2.1.1.4. The univariate analysis of the association between the duration of the current smoking episode and the risk of clearing an incident cervical infection with HPV DNA of any type.

Duration Of Current Smoking Episode	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Duration of current smoking episode (per year)[a]	328 (100.0)	269	0.97 (0.90 to 1.05)
			χ^2=0.60; 1df; p=0.44[b]
Non-smoker	192 (58.5)	158	1.00 (Reference)
(0, 24 months)[c]	41 (12.5)	36	1.14 (0.78 to 1.67)
[24, 48)[c]	36 (11.0)	28	1.09 (0.72 to 1.66)
[48, +)[c]	59 (18.0)	47	0.85 (0.61 to 1.21)
			χ^2=1.83; 3df; p=0.61[b]
Never-smoker	149 (51.2)	120	1.00 (Reference)
Ex-smoker	43 (12.0)	38	1.14 (0.78 to 1.66)
(0, 24 months)[c]	41 (12.5)	36	1.18 (0.80 to 1.74)
[24, 48)[c]	36 (11.0)	28	1.12 (0.73 to 1.73)
[48, +)[c]	59 (18.0)	47	0.89 (0.62 to 1.26)
			χ^2=2.29; 4df; p=0.68[b]

[a]Model includes a time-varying indicator variable indicating “current smoking” status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [c][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.2.1.1.5 The univariate analysis of the association between pack-years smoked and the duration of an incident cervical HPV infection of any type

This analysis was restricted to the 319 women for whom the quantity smoked was known throughout follow-up.

The univariate analysis of the association between pack-years smoked and the risk of clearing an incident cervical HPV infection of any type is described in table 14.2.1.1.5. When analysed as a continuous variable, the hazards ratio decreased by 7% (1.00-0.93) for every increase of one pack year, but the association was not statistically significant. When analysed as a categorical variable, the risk estimates showed no consistent pattern across categories, but a greater number of pack-years smoked was again associated with infections of longer duration.

Table 14.2.1.1.5. The univariate analysis of the association between pack-years smoked and the risk of clearing an incident cervical infection with HPV DNA of any type.

Pack-years Smoked	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Cumulative quantity smoked (per pack-year)	319 (100.0)	262	0.93 (0.81 to 1.07)
			χ^2=1.00; 1df; p=0.32[a]
Never-smoker	149 (46.7)	120	1.00 (Reference)
(0 , 1)[b]	75 (23.5)	65	1.25 (0.91 to 1.71)
[1 , 2)[b]	45 (14.1)	35	0.92 (0.62 to 1.36)
[2 , 3)[b]	25 (7.8)	22	0.95 (0.59 to 1.54)
[3 , +)[b]	25 (7.8)	20	0.88 (0.53 to 1.47)
			χ^2=3.24; 4df; p=0.52[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [b][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.2.1.1.6 The univariate analysis of the association between current smoking status and the duration of an incident cervical HPV infection of any type

The univariate analysis of the association between current smoking status and the risk of clearing an incident cervical HPV infection of any type is described in table 14.2.1.1.6. Current smoking status was not significantly associated with the risk of clearing an incident cervical HPV infection of any type in any of these three analyses; current ex-smokers were more likely than current smokers to clear their infection, but the difference was negligible.

Table 14.2.1.1.6. The univariate analysis of the association between current smoking status and the risk of clearing an incident cervical infection with HPV DNA of any type.

Current Smoking Status	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Never-smoker	149 (45.4)	120	1.00 (Reference)
Ex-smoker	43 (13.1)	38	1.14 (0.78 to 1.66)
Current smoker	136 (41.5)	111	1.03 (0.78 to 1.34)
			χ^2=0.46; 2df; p=0.79[a]
Non-smoker	192 (58.5)	158	1.00 (Reference)
Current smoker	136 (41.5)	111	1.00 (0.77 to 1.28)
			χ^2=0.0009; 1df; p=0.98[a]
Never-smoker	149 (45.4)	120	1.00 (Reference)
Ever-smoker	179 (54.6)	149	1.05 (0.82 to 1.35)
			χ^2=0.17; 1df; p=0.68[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.2.1.1.7 The univariate analysis of the association between current smoking intensity and the duration of an incident cervical HPV infection of any type

This analysis was restricted to the 319 (out of 328) women for whom the quantity smoked was known throughout follow-up. The model did not converge for this analysis; however, the frequency distribution of this variable and the number of clearance events in each category are shown in table 14.2.1.1.7.

Table 14.2.1.1.7. The univariate analysis of the association between current smoking intensity and the risk of clearing an incident cervical infection with HPV DNA of any type.

Current Smoking Intensity	Number of Women (%)	Number of Events
Continuous (per category)	319 (100.0)	262
Never-smoker	149 (46.7)	120
0 per day	36 (11.3)	33
1 to 9 per day	77 (24.1)	63
10 or more per day	57 (17.9)	46

14.2.1.2 HPV16 infection

The study population for this analysis comprises 90 women who had an incident cervical HPV16 infection, and who had further follow-up after the first detection of HPV16 DNA.

14.2.1.2.1 The univariate analysis of the association between a history of ever having smoked and the duration of an incident cervical HPV16 infection

The univariate analysis of the association between a history of ever having smoked and the risk of clearing an incident cervical HPV16 infection is described in table 14.2.1.2.1. The association was not statistically significant: the risk of clearing an incident cervical HPV16 infection increased by 41% (1.41-1.00) for ever- relative to never-smokers, compared to only 5% for HPV infections of any type.

Table 14.2.1.2.1. The univariate analysis of the association between a history of ever having smoked and the risk of clearing an incident cervical HPV16 infection.

Have you ever been a cigarette smoker?	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
No	38 (42.2)	31	1.00 (Reference)
Yes	52 (57.8)	45	1.41 (0.87 to 2.28)
			χ^2=2.07; 1df; p=0.15[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.2.1.2.2 The univariate analysis of the association between the time since first starting smoking and the duration of an incident cervical HPV16 infection

The univariate analysis of the association between the time since first starting smoking and the risk of clearing an incident cervical HPV16 infection is described in table 14.2.1.2.2. The association was not statistically significant.

Table 14.2.1.2.2. The univariate analysis of the association between the time since first starting smoking and the risk of clearing an incident cervical HPV16 infection.

Time Since First Starting Smoking	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per year)[a]	90 (100.0)	76	0.99 (0.87 to 1.13)
			χ^2=0.01; 1df; p=0.91[b]
Never-smoker	38 (42.2)	31	1.00 (Reference)
(0, 36 months)[c]	17 (18.9)	15	1.33 (0.70 to 2.52)
[36, 72)[c]	24 (26.7)	21	1.51 (0.83 to 2.74)
[72, +)[c]	11 (12.2)	9	1.35 (0.61 to 3.00)
			χ^2=2.23; 3df; p=0.53[b]

[a]Model includes a time-varying indicator variable indicating “ever smoker” status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [c][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.2.1.2.3 The univariate analysis of the association between the cumulative duration of smoking and the duration of an incident cervical HPV16 infection.

The univariate analysis of the association between the cumulative duration of smoking and the risk of clearing an incident cervical HPV16 infection is described in table 14.2.1.2.3. The association was not statistically significant, and all hazards ratios were small with wide confidence intervals.

Table 14.2.1.2.3. The univariate analysis of the association between the cumulative duration of smoking and the risk of clearing an incident cervical HPV16 infection.

Cumulative Duration Of Smoking	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per year)	90 (100.0)	76	1.01 (0.89 to 1.15)
			χ^2=0.03; 1df; p=0.86[b]
Never-smoker	38 (42.2)	31	1.00 (Reference)
(0, 36 months)[c]	28 (13.6)	26	1.30 (0.75 to 2.24)
[36, 72)[c]	18 (22.0)	13	1.83 (0.90 to 3.74)
[72, +)[c]	6 (4.4)	6	1.28 (0.50 to 3.31)
			χ^2=2.99; 3df; p=0.39[b]

[a]Model includes a time-varying indicator variable indicating "ever smoker" status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [c][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.2.1.2.4 The univariate analysis of the association between the duration of the current smoking episode and the duration of an incident cervical HPV16 infection

The univariate analysis of the association between the duration of the current smoking episode and the risk of clearing an incident cervical HPV16 infection is described in table 14.2.1.2.4. Overall, the association was not statistically significant, and all hazards ratios were close to the null value of 1, except the hazards ratio for women whose current smoking episode had lasted for four or more years. These women were at a nearly three-fold increased risk of clearing their infections compared to women who had never smoked, i.e. smoking had a protective effect in that smoking was associated with shorter periods of HPV16-positivity. However, the sample size in this category was small, and the confidence interval indicates that the increased risk is also consistent with the risks seen in the other smoking categories.

Table 14.2.1.2.4. The univariate analysis of the association between the duration of the current smoking episode and the risk of clearing an incident cervical HPV16 infection.

Duration Of Current Smoking Episode	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per year)[a]	90 (100.0)	76	1.08 (0.92 to 1.26)
			χ^2=0.88; 1df; p=0.35[b]
Never-smoker	38 (42.2)	31	1.00 (Reference)
Ex-smoker	16 (17.8)	16	1.24 (0.65 to 2.34)
(0, 24 months)[c]	11 (12.2)	9	1.31 (0.61 to 2.85)
[24, 48)[c]	12 (13.3)	8	1.11 (0.49 to 2.49)
[48, +)[c]	13 (14.4)	12	2.89 (1.25 to 6.68)
			χ^2=6.42; 4df; p=0.17[b]

[a]Model includes a time-varying indicator variable indicating "current smoking" status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [c][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.2.1.2.5 The univariate analysis of the association between pack-years smoked and the duration of an incident cervical HPV16 infection

This analysis was restricted to the 88 (out of 90) women for whom quantity smoked was known throughout follow-up. The univariate analysis of the association between pack-years smoked and the risk of clearing an incident cervical HPV16 infection is described in table 14.2.1.2.5. The association was not statistically significant.

Table 14.2.1.2.5. The univariate analysis of the association between pack-years smoked and the risk of clearing an incident cervical HPV16 infection.

Pack-years Smoked	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per pack-year)	88 (100.0)	74	0.98 (0.73 to 1.32)
			χ^2=0.02; 1df; p=0.89[a]
Never-smoker	38 (43.2)	31	1.00 (Reference)
(0 , 1)[b]	28 (31.8)	26	1.49 (0.86 to 2.58)
[1 , 2)[b]	10 (11.4)	6	1.38 (0.54 to 3.53)
[2 , 3)[b]	9 (10.2)	9	1.61 (0.70 to 3.69)
[3 , +)[b,c]	3 (3.4)	2	0.94 (0.22 to 4.02)
			χ^2=2.90; 4df; p=0.57[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [b][a, b) is a mathematical convention denoting the interval ≥a to <b; [c]this categorisation was retained despite the small sample size to aid comparison with the "HPV of any type" analysis.

14.2.1.2.6 The univariate analysis of the association between current smoking status and the duration of an incident cervical HPV16 infection

The univariate analysis of the association between current smoking status and the risk of clearing an incident cervical HPV16 infection is described in table 14.2.1.2.6. Current smoking status was not significantly associated with the risk of clearing an incident cervical HPV16 infection in any of the three analyses using the different definitions of current smoking status. In contrast to the analysis of any HPV type, current ex-smokers were less likely than current smokers to clear a cervical HPV16 infection.

Table 14.2.1.2.6. The univariate analysis of the association between current smoking status and the risk of clearing an incident cervical HPV16 infection.

Current Smoking Status	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Never-smoker	38 (42.2)	31	1.00 (Reference)
Ex-smoker	16 (17.8)	16	1.24 (0.65 to 2.33)
Current smoker	36 (40.0)	29	1.54 (0.90 to 2.63)
			χ^2=2.53; 2df; p=0.28[a]
Non-smoker	54 (60.0)	47	1.00 (Reference)
Current smoker	36 (40.0)	29	1.44 (0.88 to 2.36)
			χ^2=2.09; 1df; p=0.15[a]
Never-smoker	38 (42.2)	31	1.00 (Reference)
Ever-smoker	52 (57.8)	45	1.41 (0.87 to 2.28)
			χ^2=2.07; 1df; p=0.15[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.2.1.2.7 The univariate analysis of the association between current smoking intensity and the duration of an incident cervical HPV16 infection

This analysis was restricted to the 88 (out of 90) women for whom the quantity smoked was known throughout follow-up. The univariate analysis of the association between current smoking intensity and the risk of clearing an incident cervical HPV16 infection is described in table 14.2.1.2.7. The association was not statistically significant.

Table 14.2.1.2.7. The univariate analysis of the association between current smoking intensity and the risk of clearing an incident cervical HPV16 infection.

Current Smoking Intensity	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per category)[a,b]	88 (100.0)	74	1.00 (0.69 to 1.44)
			χ^2=0.0005; 1df; p=0.98[c]
Never-smoker	38 (43.2)	31	1.00 (Reference)
0 per day	14 (15.9)	14	1.24 (0.65 to 2.34)
1 to 9 per day	21 (23.9)	21	1.68 (0.89 to 3.16)
10 or more per day	15 (17.0)	15	1.37 (0.68 to 2.77)
			χ^2=2.78; 1df; p=0.43[c]

[a]Model includes a time-varying indicator variable indicating "current smoking" status; [b]the increase in risk for every increase of 1 in category of smoking intensity; [c]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.2.1.3 HPV18 infection

The sample size was too small, and in particular "clearance events" were too few to carry out an analysis of the association between smoking exposure and the duration of incident HPV18 infections.

14.2.2 Multivariate analysis of the association between smoking and the risk of clearing an incident cervical infection with HPV

Multivariate analyses of the association between smoking and the risk of clearing either an incident cervical infection with HPV of any type, or with HPV16, controlling for sexual behaviour, were undertaken. However, given that there were no significant associations to "explain" in the univariate analyses, and also that in no

instance did the hazards ratios from the univariate analysis change by more than 10% in the corresponding multivariate analyses, there is nothing further of interest to report.

Summary

There were no statistically significant associations between any of the variables measuring exposure to smoking and the risk of clearing either an incident cervical HPV infection of any type, or of clearing an incident cervical HPV16 infection; there were too few clearance events to make the analysis of the clearance of incident cervical HPV18 infections worthwhile. Multivariate analyses revealed nothing noteworthy.

14.3 THE INCIDENCE OF CERVICAL CYTOLOGICAL ABNORMALITY IN RELATION TO SMOKING HISTORY

The association between smoking and cervical cytological abnormality is primarily of interest for those cytological abnormalities which ultimately progress to high-grade CIN. Given that cytological abnormality is a surrogate endpoint for high-grade CIN, and that sufficient data relating to high-grade CIN are available, no presentation of results relating to cytological abnormality will be given here.

14.4 THE INCIDENCE OF HIGH-GRADE CIN IN RELATION TO SMOKING HISTORY

The study population for this analysis comprises 1,075 women who were HPV DNA-negative in their first evaluable virological sample, cytologically normal in their first evaluable cervical smear, and who had further follow-up after this time.

Twenty-eight women in the cohort of 1,075 were diagnosed with incident high-grade cervical intraepithelial neoplasia (CIN) during follow-up.

14.4.1 Univariate analysis

14.4.1.1 The univariate analysis of the association between smoking status at study entry and the risk of incident high-grade CIN

Of the 626 women who had never smoked prior to study entry, 14 (2%) were diagnosed with high-grade CIN during follow-up: diagnosis was made after first exposure to smoking in two (14%) of these women and before first exposure to smoking in the remaining 12 (86%). Of the 449 women who had first become smokers at some point prior to study entry, 14 (3%) were diagnosed with high-grade CIN during follow-up.

The univariate analysis of the association between smoking status at study entry and the risk of incident high-grade CIN is described in table 14.4.1.1. The association

between smoking status at study entry and the risk of incident high-grade CIN, after controlling for HPV infection, was not statistically significant, using any of the three definitions of this variable which were investigated. Women who were ex- or current smokers at study entry were at an increased risk of high-grade CIN compared to women who were never-smokers, but this increase was not significant. Women who were ex-smokers at study entry had a higher risk of high-grade CIN than women who were current smokers at study entry, but there was no significant difference between the point estimates of the hazards ratios.

A woman was categorised as an ex-smoker immediately after she stopped smoking. Varying this definition by including a “lag-time” during which ex-smokers were still classified as current smokers, may have produced different estimates of risk. The two-category variable combining ex-smokers with never-smokers provides a lower bound on the hazards ratio for a variable constructed in this way. Similarly, by combining ex-smokers with current smokers, we obtain an upper-bound. The magnitude of the hazards ratio for ex-smokers alone suggests that, however the ex-smokers are distributed between the never- and current-smokers, the resulting “lag-time” variable will not yield a statistically significant hazard ratios.

Table 14.4.1.1. The univariate analysis of the association between smoking status at study entry and the risk of incident high-grade CIN.

Smoking Status At Study Entry	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Never-smoker	626 (58.2)	14	1.00 (Reference)
Ex-smoker	93 (8.7)	3	1.77 (0.51 to 6.17)
Current smoker	356 (33.1)	11	1.31 (0.59 to 2.91)
			χ^2=0.95; 2df; p=0.62[a]
Non-smoker	719 (65.0)	17	1.00 (Reference)
Current smoker	356 (35.0)	11	1.21 (0.57 to 2.61)
			χ^2=0.25; 1df; p=0.62[a]
Non-smoker	626 (58.2)	14	1.00 (Reference)
Ever-smoker	449 (41.8)	14	1.39 (0.66 to 2.93)
			χ^2=0.76; 1df; p=0.38[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

In women who had never smoked prior to study entry, the cumulative risks of incident high-grade CIN at one, two and three years after study entry were 19.5%, 35.5% and 43.4% respectively.

14.4.1.2 The univariate analysis of the association between the age at starting smoking and the risk of incident high-grade CIN

Four hundred and forty nine (41.8%) women had regularly smoked cigarettes at some point prior to study entry; the median age at starting smoking for these women was 15 years (range 9 to 19).

The univariate analysis of the association between the age at starting smoking, as recorded at study entry, and the risk of incident high-grade CIN is described in table

14.4.1.2. The association between the age at starting smoking at study entry and the risk of incident high-grade CIN was not statistically significant. When analysed as a categorical variable, no individual category was statistically significant, and although there was a suggestion of a decreasing risk with increasing age at starting smoking, the trend was not consistent. In contrast, when analysed as a continuous variable, the point estimate of the hazards ratio suggested an increasing risk with increasing age at starting smoking. There is a danger of over-interpreting these results: numbers of events are small, all confidence intervals are wide, and no individual estimates are statistically significant; this may be the best interpretation of these results.

Table 14.4.1.2. The univariate analysis of the association between the age at starting smoking as measured at study entry and the risk of incident high-grade CIN.

Age At Starting Smoking (at study entry)	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per year)[a]	1075 (100)	28	1.07 (0.78 to 1.45)
			χ^2=0.17; 1df; p=0.68[b]
Never-smoker	626 (58.2)	14	1.00 (Reference)
≤13	100 (9.3)	4	1.74 (0.57 to 5.34)
14-15	182 (16.9)	5	1.14 (0.41 to 3.18)
≥16	167 (15.5)	5	1.48 (0.53 to 4.12)
			χ^2=1.18; 3df; p=0.76[b]

[a]Includes a variable indicating “ever smoker” status at study entry; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.4.1.3 The univariate analysis of the association between a history of ever having smoked and the risk of incident high-grade CIN

The univariate analysis of the association between a history of ever having smoked and the risk of incident high-grade CIN is described in table 14.4.1.3. Just over half of all women had smoked cigarettes at some point prior to the first diagnosis of high-grade CIN. Women who had ever smoked had a small but non-significant increase in the risk of incident high-grade CIN, compared to women who had never smoked.

Table 14.4.1.3. The univariate analysis of the association between a history of ever having smoked and the risk of incident high-grade CIN.

Have you ever been a cigarette smoker?	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
No	535 (49.8)	12	1.00 (Reference)
Yes	540 (50.2)	16	1.27 (0.60 to 2.69)
			χ^2=0.38; 1df; p=0.54[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.4.1.4 The univariate analysis of the association between the time since first starting smoking and the risk of incident high-grade CIN

At the earliest of the date of first diagnosis of high-grade CIN and the end of follow-up, median time since first starting smoking was 50.4 months (range 2.2 to 188.0) in women who had ever smoked.

The univariate analysis of the association between the time since starting smoking and the risk of incident high-grade CIN is described in table 14.4.1.4. There was a very slight and non-significant decrease in the risk of incident high-grade CIN with

increasing time since starting smoking, when this variable was analysed as a continuous variable, controlling for ever having smoked. The risk of incident high-grade CIN decreased by 6% (1.00-0.94) for every increase of one year in time since starting smoking. When the variable was analysed as a categorical covariate, with never-smokers as reference group, again the time since first starting smoking was not significantly associated with the risk of incident high-grade CIN. The trend suggested in the continuous analysis was not evident in the categorical analysis, but numbers of events were small and confidence intervals were all wide. In particular, the largest hazards ratio was seen for women for whom 72 to 84 months had elapsed since they started smoking, but the confidence interval indicates that the data are also consistent with a halving of risk.

Table 14.4.1.4. The univariate analysis of the association between the time since first starting smoking and the risk of incident high-grade CIN.

Time Since First Starting Smoking	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per year)[a]	1075 (100)	28	0.94 (0.78 to 1.13)
			χ^2=0.46; 1df; p=0.50[b]
Never-smoker	535 (49.8)	12	1.00 (Reference)
(0, 36 months)[c]	135 (12.6)	3	1.13 (0.31 to 4.11)
[36, 72)[c]	232 (21.6)	6	1.31 (0.49 to 3.51)
[72, +)[c]	173 (16.1)	7	1.31 (0.47 to 3.60)
			χ^2=0.43; 3df; p=0.93[b]
Never-smoker	535 (49.8)	12	1.00 (Reference)
(0, 36 months)[c]	135 (12.6)	3	1.14 (0.31 to 4.16)
[36, 72)[c]	232 (21.6)	6	1.33 (0.49 to 3.57)
[72, 84)[c]	51 (4.7)	3	2.07 (0.57 to 7.47)
[84, +)[c]	122 (11.3)	4	0.96 (0.28 to 3.34)
			χ^2=1.32; 4df; p=0.86[b]

[a]Model includes a time-varying indicator variable indicating “ever smoker” status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [c][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.4.1.5 The univariate analysis of the association between the cumulative duration of smoking and the risk of incident high-grade CIN

At the earliest of the date of the first diagnosis of high-grade CIN and the date of the end of follow-up, the median cumulative duration of smoking was 40.7 months (range 2.2 to 178.7) among women who had ever smoked.

The univariate analysis of the association between the cumulative duration of smoking and the risk of incident high-grade CIN is described in table 14.4.1.5. There was no statistically significant association when this variable was analysed as either a continuous or as a categorical variable. In the analysis in which cumulative duration of smoking was measured as a categorical variable, the largest hazards ratio was seen for women with a cumulative duration of smoking between 36 and 60 months, but this was not significant.

Table 14.4.1.5. The univariate analysis of the association between the cumulative duration of smoking and the risk of incident high-grade CIN.

Cumulative Duration Of Smoking	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per year)	1075 (100)	28	0.99 (0.83 to 1.18)
			χ^2=0.01; 1df; p=0.91[b]
Never-smoker	535 (49.8)	12	1.00 (Reference)
(0, 36 months)[c]	218 (20.3)	4	0.94 (0.30 to 2.92)
[36, 72)[c]	204 (19.0)	7	1.56 (0.61 to 3.98)
[72, +)[c]	104 (11.0)	5	1.30 (0.42 to 3.99)
			χ^2=1.07; 3df; p=0.79[b]

[a]Model includes a time-varying indicator variable indicating 'ever smoker' status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [c][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.4.1.6 The univariate analysis of the association between the duration of the current smoking episode and the risk of incident high-grade CIN

At the earliest of the date of the first diagnosis of high-grade CIN and the date of the end of follow-up, the median duration of the current smoking episode was 38.8 months (range 1.2 to 157.0), among women who were current smokers at that time.

The univariate analysis of the association between the duration of the current smoking episode and the risk of incident high-grade CIN is described in table 14.4.1.6. The association was not statistically significant. The suggestion of a decreasing risk of incident high-grade CIN with increasing duration of the current smoking episode seen in the analysis in which duration of current smoking episode was measured as a continuous variable, was not evident in the analysis in which this exposure variable was analysed as a categorical variable: the hazards ratios for each category were comparable, with no trend evident. Numbers of events were too small to allow for an analysis of a categorical variable including an “ex-smoker” category.

Table 14.4.1.6. The univariate analysis of the association between the duration of the current smoking episode and the risk of incident high-grade CIN.

Duration Of Current Smoking Episode	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Duration of current smoking episode (per year)[a]	1075 (100)	28	0.91 (0.75 to 1.09)
			χ^2=1.12; 1df; p=0.29[b]
Non-smoker	673 (62.6)	12	1.00 (Reference)
(0, 24 months)[c]	117 (10.9)	4	2.08 (0.67 to 6.48)
[24, 48)[c]	110 (10.2)	5	2.93 (1.02 to 8.41)
[48, +)[c]	175 (16.3)	7	1.93 (0.75 to 4.96)
			χ^2=4.83; 3df; p=0.18[b]

[a]Model includes a time-varying indicator variable indicating "current smoking" status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [c][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.4.1.7 The univariate analysis of the association between pack-years smoked and the risk of incident high-grade CIN

This analysis was restricted to the 1,038 (out of 1,075) women for whom the quantity smoked was known throughout follow-up. At the earliest of the date of the first diagnosis of high-grade CIN and the date of the end of follow-up, median pack-years smoked was 1 pack-year (range 0.1 to 8) in women who had ever smoked prior to that time.

The univariate analysis of the association between pack-years smoked and the risk of incident high-grade CIN is described in table 14.4.1.7. The association was not statistically significant. When analysed as a continuous variable, the hazards ratio

increased by 11% (1.11-1.00) for every increase of one pack year; when analysed as a categorical variable, neither the variable itself nor any individual category was significant.

Table 14.4.1.7. The univariate analysis of the association between pack-years smoked and the risk of incident high-grade CIN.

Pack-years Smoked	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Cumulative quantity smoked (per pack-year)	1038 (100)	27	1.11 (0.79-1.56)
			χ^2=0.35; 1df; p=0.55[a]
Never-smoker	535 (51.5)	12	1.00 (Reference)
(0 , 1)[b]	200 (19.3)	3	0.71 (0.20-2.53)
[1 , 2)[b]	154 (14.8)	5	1.56 (0.54-4.44)
[2 , 3)[b]	77 (7.4)	3	1.39 (0.38-5.07)
[3 , +)[b]	72 (6.9)	4	1.77 (0.53-5.90)
			χ^2=2.05; 4df; p=0.73[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [b][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.4.1.8 The univariate analysis of the association between current smoking status and the risk of incident high-grade CIN

The univariate analysis of the association between current smoking status and the risk of incident high-grade CIN is described in table 14.4.1.8. Current smoking status was significantly associated with the risk of incident high-grade CIN, with current smokers having a two-fold increased risk compared to women who were current non-smokers. There were no events among current ex-smokers, so an analysis including a three-category variable accounting for changes in smoking status could not be undertaken. Note that the third method of categorising this variable has been

presented earlier (see section 14.4.1.3), and is included here for completeness. Numbers of events were small, so this result needs to be interpreted with caution, e.g. being an ex-smoker is unlikely to be protective against high-grade CIN to the extent suggested here.

Table 14.4.1.8. The univariate analysis of the association between current smoking status and the risk of incident high-grade CIN.

Current Smoking Status	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Never-smoker	535 (49.8)	12	1.00 (Reference)
Ex-smoker	138 (12.8)	0	-
Current smoker	402 (37.4)	16	-
Non-smoker	673 (62.6)	12	1.00 (Reference)
Current smoker	402 (37.4)	16	2.21 (1.04 to 4.67)
			χ^2=4.34; 1df; p=0.04[a]
Never-smoker	535 (49.8)	12	1.00 (Reference)
Ever-smoker	540 (50.2)	16	1.27 (0.60 to 2.69)
			χ^2=0.38; 1df; p=0.54[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value

14.4.1.9 The univariate analysis of the association between current smoking intensity and the risk of incident high-grade CIN

This analysis was restricted to the 1,038 (out of 1,075) women for whom the quantity smoked was known throughout follow-up. At the earliest of the date of the first detection of HPV DNA of any type in a cervical sample and the date of the end of follow-up, the vast majority of women who were current smokers were smoking less than 19 cigarettes per day.

The univariate analysis of the association between current smoking intensity and the risk of incident high-grade CIN is described in table 14.4.1.9. When analysed as a categorical variable, the association between current smoking intensity and the risk of incident high-grade CIN was of borderline statistical significance. Note that never-smokers and those currently smoking zero cigarettes per day (current ex-smokers) had to be combined into the same category, due to there being no events in the latter group. When analysed as a continuous variable (or equivalently, when a trend test was performed), the association was highly significant.

Table 14.4.1.9. The univariate analysis of the association between current smoking intensity and the risk of incident high-grade CIN.

Current Smoking Intensity	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Current smoking intensity (per category)[a]	1038 (100)	27	2.25 (1.37 to 3.69)
			χ^2=9.79; 1df; p=0.002[b]
Non-smoker	649 (62.5)	12	1.00 (Reference)
1 to 9 per day	191 (18.4)	5	1.59 (0.59 to 4.25)
10 or more per day	198 (19.1)	10	2.87 (1.24 to 6.66)
			χ^2=5.69; 2df; p=0.06[b]

[a]The increase in risk for every increase of one in category of smoking intensity; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

Summary

In univariate analyses of the association between smoking exposure and the risk of incident high-grade CIN, only current smoking status and current smoking intensity were found to be significant predictors of risk. The risk of high-grade CIN in current smokers was twice that in current non-smokers; and increased significantly with smoking intensity.

14.4.2 Multivariate analysis

14.4.2.1 Controlling for HPV exposure

14.4.2.1.1 The multivariate analysis of the association between smoking status at study entry and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection

The multivariate analysis of the association between smoking status at study entry and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type, is described in table 14.4.2.1.1. The association between smoking status at study entry and the risk of incident high-grade CIN, after controlling for cervical HPV infection, was not statistically significant, using any of the three definitions of this variable. In the multivariate (controlled) analysis, ex-smokers had the same hazards ratio (i.e. the same risk) as current smokers, whereas they appeared to be at greater risk in the univariate analysis; estimates of hazard ratios changed only slightly in the controlled, compared to the uncontrolled, analysis for the other two variables.

Table 14.4.2.1.1. The multivariate analysis of the association between smoking status at study entry and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type.

Smoking Status At Study Entry	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Never-smoker	626 (58.2)	14	1.00 (Reference)
Ex-smoker	93 (8.7)	3	1.33 (0.37-4.76)
Current smoker	356 (33.1)	11	1.30 (0.58-2.92)
			χ^2=0.49; 2df; p=0.78[a]
Non-smoker	719 (65.0)	17	1.00 (Reference)
Current smoker	356 (35.0)	11	1.25 (0.58-2.71)
			χ^2=0.31; 1df; p=0.58[a]
Non-smoker	626 (58.2)	14	1.00 (Reference)
Ever smoker	449 (41.8)	14	1.31 (0.62-2.78)
			χ^2=0.49; 1df; p=0.48[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value

14.4.2.1.2 The multivariate analysis of the association between age at starting smoking and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection

The multivariate analysis of the association between age at starting smoking, as recorded at study entry, and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type, is described in table 14.4.2.1.2. The association was not statistically significant. In each case, the point estimate of the hazards ratio decreased compared to that obtained in the univariate analysis, but changes between these analyses were slight.

Table 14.4.2.1.2. The multivariate analysis of the association between age at starting smoking, as measured at study entry, and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type.

Age at Starting Smoking (at study entry)	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per year)[a]	1075 (100)	28	1.02 (0.97 to 1.07)
			χ^2=0.55; 1df; p=0.46[b]
Never-smoker	626 (58.2)	14	1.00 (Reference)
≤13	100 (9.3)	4	1.65 (0.53 to 5.12)
14-15	182 (16.9)	5	1.11 (0.40 to 3.09)
≥16	167 (15.5)	5	1.34 (0.47 to 3.90)
			χ^2=0.84; 3df; p=0.84[b]

[a]Includes a variable indicating "ever smoker" status at study entry; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.4.2.1.3 The multivariate analysis of the association between a history of ever having smoked and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type

The multivariate analysis of the association between a history of ever having smoked and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type, is described in table 14.4.2.1.3. The association was not statistically significant. Compared to the univariate analysis, the estimate of the hazards ratio decreased, but the change was slight.

Table 14.4.2.1.3. The multivariate analysis of the association between a history of ever having smoked and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type.

Have you ever been a cigarette smoker?	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
No	535 (49.8)	12	1.00 (Reference)
Yes	540 (50.2)	16	1.11 (0.52 to 2.37)
			χ^2=0.08; 1df; p=0.78[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.4.2.1.4 The multivariate analysis of the association between the time since first starting smoking and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type

The multivariate analysis of the association between the time since starting smoking and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type, is described in table 14.4.2.1.4. The association was no statistically significant, whether this variable was analysed as a continuous, or as a categorical, variable. Compared to the univariate analysis, the hazards ratios are closer to the null value of 1, but the changes are slight.

Table 14.4.2.1.4. The multivariate analysis of the association between the time since first starting smoking and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type.

Time Since First Starting Smoking	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per year)[a]	1075 (100)	28	0.97 (0.78 to 1.20)
			χ^2=0.10; 1df; p=0.75[b]
Never-smoker	535 (49.8)	12	1.00 (Reference)
(0, 36 months)[c]	135 (12.6)	3	1.00 (0.27 to 3.65)
[36, 72)[c]	232 (21.6)	6	1.07 (0.40 to 2.88)
[72, +)[c]	173 (16.1)	7	1.23 (0.44 to 3.45)
			χ^2=0.16; 3df; p=0.98[b]
Never-smoker	535 (49.8)	12	1.00 (Reference)
(0, 36 months)[c]	135 (12.6)	3	1.01 (0.28 to 3.69)
[36, 72)[c]	232 (21.6)	6	1.08 (0.40 to 2.90)
[72, 84)[c]	51 (4.7)	3	1.61 (0.44 to 5.93)
[84, +)[c]	122 (11.3)	4	1.01 (0.29 to 3.55)
			χ^2=0.51; 4df; p=0.97[b]

[a]Model includes a time-varying indicator variable indicating "ever smoker" status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [c][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.4.2.1.5 The multivariate analysis of the association between the cumulative duration of smoking and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type

The multivariate analysis of the association between the cumulative duration of smoking and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type, is described in table 14.4.2.1.5. The association was not statistically significant. After controlling for exposure to cervical HPV infection of any type, the hazards ratio for the "0 to 36 month" category (labelled (0, 36 months) in the table) was further from the null value of one than the uncontrolled estimate.

Table 14.4.2.1.5. The multivariate analysis of the association between the cumulative duration of smoking and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type.

Cumulative Duration Of Smoking	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Continuous (per year)	1075 (100)	28	1.04 (0.86 to 1.27)
			χ^2=0.19; 1df; p=0.66[b]
Never-smoker	535 (49.8)	12	1.00 (Reference)
(0, 36 months)[c]	218 (20.3)	4	0.70 (0.22 to 2.22)
[36, 72)[c]	204 (19.0)	7	1.48 (0.58 to 3.76)
[72, +)[c]	104 (11.0)	5	1.30 (0.41 to 4.12)
			χ^2=1.57; 3df; p=0.67[b]

[a]Model includes a time-varying indicator variable indicating "ever smoker" status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [c][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.4.2.1.6 The multivariate analysis of the association between the duration of the current smoking episode and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type

The multivariate analysis of the association between the duration of the current smoking episode and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type, is described in table 14.4.2.1.6. There was no significant association whether this variable was analysed as a continuous, or as a categorical, variable.

Table 14.4.2.1.6. The multivariate analysis of the association between the duration of the current smoking episode and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type.

Duration Of Current Smoking Episode	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Duration of current smoking episode (per year)[a]	1075 (100)	28	0.92 (0.75 to 1.12)
			χ^2=0.74; 1df; p=0.39[b]
Nonsmoker	673 (62.6)	12	1.00 (Reference)
(0, 24 months)[c]	117 (10.9)	4	1.99 (0.63 to 6.22)
[24, 48)[c]	110 (10.2)	5	2.76 (0.96 to 7.90)
[48, +)[c]	175 (16.3)	7	1.94 (0.75 to 5.03)
			χ^2=4.45; 3df; p=0.22[b]

[a]Model includes a time-varying indicator variable indicating 'current smoking' status; [b]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [c][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.4.2.1.7 The multivariate analysis of the association between pack-years smoked and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type

The multivariate analysis of the association between pack-years smoked and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type, is described in tables 14.4.2.1.7. When analysed as a continuous variable, the association between pack-years of smoking and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type, was statistically significant. However, compared to the uncontrolled analysis, the increase in the hazards ratio for every increase of one pack year approximately doubled from 11% to 24% (1.24-1.00). When analysed as a categorical variable, the risk estimates showed no consistent pattern across categories.

Table 14.4.2.1.7. The multivariate analysis of the association between pack-years smoked and the risk of incident high-grade CIN, controlling for exposure to HPV.

Pack-years Smoked	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Cumulative quantity smoked (per pack-year)	1038 (100)	27	1.24 (0.86 to 1.78)
			χ^2=1.22; 1df; p=0.27[a]
Never-smoker	535 (51.5)	12	1.00 (Reference)
(0 , 1)[b]	200 (19.3)	3	0.52 (0.14 to 1.87)
[1 , 2)[b]	154 (14.8)	5	1.71 (0.59 to 4.94)
[2 , 3)[b]	77 (7.4)	3	1.18 (0.32 to 4.37)
[3 , +)[b]	72 (6.9)	4	1.88 (0.56 to 6.37)
			χ^2=3.64; 4df; p=0.46[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [b][a, b) is a mathematical convention denoting the interval ≥a to <b.

14.4.2.1.8 The multivariate analysis of the association between current smoking status and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type

The multivariate analysis of the association between current smoking status and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type, is described in tables 14.4.2.1.8. After controlling for exposure to cervical HPV infection of any type, current smoking status remained significantly associated with the risk of incident high-grade CIN; the estimate of the hazards ratio remained essentially unchanged compared to that in the uncontrolled (univariate) analysis.

Table 14.4.2.1.8. The multivariate analysis of the association between current smoking status and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type.

Current Smoking Status	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Never-smoker	535 (49.8)	12	1.00 (Reference)
Ex-smoker	138 (12.8)	0	-
Current smoker	402 (37.4)	16	-
Non-smoker	673 (62.6)	12	1.00 (Reference)
Current smoker	402 (37.4)	16	2.16 (1.02 to 4.59)
			χ^2=4.08; 1df; p=0.04[a]
Never-smoker	535 (49.8)	12	1.00 (Reference)
Ever-smoker	540 (50.2)	16	1.11 (0.52 to 2.37)
			χ^2=0.08; 1df; p=0.78[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

14.4.2.1.9 The multivariate analysis of the association between current smoking intensity and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type

The multivariate analysis of the association between current smoking intensity and the risk of incident high-grade CIN, after controlling for exposure to cervical HPV infection of any type, is described in table 14.4.2.1.9. After controlling for exposure to cervical HPV infection of any type, despite attenuation in the estimates of the hazards ratios compared to those in the uncontrolled analysis, the associations remained statistically significant.

Table 14.4.2.1.9. The multivariate analysis of the association between current smoking intensity and the risk of incident high-grade CIN, controlling for exposure to cervical HPV infection of any type.

Current Smoking Intensity	Number of Women (%)	Number of Events	Hazards Ratio (95% CI)
Current smoking intensity (per category)	1038 (100)	27	1.94 (1.22 to 3.09)
			χ^2=8.19; 1df; p=0.004[a]
Non-smoker	649 (62.5)	12	1.00 (Reference)
1 to 9 per day	191 (18.4)	5	1.72 (0.64 to 4.63)
10 or more per day	198 (19.1)	10	2.55 (1.09 to 5.95)
			χ^2=4.66; 2df; p=0.10[a]

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value.

Summary

In univariate analyses, the risk of incident high-grade CIN in current smokers was twice that of current non-smokers; and increased significantly with current smoking intensity; in multivariate analyses controlling for past or current exposure to cervical infection with HPV16 or HPV18 and other HPV types, these hazards ratios decreased slightly, but remained statistically significant. No association was found between the risk of incident high-grade CIN and the other smoking variables.

14.4.2.2 Controlling for sexual behaviour

There were no associations between the available sexual behaviour variables and the incidence of high-grade CIN. Statistical significance is not necessarily the sole criterion for including or excluding a potential confounder from an analysis. However, it is the confounding of the relationship between smoking and high-grade CIN by cervical HPV infection which is relevant in this analysis. Although recent sexual behaviour is the strongest determinant of HPV-positivity at any point in time (e.g.

number of recent sexual partners is associated with HPV status, but not with high-grade CIN (Deacon 2000)), in this analysis I can measure HPV status directly. Therefore, controlling for cervical HPV infection status will accomplish controlling for confounding by sexual behaviour, directly.

14.4.2.3 Effect of the inclusion of women with prevalent cervical HPV infections and/or prevalent cervical cytological abnormalities in the analysis of the incidence of high-grade CIN

In the analysis presented above, all 1,075 women in the study population were recruited close to the onset of sexual activity, and were free of cytological abnormality and cervical HPV infection at study entry. Therefore, although some women are likely to have acquired and cleared a cervical HPV infection prior to study entry, the assumption has been made that in these women the first cervical HPV infection detected during the follow-up period of this study is their first such infection *overall*. The first detection of high-grade CIN during follow-up was certainly the first occurrence of this disease. The study population described in these chapters of the thesis (the 1,075 women) therefore provides the best opportunity of being able to describe the association between incident high-grade CIN and smoking. However, the number of women diagnosed with high-grade CIN in this study population is small (28 women).

Another analysis was conducted in which study entry criteria were relaxed: women with cervical cytologically abnormalities, or who were HPV-positive in their cervical,

or both, at study entry, were also included in the study population. The final study population for this analysis comprised 1,485 women, 56 of whom progressed to high-grade CIN. In this analysis, estimates of hazards ratios associated with smoking were not substantially changed, and the significant predictors identified were identical (Table 14.4.2.3).

Table 14.4.2.3. The association between smoking and the incidence of high-grade CIN in an analysis including women all women with further follow-up after study entry.

	All women (n=1485, 56 events)	
	Crude HR (95% CI)	**Adjusted HR (95% CI)[a]**
Current smoking status		
Non-smoker	1.00 (Reference)	1.00 (Reference)
Smoker	2.19 (1.29-3.74)	2.06 (1.20-3.52)
	p=0.003	p=0.007
Current smoking intensity		
Non-smoker	1.00 (Reference)	1.00 (Reference)
1-9 per day	2.04 (1.06-3.92)	1.90 (0.99-3.66)
10+	2.35 (1.27-4.37)	2.21 (1.19-4.12)
	p-trend=0.01	p-trend=0.008

[a]adjusted for exposure to cervical infection with HPV16, HPV18 and other HPV types

Summary

Conclusions remained unaltered, and estimates of association were essentially identical, in an analysis in which the study population was expanded to include women with HPV infection or cytological evidence of cervical disease at study entry.

Chapter 15

SMOKING, CERVICAL HPV INFECTION, AND THE RISK OF CERVICAL NEOPLASIA

DISCUSSION

Repeated measurements of smoking and cervical human papillomavirus (HPV) status were used to explore the impact of smoking on the acquisition and duration of incident cervical HPV infections, and to measure the risk of acquiring incident high-grade cervical intraepithelial neoplasia (CIN), in relation to changes in smoking and cervical HPV status. Several ways of measuring exposure to cigarette smoking were investigated in the analyses described in this thesis. One of the strengths of the study population and study design described in this thesis was that, by including women with *incident* cervical HPV infections, it was possible to measure the impact of changes in smoking and cervical HPV infection status during follow-up on the risk of high-grade CIN, and also to avoid any residual confounding by time from first HPV exposure.

15.1 THE INCIDENCE OF CERVICAL HPV INFECTION

HPV is a necessary, but not sufficient, cause of cervical cancer. Smoking may act as a cofactor, for example by increasing the risk of acquiring a cervical HPV infection, or by prolonging the duration of an infection which has already been acquired.

In univariate analyses of the association between smoking and the risk of acquiring an incident cervical HPV infection of any type, most of the smoking exposure variables were significant predictors. The magnitude of the risk estimates (hazards ratios) were typically small. However, it is possible that many of the significant associations merely reflect the fact that “ever-smokers” are at greater risk of infection than “never smokers”. In univariate analyses, the time since starting smoking, the cumulative duration of smoking, the duration of the current smoking episode, and pack-years of smoking, were all significant predictors of the risk of acquiring incident cervical HPV16 infection. All, however, appear to be protective: the risk of incident infection with cervical HPV16 decreased as the magnitude of each of these variables increased. There were no significant associations between smoking exposure and the risk of an incident cervical HPV18 infection, and measures of association tended to be closer to the null value of 1 than in the analysis of HPV16.

In multivariate analyses of the association between smoking exposure and the risk of an incident cervical HPV infection of any type, after controlling for sexual behaviour, amongst all available exposure variables, only cumulative duration of smoking and pack-years of smoking were found to be significant predictors. Again, both appear to be protective against cervical HPV infection. Amongst the sexual behaviour variables investigated, only lifetime number of sexual partners was found to be associated with the incidence of cervical HPV16 infection in a univariate analysis; after controlling for this variable, the duration of the current smoking episode and pack-years of smoking were both found to be significant *protective* factors. There were no significant associations between risk-factors measuring sexual behaviour and the incidence of

cervical HPV18 infection in univariate analyses, therefore multivariate analyses were not undertaken.

In the study described in this thesis, the increased risk of acquiring a cervical HPV infection of any type associated with smoking was for the most part explained by the association between smoking and sexual behaviour. These findings are consistent with those of most other longitudinal studies of incident cervical HPV infection (Ho 1998b, Moscicki 2001, Sellors 2003, Winer 2003, Minkoff 2004, Munoz 2004); although one such study reported an increased risk of incident HPV infection associated with smoking, it did not distinguish between cervical and vulvovaginal infections (Ho 1998b); another found an increased risk only in HIV-infected women (Minkoff 2004).

However, the protective effect of variables incorporating some measure of the duration of smoking, even after controlling for sexual behaviour, is more difficult to interpret. Not all studies have reported variables measuring duration of smoking. One study found an increased, but non-significant, risk associated with increasing pack-years of smoking (Moscicki 2001); another merely states that this variable was not as “highly associated” as other smoking variables, and so was not considered further (Minkoff 2004). I am reluctant to over-interpret counterintuitive results when estimates of effect for successive categories are subject to substantial fluctuation, and there is no clear or convincing trend. Another unusual finding is that ex-smokers almost always have a higher risk (based on the point estimate, not necessarily statistical significance) of incident cervical HPV infections of any type or of type 16

than current smokers. Other authors who have observed this phenomenon have also been unable to explain this.

The fact remains that these results do not support the hypothesis that smoking *increases* the risk of acquiring a cervical HPV infection.

15.2 THE DURATION OF CERVICAL HPV INFECTION

Because the cohort used in many of the analyses described in this thesis included women who tested negative for HPV DNA in cervical samples at study entry, we were able to estimate the time of onset of an episode of cervical HPV infection. This is an important consideration, because the duration of an infection cannot be measured in women who already test positive for HPV in cervical samples at study entry.

There were no statistically significant associations between any of the smoking exposure variables and the risk of clearing an incident cervical HPV infection, either of any type, or of type 16; there were too few clearance events to make the analysis of the clearance of cervical HPV18 infections worthwhile. Multivariate analyses yielded no further insight. In analyses of the duration of incident cervical HPV16 infection, although not statistically significant, hazards ratios tended to suggest that smoking increases the risk of clearing an infection, i.e. that smoking is associated with a shorter duration of cervical HPV16 infections.

Two other studies have reported that smoking has a significantly protective effect against *persistent* infection (Hildesheim 1994, Ho 1998b), but this finding is by no means consistent across studies. One study which reported that smoking prolongs the duration of HPV infection included only prevalent cases; of two studies which included both prevalent and incident cases, one found no significant association, and the other, a significant association only in HIV-negative women; and only then in those infected with high-risk HPV types (Giuliano 2002b, Richardson 2005, Koshiol 2006). If smoking *per se* does not increase the risk of the acquisition of cervical HPV infection, and if it does not prolong the duration of infection, then other mechanisms may have to be invoked to explain the excess risk of cervical neoplasia associated with smoking.

15.3 THE INCIDENCE OF HIGH-GRADE CIN

In univariate analyses of the association between smoking exposure and the risk of incident high-grade CIN, current smoking status and current smoking intensity were found to be significant predictors of risk. The risk of incident high-grade CIN in current smokers was twice that in current non-smokers; and increased significantly with smoking intensity. In multivariate analyses controlling for past or current exposure to HPV16 and HPV18 and other HPV types, these hazards ratios decreased slightly, but remained significant. No association was found with the other smoking variables. An increased risk of cervical neoplasia has occasionally been associated with smoking duration, but only following periods of exposure which are substantially in excess of those observed in our cohort**.**

The magnitude of the association between current smoking status and the risk of incident high-grade CIN was consistent with that observed in case-control and longitudinal studies restricted to women who tested positive for HPV (Deacon 2000, Castle 2002b, Moscicki 2001). The pooled IARC analyses, which included thousands of CIN3 and cancer cases, found a similar two-fold elevation of risk (International Collaboration of Epidemiological Studies of Cervical Cancer 2006a) (see section 12.2.1). However, regardless of sample size, these studies cannot reveal the *temporal* relationship between exposure and outcome, nor can they resolve the issue of residual confounding by the time from first cervical HPV exposure, even when restricted to HPV-positive women. In contrast, the study design and choice of study population described in this thesis enabled these issues to be addressed.

15.4 CHOICE OF STUDY POPULATION

The cohort study described in this thesis deliberately recruited only young women because it was the intention to recruit women who were free of both cervical disease and cervical HPV infection at study entry; only with such a study population is it possible to study the temporal relationship between exposure and incident disease. Although the women in this study were recruited immediately before, or soon after, the onset of sexual activity, the possibility cannot be ruled out that some women may have already cleared a short-lived cervical HPV infection prior to study entry. However, it *can* be confidently asserted that any high-grade CIN detected during follow-up is *incident* disease. Conclusions remained unaltered, and estimates of

association were essentially identical, when the study population was expanded to include women with cervical HPV infection, or cytological evidence of cervical disease, or both, at study entry.

15.5 CHOICE OF DISEASE ENDPOINT

CIN2 may be considered a “soft” end point, in the sense that it represents a potentially equivocal diagnosis of pre-cancer, and therefore women with CIN2 may not be at-risk of progression. CIN represents a continuum of abnormality; therefore, at one end of the spectrum (the “high” end), some cases diagnosed as CIN2 may in fact be CIN1, or at the other end (the “low” end), CIN3. CIN3 is an unambiguous diagnosis since it corresponds to abnormality extending throughout the full thickness of the epithelium, and is considered true pre-cancer (see section 2.2.1.1). However, it is current clinical practice that women with CIN2 are treated, as a result of the low, but not negligible, risk of disease progression in women with CIN2 (see sections 2.2.3 and 4.2.3). Furthermore, the incidence of high-grade CIN, as defined in this thesis, has been used as an outcome measure for clinical trials of HPV vaccines (see section 4.3).

15.6 CONTROL OF CONFOUNDING BY EXPOSURE TO CERVICAL HPV INFECTION

Cervical HPV infection is a strong predictor of the risk of progression of early cervical neoplasia to high-grade CIN. Many investigators have analysed HPV-positive and

HPV-negative women separately, or restricted the analysis to “HPV-positive” women. However, HPV status is not a static variable, nor is it clear why the risk of disease should be restricted to the period when a woman is HPV-positive, according to possibly insensitive detection methods.

The study design used by the cohort study described in this thesis enabled a more appropriate and efficient approach to the analysis of the association between smoking and high-grade CIN, controlling for exposure to cervical infection with HPV DNA of any type. HPV status was not determined at study entry only, or at diagnosis only, as other investigators have chosen, or been obliged, to do. Instead, cervical HPV status was determined prospectively at each study visit, and the analysis then incorporated these changes over time. Thus, for example, a woman who had intermittent periods of HPV-positivity during follow-up, would effectively contribute “time-at-risk” to both the exposed and unexposed groups, as appropriate. The time-varying binary indicators measuring HPV exposure status, the construction of which is described in section 13.9, ensured that the comparison group using these variables was women during the period of follow-up when they were HPV DNA-negative. Similarly for smoking status, which was also updated during follow-up. Restricting the analysis to when a woman was HPV-positive would make inefficient use of the data obtained and inevitably, and unnecessarily, introduce exposure misclassification. For example, how should a woman who went from being HPV-negative to HPV-positive, and back again, be treated?

Of course, the way in which I controlled for confounding by HPV exposure is not perfect, in at least one respect. Ideally, I would also have adjusted for exposure to other individual high-risk HPV types in the same way as I did for HPV16 and HPV18. However, numbers were insufficient to allow this to be done, even in the expanded cohort of 1,485 women. Adjustment for other aspects of cervical HPV exposure, or allowing for more complex changes in HPV exposure status, may have been possible, for example by allowing women to move between HPV-positive and HPV-negative states, by estimating cumulative exposure, or by adjusting for time since first cervical HPV infection. However, the latter variable, at least, would be redundant in this context since analyses based on the incident cohort do not suffer from confounding by time since first exposure because this occurs during follow-up, and changes in status are accounted for as they are detected.

15.7 CONTROL OF CONFOUNDING BY VARIABLES OTHER THAN CERVICAL HPV INFECTION

Confounding has to be dealt with on a study-by-study basis: the lack of confounding in one study does not rule out confounding in others. In a study intended to minimize residual confounding, it may be inappropriate to eliminate non-statistically significant variables from the final model, since confounding does not depend on considerations of statistical significance alone. My approach was to retain in the final model only those variables which were significant in multivariate analyses including smoking variables; this approach was applied consistently for all analyses. The analyses in which the inclusion of a potential confounder changed an estimated hazards ratio by

more than 10%, regardless of the statistical significance of the variable itself, were also those in which that confounder was a significant predictor, and for which results have been presented. The exception was age at first sexual intercourse in the analysis of high-grade CIN. Age at first intercourse was not a significant independent predictor of high-grade CIN. However, inclusion of this variable in a model also controlling for exposure to cervical HPV infection, yielded greater point estimates of association and corresponding increases in significance for all of the smoking variables, but without altering conclusions. Because the study hypothesis relates to the association between smoking and high-grade CIN, controlling for exposure to cervical HPV infection, I decided that presenting results which also controlled for age at first sexual intercourse detracted from the clarity of the presentation.

15.8 MECHANISTIC EXPLANATION FOR THE ROLE OF SMOKING IN THE AETIOLOGY OF CERVICAL DISEASE

I stated earlier that smoking may act as a cofactor, either by increasing the risk of acquiring a cervical HPV infection, or by prolonging the duration of an infection. This is not to rule out the possibility that other mechanisms exist by which cervical HPV infection might influence disease progression. Alternative mechanisms have not been well defined, nor will they be unless and until existing explanations are shown to be inadequate. There is no incentive for pursuing mechanistic explanations unless and until it can first be shown that the association between smoking and cervical neoplasia is not simply a consequence of residual confounding from time from first cervical HPV exposure. However, having now shown that smoking does not increase

the risk of acquiring cervical HPV infection, once sexual behaviour is controlled for; nor that it prolongs the duration of cervical HPV infection, we are obliged to speculate regarding other, as yet inadequately characterised, mechanisms. The longitudinal studies necessary to show that such smoking-related changes increase the risk of disease initiation, or disease progression, have yet to be undertaken.

The evidence for a possible impact of smoking on the humoral and local immune responses to cervical HPV infection is conflicting. One study reported that HPV16-infected women were less likely to have type-specific antibodies if they were smokers; another, that a sequential rise in HPV16 antibody levels was more common in smokers than in non-smokers (Silverberg 2006, Wiley 2006). Whereas some investigators have concluded that smoking reduces the number of intraepithelial Langerhans' cells in the uterine cervix, others have found no association (Campaner 2006, Nadais 2006).

15.9 CONCLUSION

Although cigarette smoking by adolescents continues to be an important public health problem, the modest risk of cervical neoplasia associated with smoking is unlikely to persuade many to quit. Early age at smoking initiation is reported to increase the risk of lung and breast cancer, independent of smoking duration or intensity (Wiencke 1999, Gram 2005). Although the interdependent effects of age at smoking initiation, duration of smoking, time since smoking cessation, and calendar age, are not easily disentangled, such an effect is biologically plausible because

these tissues continue to grow and develop into early adulthood (Leffondre 2002, Wiencke 2002). Empirical evidence of the long-term effects on these tissues of early smoking initiation will be difficult to obtain. However, the cervix can be sampled more easily than many other sites (in terms of both accessibility and the acceptability of the sampling technique), and were smoking-induced DNA damage shown to precede the development of a pre-malignant phenotype at this site, then it would provide more compelling evidence that the harmful effects of smoking can occur earlier than is commonly believed.

15.10 PUBLICATION

A paper based on the work described in chapters 12, 13, 14 and 15 of this thesis was published in January 2010 (Collins 2010: see appendix 4).

Chapter 16

THE KINETICS OF THE HUMORAL IMMUNE RESPONSE TO INCIDENT CERVICAL HPV16 AND HPV18 INFECTIONS

INTRODUCTION

Orientation

In this chapter I:

- describe relevant aspects of the immune system
- describe the immune response to HPV infection
- provide an overview of the evidence in the literature relating to the kinetics of the immune response to cervical HPV infection
- describe relevant aspects of the measurement of the humoral immune response to incident cervical HPV16 and HPV18 infections

16.1 BACKGROUND

The kinetics of the humoral immune response following an incident cervical HPV infection in naïve women, i.e. women who have never been exposed to HPV infection, are not well characterised. There are at least two major difficulties in arriving at such a characterisation.

First, because cervical HPV infection is an asymptomatic condition, its time of onset can never be precisely determined. However, its time of onset can be approximated

by making frequent observations at short time intervals in women who are unexposed to cervical HPV infection at recruitment. These women should be recruited soon after they have first had sexual intercourse, because the longer the time interval between first sexual intercourse and study entry, the more likely it is that a woman will have acquired, and then cleared, at least one cervical HPV infection during this time (Woodman 2001, Collins 2002).

The second difficulty relates to the assay used to measure the immune response to HPV infection. The most commonly used assay is the virus-like particle-based enzyme-linked immunosorbent assay, the VLP-ELISA. The integrity of the VLP-ELISA depends on the maintenance of intact virus-like particles (VLP); disrupted or incorrectly-folded VLPs may lead to the detection of non-neutralizing and cross-genotype-reactive antibodies, thus complicating the interpretation of results. Competitive radioimmunoassays (cRIA) which specifically measure HPV type-specific neutralizing antibodies have also been developed. The cRIA is more sensitive than the ELISA, and is less likely to be influenced by impurities in the sample. However, like any serological assay based on competition, it will fail to detect serum antibodies that which to bind to, or compete with, the epitopes that bind the competing antibody.

A sensitive and specific assay capable of measuring the functionally-relevant HPV type-specific neutralizing antibodies is required.

16.2 THE IMMUNE SYSTEM

The immune response can be dichotomised into *humoral immunity* and *cell-mediated immunity*: humoral immunity refers to those aspects of the immune response which are mediated by antibodies; cell-mediated immunity refers to those aspects which are mediated by cells, such as T-lymphocytes, macrophages and natural killer cells. Cell-mediated immunity is directed primarily against intracellular pathogens and is very effective at removing virus-infected cells, and intracellular bacteria, but also provides defence against fungi, protozoans, and larger parasites (Abbas 2003, Lydyard 2004, Stanley 2006).

The immune system has both non-specific and specific components, *innate immunity* and *adaptive immunity*, respectively. These components function cooperatively. Innate immunity provides an early and rapid response to “insult”, or more specifically to infection, but it is not very specific, i.e. it may not distinguish subtle differences between different pathogens, and it responds in essentially the same way to repeated infections of the same type. In contrast, although the adaptive, or *acquired*, immune response takes longer to respond to antigen than the innate immune response, it is more powerful, versatile, specific, and “remembers” prior exposures. The adaptive immune response increases in magnitude and capability with successive exposures to the same antigen. The adaptive immune response, as with the immune response generally, has two branches, cell-mediated immunity and humoral immunity (Abbas 2003, Lydyard 2004).

16.2.1 Cell-mediated immunity

The principal cells involved in cell-mediated immunity are T-lymphocytes, or T-cells. The principal mechanisms of innate immunity against viruses are inhibition of infection by type 1 interferons, and natural-killer-cell-mediated killing of infected cells (Abbas 2003, Andersen 2006, Lydyard 2004). In the cohort study described in this thesis, cell-mediated immunity to cervical HPV infection was not measured and so will not be considered further.

16.2.2 Humoral immunity

The principal components of the adaptive immune response are antibodies produced by B-lymphocytes, and by plasma cells, which also stimulate the production of chemical mediators of the immune response, such as complement and interferon. The principal cells involved in humoral immunity are B-lymphocytes, or B-cells, which produce large quantities of antibodies. Antibodies are highly-specific antigen-binding molecules, which either remain attached to the B-cell as a receptor, or are released into the circulation to mediate elimination of antigen.

There are five classes of antibody, each of which also has two or more subclasses: IgM (immunoglobulin M), IgD, IgA, which provides the first line of defence at mucosal surfaces, IgE, and IgG. IgG is found throughout the body, and is the main defence against blood-borne infectious agents.

Antibodies can act directly, i.e. on their own independently of a cellular intermediary, to neutralize certain antigens, such as viruses; they can coat (“opsonise”) a target antigen, and thereby enhance the activity of phagocytes; or lead to the direct killing of cells *via* antibody-dependent cellular cytotoxicity. In general, adaptive immunity against viral infections is mediated by, antibodies which block virus binding and entry into host-cells, and by cytotoxic T-lymphocytes, which eliminate the infection by killing infected cells (Abbas 2003, Lydyard 2004).

16.3 DESCRIPTION OF THE IMMUNE RESPONSE TO CERVICAL HPV INFECTION

Infection with an oncogenic HPV type (e.g. types 16 and 18) is a *necessary* but not *sufficient* cause of cervical cancer. Cervical HPV infection is extremely common, but most infections, particularly in young women, appear to be transient and of short duration (Woodman 2001). The immune system is clearly important in the control of cervical HPV infections, a fact which can be deduced directly from the increased incidence and duration of cervical HPV infections and squamous intraepithelial lesions of the cervix in immunosuppressed women (Vernon 1994, Zur Hausen 2002, Baseman 2005). Immune responses against HPV during the early stages of infection may be important in preventing reinfection, inhibiting viral spread, and preventing viral persistence (De Gruijl 1997). However, the kinetics of the immune response to cervical HPV infections have yet to be characterised.

16.3.1 Humoral immunity to cervical HPV infection

The precise role of the humoral immune response in the mediation of cervical HPV infection, or of cervical neoplasia, has yet to be established, but has now been investigated in several studies.

16.3.1.1 Cross-sectional studies examining humoral immunity to cervical HPV infection

Evidence from cross-sectional studies regarding an association between cervical HPV infection and the immune response to that infection is ambiguous, since the cross-sectional study design does not enable the determination, with certainty, of the order in which exposure and outcome arrived. Cross-sectional studies will therefore not be considered further in this chapter.

16.3.1.2 Longitudinal studies examining humoral immunity to cervical HPV infection

Only longitudinal study designs, such as that used in the cohort study described in this thesis, are capable of defining the kinetics of the immune response to incident cervical HPV infection. In this section, I describe reports from longitudinal studies which have investigated some aspect of the immune response to cervical HPV infection; results relating to disease endpoints are not considered, since these are not the subject of analyses in this thesis and so cannot be put into context. I focus on

a subset of mature longitudinal studies: the names given to these studies are for convenience of reference only, and are not necessarily those with which the principal investigators endowed them.

16.3.1.2.1 The Amsterdam cohort study

This was a prospective cohort study describing the development, persistence, and progression, of CIN lesions in relation to the natural history of HPV infection, including the humoral immune response to that infection. The cohort consisted of 353 women aged 18 to 55 years, who were routinely referred to a single colposcopy clinic in Amsterdam, The Netherlands. All women had mild to moderate or severe dyskaryosis at study entry and were followed at intervals of three to four months until they *clinically* progressed to CIN3. PCR was used to test for 14 high-risk HPV types.

Results from this cohort study have now appeared in several publications (Remmink 1995, De Gruijl 1996a, Rozendaal 1996, De Gruiijl 1996b, De Gruijl 1997, Bontkes 1998a, Bontkes 1998b, Nobbenhuis 1999, Bontkes 1999a, Bontkes 1999b, De Gruijl 1999, Bontkes 2000, Rozendaal 2000, VanDuin 2002). One of these investigated the humoral immune response against HPV16 VLPs, and its relationship with the clearance of HPV16 infection and cervical lesions (Bontkes 1999b). Cervical mucus samples from 125 women, and plasma samples from 100 women, respectively, were tested for IgA and IgG antibodies, using HPV16 L1/L2 VLP-based ELISA. Women were more likely to have a humoral IgA response when HPV16 DNA was not detected in the smear, than when it was, although this difference was not statistically

significant; in contrast, women were more likely to have a humoral IgG response when HPV16 DNA was detected.

Although the sample size was relatively large, this was essentially a cross-sectional analysis of a longitudinal study and therefore has the weaknesses of a cross-sectional study rather than the strengths of a longitudinal study. There are also many design flaws with this study: for example, it was not certain that women were free of the primary outcome at study entry, a pre-requisite for a cohort study; women were reported as being disease-free at study entry, yet a fifth were cytologically abnormal at this time; and the cytological classification of some women was inaccurate (Woodman 2002) (see section 20.3.2.2 for further discussion of some of the flaws with this study). In the study population used in the analyses described in this section of the thesis, all women were unambiguously free of cervical cytological abnormalities at study entry.

16.3.1.2.2 The San Francisco cohort study

This was a prospective cohort study investigating the early natural history of cervical HPV infection in women aged between 13 and 22 years recruited from one family planning clinic and one student health center in the San Francisco Bay area, USA. Several methods were used for the detection of HPV DNA in this study, with, for example, PCR used on only selected subsets of women.

Results from this cohort study have now appeared in several publications (Moscicki 1998, Moscicki 1999, Scott 1999, Nakagawa 2000, Moscicki 2001, Nakagawa 2002, Moscicki 2004). One paper describes temporal changes in the humoral immune response to HPV16, using serial virus-like particle (VLP) antibody assays (Nakagawa 2002). However, this description is based on only eight women: it is at best suggestive. Four women had two HPV16 DNA-positive episodes; for two of these women, the second episode “occurred in the presence of high titers of HPV16 VLP antibody”. The authors say “this brings into question the protective role of humoral immunity in preventing repeated infection”. They also found that all four of these women rapidly became HPV16 DNA-negative following the start of the second HPV16 episode, “suggesting the presence of immunological memory”.

16.3.1.2.3 The New Brunswick cohort study

This was a prospective cohort study investigating the natural history of cervical HPV infection, which recruited 608 students (mean age 20 years) from Rutger’s University in New Brunswick, USA. Women were followed-up for a maximum of three years, at intervals of six months. At each visit, a cervicovaginal lavage was done for use in HPV testing; PCR and southern blot were used to test for HPV DNA. Blood samples for the preparation of serum were collected at the annual visits initially, but also at the six-month visits in the last year of the study.

Results from this cohort study have now appeared in several publications (Burk 1996b, Ho 1998b, Morrison 1998, Ho 2002, Kahn 2002a, Kahn 2002b, Ho 2004).

One report from this study investigated whether host immunity protected against subsequent infection, in a subset of 247 women who had further follow-up after their initial incident, or prevalent, HPV infection (Ho 2002). In this report, the authors found that the detection of persistent antibodies to HPV16 VLPs was associated with a reduced risk of subsequent HPV infection, and that type-specific persistent infection had a protective effect on subsequent infection with a HPV16-related type. A later report examined seroconversion to HPV16 VLPs (n=403, e (number of events)=25) and clearance (duration) of seroconversion to HPV16 VLPs (n=87, e=37) (Ho 2004). Women infected with HPV16 were more likely to seroconvert than those who were HPV-negative, or than those who were infected with other HPV types. There was no difference in the risk of seroconversion for women with "persistent" infection with HPV16, compared to those without. Among women who became seropositive, those who had no subsequent HPV infection had the highest chance of becoming seronegative; persistent HPV16 infection did not significantly alter the risk of becoming seronegative.

This cohort study fulfils at least one of the ideal criteria required of a study designed to investigate the humoral immune response to HPV, i.e. frequent observations at short intervals in women who are recruited soon after first sexual intercourse. However, this study did not separately distinguish cervical from other lower genital tract infections (by design). Therefore, notwithstanding its relatively large sample size, this study can make only a limited contribution to the study of the humoral response to *cervical* HPV infections. In contrast, the study described in this thesis examined the humoral immune response to cervical HPV infections specifically.

16.3.1.2.4 The Guanacaste cohort study

This was a population-based cohort study of the natural history of cervical neoplasia, which recruited over 10,000 women of all ages from Guanacaste, a rural province of Costa Rica with consistently high rates of invasive cervical cancer. Women were followed-up at intervals of between six months and one year, for a period of several years (this study may be ongoing at the present time (August 2008)). The presence of HPV DNA was detected using PCR.

Results from this cohort study have now appeared in several publications (Herrero 1997, Ung 1999, Hildesheim 2001, Wang 2001, Gravitt 2003, Wang 2003, Castle 2004, Viscidi 2004, Wang 2004a, Wang 2004b, Castle 2005a, Castle 2005b, Castle 2006). In reports from this study investigating the immune response to HPV infection, only the sera samples taken at study entry and at the end of follow-up, five to seven years later, have been analysed. The investigators have reported that 55% of women who were HPV16-seropositive at baseline “remained” seropositive five to seven years later, and that there was an association between viral load at baseline and the risk of seroconversion to HPV16 at the end of follow-up (Wang 2004a). Seropositivity to either HPV16, HPV18, or HPV31, did not protect against subsequent infection with the specific type, or either of the other types; in fact, there was a significantly increased risk of infection with HPV18-related types among women seropositive for HPV16 or HPV31 at study entry (Viscidi 2004).

It is difficult to interpret these results: with such a long time interval between measurements of serostatus, these analyses cannot eliminate the possibility that women reverted to being seronegative in between the two measurements. The study described in this thesis made frequent measurements of serostatus at short time intervals, and was thus in a better position to ensure that any change in serostatus would be observed.

16.3.1.2.5 The Copenhagen cohort study

This was a population-based cohort study of women randomly selected from the general population of Copenhagen, Denmark. Women made two visits, the second "about" two years after the recruitment visit. Cervical samples for HPV DNA testing and blood samples for serological testing, were taken at both visits.

Results from this cohort study have now appeared in several publications (Kjaer 1996, Kjaer 1997, Kjaer 1998, Kjaer 2001, Kjaer 2002, Pastrana 2004). One report describes data relating to 100 women who were virgins at study entry, and 105 monogamous women (mean age at recruitment approximately 24 years) (Kjaer 2001). Among initially virgin women who first experienced sexual intercourse during the study period, nine had become HPV16 DNA-positive by the second of their two visits, of whom seven (78%) were seropositive to HPV16 VLPs; in contrast, only two of 13 women who had acquired an infection with HPV DNA of another type during follow-up seroconverted, as did one woman who was not HPV16 DNA-positive, but was positive to HPV DNA of another type, at study entry. Among women who were

monogamous at study entry, five had become HPV16 DNA positive at the second visit, of whom four (80%) were seropositive to HPV16 VLPs; seroconversion was also observed in three women who became positive for other types during follow-up, and at study entry, or during follow-up, in nine women in whom no HPV DNA was ever detected. This is difficult to interpret since there were only two visits, with a substantial interval between them: it is not possible to be confident about temporality with such a large time interval. In contrast, the ability to demonstrate temporality was one of the clear strengths of the design of the study described in this thesis.

16.3.1.2.6 The Swedish Schoolgirl cohort study

The was a cohort study (Andersson-Ellstrom 1996a, Andersson-Ellstrom 1996b) which recruited 98 women aged 15 to 17 years who were attending a secondary school in Karlstad, Sweden. Women were seen at intervals of six months. PCR was used to detect HPV DNA.

One report on this study used HPV L1/L2 VLP ELISA to detect IgG antibodies to HPV16 and HPV33 VLPs (interestingly, “disrupted” BPV capsids were also included as negative-controls: see later) (Andersson-Ellstrom 1996b). Seven women who were initially HPV16 DNA-negative and seronegative to HPV16 VLPs, seroconverted during follow-up; five of these women were HPV16 DNA-positive at their final visit, including one woman known to be HPV16 DNA-positive at her second visit. Six women who were initially HPV16 DNA-negative and seronegative to HPV16 VLPs, and who subsequently acquired HPV16 DNA, also seroconverted during follow-up

(these two groups are not mutually exclusive). Unfortunately, despite the longitudinal nature of this study, and the presentation of results for individuals, sufficient details are not provided to establish the temporal relationship between the acquisition of HPV DNA and seroconversion. As discussed above, the study described in this thesis was in an advantageous position to establish temporality.

16.3.1.2.7 The University of Washington cohort study

This was a cohort study which recruited 603 female students who attended the University of Washington (UOW), Washington, USA, between 1990 and 1997. There are now several reports on this cohort, including five which focus on some aspect of the immune response to incident HPV infections (Xi 1995, Carter 1996, Xi 1997, Carter 2000, Hagensee 2000, Thomas 2000, Xi 2002, Onda 2003, Winer 2003, Buchinsky 2006). The UOW cohort study is of particular interest because, as well as fulfilling the ideal criteria required of a study designed to investigate the humoral immune response to HPV, it is possibly the most comparable cohort study to the study reported in this thesis, the Birmingham Brook Clinic cohort.

16.4 COMPARISON OF THE "UNIVERSITY OF WASHINGTON" AND "BIRMINGHAM BROOK CLINIC SEROLOGY" COHORTS

The final UOW cohort comprised 603 women, not all of whom contributed to the longitudinal analyses due to the inevitable exclusion of women from analyses

because of an imperfect correspondence between the available DNA and serum samples. For example, only 528 women entered one analysis (Xi 2002).

The Birmingham Brook Clinic cohort study enrolled 2,011 women. The following women were then excluded: those who had an abnormal smear, or tested positive for HPV DNA in a cervical sample, at enrolment; and those who did not provide two or more sera, and cervical, samples which could be tested both cytologically and virologically. This left 554 women available for a longitudinal study of the immune response to incident cervical HPV infections. For ease of reference, these 554 women will be referred to as the Birmingham Brook Clinic Serology cohort (BBS).

16.4.1 Sexual experience

The median age at study entry was 19 years in the UOW cohort and 18 years in the BBS cohort. The median number of years of sexual activity prior to study entry was two for members of the BBS cohort; this statistic is not reported for the UOW cohort. One hundred and seventy one (28.2%) women in the UOW cohort were virgins at enrollment; all women in the BBS cohort were already sexually active at study entry. The median number of sexual partners at study entry was two in those women in the UOW cohort who were already sexually active at study entry, and two in the BBS cohort.

16.4.2 Frequency of sampling and duration of follow-up

The median time interval between visits in the UOW cohort was 4.7 months, compared to 7 months in the BBS cohort; the median duration of follow-up in studies reporting on the immune response in the UOW cohort ranged from 26 to 39 months; it was 29 months for the BBS cohort.

16.4.3 The prevalence of HPV infection and cervical cytological abnormality at study entry

The prevalence of HPV infection in the UOW cohort was 19.7%; that of cervical cytological abnormality is not reported, nor is the presence of cytological abnormality listed among the exclusion criteria. All women in the BBS cohort tested negative for HPV DNA in cervical samples (prevalence 0%) and had a normal cervical smear at study entry.

16.4.4 The number of incident HPV infections

In the UOW cohort, the number of women with incident HPV16 infections included in the various analyses of the immune response ranged from 25 to 67; the only analysis of incident HPV18 infections included 35 women. The BBS cohort included 72 women with incident cervical HPV16 infections and 40 with incident cervical HPV18 infections.

16.4.5 Detection of types other than HPV16 prior to the first detection of HPV16

Approximately 43% of women with an incident HPV16 infection tested positive for a type other than HPV16 at enrolment in the UOW cohort; all women in the BBS cohort were HPV DNA-negative in cervical samples at enrolment. After enrolment, 23% of women in the BBS cohort with an incident cervical HPV16 infection tested positive for a type other than HPV16 before HPV16 was first detected; this statistic is not reported for the UOW cohort.

16.4.6 Availability of sera samples

The mean number of sera samples available for women in the UOW cohort was 6.6, compared to 4.2 for the BBS cohort. The mean number of sera samples available for women in the BBS cohort with an incident cervical HPV16 infection was 5.3, and 5.8 for those with an incident cervical HPV18 infection.

16.4.7 Type of infection

Regrettably, none of the papers describing the immune response to incident HPV infection among the UOW papers, distinguishes between women with cervical and vulval HPV infections. In contrast, the BBS cohort includes only women with incident *cervical* HPV infections (intentionally: infections at other sites were not tested for explicitly).

16.5 DISCUSSION OF LONGITUDINAL STUDIES

Unlike the BBS cohort, the UOW cohort does not describe the kinetics of the immune response to incident cervical HPV infections *specifically*. The estimated time to seroconversion following an incident HPV16 infection also varies across the different reports based on the UOW cohort, with differences attributed to changes in sample size. Alternatively, the variation may reflect differences between the reports in terms of the methods used to measure or analyse time to seroconversion. It is also impossible to interpret reported associations between the occurrence of cervical cytological abnormality and seroconversion following an incident cervical HPV infection, given that an unspecified, but substantial, number of these infections did not involve the cervix. The BBS cohort includes a substantial number of women with incident cervical HPV16 and/or HPV18 infections; it is also a potentially “cleaner” cohort for investigating the immune to cervical HPV infections because it includes only women with such infections, all of whom tested negative for HPV DNA and had a normal cervical smear at study entry. The analysis of the immune response to cervical HPV infection in the cohort described in this thesis is therefore worth pursuing.

16.5.1 Need for further studies

Some of the cohort studies described above have now been analysed and reported extensively. While most of the cohorts are large, the study population for the analysis of the immune response to HPV infection has often been quite small. Overall, these

studies do not provide convincing evidence of temporality, i.e. that the immune response *follows*, rather than *precedes*, the detection of HPV DNA; and some of the more mature studies do not separately distinguish cervical from other lower genital tract infections. However, the main criticism which can be levelled at all of the studies described above relates to the assay used to measure the immune response. Measuring the HPV-specific immune response requires prior knowledge of the types of immunity that may be important, and different immunological assays often produce conflicting results (Crosbie 2006). In all of the studies discussed above, VLP-based ELISA was used to detect antibodies to HPV infection: however, this assay may not measure the functionally-relevant aspect of the humoral immune response, the *neutralizing antibody response*.

16.6 THE NEUTRALIZING ANTIBODY RESPONSE TO HPV

It is currently assumed that the protective humoral response to HPV infection is mediated *via* neutralizing antibodies (Doorbar 2005). Neutralizing antibodies prevent the infection of susceptible cells by HPV, by preventing the HPV virions from attaching to cellular receptors, thereby denying them access to the cell and therefore also denying them the opportunity to replicate. In natural infection in animals, the generation of anti-L1 (the major capsid protein: see section 3.1.2.1) neutralizing antibodies is associated with protection against challenge with infectious virus (Stanley 2006). However, a newly synthesised HPV virion is not released from a cell until it reaches the epithelial surface (Doorbar 2005), and is thus protected from interference by the humoral immune system of the host throughout the majority of its

life cycle. It is therefore unlikely that antibodies alone, neutralizing or otherwise, will be sufficient to clear an HPV infection once it has become established.

16.7 MEASUREMENT OF THE HUMORAL IMMUNE RESPONSE

16.7.1 HPV neutralization assays

Neutralization assays are serological assays which measure the neutralizing-antibody response to a virus. Some neutralization assays employ real viruses, but most use surrogates, in particular the virus-like particle-based enzyme-linked immunosorbent assay (VLP-based ELISA).

16.7.1.1 VLP-based ELISAs

Enzyme-linked immunosorbent assay (ELISA) is a widely used immunoassay for the detection of antigen. VLP-based ELISA's are so-called because they use a virus like particle (VLP) as the target antigen. The capsid of every HPV type consists of two structural proteins, L1 and L2, which, although homologous across types, are nevertheless specific to a given type (see section 3.1.2.1). VLP are manufactured *in vitro* from these structural proteins: they resemble the virus from which they are derived, but contain no viral genome, and are therefore non-infectious. VLPs can, however, display conformationally-dependent surface epitopes, and so are expected to generate a similar immune response to the real virus which they mimic.

16.7.1.1.1 VLP-based ELISAs: Advantages

VLP-based ELISA is now a well-established technique: it is sensitive, rapid to perform, and inexpensive. VLP-based ELISA is an economical assay for testing samples on a large scale, high-throughput, testing.

16.7.1.1.2 VLP-based ELISAs: Disadvantages

While VLP can be produced in large quantities using commercially-available expression systems, manufacturing stable VLP for some types of HPV, in particular for HPV18, is still not straightforward. VLP-based ELISA are reliant upon intact high-quality VLP; disrupted or incorrectly-folded VLP may lead to the detection of non-neutralizing and cross-genotype-reactive antibodies. Cross-reactions among phylogenetically-similar papillomaviruses have been reported, and multiple cross-reactions between distinct types have been observed in a highly-infected population (Coursaget 2003). It is therefore potentially difficult to determine whether the assay is measuring a neutralizing, or a non-neutralizing, immune response, and to distinguish between an immune response to the target HPV type, to one of its closest genotypes, or to another HPV type.

16.7.1.2 Pseudovirus-based assays

Early neutralization assays were technically demanding, expensive, low-throughput, and the limited in terms of the number of HPV types which could be assayed. Most assays now rely on the production of pseudovirions *in vitro* (Roden 1996, Unckell 1997, Kawana 1998, Stauffer 1998, Touze 1998, Rossi 2000), yielding assays which are comparatively high-throughput (Pastrana 2004, Buck 2005). In theory, almost any type of HPV-pseudovirion, i.e. a PsV which mimics any of the HPV types, can be produced, although some types are easier to produce and more stable than others.

Pseudovirions are similar to VLP, in that they are constructed of the L1 and L2 proteins of the relevant virus, but they differ from VLP in that they also package a “genome”, which usually encodes a reporter or marker gene. Pseudovirions are infectious in the sense that they are capable of attaching themselves to the appropriate cell-surface receptors of a susceptible cell and delivering the pseudovirions genome into that cell, where gene expression of the genes encoded by the pseudovirus genome can occur. Measurement of the expression of reporter molecules allows quantification of how efficient an “infection” has been, and consequently how much neutralizing antibody is present in a test serum.

16.7.1.2.1 Pseudovirus-based assays: Advantages

Pseudovirus-based assays measure the most functionally-relevant aspect of the immune response to viruses. There is no doubt that an appropriately conducted pseudovirus-based assay specifically measures neutralizing, as opposed to non-neutralizing, antibodies: unlike VLP-based assays, the “genes” of a pseudovirion will only be expressed if the pseudovirus enters (infects) the cell, and co-opts the cellular DNA-replication machinery so that the gene which produces the marker molecule can be expressed. This assay is sensitive, and HPV type-specific.

16.7.1.2.2 Pseudovirus-based assays: Disadvantages

Neutralization assays are a relatively new technology. At present, they are cumbersome, time consuming, technically difficult to establish, i.e. get to work correctly at first, and perform, and are expensive.

16.7.1.3 Justification for the assay used in this study

In this study, an adaption of a neutralizing antibody assay originally developed by Pastrana *et al.* (Pastrana 2004) was used. This assay is described in detail in chapter 17. This assay is as sensitive as the commonly used VLP-based ELISA, is more HPV type-specific, and specifically measures neutralizing antibodies, the functionally-relevant antibodies.

Having first characterised the reproducibility of the neutralizing antibody assay, I then describe the kinetics of the neutralizing antibody response in a cohort of young women who were recruited soon after first sexual intercourse, and who first tested positive in cervical samples for HPV16 DNA, or HPV18 DNA, or both, during follow-up.

Chapter 17

THE KINETICS OF THE HUMORAL IMMUNE RESPONSE TO INCIDENT CERVICAL HPV16 AND HPV18 INFECTIONS

METHODS

Orientation

In this chapter I:

- describe the laboratory methods and procedures used to establish and carry out a neutralizing antibody assay
- describe the way in which the neutralizing antibody assay was carried out in practice
- provide a numerical example to illustrate the data produced by the neutralizing antibody assay and the method used to calculate the magnitude of the neutralizing antibody response
- describe the statistical methods used to measure intra- and inter-assay reliability

17.1 SEROLOGICAL SAMPLES

At each visit, women were requested to provide a blood sample. Serum was collected from each blood sample, by centrifugation, and stored for future analysis.

17.2 THE DETECTION OF NEUTRALIZING ANTIBODIES TO HPV16 AND HPV18

Neutralizing antibody assays detect functionally relevant antibodies which block infection of susceptible target cells. The assay used in this analysis employed structures which “simulate” the viruses of interest, called pseudovirions (PsV); these are essentially virus-like particles which also encapsidate a reporter-gene. Serum samples were tested for their ability to block the infection of target cells by these PsV.

In this study we have used an adaption of a neutralizing antibody assay originally developed by Pastrana *et al*. (Pastrana 2004). A description of the laboratory methods used in conducting this assay is given below; a formal technical description appears in appendix 5.

17.2.1 Preparation of pseudovirions

To manufacture HPV type-specific PsV, 293TT cells were transfected with three plasmid vectors (a plasmid is a small extra-chromosomal DNA molecule which is capable of independent replication): one each encoding the major (L1) and minor (L2) capsid proteins of the relevant HPV type (see section 3.1.2.1), and one encoding a “reporter” molecule, in this case secreted (embryonic) alkaline phosphatise (SEAP). For reasons which will be explained subsequently, PsV of bovine papillomavirus type 1 (BPV1) were also manufactured, using two plasmid vectors: one encoding *both* of the capsid proteins of BPV1, and one encoding SEAP.

HPV16, HPV18 and BPV1 PsV were thus manufactured in essentially identical, but separate, processes.

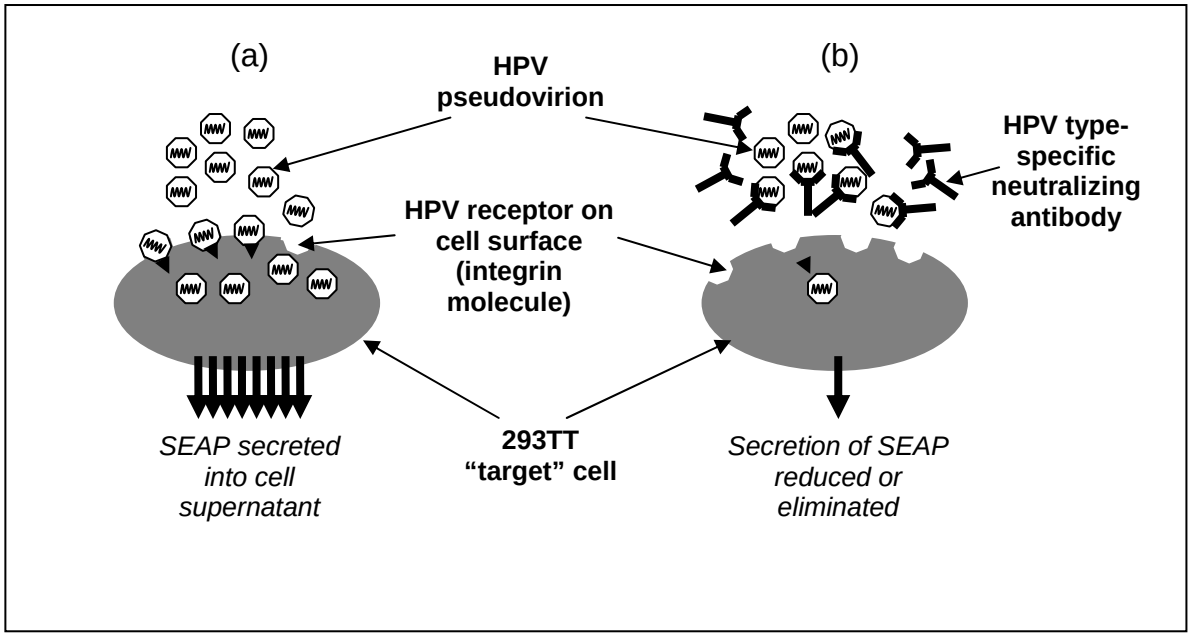
Once a cell has been transfected with a plasmid, it will begin to express the gene encoded by that plasmid. In this case, this will lead to the expression of the L1 and L2 proteins of the viruses of interest. The L1 and L2 proteins self-assemble into a PsV consisting of a capsid encapsulating the plasmid encoding SEAP. It is this SEAP-encoding piece of DNA which distinguishes a PsV from a virus-like particle (VLP): a VLP consists of an empty capsid, whereas a PsV also contains a DNA “genome”. Having been transfected, and the PsV allowed to assemble and mature, the 293TT cells are then disrupted to release the PsV, which are then isolated and purified.

17.2.2 The Neutralizing antibody assay

The principle behind the assay is illustrated in figure 17.2.2. 293TT cells are used as “target” cells in this assay (as well as being used in the manufacture of PsV). These cells possess the appropriate cell-surface receptors for infection by HPV, and therefore also the PsV which are designed to mimic HPV. If a solution of PsV and target cells is prepared, the PsV will “infect” the target cells and start to express the PsV genes, which in this case means the expression of SEAP (figure 17.2.2 part a). The target cells secrete SEAP into the surrounding culture medium. Quantitation is then possible, following the addition of a substrate for alkaline phosphatise, which produces a luminescent product which can be measured. However, if antibodies to

the virus mimicked by the PsV, or more specifically a serum sample which is suspected of containing such antibodies, is also included in the solution being tested, then these antibodies can potentially “neutralize” the PsV before they infect the target cells, thereby preventing, or reducing, the expression of SEAP (figure 17.2.2 part b).

Figure 17.2.2. The principle behind the neutralizing antibody assay.



Developing and optimising the neutralizing antibody assay in the laboratory in which the assay was implemented took several months. Setting-up the assay required the concentrated effort of a laboratory technician for most of a working day; which was then followed by three days of incubation i.e. the assay components were mixed and were allowed to interact for three days; and finally, the processing of the well-plates for reading again consumed a substantial portion of a working day. At any given time, only a single laboratory technician was employed in conducting this assay. The

extensive amount of work required in the first and third steps of the assay, outlined above, severely limited the number of samples which could be analysed at any one time. The amount of time required to conduct the assay from start to finish, and the constraints of the working week, further limited the throughput of this assay.

17.2.3 Determination of the amount of pseudovirus to use in the neutralizing antibody assay

One of the final stages in the development of the neutralizing antibody assay was to determine the amount of PsV to use: if too many PsV were used, even sera containing high antibody titres would be unable to neutralize the PsV; if too few were used, the PsV would be neutralized by sera with even very low antibody titres.

The production, purification and determination of the amount of PsV to use in the assays for each of the three viruses involved, i.e. HPV16, HPV18 and BPV1, consumed five months.

17.3 CONDUCTING THE NEUTRALIZING ANTIBODY ASSAY

To measure the neutralizing antibody response to HPV16 or HPV18 using the selected neutralizing antibody assay, it was necessary to conduct *two* assays simultaneously: one for the specific HPV type of interest, and another for the “non-specific immune response”, which involved testing the sample for the neutralizing

antibody response to BPV1. Both assays were conducted in the same way, and the results interpreted in conjunction.

17.3.1 Well-plate configuration

The assay was conducted using a 96-well well-plate, comprising a grid of wells in an 8-row-by-12-column configuration. Within laboratory limits, each non-empty well on the well-plate contained identical numbers of 293TT target cells, but the non-empty wells differed with respect to other factors (figure 17.3.1):

1) Empty wells.

The perimeter wells, i.e. the wells closest to the edge of the well-plate, were left empty, to avoid evaporation. Clearly, this restricted the number of informative wells on each well-plate in each assay, but was unavoidable since perimeter wells often produce erratic results with this assay.

2) “Sample wells”.

Each study sample was tested using all six of the internal wells of one column on the well-plate. These wells contained target cells and PsV in identical concentrations, but with increasing dilutions of study sample serum moving down the column. These dilutions were chosen (following the developmental assays described subsequently) because they were considered to cover the range within which the titres of the study

samples were likely to lie: the first well was at a dilution of 1-in-40, with three-fold increases in dilution in successive wells, i.e. dilutions of 1-in-120, 1-in-360, 1-in-1080, 1-in-3240, and 1-in-9720. Assuming that the sample contains neutralizing antibodies, then moving down the column, wells contain progressively fewer antibodies and are therefore less likely to be able to neutralize the PsV. Therefore, the production of SEAP in wells will increase moving down the column, and the signal, i.e. the amount of SEAP measured, will increase.

The minimum dilution used was 1-in-40 because Pastrana *et al.* (Pastrana 2004) found that below a dilution of 1-in-40, there was a high non-specific background reaction (non-specific inhibition). Therefore a titre of less than 1-in-40 is essentially "negative", whereas a titre of 1-in-40 or above is "positive", given a satisfactory comparison with the BPV1 titre (see section 7.3.2 point 6).

Given that only the internal wells of the well-plate could be used, i.e. the wells other than the perimeter wells, and that each sample occupied all of the internal wells in a given column, each 96-well well-plate could be used to test a maximum of 10 study samples. In practice, this number was further reduced by the need to include wells used for purposes other than testing samples directly.

Note that it is not a requirement of this neutralizing antibody assay that all dilutions of the same sample must be analysed in the same column: the wells used for the different dilutions can appear in any order on the same well-plate. But for practical

reasons, it is simpler to have one column correspond to the same sample, and for consecutive wells to have consecutive dilutions.

3) Positive- and negative-control wells.

These wells contain target cells and PsV in identical concentrations, but with increasing dilutions of control sample serum moving down the column. Ideally a standard (validated) set of controls with known titres would be used (e.g. controls supplied by an external agency), but internal controls can be used if they can be validated.

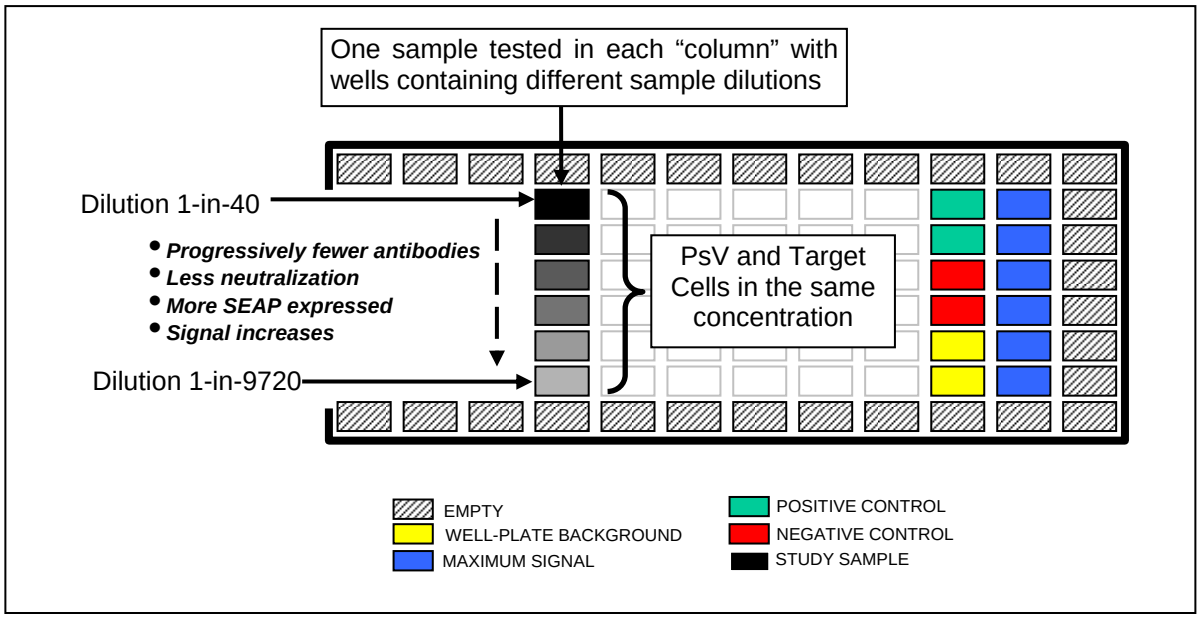
4) Maximum signal wells

These wells contain target cells and PsV in identical concentrations. These wells do not contain any antibodies, and therefore the expression of SEAP in these wells is the maximum attainable, since the PsV can “infect” the target cells unimpeded. The mean signal from all relevant wells is used in calculations (see section 17.3.2).

5) Background wells.

These wells contain identical concentrations of target cells only. They provide a background signal for the well-plate i.e. a measure of the natural production of SEAP by 293TT cells.

Figure 17.3.1. Well-plate configuration for conducting the neutralizing antibody assay. Three-fold dilutions assumed.



17.3.2 The result of the assay: determination of the 50% neutralization titre

The measurement, or “result”, produced by this neutralizing antibody assay is called the “50% neutralization titre” (or just the “titre” for short).

The result of the assay is based on the relative reduction in the expression of SEAP in the wells with serum added (sample wells), compared to the maximum expression possible. The 50% neutralization titre is the lowest dilution which causes at least a 50% reduction in SEAP expression compared to the maximum (corrected) SEAP expression possible.

The scale used to measure titres employing a well-plate with the configuration in figure 17.3.1 is an ordinal categorical scale with seven categories, one for each

dilution including a “negative” category. With the dilutions ultimately used, these categories were: negative, 1-in-40, 1-in-120, 1-in-360, 1-in-1080, 1-in-3240, and 1-in-9720. In theory, linear interpolation could also be used to obtain an “exact” measurement of the titre measured on a “continuous scale.

The result is calculated as follows:

1) The mean signal from all of the well-plate “background wells” is calculated.

2) The signal from each of the remaining non-empty wells on the well-plate is “corrected for background”, by subtracting the mean signal calculated from the background wells in step 1. In practice, the magnitude of the signals from these wells was small compared to those from the rest of the wells on the well-plate, and so this step was found to make little difference in practice.

3) The maximum possible signal is determined by calculating the mean corrected signal from all of the “maximum signal wells”.

4) The corrected signal from each sample-well is divided by the maximum possible signal. The “50% neutralization titre” is then the value of the highest dilution at which a 50% reduction in signal is first achieved.

5) For the positive- and negative-controls, calculate the 50% neutralization titre as in step 4. The results of the assay, for all samples tested on the same well-plate, will be

rejected if the well-plate is judged to have “failed” according to the known results for the positive- and negative-controls. A certain amount of fluctuation was apparent with this assay in practice. Therefore, two negative-control wells were included on each well-plate, with the well-plate being accepted if at most one negative-control was erroneously positive.

6) Compare the observed 50% neutralization titre against HPV16 or HPV18, with the observed 50% neutralization titre against BPV1.

BPV1 pseudovirus was included as a specificity control. Human exposure to BPV1 is assumed to be rare, certainly within the context of the study described in this thesis i.e. young women attending an urban family planning clinic are considered unlikely to have been exposed to a virus carried by *cattle*. Therefore, sera samples from humans are not expected to display substantial anti-BPV1 neutralizing activity. Consequently, any observed inhibition of the activity of BPV1 PsV will not result from cross-neutralizing antibodies against HPV16 or HPV18, but rather from non-specific factors present in sera. Pastrana *et al*. (Pastrana 2004) found that all but one of the 60 unvaccinated human subjects they tested had a titre against BPV1 of at most 1-in-40; the exception being a subject with a titre of 1-in-160. They therefore recommend that a serum sample should only be considered positive for HPV type-specific neutralizing antibodies if its HPV type-specific neutralization titre is greater than 1-in-40 and also at least one titre category greater than that found in the BPV1 assay.

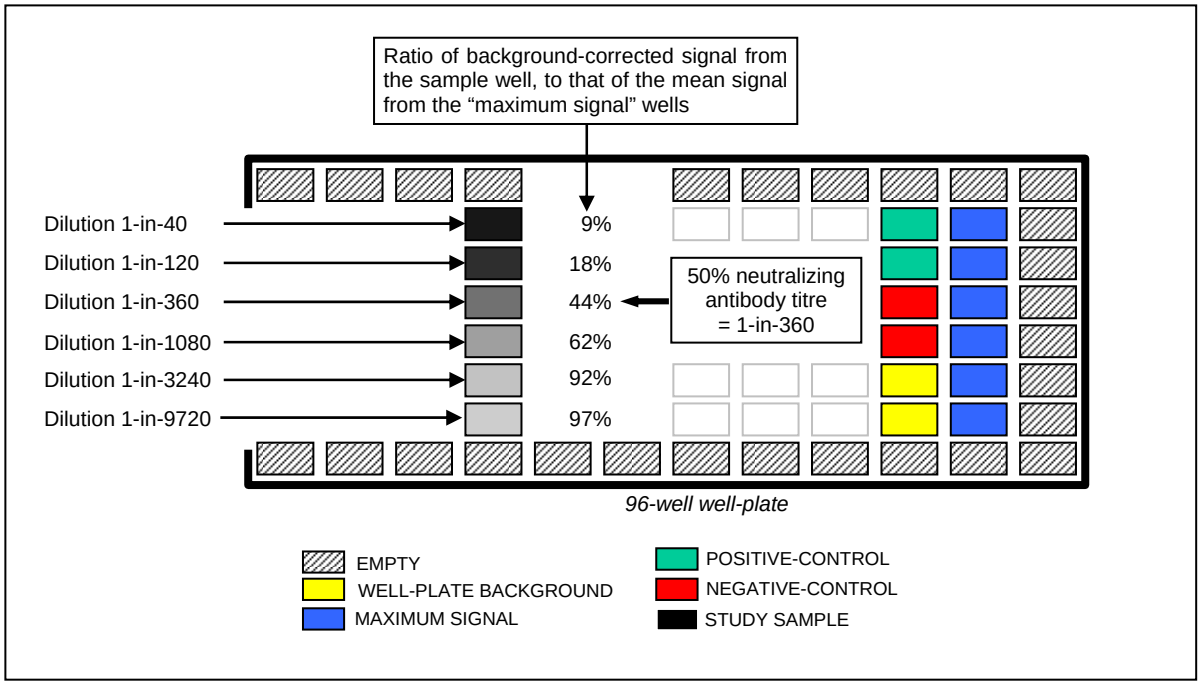
Therefore, although one 96-well well-plate can be used to test a maximum of eight study samples for neutralizing antibodies to a given HPV type (with the configuration given in figure 17.3.1), this well-plate must be “supported” by an additional well-plate testing the samples for neutralizing antibodies to BPV1. Provided samples are tested contemporaneously for neutralizing antibodies to both HPV16 and HPV18 (on separate well-plates), a single BPV1 assay can be used to support both type-specific assays. Hence, testing *N* well-plates containing study samples implies running *3N* well-plates in total i.e. no matter how many HPV types a sample is tested for, only one test for the non-specific neutralizing response needs to be undertaken.

An illustrative example of the calculation of the 50% neutralizing antibody titre is shown in figure 17.3.2. In this case, the well at a dilution of 1-in-360 yields a signal which is 44% of the maximum signal possible i.e. a decrease of 56%; the well at a dilution of 1-in-1080 yields a signal which is 62% of the maximum i.e. a decrease of only 38%. The 50% neutralization titre, the point at which 50% inhibition of the expression of SEAP was achieved, is therefore 1-in-360. Note that this implies that the “true” titre could be as low as 1-in-121 or as high as 1-in-1079, i.e. just above the next lowest titre, or just below the next highest titre; interpolation to identify the “true” titre (a value measured on a continuous scale) could be performed if required.

Provided the 50% neutralization titre against BPV1 is either negative, 1-in-40, or 1-in-120, then the sample illustrated in figure 17.3.2a is seropositive for neutralizing antibodies against the specific HPV type, and we conclude that the sample has a titre of 1-in-360 for reducing the expression of SEAP by 50%. If, however, the 50%

neutralization titre against BPV1 had been 1-in-360 or greater, then the conclusion would be that this sample is seronegative for neutralizing antibodies against the specific HPV type.

Figure 17.3.2a. Calculation of the 50% neutralizing antibody titre. Three-fold dilutions used.



A numerical example of this assay is shown in figure 17.3.2b. This figure presents the observed signals from all of the wells on the 96-well well-plate for a HPV16 assay (plate 2 of the 25.08.2006 assay).

In this example, the signals from the perimeter wells (the grey wells) are the lowest signals on the well-plate, but vary a great deal. These signals are not interpreted further.

The mean background signal for this well-plate is 1,745 (the mean of 1,600 and 1,890, the yellow wells). The magnitude of this signal is thus smaller than the signal from all other non-empty wells on the well-plate and, except when compared to some wells in the first two “rows”, is essentially negligible.

The mean signal from the six maximum signals wells is 134,793, or 133,048 when corrected for plate background. Hence for 50% neutralization, the “target” for study samples is to reduce this signal by at least half i.e. to a corrected signal of 66,524.

Figure 17.3.2b. Numerical example of the calculation of the 50% neutralizing titre: observed signals. Three-fold dilutions used. Grey cells correspond to empty wells on the well-plate; blue to maximum signal wells; yellow to well-plate background wells; green and red to positive- and negative-control wells, respectively; white cells correspond to study samples, with all wells in the same column being used to test the same sample at different dilutions, starting at 1-in-40 in row 2 and proceeding in multiples of three until 1-in-9720 is reached in row 7.

90	100	150	210	200	220	140	270	310	240	420	310
220	44240	2420	10160	2510	1810	2200	13360	124430	14870	165880	320
120	89050	8670	31840	9020	24510	22220	41640	119880	57230	146380	390
220	120580	25290	37740	31210	60340	33990	66140	133070	78210	133410	500
340	97810	43110	86220	59660	79350	63060	76090	130290	85760	108650	580
320	154460	73730	103210	105320	101390	81050	116350	111520	1890	114670	510
300	130900	98520	132600	105320	111230	121380	103320	110170	1600	139770	380
180	300	410	410	510	650	510	440	500	440	320	400

The ratio of the background-corrected signal to the mean corrected maximum signal for each of the wells used to test the study samples is shown in figure 17.3.2c, expressed as a percentage.

The relative signals from the negative-control wells were 57% and 63%. Therefore, in neither well was SEAP expression reduced by at least 50%, and so both are

seronegative, and the well-plate is not rejected on the basis of this criterion, and the results from this well-plate are accepted.

The positive-control wells were at a dilution of 1-in-360 and 1-in-1080. The relative signals from these wells were 10% and 42%, respectively. Therefore, SEAP expression was reduced by greater than 50% in both of these wells, and so the titre for the positive-control is at least 1-in-1080 i.e. the positive-control *is* positive, and the well-plate is not rejected on the basis of this criterion.

This well-plate was not rejected according to failure to pass either the positive- or negative-control criteria: therefore, the results (titres) of all samples tested on this well-plate were accepted as valid.

Sample 8 appears to contain no neutralizing antibodies. Even at the most concentrated dilution of 1-in-40, the signal is 92% of the maximum signal i.e. SEAP expression is essentially at a maximum. The expression at the other dilutions shows some fluctuation, but is close to 100%: the greatest signal reduction is 19% at a dilution of 1-in-9720.

Samples 1 to 7 are apparently seropositive for neutralizing antibodies. For each of these samples, the relative signal increases as the dilution increases. Although increases are not monotonic, wells at lower dilutions tend to produce greater relative signal reductions than those at higher dilutions. In each case, at least one well corresponds to a reduction in relative signal of 50% or more. Although the observed

50% neutralizing antibody titres for samples 3 and 7 are both 1-in-360, sample 3 produced a 73% relative reduction in signal at this dilution, compared to only a 52% reduction for sample 7. This suggests that the titre of sample 3 is in fact higher than that of sample 7. Similarly for samples 2 and 6.

Now, the 50% neutralizing antibody titres against BPV1 for the samples ranges from negative to 1-in-360. However, for most samples the BPV1 titre is less than the observed HPV16 titre. The exception is sample 1, for which there is an observed titre of 1-in-40 against HPV16, but 1-in-360 against BPV1, i.e. the observed titre against HPV16 is *not* at least three times that against BPV1, and therefore this sample is seronegative for neutralizing antibodies against HPV16.

Figure 17.3.2c. Numerical example of the calculation of the 50% neutralizing antibody titre: background-corrected signals relative to mean corrected maximum signals. Grey cells correspond to the observed 50% neutralization titre.

Dilution	**Sample number**							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1-in-40	32%	1%	6%	1%	0%	0%	9%	92%
1-in-120	66%	5%	23%	5%	17%	15%	30%	89%
1-in-360	89%	18%	27%	22%	44%	24%	48%	99%
1-in-1080	72%	31%	63%	44%	58%	46%	56%	97%
1-in-3240	115%	54%	76%	78%	75%	60%	86%	83%
1-in-9720	97%	73%	98%	78%	82%	90%	76%	81%
Observed 50% titre	1-in-40	1-in-1080	1-in-360	1-in-1080	1-in-360	1-in-1080	1-in-360	Negative
BPV 50% titre	1-in-360	Negative	1-in-120	Negative	1-in-40	1-in-120	1-in-40	1-in-40
50% NA titre	Negative	1-in-1080	1-in-360	1-in-1080	1-in-360	1-in-1080	1-in-360	Negative

17.4 STATISTICAL METHODS

17.4.1 The measurement of intra- and inter-assay reliability

To measure within-assay (intra-assay) reliability, samples were tested in duplicate on the same well-plate; to measure between-assay (inter-assay) reliability, the analysis of the samples on these well-plates was repeated, with the well-plate kept in the same configuration. Estimates of intra- and inter-assay reliability were then obtained using weighted kappa statistics.

To assess the extent to which a given measurement on a subject is reliable, it is necessary to make repeated measurements on a number of subjects: the observed degree of agreement provides an *upper bound* on the degree of reliability of the measurements (Fleiss 2003).

In any set of repeated observations on the same subject, some observations will be in agreement purely by chance. The kappa statistic is a chance-corrected measure of interrater agreement, taking the value one for complete agreement, values close to zero for no agreement, and values less than zero when observed agreement is less than that expected by chance; the minimum value depends on marginal proportions, but is not important in this context.

The unweighted kappa statistic depends on exact agreement only: it is the difference between the observed agreement (O) and the agreement expected by chance (E),

as a proportion of the maximum possible excess agreement above that expected by chance (1-E), i.e. unweighted kappa = (O – E) / (1 – E). When measurements are made on a categorical scale with more than two categories, and the categories are ordered, a difference of one category between two measurements is less serious than a difference of two or more categories. The weighted kappa statistic is a statistic which incorporates this aspect of the measurements, with values interpreted in the same way as values for the unweighted kappa statistic.

Landis and Koch (Landis 1977) give ranges for interpreting the strength of agreement based on kappa: less than 0.00 is **"**poor", 0.00 to 0.20 is "slight", 0.21 to 0.40 is "fair", 0.41 to 0.60 is "moderate", 0.61 to 0.80 is **"**substantial", and 0.81 to 1.00, is "almost perfect". Although the authors are clear in their paper that this categorization is arbitrary (clearly, it is obtained by dividing the range 0 to 1 into five equal parts) and intended for the specific purpose described in their paper, this categorization (which I will refer to as the "Landis scale") is now frequently used to report levels of agreement (including for the continuous analogue of the kappa statistic, the intracluster correlation coefficient). In essence, values of kappa of 0.75 or greater may be taken to represent excellent agreement beyond chance; values below 0.40 represent poor agreement beyond chance; and values between 0.40 and 0.75 represent fair to good agreement beyond chance (Fleiss 2003). Kappa does not have a probabilistic interpretation: even with a high value of kappa for a particular measurement technique, one has no quantifiable guarantee that if any given observation were repeated, one would obtain an essentially identical observation. However, a large value of kappa implies that an observed difference in

measurements made on the same individual, for instance at two different points in time, is a true difference rather than an artefact of the measurement technique: the higher the value of kappa, the greater is the reassurance that observed differences are real.

17.4.2 Further statistical analyses

Given the small sample size ultimately available, and the small number of serum samples involved, only a descriptive analysis was appropriate.

Chapter 18

THE KINETICS OF THE HUMORAL IMMUNE RESPONSE TO INCIDENT CERVICAL HPV16 AND HPV18 INFECTIONS

RESULTS

Orientation

In this chapter I:

- describe the laboratory and statistical analyses undertaken to establish a working assay for the analysis of the neutralizing antibody response to infection with HPV16 and HPV18
- determine the reproducibility of the neutralizing antibody assay
- describe the practical challenges involved with the use of the neutralizing antibody assay
- describe changes over time in the neutralizing antibody response against HPV16 and HPV18

18.1 RESTATEMENT OF OBJECTIVES

To investigate how the natural history of cervical HPV infection varies with the humoral immune response. Specifically:

- To evaluate the performance of an assay for the measurement of the neutralizing antibody response to infection with HPV16 and/or HPV18.

- To describe the kinetics of the neutralizing antibody response in a cohort of young women who were recruited soon after first sexual intercourse, and who first tested positive for HPV16 DNA, or HPV18 DNA, or both, in cervical samples during follow-up.

18.2 DEVELOPMENT AND VALIDATION OF AN ASSAY USED FOR THE MEASUREMENT OF THE NEUTRALIZING ANTIBODY RESPONSE TO CERVICAL HPV16 AND/OR HPV18 INFECTIONS

In the study described in this thesis, an adaption of a neutralizing antibody assay originally developed by Pastrana *et al*. (Pastrana 2004) was used. The development, validation, and implementation of the assay in the laboratory used to undertake the assay described in this thesis, occurred in three stages. The first stage was a laboratory process to establish the procedures for the production and purification of pseudovirions, and for the determination of the amount of pseudovirions to use in the assays, for each of the three viruses involved (HPV16, HPV18 and BPV1). This stage did not involve the candidate directly. The second stage involved applying, and evaluating, the techniques established in the first stage, to samples from women in the cohort described in this thesis: the objectives of this stage were to refine the procedures for the assay, including the determination of the optimal configuration of the well-plates; to determine whether the technique was reproducible; and to identify potential controls for use in future assays of samples from the study population described in this thesis. The third stage was the testing of study samples from cohort

members: this stage could only begin when assay development reached a point at which the technique could be applied with confidence to the testing of study samples, a limited resource.

18.2.1 Selection of study samples for use in the development and validation of the neutralizing antibody assay

The measurement of intra- and inter-assay reliability used 42 sera samples taken from 21 women in the full study cohort, and a further seven sera samples which had been included in the WHO collaborative study on the standardization of the detection of antibodies to human papillomaviruses (Ferguson 2006, hereafter referred to as the “WHO collaborative study”) (kindly supplied by Morag Ferguson). The frequency distribution of the number of sera samples tested per woman during the developmental stages of the neutralizing antibody assay is given in table 18.2.1a.

The samples from the study described in this thesis are a limited resource, and so must be used sparingly with future uses kept in mind, both known and as yet unperceived. The 42 sera samples from the study cohort women were therefore chosen with the larger aims of the longitudinal study described in this thesis, kept in view. Most women in the analysis used to develop and evaluate the neutralizing antibody assay contributed only a single sample because they only ever *had* one evaluable serum sample, or because they were from the “prevalent cohorts” i.e. women who were HPV DNA-positive, or cytologically abnormal in the cervical sample taken at study entry, or both. Such women already have cervical disease, or

are already infected with cervical HPV infection, and are therefore less valuable in an analysis of *incident* cervical HPV infection or disease. The women contributing more than one serum sample to this analysis tended to be those used in an attempt to identify suitable internal positive- and negative-controls (see section 18.2.3).

Table 18.2.1a. Frequency distribution of the number of serum samples tested per woman in the development and validation of the neutralizing antibody assay.

Number of samples tested per woman	Number of women
1	21[a]
2	1
3	1
4	4
7	1

[a]including the samples from the WHO collaborative study.

The cytological samples corresponding to the sera samples tested had been found to test positive for HPV16 DNA, HPV18 DNA, or both, using the GP5+/GP6+ system, either alone or in the presence of other HPV types; or to test negative for HPV (table 18.2.1b). The seven sera samples from the WHO collaborative study had been taken from women who were described as having “good histories of benign natural monovalent infections” (three with HPV16, two with HPV18, and one with both), with the other described as negative; however, the contemporaneous cervical HPV DNA status of these samples was unknown.

Table 18.2.1b. Frequency distribution of the cervical HPV infection status of the cervical cytological samples taken contemporaneously with the sera samples used in the development and validation of the neutralizing antibody assay.

HPV DNA status and types detected	Number of women
Negative	15
Positive	26
Missing[a]	1
Unknown[b]	7
16	9
18	6
16;18	4
6/11;16;18;52	1
18;33	1
56[c]	1
6/11;16;33	1
16;58	1
16;18;52	1
16;31	1

[a]The cytological sample taken on the date corresponding to the sera sample was missing; [b]these are the samples from the WHO collaborative study; [c]type obtained by sequencing.

18.2.2 Development and refinement of the neutralizing antibody assay: sera samples used

During the development and refinement of the neutralizing antibody assay, the assay was used to test a limited number of samples on several occasions. The frequency distribution of the number of occasions on which sera samples were tested is shown in table 18.2.2a.

Table 18.2.2a. Frequency distribution of the number of times sera samples were tested during the development and validation of the neutralizing antibody assay.

Number of times tested	Number of samples	
	HPV16[a]	HPV18[a]
0	–	12
2	8	11
4	27	13
6	1	1
8	2	4
10	3	–
14	1	1

[a]HPV type tested for.

During the developmental stage of the neutralizing antibody assay, samples were tested in duplicate on the same well-plate to enable the measurement of intra-plate reliability, as described subsequently. The lowest dilution used in all assays was 1-in-40, as recommended by Pastrana *et al*. (Pastrana 2004). However, the increments in dilution evaluated were two, three and four. This means that, even with a perfectly reliable (reproducible) assay, although the results of the repeated tests of the same sample can be compared for *consistency*, the observed titres can not be expected to be *identical*.

For example, possible titres for a sample tested using a dilution of two are negative, 1-in-40, 1-in-80, 1-in-160, 1-in-320, 1-in-640, and 1-in-1280; possible titres for a sample tested using a dilution of three are 0, 1-in-40, 1-in-120, 1-in-360, 1-in-1080, 1-in-3240, and 1-in-9720. A sample which has a titre of 1-in-360 using a dilution of three is consistent with a titre of either 1-in-320 or 1-in-640 using a dilution of two; it is *not* consistent with a titre of 1-in-1280 since then we would have expected a titre of 1-in-1080 using a dilution of three.

The frequency with which each serum sample was tested for neutralization against HPV16 at each dilution increment, is shown in table 18.2.2b; the corresponding frequency distribution for HPV18 is shown in table 18.2.2c.

Table 18.2.2b. Frequency distribution of the number of times with which each serum sample was tested for neutralization against HPV16, and the dilution increment used. For example, three samples were tested for HPV16 twice in assays using two-fold dilutions, six times using three-fold dilutions, and twice using four-fold dilutions.

Number of times tested and **dilution used**											**Number of samples**
2		**3**						**4**			
0	*2*	*0*	*2*	*4*	*6*	*8*	*10*	*0*	*2*	*4*	
	X		X						X		1
	X			X					X		1
	X				X				X		3
	X						X		X		1
	X	X							X		6
X				X				X			10
X				X						X	1
X		X							X		8
X		X								X	11

Table 18.2.2c. Frequency distribution of the number of times with which each serum sample was tested for neutralization against HPV18 and the dilution increment used. For example, 10 samples were tested for HPV18 four times in assays using three-fold dilutions, but not at all in any of the assays which used two- or four-fold dilutions.

Number of times tested and **dilution used**											**Number of samples**
2		**3**						**4**			
0	*2*	*0*	*2*	*4*	*6*	*8*	*10*	*0*	*2*	*4*	
X		X						X			12
	X	X						X			4
	X		X					X			1
	X			X					X		1
	X				X			X			2
	X				X				X		1
	X						X		X		1
	X	X							X		2
X				X				X			10
X				X					X		1
X		X							X		7

18.2.3. Development and refinement of the neutralizing antibody assay: developmental analyses

Twelve analyses were conducted during the developmental stage of the neutralizing antibody assay (recall that an “analysis” consisted of a group of well-plates tested at the same time). In all analyses, the overriding objective was to determine whether the assay was reproducible, but different analyses also had additional practical objectives. The analyses and their specific objectives are adumbrated in table 18.2.3a. Note that, had the technique proved reproducible at the first attempt, it is likely that at least four, but less than 12, analyses would have been undertaken prior to proceeding to stage three of the analysis.

Table 18.2.3a. Analyses undertaken during the developmental stage of the neutralizing antibody assay, and their specific objectives.

Analysis	Objective(s) and comments
1	To measure intra-assay reliability; first use of study samples
2	To measure intra- and inter-assay reliability; repeat of analysis 1
3	To identify potential internal positive and negative-controls
4	To identify potential internal positive and negative-controls
5	To identify potential internal positive and negative-controls
6	To evaluate the effect of refining the HPV18 assay; first use of controls from the WHO collaborative study
7	Training for a new laboratory technician; repeat of analysis 6 for HPV18 only
8	To evaluate the effect of refining the HPV18 assay
9	To measure intra- and inter-assay reliability; repeat of analysis 8
10	To evaluate the effect of refining the HPV18 assay; to measure intra- and inter-assay reliability; repeat of analysis 8
11	To identify potential internal and external positive- and negative-controls for use during testing of the study samples; to measure intra-assay reliability
12	To identify potential internal and external positive- and negative-controls for use during the testing of study samples; to measure intra- and inter-assay reliability; repeat of analysis 11

As well as varying in their objectives, these analyses involved different well-plate configurations: these are described in table 18.2.3b. In most analyses, the neutralizing antibody response to both HPV16 and HPV18 infection was determined, but two analyses measured the response to HPV16 only, and one to HPV18 only. The number of well plates which could be set-up and read by a single laboratory technician in one analysis was limited to three for each of the virus types i.e. a total of nine well-plates in any single analysis. Each sample was measured in duplicate on each well-plate, therefore a maximum of four samples could be measured on each well-plate. In the first two analyses, only 10 samples were analysed on each well-plate; thereafter, 12 samples were analysed on each well-plate in all analyses. Initially, the increment in dilutions used was four (i.e. dilutions of 0, 1-in-40, 1-in-160

etc.). Once experience had been gained with the assay, and the likely range of titres for study samples established, the increment was altered, first to two (i.e. dilutions of 0, 1-in-40, 1-in-80, etc.), and then to three (0, 1-in-40, 1-in-120 etc.); after analysis 5, all subsequent analyses used an increment of three. On each well-plate, the aim was to measure as many study samples as possible, whilst still allowing some wells to be used for measuring well-plate-specific data, i.e. the background for the well-plate, the maximum signal possible for the well-plate, and the signal for controls (see section 17.3.1). In practice, this meant that two “columns” were available for well-plate-specific data on each well-plate. The number of wells used for each of these purposes was varied in the first few analyses. Once it had been observed that the background signal was always very low, the number of wells used for measuring the background signal was reduced from four to two; the use of only one well in analysis 4 was not repeated, since it was considered prudent to allow for at least some variability in estimating all signals and thereby to enable an aberrant measurement to be identified. Given that the (corrected) maximum signal is used as a comparator for all other wells, most available wells were used for measuring this parameter to enable a more precise, or stable, estimate to be obtained. Ultimately, once the testing of study samples began, all usable wells in one column, i.e. six wells, were used for measuring the maximum signal. In most laboratory analyses, positive- and negative-controls are included to rule out systematic errors: however, as will be discussed subsequently, only analyses 3 and 4 included such controls during the developmental stage of the assay.

Table 18.2.3b. Well-plate configurations used in the twelve analyses undertaken during the developmental stage of the neutralizing antibody assay.

Analysis	HPV[a]	Sample size	Number of well plates	Dilutions	Number of maximum signal wells	Number of background wells	Number of control wells[b]
1	16+18	10	9	4	8	4	0
2	16+18	10	9	4	8	4	0
3	16	12	6	4	4	2	4+/2–
4	16+18	12	6	4	5	1	4+/2–
5	16+18	12	9	2	10	2	0
6	16+18	12	9	3	10	2	0
7	18	12	6	3	10	2	0
8	16+18	12	9	3	10	2	0
9	16+18	12	9	3	10	2	0
10	16+18	12	9	3	10	2	0
11	16+18	12	9	3	10	2	0
12	16+18	12	9	3	10	2	0

[a]HPV type assayed; [b]No. For positive-/negative-controls.

A brief description of some of the issues raised by each of these analyses follows.

18.2.3.1 Developmental analyses 1 and 2

These two analyses were the first attempts to measure intra- and inter-assay reliability, and will be described in greater detail than the others to illustrate the data produced by the assay and the method of analysis. Summary tables for all of the developmental analyses appear at the end of this section, but in this subsection I provide a more full description of the type of data produced by the assay, and the approach adopted for the analysis of the data collected for the study described in this thesis.

In analyses 1 and 2, a total of ten sera samples from women in the full study cohort of 2,011 women were analysed for their neutralizing antibody response to both

HPV16 and HPV18, with each sample analysed in duplicate on the same well-plate. Although not a requirement of the assay, in practice the same sample was analysed in contiguous columns of the well-plate. In an attempt to obtain a mixture of seropositive and seronegative results for both types in these initial analyses, the cytological samples corresponding to the sera samples were deliberately chosen to be a mixture of HPV16 DNA- and/or HPV18 DNA-positive, and HPV DNA-negative samples. Well-plates for analysis 1 were set-up on Monday 7th November 2005, with SEAP-expression signals being measured ("read") the following Thursday; those for analysis 2 were set-up on Tuesday 8th November 2005, and read the following Friday, a clear demonstration of how the working week limited the number of assays which could be performed by a *single* laboratory technician.

The sera samples used in analyses 1 and 2, and the HPV DNA status of the corresponding cytological samples, are shown in table 18.2.3.1a.

Table 18.2.3.1a. Sera samples, and the HPV DNA status of the corresponding virological samples, used in analyses 1 and 2

Neutralizing antibody assay identifier	Serum sample study number	HPV DNA status of corresponding virological sample
1	1619.1	HPV16
2	1688.1	HPV18
3	1675.1	HPV-negative
4	1766.3	HPV16+HPV31
5	1239.1	HPV16
6	1161.2	HPV16+HPV58
7	965.1	HPV16+HPV6/11+HPV33
8	974.1	HPV18
9	126.1	HPV16
10	1008.1	HPV16

The 50% neutralization titres against HPV16 and HPV18 for each replicate of each sample from analyses 1 and 2 are shown in table 18.2.3.1b. Note that the lowest non-zero dilution used was 1-in-40, with increments in multiples of four.

Titres ranged from zero (sernegative) to 1-in-2560 for HPV16, and from zero to 1-in-640 for HPV18. Sixteen of the 40 BPV1 results were non-zero (i.e. non-seronegative), with a maximum titre of 1-in-160; nine non-zero HPV16 titres were changed to zero (seronegative) following comparison with the BPV1 titre, as were three non-zero HPV18 titres; one replicate from one analysis had a BPV1 titre of 1-in-40 in the absence of any titre against HPV16 or HPV18; two results indicated a titre against BPV1 of 1-in-160 (BPV1 results not illustrated).

Table 18.2.3.1b. The 50% neutralization titres against HPV16 and HPV18 for each replicate of each sample from analyses 1 and 2. AnRm denotes analysis n (n=1,2) replicate m (m=1,2). Numbers in the table are inverse titres, except for "0", which is "seronegative".

Sample[a]	HPV16[b]				HPV18[b]			
	A1R1	A1R2	A2R1	A2R2	A1R1	A1R2	A2R1	A2R2
1	160	160	160	160	0	0	160	0
2	0[c]	0[c]	0[c]	0[c]	0	0	0	0
3	40	0	40	0	640	640	640	640
4	40	40	40	40	0	0	0	0
5	40	0[c]	40	40	0	0	0	0
6	0	0[c]	0[c]	0[c]	0	0	0[c]	0[c]
7	640	640	2560	2560	0	0	0	160
8	0	40	0	0	40	40	0[c]	0
9	0[c]	160	160	160	640	640	640	640
10	0	0	0	40	0	0	40	40

[a]Corresponds to "neutralizing antibody assay identifier" in table 18.2.3.1a; [b]zero means seronegative for neutralizing antibodies; [c]these samples are seronegative after comparison with the non-specific neutralizing response obtained for BPV1.

In table 18.2.3.1b, there appears to be reasonable agreement between replicates and between analyses for each virus type. Agreement is perfect between replicates 1 and 2 for HPV18 in analysis 1, but for all other comparisons there is some discrepancy. The greatest discrepancy between titres for the same sample is a difference of two titre categories, but there are also several discrepancies of only one titre category. Tables 18.2.3.1c and 18.2.3.1d present estimates of weighted kappa statistics measuring agreement between each possible pair of replicates, for HPV16 and HPV18, respectively. Agreement ranges from "moderate" to "almost perfect" for HPV16, and from "substantial" to perfect for HPV18, using the Landis scale (see section 17.4.1).

Comparisons between replicates within the same analysis yield measurements of intra-assay (within-assay) reliability. For HPV16, intra-assay reliability is "moderate"

in analysis 1, but “almost perfect” in analysis 2. For HPV18, intra-assay reliability is perfect in analysis 1, and “substantial” in analysis 2.

Comparisons between replicates in *different* analyses yield measurements of inter-assay (between-assay) reliability: there are four ways in which a replicate from the first analysis can be paired with a replicate from the second (A1R1-A2R1, A1R1-A2R2, A1R2-A2R1, and A1R2-A2R2), hence there are four (dependent) estimates of inter-assay reliability. For HPV16, values of weighted kappa statistics measuring inter-assay reliability indicate reliability which ranges from “moderate” to “substantial”. For HPV18, values of weighted kappa statistics measuring inter-assay reliability are consistently 0.66 (to two decimal places), corresponding to “substantial” agreement. Ignoring the extreme result, intra- and inter-assay reliability are identical for HPV18.

Table 18.2.3.1c. Estimates of weighted kappa statistics measuring agreement between each possible pair of replicates included in analyses 1 and 2: HPV16. Shaded cells indicate inter-assay measurements, unshaded cells intra-assay measurements.

	Analysis 1 Replicate 2	Analysis 2 Replicate 1	Analysis 2 Replicate 2
Analysis 1 Replicate 1	0.53	0.74	0.57
Analysis 1 Replicate 2	–	0.67	0.67
Analysis 2 Replicate 1	–	–	0.84

Table 18.2.3.1d. Estimates of weighted kappa statistics measuring agreement between each possible pair of replicates included in analyses 1 and 2: HPV18. Shaded cells indicate inter-assay measurements, unshaded cells intra-assay measurements.

	Analysis 1 Replicate 2	Analysis 2 Replicate 1	Analysis 2 Replicate 2
Analysis 1 Replicate 1	1.00	0.66	0.66
Analysis 1 Replicate 2	–	0.66	0.66
Analysis 2 Replicate 1	–	–	0.67

The comparison with the least agreement is that between replicates 1 and 2 from analysis 1 of HPV16 (kappa=0.53). Figure 18.2.3.1 presents a cross-tabulation of the results for these two replicates.

Figure 18.2.3.1. Comparison of the 50% neutralization titres against HPV16 for both replicates in analysis 1. Inverse titres are given in the table for ease of display, except for 0 which is "seronegative". Shaded cells correspond to perfect intra-assay agreement.

		Analysis 1 Replicate 1						
	Titre	0	40	160	640	2560	10240	40960
Analysis 1 Replicate 2	0	3	1[a]	1[b]	–	–	–	–
	40	2[a]	1	–	–	–	–	–
	160	–	–	1	–	–	–	–
	640	–	–	–	1	–	–	–
	2560	–	–	–	–	–	–	–
	10240	–	–	–	–	–	–	–
	40960	–	–	–	–	–	–	–

[a]Titre discrepant by one category; [b]titre discrepant by two categories

Ideally, all results would lie on the leading diagonal of table 18.2.3.1e. However, three results are discrepant by one titre category, and one is discrepant by two categories, yielding a titre of 1-in-160 for replicate 1, but a negative result for replicate 2. The weighted kappa statistic measuring agreement in this table is 0.53 (95% CI 0.05 to 1.00), indicating "moderate" agreement only; kappa is significantly

different from zero at the 5% two-sided significance level. The 95% confidence interval and test of statistical significance are reliant on large sample approximations, which, however, are unlikely to be met with a sample size of 10. Even if the approximation *is* valid, this confidence interval suggests that the data is consistent with the full range of agreement which is possible, from none to “almost perfect”.

For HPV16, the results of these first two analyses were certainly encouraging, but achieved “almost perfect” agreement for only one comparison. For HPV18, the well-plate signals were found to be somewhat erratic, not withstanding the reasonable values of kappa statistics obtained. It was suggested that the HPV18 pseudovirions (PsV) were not maturing sufficiently i.e. PsV were self-constructing correctly, but were not maintaining their structural integrity. Although the HPV16 PsV did not, apparently, suffer from this problem, further refinement of both assays was felt to be required.

18.2.3.2 Developmental analyses 3 and 4

As well as refining the assay, and gaining more experience in its use and interpretation, it was necessary to try and identify suitable positive- and negative-controls for use when testing study samples. The criteria for a suitable positive-control was that it produce a high-titre, which was towards the middle of the possible range of titres i.e. not too extreme. The ideal criteria for a suitable negative control was that it was consistently seronegative.

To identify potential positive controls, and in the absence of better criteria, all women who did not belong to any of the “important” sub-cohorts (women who were free of cervical HPV infection and cervical disease at study entry and who had further follow-up: for these women, ideally all samples would be used as little as possible) were identified. Those women with a lengthy sequence of HPV16 DNA- or HPV18 DNA-positive virological samples and who had sera samples available, were then identified, and three of each selected. All samples for the women identified were tested during analyses 3 and 4. The possible flaw with this approach was that women who had a long sequence of seropositive results may have had such a sequence because they did not, or could not, produce a neutralizing antibody response to their infection, thus defeating the purpose of this approach.

For negative-controls, all women who never acquired cervical HPV infection of any type, or cervical cytological abnormality, during follow-up, and who had sera samples available were identified, and one serum sample selected. Also, one woman who had several sera samples available prior to first sexual intercourse was selected. Again, all samples for the women identified were tested during analyses 3 and 4.

These analyses included wells for positive- and negative-controls. The positive-control used was the serum sample with sample ID 965.1, which had given a titre of 1-in-640 in both analyses 1 and 2. The negative-control used was the serum sample with sample ID 1675.1, which had given two titres of 1-in-40, as well as two seronegative results, in analyses 1 and 2. Clearly, this was therefore a less-than-ideal negative-control.

Analysis 3 was restricted to testing for the neutralizing antibody response to HPV16 only, whilst the production of HPV18 PsV was re-examined; analysis 4 then included testing for HPV18 once again.

In analysis 3, all sera samples from the HPV16 DNA-positive women were found to be seropositive to HPV16, with the exception of one sample which had a titre of 1-in-40 for one replicate, but was seronegative for the other; the greatest titre observed was 1-in-640. All sera samples from the HPV DNA negative and cytologically normal woman were seronegative in all replicates; three replicates had titres of 1-in-40 for the woman who was a virgin during the period when the sera samples were taken. Three replicates disagreed by one titre category. The weighted kappa statistic measuring intra-assay reliability for this analysis was 0.87, indicating “almost perfect” agreement.

In analysis 4, the value of the weighted kappa statistic measuring intra-assay reliability for HPV16 was 0.85, indicating “almost perfect” agreement. For HPV18, the value of the weighted kappa statistic measuring intra-assay reliability was 0.80, again indicating “almost perfect” agreement, but it was nevertheless felt that this assay remained unstable (this was the first time a discrepancy of two titres had been seen).

It became apparent that the use of only one well for measuring the background signal from a well-plate was ill-advised. One of the three well-plates used to measure BPV1 had a very large background signal which was three to four times higher than

had been observed previously. In the absence of another well measuring this signal, it was not possible to state that this unusual signal was an extreme value, and therefore that it could be disregarded.

18.2.3.3 Developmental analysis 5

Instead of four-fold dilutions (1-in-40, 1-in-160, 1-in-640, 1-in-2560, 1-in-10240, 1-in-40960), as used in all previous analyses, this analysis used two-fold dilutions (1-in-40, 1-in-80, 1-in-160, 1-in-320, 1-in-640, 1-in-1280). This was not a *direct* repeat of an earlier analysis, but all samples involved *had* been tested previously. Seven of the samples tested in this analysis were potential HPV16-positive controls, two were potential HPV18 positive-controls, and three were potential negative-controls for both types. No non-specific assay (BPV1) controls were included.

The HPV16 assay appeared to be very reliable once again. Taken at face value, the dilutions chosen for this analysis did not cover the range of interest for HPV18 (all observed titres were at 1-in-1280, the maximum possible with this dilution), so the results were largely non-informative: all samples were either HPV18-seronegative or HPV18-seropositive at the maximum titre. However, the assay was again unstable: several sample wells gave signals either less than, or comparable in magnitude to, that from the background wells.

18.2.3.4 Developmental analyses 6 and 7

Attention was now focussed on producing a working assay for HPV18, and so HPV16 was not tested for in analyses 6 and 7. Included in these analyses were five samples from the study cohort described in this thesis, together with seven samples supplied from the WHO collaborative study for use as potential controls. Analysis 7 was a direct repeat of analysis 6, with the additional aim of training a laboratory technician (K Wen) who was new to this particular assay (he was otherwise very experienced), in performing the assay, under the guidance of the previous technician (G Ryan). The new technician would undertake all subsequent analyses, including, ultimately, those testing study cohort samples.

In both analyses 6 and 7, the laboratory scientists were concerned about observed signal variability, and recommended that the results be discarded (nevertheless, results for analyses 6 and 7 appear in tables 18.2.3.5a and 18.2.3.5b).

18.2.3.5 Developmental analyses 8 to 12

The laboratory scientists now believed that they had resolved the issue with the HPV18 assay, whatever its true nature. Analyses 8 and 9 were therefore the second attempt to estimate intra- and inter-assay reliability for both HPV16 and HPV18, and analysis 10 combined with analysis 8 was the third attempt. The samples included in these analyses, some of which had been analysed previously, were chosen in the hope that they would yield a range of titres against HPV18, to avoid the problem

found in analysis 5. The belief was that the samples chosen previously may have been responsible for the aberrant results observed, to some extent.

Unfortunately, the HPV18 results from analysis 9 again had to be discarded: the laboratory scientists stated that they had been “suspicious” of analysis 9 for HPV18 prior to obtaining any results for HPV18, and were not confident about the HPV16 results either. The estimate of the weighted kappa statistic measuring intra-assay reliability for analysis 8 was also unimpressive.

Based on analysis 10, reliability of the HPV16 assay was “almost perfect”, with only one intra-assay discrepancy between replicates, which was of one titre category only; also, at least one titre per sample from analysis 8 was consistent with at least one titre for that sample from analysis 10.

Results from the HPV18 assay were more variable than those from the HPV16 assay. Eleven of the twelve samples had discrepancies of one titre category between replicates in analysis 10, compared with five of 12 in analysis 8. However, at least one titre per sample from analysis 10 was consistent with at least one titre for that sample from analysis 8, and there were no discrepancies of two titres categories or more.

It was now concluded that the greater variability of the HPV18 assay was an inherent feature of the assay. The laboratory scientists were convinced that the HPV18 assay had reached the limits of possible refinement. Given the difficulties with analysis 8,

and the requirement, which still remained, of identifying potential positive and negative controls, two further analyses were undertaken. The developmental phase of the assay was then considered complete.

A summary of the results regarding intra-assay reliability of the developmental analyses is given in tables 18.2.3.5a and 18.2.3.5b, for HPV16 and HPV18 respectively. For HPV16, estimates of the weighted kappa statistic ranged from 0.56, corresponding to “moderate” agreement, to 0.95, for “almost perfect” agreement on the Landis scale; in all analyses, observed agreement was greater than 92%, consistent with most discrepancies being of only one titre category. In eight of the eleven analyses, estimates of the weighted kappa statistic suggest “almost perfect” agreement, according to the Landis scale. The analyses with the worst intra-assay reliability for HPV16 (analyses 1 and 8), were the first conducted by each of the two laboratory technicians involved in the validation of the assay (although the same result was not seen for the HPV18 assay).

Ignoring the extreme result from analysis 1, for HPV18 estimates of the weighted kappa statistic ranged from 0.52, corresponding to “moderate” agreement, to 0.90, “almost perfect” agreement; in all analyses, observed agreement was at least 85%. In four of the eight analyses, kappa values suggest “almost perfect” agreement according to the Landis scale. The HPV18 assay is clearly more erratic than the HPV16 assay, with a substantial number of discrepancies between titres, including several of two titre categories.

Table 18.2.3.5a. Estimates of weighted kappa statistics measuring intra-assay reliability for the developmental analyses of the neutralizing antibody assay: HPV16.

Analysis	Sample size[a]	Dilution	Number of discrepancies[b]	Observed agreement (%)	Expected agreement (%)	Weighted kappa
1	10	4	3(1)	93	84	0.57
2	10	4	2(1)	97	82	0.82
3	12	4	2(1)	97	78	0.87
4	12	4	3(1)	96	72	0.85
5	12	2	3(1)	96	55	0.91
6	12	3	1(1), 1(2)	96	67	0.87
7[c]	12	3	NA	NA	NA	NA
8	12	3	6(1)	92	81	0.56
9	12	3	4(1)	94	79	0.73
10	12	3	1(1)	99	85	0.91
11	12	3	1(1)	99	71	0.95
12	12	3	2(1)	97	74	0.89

[a]Number of sera samples tested; [b]number of discrepancies between replicates, with the number of titre categories difference in brackets e.g. in analysis 6 there was one difference of one titre, and one of two titres; [c]tested for HPV18 only

Table 18.2.3.5b. Estimates of weighted kappa statistics measuring intra-assay reliability for the developmental analyses of the neutralizing antibody assay: HPV18.

Analysis	Sample size	Dilution	Number of discrepancies[a]	Observed agreement (%)	Expected agreement (%)	Weighted kappa
1	10	4	0	100	82	1.00
2	10	4	1(1), 2(2)	92	79	0.60
3	12	4	NA	NA	NA	NA
4	12	4	3(1), 1(2)	93	65	0.80
5	12	2	NA	NA	NA	NA
6	12	3	4(1)	94	58	0.87
7	12	3	NA	NA	NA	NA
8	12	3	5(1)	93	67	0.79
9	12	3	5(1), 2(2)	88	65	0.64
10	12	3	11(1)	85	68	0.52
11	12	3	3(1), 1(2)	93	62	0.82
12	12	3	3(1)	96	59	0.90

[a]Number of discrepancies between replicates, with the number of titre categories difference in brackets e.g. in analysis 9 there were five differences of one titre, and two of two titres.

Summary

An adaption of a neutralizing antibody assay originally developed by Pastrana *et al*. (Pastrana 2004) was adapted for use in the measurement of the humoral immune response to HPV16 and HPV18 infections in sera samples. During the development and refinement of this assay, it was used to test a limited number of sera samples on several occasions: 42 sera samples taken from 21 women, and a further seven sera samples from an external source, were tested (often several times) in 12 analyses. An optimum well-plate configuration had been determined and the assay was considered to have reached the stage where the adaptation was believed to have been successful, and the reliability of the assay could be formally evaluated.

18.3 RELIABILITY OF THE NEUTRALIZING ANTIBODY ASSAY

To provide definitive measurements of the intra- and inter-assay reliability (within- and between-assay reliability, respectively) of the neutralizing antibody assay, and to confirm the viability of potential positive- and negative-controls identified for use during the testing of the samples from the study cohort described in this thesis, two final analyses were undertaken (analyses D1 and D2). These analyses provided two estimates of intra-assay reliability, and four (correlated) estimates of inter-assay reliability.

The configuration used for well-plates used is shown in figure 18.3. In analyses D1 and D2, the same set of 12 sera samples was tested, with four samples tested in duplicate in contiguous columns on each of the three well-plates analysed in each assay. This example shows samples 1 to 4; the second plate was used to test samples 5 to 8, and the third 9 to 12. The lowest dilution of serum was 1-in-40, increasing in multiples of three to a maximum of 1-in-9720. Ten wells containing

target cells and PsV in identical concentrations were used to measure the maximum signal. Two wells, containing target cells alone in identical concentrations, were used to measure the well-plate background signal.

Figure 18.3. The well-plate configuration for each of the well-plates used in the analyses to demonstrate the intra- and inter-assay reliability of the neutralizing antibody assay. Sn is sample number n, n=1 to 4. 1-in-D is the dilution of sera samples. Wells labelled "Cells+PsV" are wells used to measure maximum signal; those labelled "Cells alone" are wells used to measure the well-plate background signal.

	S1 1-in-40	S1 1-in-40	S2 1-in-40	S2 1-in-40	S3 1-in-40	S3 1-in-40	S4 1-in-40	S4 1-in-40	Cells + PsV	Cells + PsV	
	S1 1-in-120	S1 1-in-120	S2 1-in-120	S2 1-in-120	S3 1-in-120	S3 1-in-120	S4 1-in-120	S4 1-in-120	Cells + PsV	Cells + PsV	
	S1 1-in-360	S1 1-in-360	S2 1-in-360	S2 1-in-360	S3 1-in-360	S3 1-in-360	S4 1-in-360	S4 1-in-360	Cells + PsV	Cells + PsV	
	S1 1-in-1080	S1 1-in-1080	S2 1-in-1080	S2 1-in-1080	S3 1-in-1080	S3 1-in-1080	S4 1-in-1080	S4 1-in-1080	Cells + PsV	Cells + PsV	
	S1 1-in-3240	S1 1-in-3240	S2 1-in-3240	S2 1-in-3240	S3 1-in-3240	S3 1-in-3240	S4 1-in-3240	S4 1-in-3240	Cells + PsV	Cells + PsV	
	S1 1-in-9720	S1 1-in-9720	S2 1-in-9720	S2 1-in-9720	S3 1-in-9720	S3 1-in-9720	S4 1-in-9720	S4 1-in-9720	Cells alone	Cells alone	

The samples tested are presented in table 18.3a. The samples used comprised three potential internal HPV16 and/or HPV18 positive-controls, two potential internal HPV16 *and* HPV18 negative-controls, including one from a woman who was a virgin when the serum sample was taken, and seven potential positive- or negative-controls supplied from the WHO collaborative study. All samples had been tested previously on at least one occasion.

Table 18.3a. Sera samples, and the HPV DNA status of the corresponding virological samples, used in the analyses to demonstrate the reliability of the neutralizing antibody assay

Assay identifier	Serum sample identifier	HPV DNA status of corresponding virological sample
1	WHO1	Unknown
2	WHO2	Unknown
3	WHO3	Unknown
4	WHO4	Unknown
5	WHO5	Unknown
6	WHO6	Unknown
7	WHO7	Unknown
8	277.4	HPV-negative
9	368.7	HPV18
10	894.1	Result missing
11	1621.3	HPV16
12	1621.5	HPV-negative

The 50% neutralizing antibody titres against HPV16 and HPV18 for each replicate of each sample from analyses D1 and D2, are shown in table 18.3b. Note that the lowest dilution used in these analyses was 1-in-40, with increments in multiples of three.

Table 18.3b. The 50% neutralizing antibody titres (inverse) against HPV16 and HPV18 for each replicate of each sample from the final developmental analyses for the neutralizing antibody assays. DnRm denotes analysis Dn (n=1,2) replicate m (m=1,2).

Sample	HPV16[a]				HPV18[a]			
	D1R1	D1R2	D2R1	D2R2	D1R1	D1R2	D2R1	D2R2
1	1080	1080	360	360	0	0	0	0
2	0	0	0	0	0	0	0	0
3	360	360	360	120	40	0	0	0
4	40	0[b]	0[b]	40	360	360	120	120
5	0	40	0	0	0	0	0	0
6	0	0	0	0	40	40	40	0
7	0	0	0	0	1080	3240	1080	3240
8	0	0	0	0	40	0	0	0
9	40	0[b]	0	0	1080	1080	1080	1080
10	0	0	0	0	0	0	0	0
11	1080	1080	1080	1080	3240	3240	9720	9720
12	1080	1080	1080	360	1080	9720	3240	9720

[a]Zero means seronegative for neutralizing antibodies; [b]these samples are seronegative after comparison with the non-specific neutralizing antibody response obtained for BPV1.

18.3.1 Intra-assay reliability

18.3.1.1 HPV16

Estimates of weighted kappa statistics measuring intra-assay (within-assay) reliability for the neutralizing antibody response to HPV16 from analyses D1 and D2 are presented in table 18.3.1.1.

In analysis D1 there were three discrepancies between replicates in their measured neutralizing antibody response to HPV16, all of one titre category: the estimate of the weighted kappa statistic was 0.86 (95% CI 0.70 to 1.00), indicating “almost perfect” intra-assay agreement.

In analysis D2, there were also three discrepancies of only one titre category, with one of these samples also being discrepant in analysis D1: the estimate of the weighted kappa statistic was 0.84 (95% CI 0.69 to 0.99), indicating “almost perfect” intra-assay agreement.

Note that in analysis D1, before accounting for the non-specific immune response (i.e. that to BPV1), there was only one discrepancy between observed titres, compared with two in analysis D2 (table 18.3c). As well as HPV16 and HPV18, samples were also tested in duplicate for BPV1, the results of which are also subject to fluctuation. In analysis D1, all BPV1 results were seronegative, with the exception of four which gave titres of 1-in-40: in each of these cases, the seropositive results were from one replicate only. In analysis D2, only one sample was seropositive for neutralizing antibodies against BPV1, at a titre of 1-in-40. Ignoring the impact of accounting for the non-specific immune response, estimates of the weighted kappa statistic were 0.95 and 0.89 from analyses D1 and D2, respectively.

Table 18.3.1.1. Estimates of weighted kappa statistics measuring intra-assay reliability for the neutralizing antibody response to HPV16 from analyses D1 and D2.

Analysis	Discrepancies[a]	Weighted kappa
D1	3(1)	0.86
D2	3(1)	0.84

[a]Number of discrepancies between replicates, with the difference in terms of the number of titre categories, in brackets.

18.3.1.2 HPV18

Estimates of weighted kappa statistics measuring intra-assay (within-assay) reliability for the neutralizing antibody response to HPV18 from analyses D1 and D2 are presented in table 18.3.1.2.

In analysis D1 there were four discrepancies between replicates in their measured neutralizing response to HPV18, three of only one titre category, and one of two: the estimate of the weighted kappa statistic was 0.82 (95% CI 0.68 to 0.96), indicating “almost perfect” intra-assay agreement.

In analysis D2, there were three discrepancies, all of only one titre category, with two of these samples also being discrepant in analysis D1: the estimate of the weighted kappa statistic was 0.90 (95% CI 0.80 to 0.99), indicating “almost perfect” intra-assay agreement.

Accounting for the non-specific immune response to BPV1 did not change the measured titres against HPV18 in either of the two analyses.

Table 18.3.1.2a. Estimates of weighted kappa statistics measuring intra-assay reliability for the neutralizing antibody response to HPV18 from analyses D1 and D2.

Analysis	Discrepancies[a]	Weighted kappa
D1	3(1), 1(2)	0.82
D2	3(1)	0.90

[a]Number of discrepancies between replicates, with the difference in terms of the number of titre categories, in brackets.

18.3.2 Inter-assay reliability

Measurements of inter-assay reliability were obtained for each of the four ways in which a replicate from analysis D1 could be paired with a replicate from analysis D2; hence these two analyses yielded four correlated estimates of weighted kappa for measuring inter-assay reliability. I have been unable to identify a statistical method which will combine these four estimates into a single estimate.

18.3.2.1 HPV16

Estimates of weighted kappa statistics measuring inter-assay (between-assay) reliability for the neutralizing antibody response to HPV16 from analyses D1 and D2 are presented in table 18.3.2.1.

For HPV16, values of weighted kappa statistics measuring inter-assay reliability indicate reliability which ranges from “substantial” to “almost perfect”. The comparison between replicate 2 from analysis D1 with replicate 2 from analysis D2

included five discrepant results, all of one titre category; replicate 2 from analysis D1 had only two discrepancies of one titre category when compared against replicate 1 from analysis D2.

Table 18.3.2.1a. Estimates of weighted kappa statistics measuring agreement between measurements of the neutralizing antibody response to HPV16, for each possible pair of replicates included in analyses D1 and D2.

	Analysis D2 Replicate 1	Analysis D2 Replicate 2
Analysis D1 Replicate 1	0.85	0.80
Analysis D1 Replicate 2	0.90	0.75

18.3.2.2 HPV18

Estimates of the weighted kappa statistics measuring inter-assay (between-assay) reliability for the neutralizing antibody response to HPV18 from analyses D1 and D2 are presented in table 18.3.2.2.

For HPV18, values of weighted kappa statistics measuring inter-assay reliability indicate reliability which ranges from “substantial” to “almost perfect” according to the Landis scale. The comparison between replicate 1 from analysis D1 and replicate 2 from analysis D2 includes seven discrepant results, six of one titre category, and one of two; replicate 2 from analysis D1 has only three discrepancies of one titre category when compared against replicate 2 from analysis D2.

Table 18.3.2.2. Estimates of weighted kappa statistics measuring agreement between measurements of the neutralizing antibody response to HPV18, for each possible pair of replicates included in analyses D1 and D2.

	Analysis D2 Replicate 1	Analysis D2 Replicate 2
Analysis D1 Replicate 1	0.81	0.72
Analysis D1 Replicate 2	0.86	0.90

Summary

The reproducibility of the HPV16 assay was excellent: values of the intra- and inter-assay weighted kappa statistics were 0.95 and 0.89, respectively. Discrepancies between titres from repeated measurements of the same sample were at most one titre category. The HPV18 assay was clearly more variable, and therefore less reliable, than the HPV16 assay: the estimates of the weighted kappa statistics measuring intra- and inter-assay reliability were, however, both 0.90.

18.3.2.3 Identification of positive- and negative-controls

All of the sera samples supplied from the WHO collaborative study were tested on three occasions, in duplicate. Results from this testing are shown in figure 18.3.2.3. This figure sheds further light on the reproducibility of the assay. In this figure, the darkest boxes were from analysis R1, which will be discussed in section 18.5.2.

On the basis of these results, sample WHO1 was selected for use as the HPV16 positive control when testing patient samples from the cohort described in this thesis. Sample WHO7 was used as the HPV18 positive-control; and sample WHO5 was used as the negative-control for all viral types.

Figure 18.3.2.3. Results of the repeated testing of sera samples supplied from the WHO collaborative study. Cross-hatched boxes indicate sample results changed to 0 (seronegative) after comparison with the BPV1 result. Sample IDs appear in the centre of the figure. Inverse titres are shown, except for 0 which is seronegative.

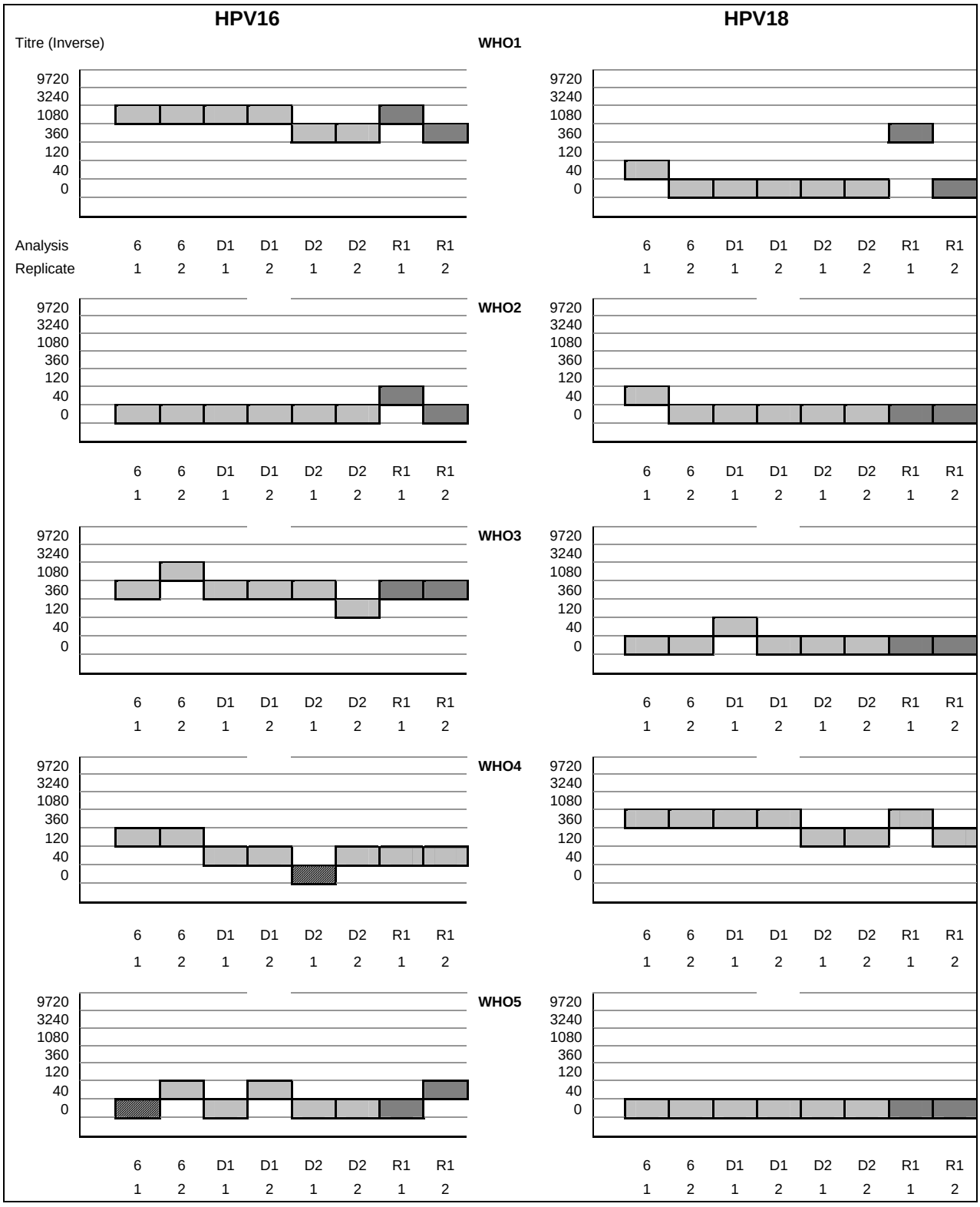
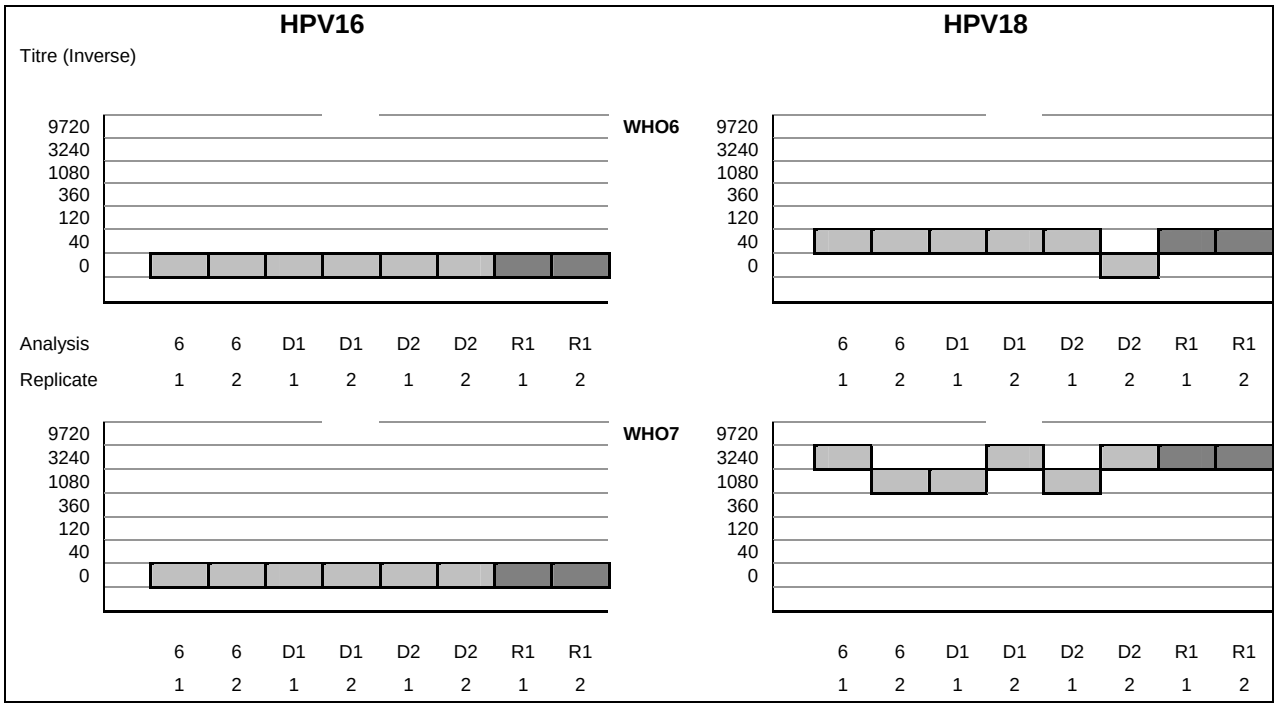


Figure 18.3.2.3 (continued). Results of the repeated testing of sera samples supplied from the WHO collaborative study. Cross-hatched boxes indicate sample results changed to 0 (seronegative) after comparison with the BPV1 result. Sample IDs appear in the centre of the figure. Inverse titres are shown, except for 0 which is seronegative.



Summary

Following twelve developmental and confirmatory analyses, carried out over a period of five months, substantial experience had been gained in the use of the neutralizing antibody assay, an optimum well-plate configuration had been determined, suitable positive- and negative-controls had been identified, and, most importantly, adequate reliability of the assay had been demonstrated. The analysis of samples from the cohort described in this thesis could now proceed with confidence.

18.4 INTENDED STUDY POPULATION FOR THE ANALYSIS OF THE KINETICS OF THE NEUTRALIZING ANTIBODY RESPONSE TO INCIDENT CERVICAL HPV16 AND HPV18 INFECTIONS

The neutralizing antibody assay was next used to describe the kinetics of the neutralizing antibody response to incident cervical HPV16 and HPV18 infections in a cohort of young women who were recruited soon after first sexual intercourse.

All women within the full study cohort (n=2,011) who were HPV DNA-negative and cytologically normal at study entry, who had further follow-up after this time, and who first tested positive in cervical cytological samples for HPV16 DNA, or HPV18 DNA, or both, during follow-up, were identified. For the purposes of identifying the study population, the method used for the detection of HPV DNA was general primer (GP5+/GP6+)-mediated PCR, as described previously (see section 5.5.4.1).

One hundred and fifty five women fulfilled these criteria: 91 women acquired an incident cervical HPV16 infection, 45 acquired an incident cervical HPV18 infection, and 19 acquired both, at some point prior to the end of follow-up, either contemporaneously or otherwise. When appropriate, these three cohorts will be referred to as the “HPV16-only”, “HPV18-only” and “HPV16-and-HPV18” cohorts, respectively. When combined by type, the 110 women who acquired an incident cervical HPV16 infection during follow-up will be referred to as the “HPV16 cohort”; similarly, the 64 women who acquired an incident cervical HPV18 infection during follow-up will be referred to as the “HPV18 cohort”.

18.4.1 Characteristics of the intended study population

Baseline characteristics. The mean age (standard deviation) at study entry of women with an incident cervical HPV16, or HPV18, infection was 17.4 (1.2) and 17.5 (1.2) years, respectively. The median age at first sexual intercourse of the women in these two groups was 16 years, and the median duration of sexual activity before study entry was 19.6 months, and the median number of sexual partners before study entry was two.

Time on study. Time on study (follow-up) was measured from the date of the first evaluable virological sample, with censoring on the earliest of date of: diagnosis of high-grade CIN; treatment; or last smear. The median time on study for women with an incident HPV16 or HPV18 infection was 47 months (range 6 to 105) and 50 months (6 to 107), respectively. The median time interval between study visits was 6.4 months.

Availability of virological samples. The median numbers of virological samples available for HPV DNA testing were 7 (range 2 to 14) and 8 (2 to 14) for women with cervical HPV16 and HPV18 infections, respectively.

18.4.2 The number of sera samples available for testing for each woman according to cohort

The frequency distribution of the number of sera samples available for each woman, according to (GP5+/GP6+) cohort, is presented in table 18.4.2; this table includes multiple sera samples from the same visit.

There were 19 women in the HPV16+HPV18 cohort. The median number of sera samples available for each woman was 4 (range 0 to 11); three women had no sera samples. The total number of sera samples available was 75; one woman had duplicate samples from the same visit, so 74 sera samples were available from distinct occasions.

There were 91 women in the HPV16-only cohort. The median number of sera samples available for each woman was 3 (range 0 to 15); 25 women had no sera samples. The total number of sera samples available was 317; five women had a duplicate sample from the same visit, so 312 sera samples were available from distinct occasions.

There were 45 women in the HPV18-only cohort. The median number of sera samples available for each woman was 3 (range 0 to 14); 13 women had no sera samples. The total number of sera samples available was 172; six women had a duplicate sample from the same visit, so 166 sera samples were available from distinct occasions.

Table 18.4.2. Frequency distribution of the number of sera samples available for each woman in the HPV16-and-HPV18, HPV16-only, and HPV18-only, cohort[a].

Number of sera samples	Cohort: Number of women			Total
	HPV16+HPV18 (n=19)	HPV16-only (n=91)	HPV18-only (n=45)	
0	3	25	13	41
1	3	10	6	19
2	1	9	2	12
3	1	11	3	15
4	3	7	4	14
5	2	4	4	10
6	2	5	3	10
7	2	7	2	11
8	1	5	2	8
9	0	1	1	2
10	0	2	2	4
11	1	3	0	4
13	0	1	2	3
14	0	1	1	2

[a]e.g. three women in the HPV16+HPV18 cohort had four sera samples, compared with seven and four in the HPV16-only and HPV18-only cohorts, respectively.

Summary

All women who were HPV DNA-negative in cervical cytological samples and cytologically normal in cervical smears at study entry, who had further follow-up after this time, and who first tested positive for HPV16 DNA, or HPV18 DNA, or both (using GP5+/GP6+ mediated PCR), during follow-up, were identified: 155 women were identified, 91 of whom acquired an incident cervical HPV16 infection, 45 acquired HPV18 infection, and 19 acquired both. Mean age (standard deviation) at study entry of women with HPV16, or HPV18, infections was 17.4 (1.2) and 17.5 (1.2) years, respectively; median age at first sexual intercourse in both groups was 16 years; the median duration of sexual activity before study entry was 19.6 months; and the median number of sexual partners before study entry was two. The median time on study for women with HPV16 or HPV18 infection was 47 months (range 6 to 105) and 50 months (6 to 107), respectively. The median interval between study visits was 6.4 months. The median numbers of virological samples available for HPV DNA testing were 7 (range 2 to 14) and 8 (2 to 14) for women with cervical HPV16 and HPV18 infections, respectively.

18.5 TESTING OF SERA SAMPLES

18.5.1 Configuration of well-plates when conducting the neutralizing antibody assay

The configuration of the well-plates used to test study samples is shown in figure 18.5.1. The configuration of the well-plates was identical for each of the virus types tested for, except with respect to the positive control wells.

Twenty four samples were analysed in each assay, with one assay consisting of a total of nine well-plates, i.e. three well-plates used to test samples for their neutralizing antibody response against each of HPV16, HPV18, and BPV1. Each well-plate was used to test eight study samples. For example, figure 18.5.1 shows the configuration of the well-plate used to test samples 1 to 8.

Two wells containing identical concentrations of target cells and PsV were used for testing a positive-control: on the HPV16 and HPV18 well-plates, one of these two wells had a serum dilution of 1-in-360, and the other 1-in-1080. These dilutions were used since both positive-controls, i.e. the HPV16 and the HPV18 positive-controls, were found to have 50% neutralizing antibody titres of at least 1-in-360, and so should always have given a seropositive result when the assay was working correctly. Clearly, the positive control used on the HPV16 well-plate was different from that used on the HPV18 well-plate. The BPV1 well-plate also had two “positive-control” wells, but both at a dilution of 1-in-360: one used to test the HPV16 positive-

control, the other the HPV18 positive-control. Two wells were used for testing a negative-control, both at a serum dilution of 1-in-40.

Note: The controls used on the BPV1 well-plates were effectively redundant, since no controls for BPV1 were available, and since the number of wells available on a well-plate for measuring the signals from controls was severely limited. This meant that, in practice, the results from the controls on the BPV1 well-plate were disregarded; and, in contrast to "samples wells", where a result was only positive if it had a non-zero titre at least three times greater than the BPV1 titre (see section 17.3.2), the results from the "control wells" on the HPV16 and HPV18 well-plates were interpreted at face value, without a supporting BPV1 result.

Figure 18.5.1. The well-plate configuration for each of the well-plates used in the analysis to determine the reliability of the neutralizing antibody assay. Sn is sample number n, n=1 to 8. White cells correspond to empty wells on the well-plate. Blue cells are wells used to measure the maximum signal. Yellow cells are used to measure well-plate background. Green and red cells measure positive- and negative-control signals, respectively. Interior cells correspond to wells used to test study samples, with all wells in the same column being used to test the same sample at different dilutions: dilutions start at 1-in-40 in row 2, and proceed in multiples of three until 1-in-9720 is reached in row 7.

	S1 1-in-40	S2 1-in-40	S3 1-in-40	S4 1-in-40	S5 1-in-40	S6 1-in-40	S7 1-in-40	S8 1-in-40	Positive 1-in-360	Cells + PsV	
	S1 1-in-120	S2 1-in-120	S3 1-in-120	S4 1-in-120	S5 1-in-120	S6 1-in-120	S7 1-in-120	S8 1-in-120	Positive 1-in-1080	Cells + PsV	
	S1 1-in-360	S2 1-in-360	S3 1-in-360	S4 1-in-360	S5 1-in-360	S6 1-in-360	S7 1-in-360	S8 1-in-360	Negative 1-in-40	Cells + PsV	
	S1 1-in-1080	S2 1-in-1080	S3 1-in-1080	S4 1-in-1080	S5 1-in-1080	S6 1-in-1080	S7 1-in-1080	S8 1-in-1080	Negative 1-in-40	Cells + PsV	
	S1 1-in-3240	S2 1-in-3240	S3 1-in-3240	S4 1-in-3240	S5 1-in-3240	S6 1-in-3240	S7 1-in-3240	S8 1-in-3240	Cells alone	Cells + PsV	
	S1 1-in-9720	S2 1-in-9720	S3 1-in-9720	S4 1-in-9720	S5 1-in-9720	S6 1-in-9720	S7 1-in-9720	S8 1-in-9720	Cells alone	Cells alone	

18.5.2 Difficulties encountered using the neutralizing antibody assay

Each well-plate in the configuration shown in figure 18.5.1 could be used to test a total of eight study sera samples. In filling the available spaces on the well-plates, each of the three subcohorts (i.e. the HPV16+HPV18, HPV16-only, and HPV18-only, cohorts) was treated separately. Table 18.5.2 describes the number of analyses which would have been required to analyse all available sera samples in the relevant study cohorts. A total of 216 well-plates in 72 analyses would have been required. At a rate of two analyses a week, with no unforeseen problems, this would have taken eight or nine months.

Table 18.5.2. The number of neutralizing antibody analyses required to analyse all available sera samples in the HPV16-only and HPV18-only subcohorts. Each potential analysis consists of nine sets of well-plates.

Subcohort	Number of sera samples	Number of analyses required	Number of well-plates required
HPV16-only	317	40	120
HPV18-only	172	22	66
HPV16+HPV18	75	10	30
Total	564	72	216

Women were prioritised according to the order in which they would have their sera samples tested. These priorities changed during the course of testing due to various external influences (the availability of HPV viral load measurements, conference deadlines, etc.). Initially, the intention was to test every serum sample from every woman within the three relevant subcohorts; therefore, the changing priorities were not perceived as presenting any threat to the study. However, ultimately, testing was discontinued before all women had been tested.

Difficulties arose with the assay during the testing of sera samples from the cohort described in this thesis; and conducting two analyses per week proved unsustainable, due to the strain it placed on the (single) laboratory scientist conducting the assay. Ultimately, this led to the decision to cease testing any further study samples using this assay.

According to the original developers of this neutralizing antibody assay, PsV, once manufactured, could be snap-frozen at -80^{o}C and stored for a period of two years: this was not confirmed in our laboratory. Within a year of manufacture, the “activity”

of the BPV1 PsV was found to have deteriorated to a point at which they could no longer be relied upon to work correctly in the assay. The records of the refrigerator used to store the stock of BPV1 PsV did not reveal any concerns: it was speculated that the BPV1 PsV had become denatured to some extent, at least within subpopulations of the PsV.

It was predicted that the BPV1 PsV could still be used, but at a higher dilution than had been used previously. Initially, the dilution used had been 1-in-800; now, however, it was proposed to use a dilution of 1-in-25,600. The use of BPV1 PsV at higher dilutions was validated by testing previously tested samples, measuring intra- and inter-assay reliability, and confirming that results were consistent with those obtained previously. The opportunity was also taken to identify a new negative-control, since the existing negative-control (sample WHO2) was nearing exhaustion. This approach failed. It was therefore necessary to manufacture *new* BPV1 PsV, determine the amount to use in each assay, confirm the reproducibility of the assays using these new BPV1 PsV, and confirm that results were consistent with those obtained previously (analysis R1). Subsequently, at different stages during the period when study samples were being tested, the HPV16 and HPV18 PsV were also found to have deteriorated to a point where they could no longer be relied upon to work correctly in the assay, and so the same time-consuming procedures were required to re-manufacture and re-evaluate the new HPV16 and HPV18 PsV.

The original laboratory kits used in these assays, which cost £700 each, were manufactured by BD Biosciences: one kit contained enough reagent to analyse

eighteen 96-well well-plates, with no leeway for error. One kit could therefore be used to test 48 study samples, assuming that eight study samples were tested on each well-plate. This meant that one kit was exhausted every week initially, and every two weeks subsequently. During the period when samples from the cohort described in this thesis were being tested, BD Biosciences was bought by Clontech, and the original kits were no longer available. The “new” kit offered by Clontech, which cost £470, was sufficient for conducting 300 reactions (one reaction=one well). That being said, the catalogue number for ordering these kits was the same as it was previously, and the catalogue description for the old kit also seemed to suggest that it was only sufficient for conducting 300 reactions. So, on the face of it, it appeared that the new and original kits were equivalent, if now more expensive. However, it was necessary to confirm this. The new kit was evaluated in the same way as the original kit: intra- and inter-assay reliability were measured by repeating analyses D1 and D2 from the developmental stage with the assay, and comparing the new results with those taken previously.

18.5.3 Study population tested

Due to competing priorities within the study, the decision was made to stop testing samples when equal numbers of women from each of the three relevant study subcohorts (the HPV16-only, the HPV18-only and the HPV16+HPV18 cohorts) had been tested: the intention was that testing would resume at some point in the future. Since all women with sera samples available in the HPV16-and-HPV18 cohort had been tested when this decision was taken, this determined the total number of

women from each of the remaining two subcohorts which would be tested. Subsequently, it was decided that sufficient data was available to draw conclusions about the utility of the neutralizing antibody assay, and to provide a description of the kinetics of the immune response following acquisition of the relevant type-specific cervical HPV infections. However, as a result of the cumulative difficulties encountered with this expensive, and difficult to perform assay, no further testing was undertaken, and none is planned.

The final population tested for their neutralizing antibody response to cervical HPV16 and HPV18 infections comprised 48 women. Table 18.5.3 describes the relationship between the study population it was the *intention* to test and the population *actually* tested. Forty eight (42%) women with sera samples available were tested for their neutralizing antibody response: all women in the HPV16+HPV18 cohort who could have been tested *were* tested, compared with just over half of the women in the HPV18-only cohort, and slightly more than a fifth of women in the HPV16-only cohort.

Table 18.5.3. Relationship between the study population it was the intention to test and the population actually tested for their neutralizing antibody response to incident cervical HPV16 or HPV18 infections, or both. Frequencies are numbers of women.

Population	Subcohort			Total
	HPV16-only	HPV18-only	HPV16-and-HPV18	
Intended study population	91	45	19	155
Number with serum samples available	66	32	16	114
Population actually tested	15	17	16	48

18.5.3.1 Sera samples available

The availability of sera samples from each study visit for each woman in the study population are illustrated in table 18.5.3.1. As an example, the woman with study number (ID) 350 belongs to the HPV16+HPV18 cohort, and HPV18 was first detected at an earlier visit than HPV16 (hence 18-16). She made 14 visits to the study: the maximum number of visits made by any woman was 16, hence this woman has her last two (16-14) cells “blacked-out”. For five of her 14 visits, at least one serum sample was available for testing; for the remaining nine visits, no sera samples were available. For her fourth visit, there were two sera samples available. For her tenth visit, there was one serum sample available, but this was tested twice for the neutralizing antibody response to HPV16.

The median number of study visits was 9 (range 2 to 16). The median number of study visits for which a serum sample was available was 5 (range 1 to 14), with nine women providing a serum sample on only a single occasion. The median number of study visits for which no serum sample was available was 3 (range 0 to 11), with only five women having a “complete” set of sera samples, i.e. a serum sample taken at every possible study visit. However, this latter result should be interpreted as missed opportunities at measurement, rather than as missing samples *per se*.

Table 18.5.3.1. The availability of sera samples from each study visit for each woman in the population tested for their neutralizing antibody response to HPV16 and HPV18. ID=study number; C=Cohort (the order in which HPV16 and HPV18 were detected e.g. 16-18 means 16 followed by 18, whereas 16+18 means both types were first detected at the same time); V=Number of study visits; S=Number of visits for which a serum sample is available; M=number of visits with no serum sample. A white cell indicates a visit for which at least one serum sample was available; a grey cell indicates a visit for which no serum sample was available; a black cell indicates that the woman had reached the end of follow-up; 2S=Two sera samples available for this visit; T2=one serum sample available which was tested twice.

ID	C	V	S	M	Visit number															
					1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
2	16-18	10	5	5																
27	18	14	9	5			2S													
29	18	12	10	2																
43	18	10	7	3																
69	16	16	7	9																
92	16-18	5	5	0																
94	18	2	1	1																
118	16-18	9	3	6																
128	18-16	16	11	5																
147	16-18	11	1	10																
150	16-18	12	9	3																
189	16	9	8	1																
240	16-18	7	7	0																
244	16	9	7	2																
247	16+18	9	7	2																
292	18	16	14	2					T2	T2										
330	16	5	3	2																
350	18-16	15	5	10				2S						T2						
360	16	4	1	3																
379	16-18	13	5	8																
381	18	11	5	6																
393	16+18	7	4	3																
446	18	8	7	1																
490	18	9	8	1																
515	16	6	6	0			2S													
524	18	8	1	7																
561	18	3	3	0																
590	16	8	7	1																
692	18	8	3	5																
744	18	8	6	2																
761	16	2	2	0																
938	18	13	12	1											2S					
941	16	8	6	2																
997	18-16	9	4	5																
1011	16	11	7	4								2S								
1016	18	8	2	6																
1268	16-18	9	1	8																
1278	18	6	5	1	2S															
1347	16	2	1	1																
1365	16	4	1	3																
1367	16	12	8	4																
1413	16	7	1	6																
1430	18	9	8	1																
1540	18-16	9	4	5																
1551	16+18	12	8	4																
1604	16	12	1	11																
1803	16+18	5	2	3																
1814	18	12	6	6																

18.5.3.2 Practical issues arising during the testing of the final study population

The analysis of all sera samples from the women comprising the final population tested required 11 analyses, excluding those discarded due to problems with the assay, as described previously (section 18.5.2.1). Each well-plate could be used to test a total of eight samples from the relevant study subcohorts. In allocating samples to the available spaces on the well-plates, each of the three relevant subcohorts (the HPV16-only, HPV18-only and the HPV16-and-HPV18 cohorts) were treated separately, as far as possible. Despite the relatively large value of the weighted-kappa statistic measuring inter-assay reliability, the first criteria for allocating samples to well-plates was to test all samples from the same woman on the same well-plate, where possible. However, five women had samples tested on two different well-plates.

A total of 264 sera samples from 254 study visits made by 48 women were tested; six women had two sera samples from the same visit for one of their visits; and for three samples, the opportunity was taken to repeat the testing of the same sample. For one sample (ID 92.8), there was no corresponding visit: this sample was mislabelled, and was excluded from all analyses.

The status of the controls on the well plates, the requirement for a BPV1 titre against which to judge a type-specific titre, and the fact that testing was discontinued before all women had been tested, all had an impact on the size of the final study population tested.

All well-plates were read automatically, with numerical signal measurements for each well-plate being output to a computer file. Unfortunately, for one analysis, the same output file was inadvertently saved twice under different file names. This meant that the BPV1 titre for the samples tested on those well-plates were unknown; given that titres against BPV1 of 1-in-120 and 1-in-360 had been observed in women from the cohort described in this thesis, there was no choice but to exclude the two women involved (study IDs 69 and 150), from *all* analyses.

For all well-plates, testing for either HPV16 or HPV18, the positive-controls were always seropositive at a titre of at least 1-in-360, as required. All well-plates were therefore acceptable on the basis of the positive-control result.

The negative-controls were interpreted together: a well-plate was considered to have a seronegative result for the negative-control if at least one of the negative-control wells indicated a seronegative result. For well-plate 7-2-190606, both negative control wells in the HPV16 analysis indicated a seropositive result; for well-plates 7-1-190606 and 10-1-040806, both negative-control wells in the HPV18 analysis indicated a seropositive result. These well-plates were considered to have “failed”, and the results for the study samples on these well-plates were therefore rejected. The intention was to complete testing of all eligible study samples once, before repeating the testing of samples whose first result had been rejected due to failure of the well-plates on which they were tested. However, as stated previously, no further testing was undertaken. Therefore, one woman (ID 490) was excluded from all

analyses of HPV16, but not of HPV18; and three women (ID's 29, 244, and 360) were excluded from all analyses of HPV18, but not of HPV16.

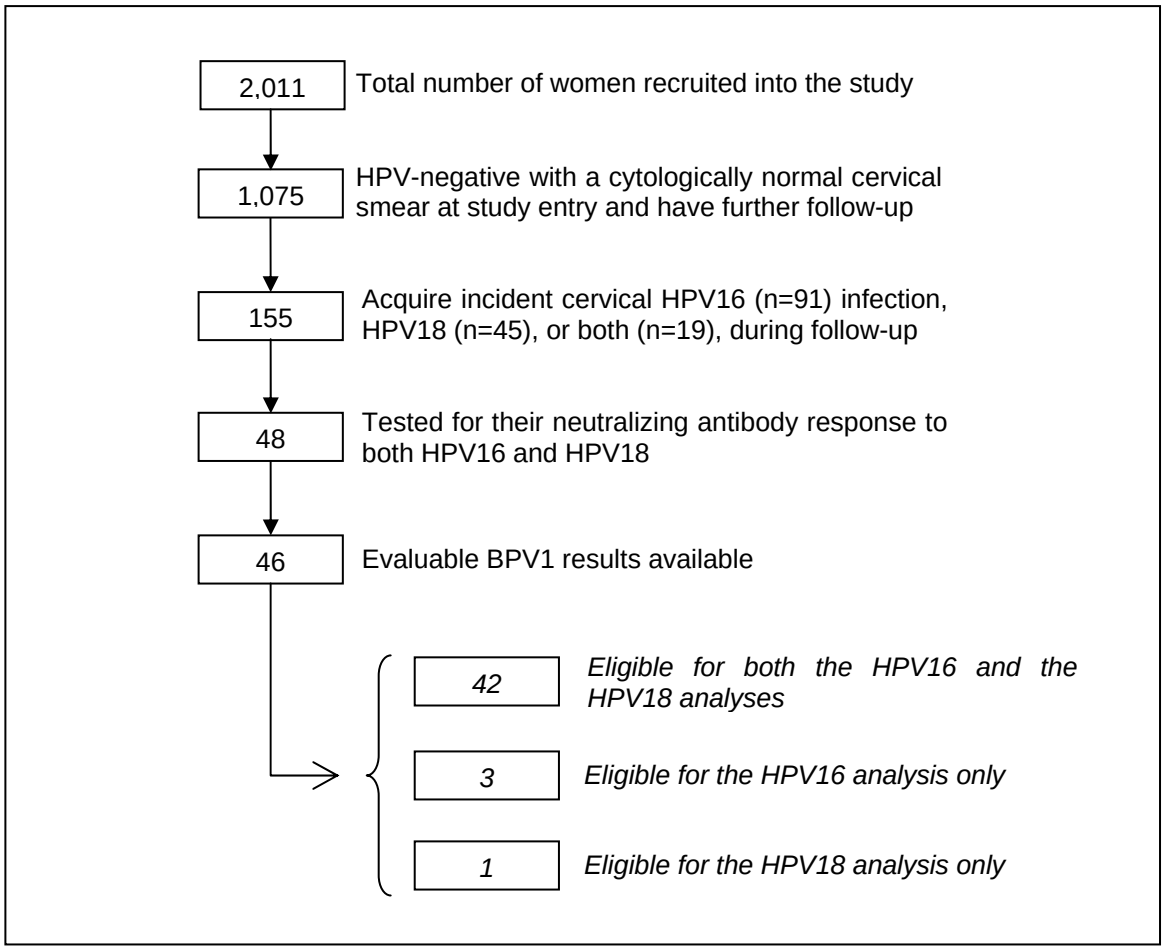
Summary

Difficulties arose with the assay, so testing stopped before all eligible women had been tested. However, sufficient data was available to draw conclusions about the utility of the assay, and to describe the kinetics of the immune response following the acquisition of the relevant type-specific cervical HPV infections. The final population tested comprised 48 women out of a possible 114 who had sera samples available. The median number of study visits was 9 (range 2 to 16); the median number of visits for which a serum sample was available was 5 (range 1 to 14). The analysis of all sera samples from the women comprising the final population tested required 11 analyses in which 264 samples from 254 visits were tested.

18.6 FINAL STUDY POPULATION FOR THE ANALYSIS OF THE KINETICS OF THE NEUTRALIZING ANTIBODY RESPONSE TO INCIDENT CERVICAL HPV16 AND HPV18 INFECTIONS

The final study population for the analysis of the kinetics of the neutralizing antibody response to incident cervical HPV16 and HPV18 infections comprised 45 women for the analysis of HPV16, and 43 for the analysis of HPV18 (42 women were common to both groups). The derivation of the final study population from the full cohort of 2,011 women is shown in figure 18.6. For ease of reference, in this subsection, the 45 (42+3) women will be referred to as the "HPV16 study population", and the 43 (42+1) women as the "HPV18 study population".

Figure 18.6. Derivation of the final study population from the cohort of all women recruited to the study.



18.6.1 Characteristics of the final study population

All women included in the final study populations for the analysis of the kinetics of the neutralizing antibody response to incident cervical HPV16 and HPV18 infections were sexually experienced to some extent. The median interval between the date of first sexual intercourse and the date the first serum sample was taken was 542 days for women in the HPV16 study population, and 503 days for women in the HPV18

study population; the median number of sexual partners prior to the date of the first available serum sample was 2 (range 1 to 9) for both study populations.

For both the HPV16 and HPV18 study populations: the median number of study visits was 9 (range 2 to 16); the median number of study visits for which at least one serum sample was available was 5 (range 1 to 14); the median number of study visits for which no serum sample was available was 3 (range 0 to 11), with only five women having a “complete” set of sera samples, i.e. a serum sample for every possible study visit. However, this latter result should be interpreted as missed opportunities at measurement, rather than as missing samples *per se*.

In the HPV16 population, 36 women provided at least two sera samples: the median time interval between the date of the first and last sera samples was 1,341 days. In the HPV18 population, 35 women provided at least two sera samples: the median time interval between the dates of the first and last sera samples was 1,329 days. In both study populations, for all women the first serum sample available was that taken at study entry; for 14 women, the last serum sample available was that taken at the last visit.

Summary

The final study population for the analysis of the kinetics of the neutralizing antibody response to incident cervical HPV16 and HPV18 infections comprised 45 women for the analysis of HPV16, and 43 for the analysis of HPV18 (with 42 common to both groups). The median interval between first sexual intercourse and the first serum sample was taken was 542 days for women in the HPV16 study population, and 503 days for women in the HPV18 study population; the median number of sexual partners prior to the date of the first available serum sample was 2 (range 1 to 9) for both study populations. For both study populations: the median number of study visits was 9 (2 to 16); the median number of study visits for which at least one serum sample was available was 5 (1 to 14). In the HPV16 population, 36 women provided at least two sera samples: the median interval between the dates of the first and last sera samples was 1,341 days. In the HPV18 population, 35 women provided at least two sera samples: the median interval between the dates of the first and last sera samples was 1,329 days.

18.7 TYPE-SPECIFIC DISTRIBUTION OF NEUTRALIZING ANTIBODY TITRES

The result of this assay is the 50% neutralization titre, the lowest dilution which causes at least a 50% reduction in SEAP expression (see section 17.3.2).

18.7.1 Cross-sectional analysis of the neutralizing antibody response to HPV16

Forty-five women were tested for their neutralizing antibody response to HPV16. The observed 50% neutralization titre against HPV16 for all sera samples from every woman in the HPV16 study population are shown in table 18.7.1a. The frequency distribution across all visits of the observed 50% neutralization titres is shown in table 18.7.1b. In these women with *natural* infections (as opposed to infections seen in vaccinated women), all possible titres were observed at least once. Three women had a maximum titre of 1-in-9720, including one woman (ID 938) who appears to

have maintained this titre for at least three visits, corresponding to a period of 371 days.

Table 18.7.1a. The frequency distribution of the observed 50% neutralizing antibody titre against HPV16 across all visits for all women. Where multiple sera samples were available, or the same serum sample was tested more than once, the greatest titre observed was taken as the result of testing that sample.

Titre	Number of visits
Seronegative	156
1-in-40	12
1-in-120	17
1-in-360	15
1-in-1080	12
1-in-3240	14
1-in-9720	4
Total	230

Table 18.7.1b. The observed 50% neutralizing antibody titre (inverse) against HPV16 of sera samples from each study visit for each woman in the study population. ID=study number; C=Cohort (the order in which HPV16 and HPV18 were detected e.g. 16-18 means 16 followed by 18, whereas 16+18 means both types were first detected at the same time). A white cell indicates a visit for which a serum sample was available; a grey cell indicates a visit for which no serum sample was available; a black cell indicates that the woman had reached the end of follow-up. A cell which is split indicates where two sera samples were available for that visit, or where one serum sample was available which was tested twice. A black outline to a cell indicates that the cytological sample taken at that visit was positive for HPV16 DNA according to the GP5+/GP6+ system. A titre of zero indicates a seronegative result.

ID	C	Visit															
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*
2	16-18	0		0			0		0	40							
27	18	0	0	0 0	0		0	0				0	0	0			
29	18	120	0	0	0	0	0	40	0	0		0					
43	18	0	0	0		120		0	0		0						
92	16-18	3240	3240	3240	1080	3240											
94	18	0															
118	16-18	0		0				0									
128	18-16	0	0	0	0	0		0		0	0	0		0	0		
147	16-18		0														
189	16	0	0	0	0	0	0		0	0							
240	16-18	0	0	3240	3240	1080	3240	9720									
244	16	0	0	0	0	0	0		0								
247	16+18	40	0	0		0	120	0	120								
292	18	0	0	0	0	0 0	0 0	0	0	0	0	0		0	0		0
330	16	0	120	360													
350	18-16	0		0	120 360				120		40 120						
360	16			0													
379	16-18	0	0	360		360	360										
381	18	40	0	0		40	0										
393	16+18	0	0	0	40												
446	18	0	0	40	0	0		0	120								
515	16	360	360	0 360	120	120	120										
524	18						0										
561	18	0	9720	1080													
590	16	0	1080	360	1080	360	1080	360									
692	18	0	0			0											
744	18	0			0	0	0	0	0								
761	16	0	0														
938	18	0	0	360	1080	3240	3240	9720		9720	3240	3240	3240	3240	3240		
941	16	0		0	0	0	0	0									
997	18-16	0				0		0	0								
1011	16	0	360	1080	3240	1080			1080 1080	1080							
1016	18	0				0											
1268	16-18	40															
1278	18	0 0	0	0	0	0	0										
1347	16	0															
1365	16	0															
1367	16	0	0	0		40			0	0	0	0					
1413	16	0															
1430	18	0	120		0	0	0	0	40	0							
1540	18-16	0	0			360	1080										
1551	16+18	0	0				120	120	120	40	40		120				
1604	16								360								
1803	16+18	0				0											
1814	18	0	0				0	0	0		0						

The frequency distribution of the number of visits for which the cytological sample taken at that visit was positive for HPV16 DNA, and the number of HPV16 DNA-positive samples for which there was a corresponding serum sample is shown in table 18.7.1c (note also that such visits are distinguished by a black outline to the cell in table 18.7.1b). Twenty nine women had one or more HPV16 DNA-positive samples during follow-up; just over half had only one such sample, but one woman had six. Twenty five of these women had 37 sera samples taken at the same time as a HPV16 DNA-positive sample; 25 visits for which the cervical cytological sample was HPV16 DNA-positive had no corresponding serum sample.

Table 18.7.1c. Frequency distribution of the number of visits for which the cervical cytological sample taken at that visit was positive for HPV16 DNA, and the number of HPV16-positive samples for which there was a corresponding serum sample. Example: five women had a total of three HPV16 DNA-positive samples during follow-up; however, for one woman there were no corresponding sera samples from two of those visits, one woman had serum samples corresponding to two of those positive samples but not the third, and the remaining three women had sera samples for all three visits.

Number of cervical cytological samples positive for HPV16 DNA	**Number of women (n=45)**	**Number of HPV16 DNA-positive cervical cytological samples for which there was a corresponding serum sample**						
		0	*1*	*2*	*3*	*4*	*5*	*6*
0	16	16	–	–	–	–	–	–
1	15	6	9	–	–	–	–	–
2	3	1	0	2	–	–	–	–
3	5	0	1	1	3	–	–	–
4	4	2	1	0	1	0	–	–
5	1	0	0	0	0	1	0	–
6	1	0	0	0	0	1	0	0

The frequency distribution of the 50% neutralization titres of the sera samples taken at visits for which the corresponding cervical cytological sample was positive for HPV16 DNA is shown in table 18.7.1d. Most sera samples were seronegative for

neutralizing antibodies. The frequency distribution of the 50% neutralization titres of the sera samples taken at every visit, stratified by the HPV16 DNA status, and the study cohort to which the woman belonged, is shown in table 18.7.1e.

Table 18.7.1d. Frequency distribution of the 50% neutralization titres of the sera samples taken at visits for which the corresponding cervical cytological sample was positive for HPV16 DNA.
Where multiple sera samples were available, or the same serum sample was tested more than once, the greatest titre observed was taken as the result of testing that sample.

Titre	Number of sera samples
Seronegative	21
1-in-40	1
1-in-120	3
1-in-360	6
1-in-1080	2
1-in-3240	4
1-in-9720	0
Total	37

18.7.2 Cross-sectional analysis of the neutralizing antibody response to HPV18

The observed 50% neutralization titre against HPV18 of all sera samples from every woman in the study population are shown in table 18.7.2a. The frequency distribution across all visits of the observed titres is shown in table 18.7.2b. In these women with *natural* infections, all possible titres were observed at least once. Two women had the maximum possible titre of 1-in-9720, on one occasion.

Table 18.7.1e. Frequency distribution of the 50% neutralizing antibody titres of the sera samples taken at every visit, stratified by the HPV16 DNA status of the cervical cytological sample and the study cohort to which the woman belongs.

Titre	HPV16 DNA-positive samples				HPV16 DNA-negative samples				All samples
	HPV16-only cohort	HPV16+ HPV18	Total		HPV16-only cohort	HPV18-only cohort	HPV16+ HPV18	Total	
Seronegative	12 (67%)	9 (47%)	21 (57%)		24 (62%)	78 (79%)	33 (60%)	135 (70%)	156 (68%)
1-in-40	0 (0%)	1 (5%)	1 (3%)		1 (3%)	5 (5%)	5 (9%)	11 (6%)	12 (5%)
1-in-120	2 (11%)	1 (5%)	3 (8%)		2 (5%)	4 (4%)	8 (15%)	14 (7%)	17 (7%)
1-in-360	2 (11%)	4 (21%)	6 (16%)		6 (15%)	1 (1%)	2 (4%)	9 (5%)	15 (7%)
1-in-1080	1 (6%)	1 (5%)	2 (5%)		6 (15%)	2 (2%)	2 (4%)	10 (5%)	12 (5%)
1-in-3240	1 (6%)	3 (16%)	4 (11%)		0 (0%)	6 (6%)	4 (7%)	10 (5%)	14 (6%)
1-in-9720	0 (0%)	0 (0%)	0 (0%)		0 (0%)	3 (3%)	1 (2%)	4 (2%)	4 (2%)
Total	18	19	37		39	99	55	193	230

Table 18.7.2a. Frequency distribution of the observed 50% neutralizing antibody titres against HPV18 across all visits for all women. Where multiple sera samples were available, or the same serum sample was tested more than once, the greatest titre observed was taken as the result of testing that sample.

Titre	Number of visits
Seronegative	170
1-in-40	9
1-in-120	6
1-in-360	6
1-in-1080	12
1-in-3240	15
1-in-9720	2
Total	220

Table 18.7.2b. The observed 50% neutralizing antibody titre (inverse) against HPV18 of sera samples from each study visit for each woman in the study population. ID=study number; C=Cohort (the order in which HPV16 and HPV18 were detected e.g. 16-18 means 16 followed by 18, whereas 16+18 means both types were first detected at the same time). A white cell indicates a visit for which a serum sample was available; a grey cell indicates a visit for which no serum sample was available; a black cell indicates that the woman had reached the end of follow-up. A cell which is split indicates where two sera samples were available for that visit, or where one serum sample was available which was tested twice. A black outline to a cell indicates that the cytological sample taken at that visit was positive for HPV18 DNA according to the GP5+/GP6+ system. A titre of zero indicates a seronegative result.

ID	C	Visit															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
2	16-18	0		0			0		0	0							
27	18	0	0	0 0	0		0	0				0	0	0			
43	18	1080	360	40		120		120	120		360						
92	16-18	0	0	0	0	0											
94	18	0															
118	16-18	0		0				0									
128	18-16	0	0	0	120	0		0		0	0	0		0	40		
147	16-18		0														
189	16	0	0	0	0	0	0		0	0							
240	16-18	40	0	9720	1080	1080	1080	360									
247	16+18	0	0	0		0	0	0	0								
292	18	0	0	3240	1080	3240 3240	3240 1080	3240	3240	3240	1080	3240		1080	3240		3240
330	16	0	0	0													
350	18-16	0		0	0 0				0		0 0						
379	16-18	0	0	0		0	360										
381	18	0	0	0		0	40										
393	16+18	0	0	0	0												
446	18	0	0	40	3240	1080		1080	360								
490	18	0	0	0	0	0	0	0	1080								
515	16	0	0	1080 0	0	0	0										
524	18						0										
561	18	0	0	0													
590	16	0	0	0	0	0	0	0									
692	18	0	0			0											
744	18	0			3240	9720	3240	3240	3240								
761	16	0	0														
938	18	0	0	0	0	0	0	0		0	0	120 0	0	0			
941	16	0		0	0	0	0	0									
997	18-16	40				0		0	0								
1011	16	0	0	0	0	0			0 0	0							
1016	18	0				0											
1268	16-18	0															
1278	18	40 0	0	0	0	1080											
1347	16	0															
1365	16	0															
1367	16	40	0	0		0			0	0	0	0					
1413	16	120															
1430	18	0	0		0	0	0	0	40	0							
1540	18-16	0	0			3240	360										
1551	16+18	0	0				0	0	0	0	0		0				
1604	16								0								
1803	16+18	0				0											
1814	18	0	0				0	0	0		0						

The frequency distribution of the number of visits for which the cervical cytological sample taken at that visit was positive for HPV18 DNA, and the number of HPV18 DNA-positive samples for which there was a corresponding serum sample, is shown in table 18.7.2c (note also that such visits are distinguished by a black outline to the cell in table 18.7.2b). Thirty-one women had one or more HPV18 DNA-positive samples during follow-up; the vast majority had only one such sample, but one woman had five. Twenty three of these women had 33 sera samples taken at the same time as a HPV18 DNA-positive sample; 20 visits for which the cervical cytological sample was HPV18 DNA-positive had no corresponding serum sample.

Table 18.7.2c. Frequency distribution of the number of visits for which the cervical cytological sample taken at that visit was positive for HPV18 DNA, and the number of HPV18-positive samples for which there was a corresponding serum sample. Example: 26 women had only one HPV18 DNA-positive sample during follow-up; however, for eight women there were no corresponding sera samples from that visit, but the remaining 18 did have a serum sample from that visit.

Number of samples positive for HPV18	**Number of women (n=43)**	**Number of HPV18-positive samples for which there was a corresponding serum sample**					
		0	*1*	*2*	*3*	*4*	*5*
0	12	12	–	–	–	–	–
1	26	8	18	–	–	–	–
2	1	0	0	1	–	–	–
3	2	0	0	0	2	–	–
4	1	0	0	1	0	0	–
5	1	0	0	0	0	0	1

The frequency distribution of the 50% neutralizing antibody titres of the sera samples taken at visits for which the corresponding cervical cytological sample was positive for HPV18 DNA, is shown in table 18.7.2d. Most sera samples were seronegative for neutralizing antibodies. The frequency distribution of the 50% neutralizing antibody

titres of the sera samples taken at every visit, stratified by the HPV18 DNA status and the study cohort to which the woman belonged, is shown in table 18.7.2e.

Table 18.7.2d. The frequency distribution of the 50% neutralization titres of the serum samples taken at visits for which the corresponding cervical cytological sample was positive for HPV18 DNA. Where multiple sera samples were available, or the same serum sample was tested more than once, the greatest titre observed was used as the result of testing that sample.

.

Titre	Number of samples
Seronegative	22
1-in-40	1
1-in-120	1
1-in-360	1
1-in-1080	0
1-in-3240	6
1-in-9720	2
Total	33

Summary

In the 45 and 43 women who were tested for their neutralizing antibody response to HPV16, and HPV18, respectively, all possible titres were observed at least once. Twenty nine women had one or more HPV16 DNA-positive samples during follow-up; 25 of these had 37 sera samples taken at the same time as a HPV16 DNA-positive sample. Thirty-one women had one or more HPV18 DNA-positive samples during follow-up; 23 of these had 33 sera samples taken at the same time as a HPV18 DNA-positive sample. For both HPV types, most sera samples were seronegative for neutralizing antibodies.

Table 18.7.2e. Frequency distribution of the 50% neutralizing antibody titres of the sera samples taken at every visit, stratified by the HPV18 DNA status of the cervical cytological sample and the study cohort to which the woman belongs.

Titre	HPV18 DNA-positive samples				HPV18 DNA-negative samples				All samples
	HPV18-only cohort	HPV16+ HPV18	Total		HPV16-only cohort	HPV18-only cohort	HPV16+ HPV18	Total	
Seronegative	13 (59%)	9 (82%)	22 (67%)		48 (94%)	49 (65%)	51 (84%)	148 (79%)	170 (77%)
1-in-40	1 (5%)	0 (0%)	1 (3%)		1 (2%)	4 (5%)	3 (5%)	8 (4%)	9 (4%)
1-in-120	1 (5%)	0 (0%)	1 (3%)		1 (2%)	3 (4%)	1 (2%)	5 (3%)	6 (3%)
1-in-360	0 (0%)	1 (9%)	1 (3%)		0 (0%)	3 (4%)	2 (3%)	5 (3%)	6 (3%)
1-in-1080	0 (0%)	0 (0%)	0 (0%)		1 (2%)	8 (11%)	3 (5%)	12 (6%)	12 (5%)
1-in-3240	6 (27%)	0 (0%)	6 (18%)		0 (0%)	8 (11%)	1 (2%)	9 (5%)	15 (7%)
1-in-9720	1 (5%)	1 (9%)	2 (6%)		0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (1%)
Total	22	11	33		51	75	61	187	220

18.8 TIME TO SEROCONVERSION TO HPV16 AND HPV18

Changes in the neutralizing antibody response to HPV16 and HPV18 over time are presented in figure 18.8 for all women tested for their neutralizing antibody response to HPV16 and/or HPV18. For ease of comparison, all figures are plotted on the same horizontal (time) and vertical scales (inverse of the 50% neutralizing antibody titre). In each case, time is measured from the date of the first evaluable smear. For aesthetic reasons, if a sample from a visit was not available, not tested, or not evaluable, the previously measured value of neutralizing antibody response was plotted (observed values are plotted with a closed plotting symbol, backwards-interpolated values are plotted with an open symbol) to form a left-continuous curve. Unique identifiers for women (study numbers) and the HPV type being analysed in the figure appear in the top right-hand corner of each of the charts. Just above the time axis, acquisition of a GP5+/GP6+ positive sample of the relevant HPV type is indicated by a red triangle (note that all women who were tested for their neutralizing antibody response to cervical HPV infection, were tested for both HPV16 and HPV18; some women were positive for only HPV16 DNA, or for only HPV18 DNA, hence some of these charts have no red triangle).

Figure 18.8. Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time. The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months) and vertical scale (inverse of the 50% neutralizing antibody titre). A closed (open) plotting symbol indicates that a sample was taken (was not taken, or was unavailable): the 50% neutralizing antibody titre retains it's last known value until updated. Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

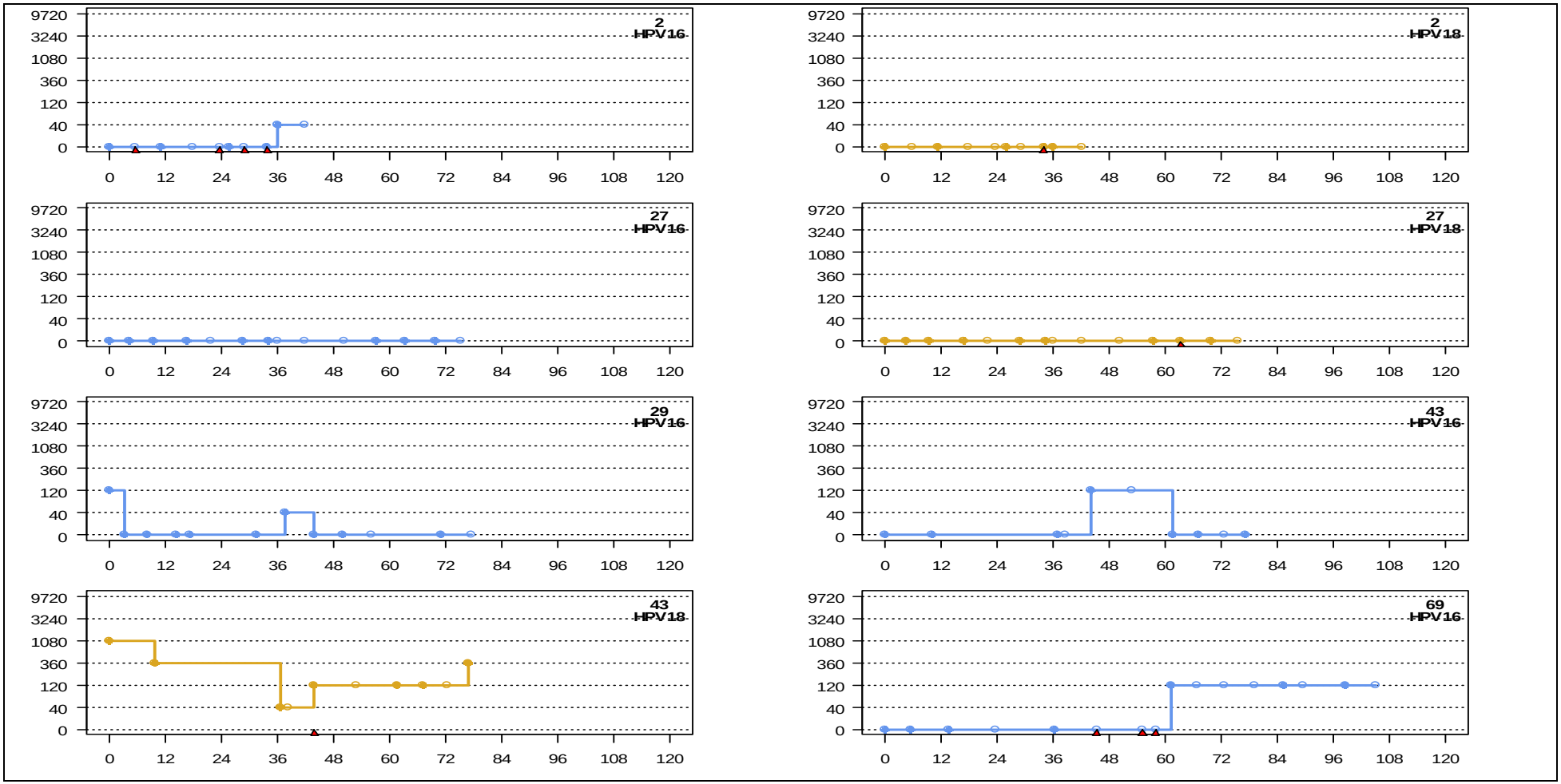


Figure 18.8 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time. The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months) and vertical scale (inverse of the 50% neutralizing antibody titre). A closed (open) plotting symbol indicates that a sample was taken (was not taken, or was unavailable): the 50% neutralizing antibody titre retains it's last known value until updated. Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

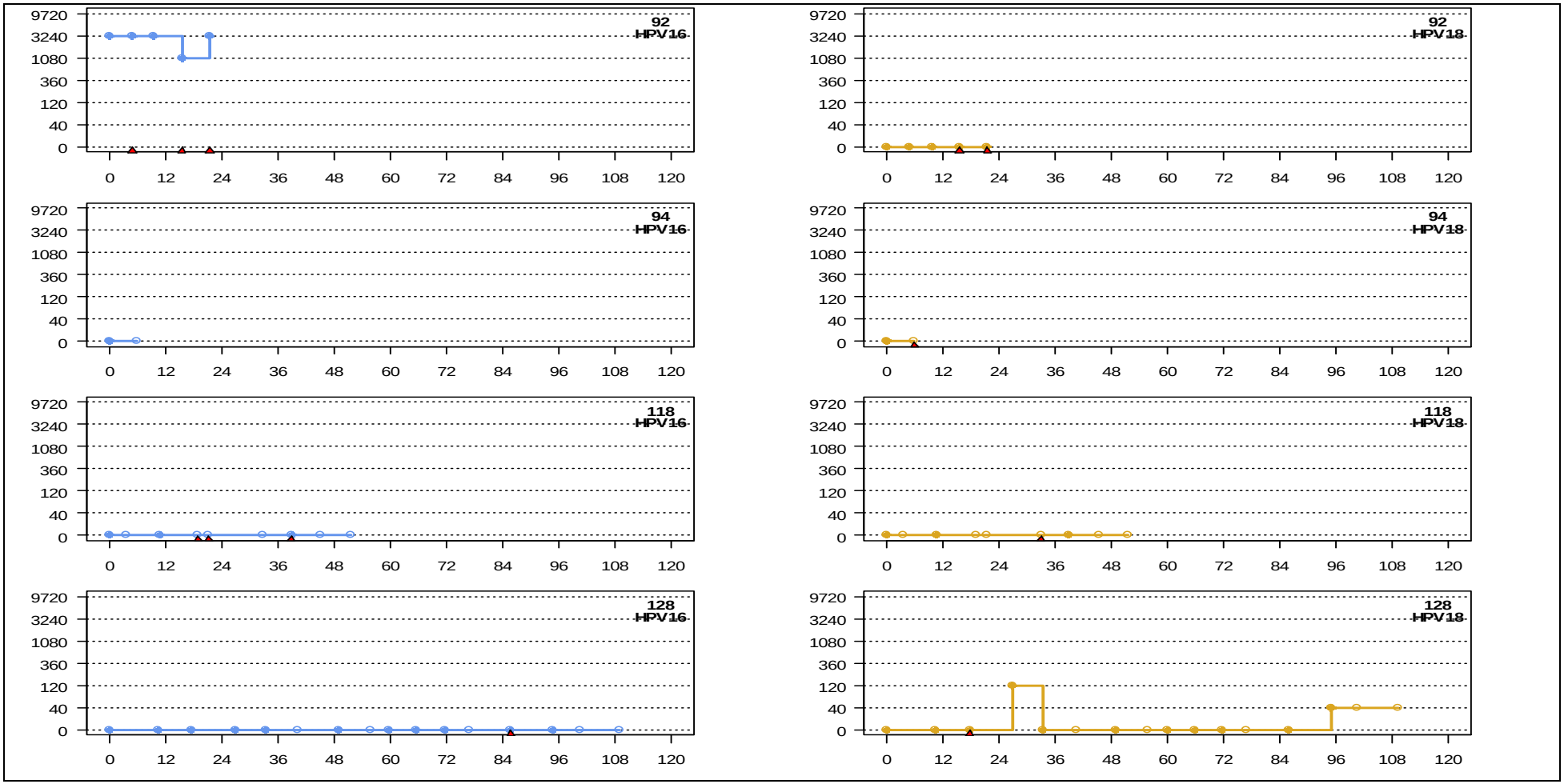


Figure 18.8 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time. The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months) and vertical scale (inverse of the 50% neutralizing antibody titre). A closed (open) plotting symbol indicates that a sample was taken (was not taken, or was unavailable): the 50% neutralizing antibody titre retains it's last known value until updated. Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

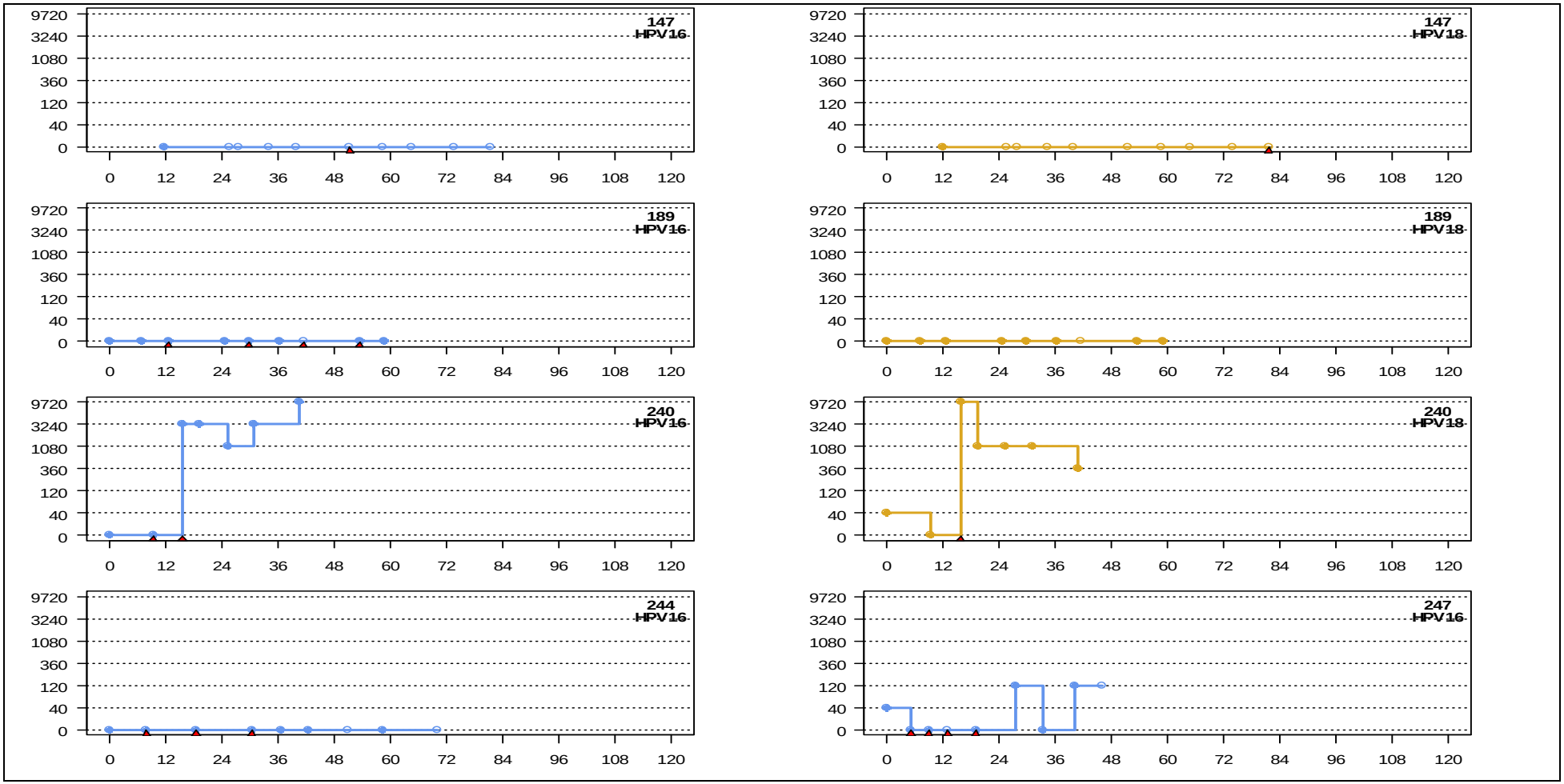


Figure 18.8 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time. The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months) and vertical scale (inverse of the 50% neutralizing antibody titre). A closed (open) plotting symbol indicates that a sample was taken (was not taken, or was unavailable): the 50% neutralizing antibody titre retains it's last known value until updated. Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

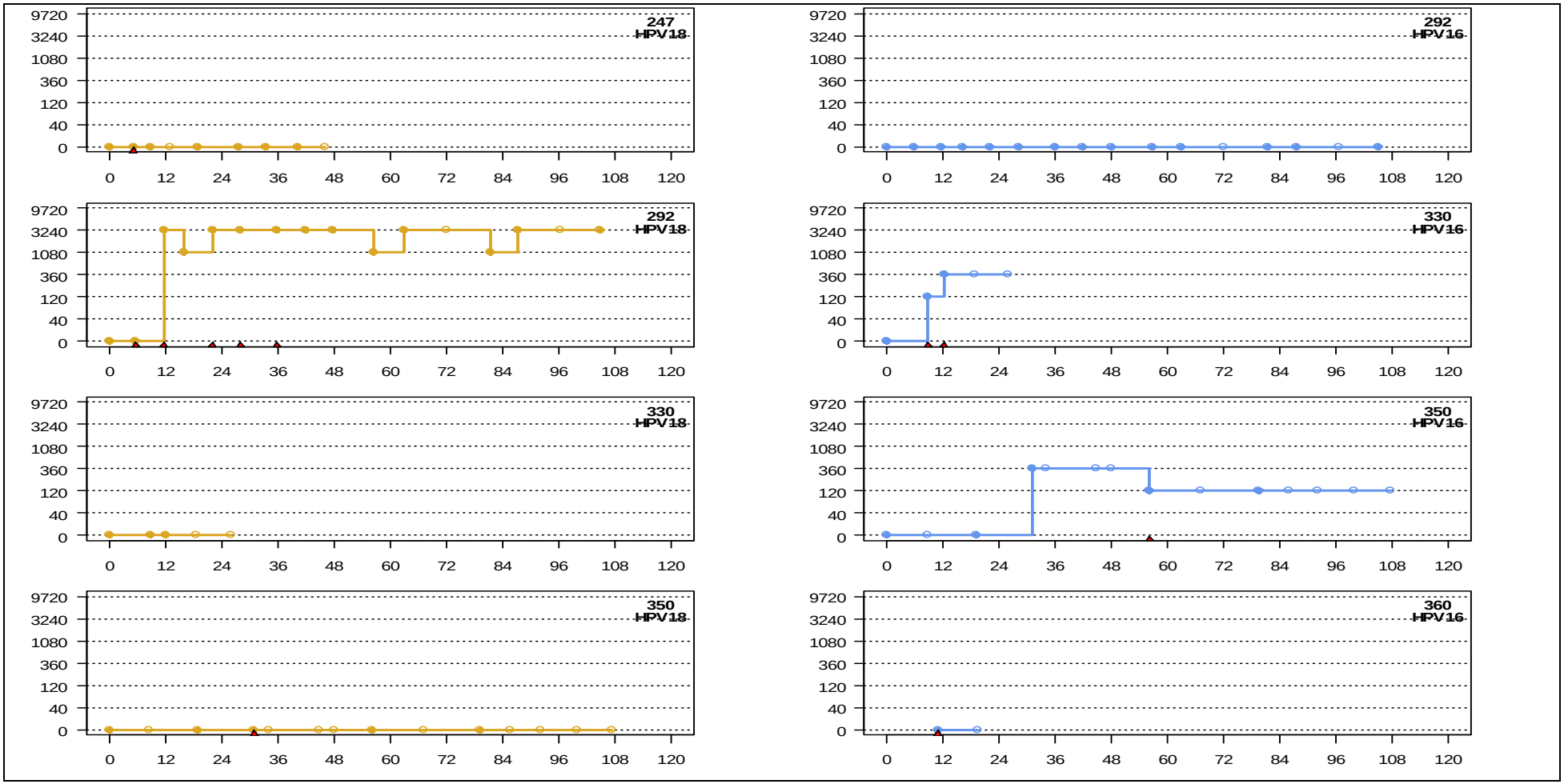


Figure 18.8 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time. The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months) and vertical scale (inverse of the 50% neutralizing antibody titre). A closed (open) plotting symbol indicates that a sample was taken (was not taken, or was unavailable): the 50% neutralizing antibody titre retains it's last known value until updated. Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

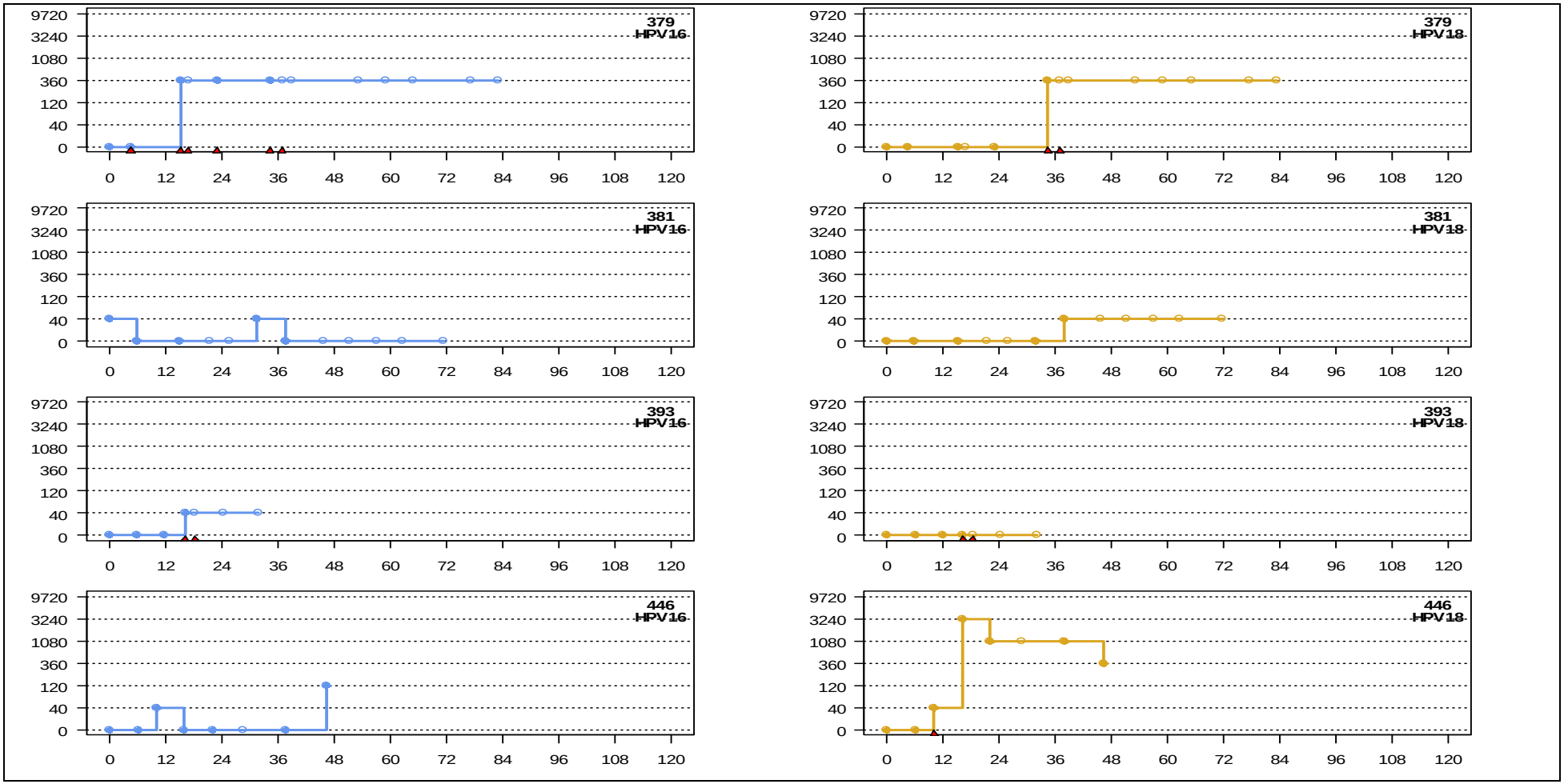


Figure 18.8 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time. The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months) and vertical scale (inverse of the 50% neutralizing antibody titre). A closed (open) plotting symbol indicates that a sample was taken (was not taken, or was unavailable): the 50% neutralizing antibody titre retains it's last known value until updated. Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

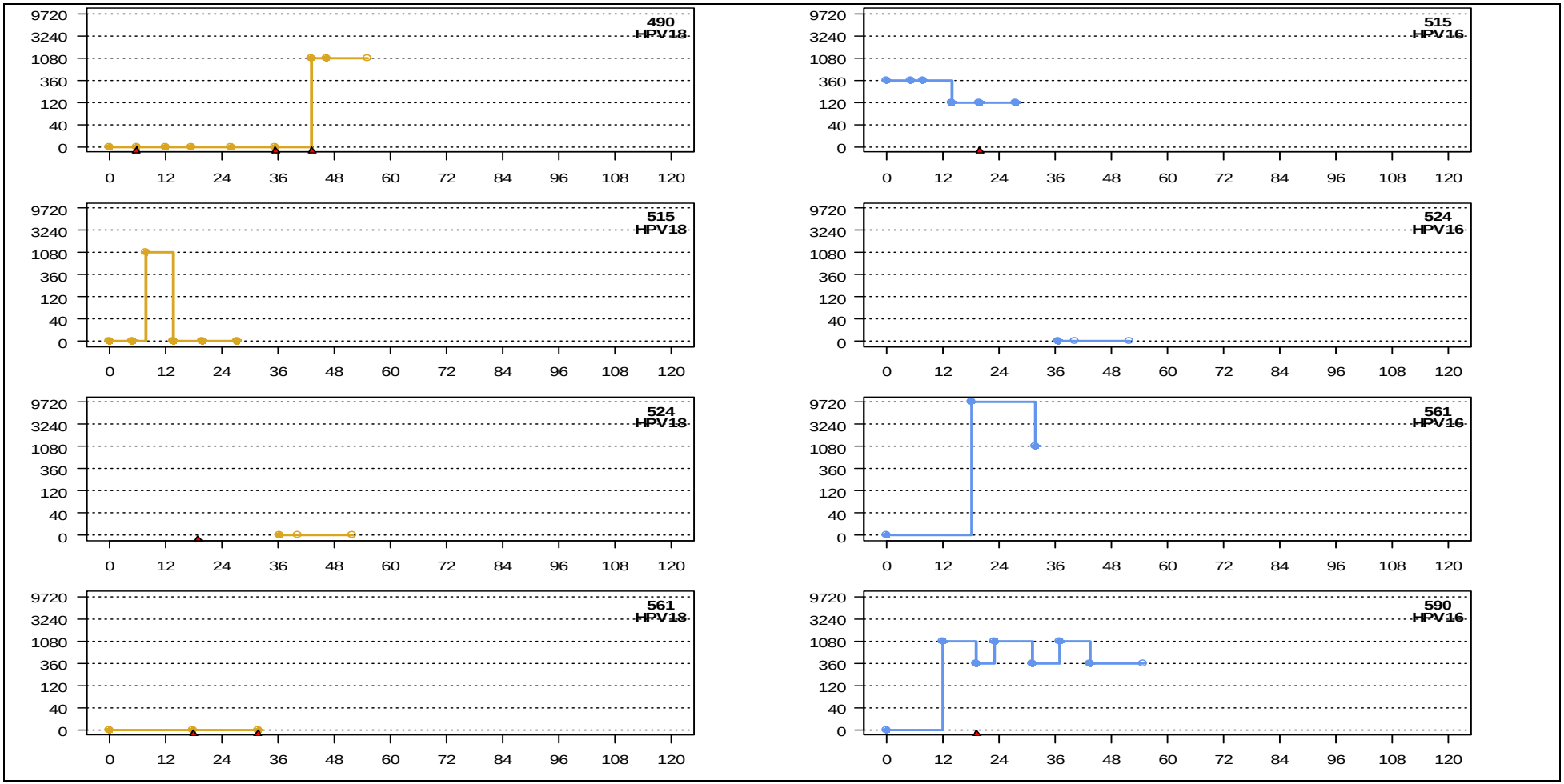


Figure 18.8 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time. The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months) and vertical scale (inverse of the 50% neutralizing antibody titre). A closed (open) plotting symbol indicates that a sample was taken (was not taken, or was unavailable): the 50% neutralizing antibody titre retains it's last known value until updated. Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

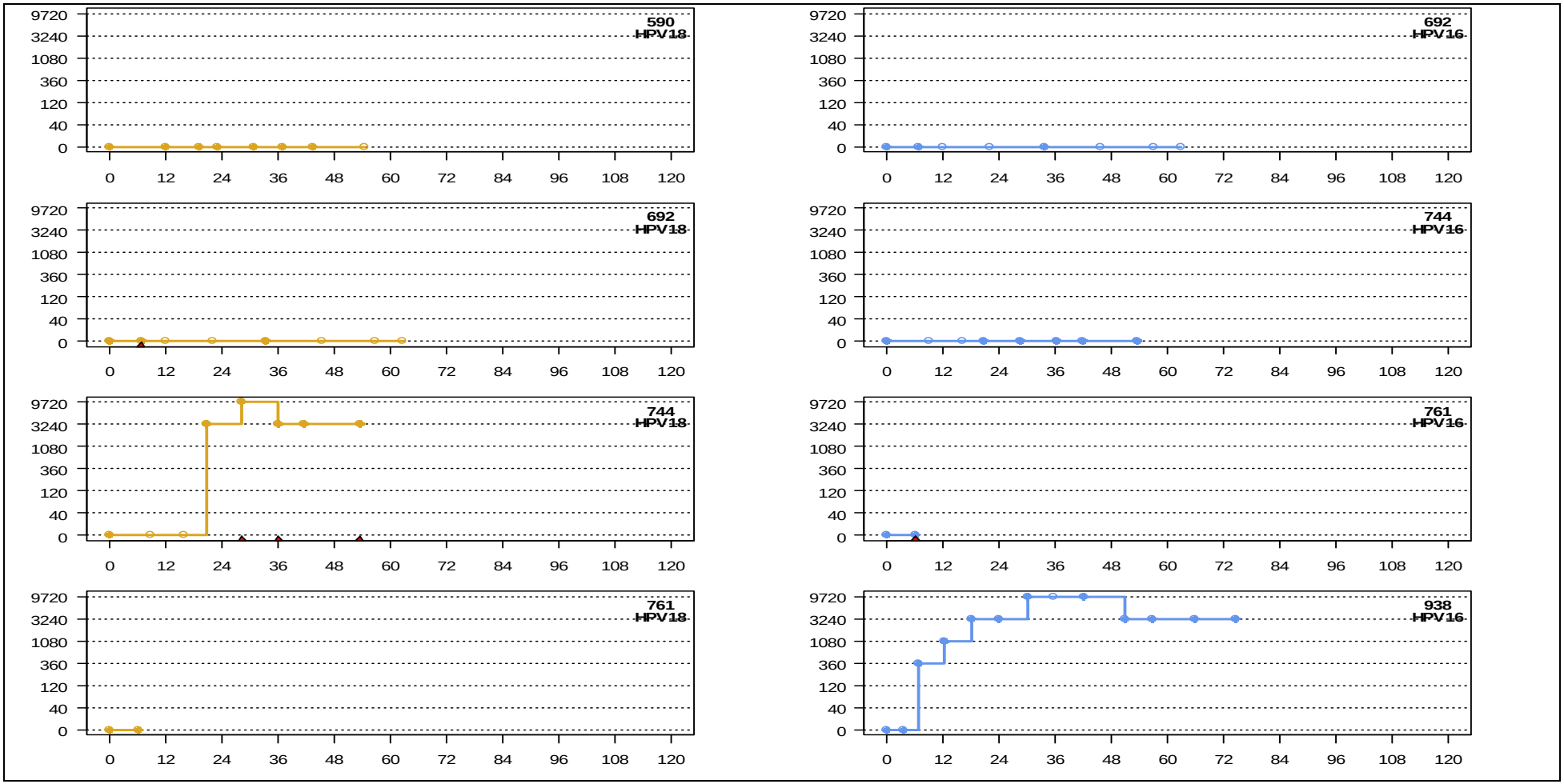


Figure 18.8 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time. The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months) and vertical scale (inverse of the 50% neutralizing antibody titre). A closed (open) plotting symbol indicates that a sample was taken (was not taken, or was unavailable): the 50% neutralizing antibody titre retains it's last known value until updated. Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

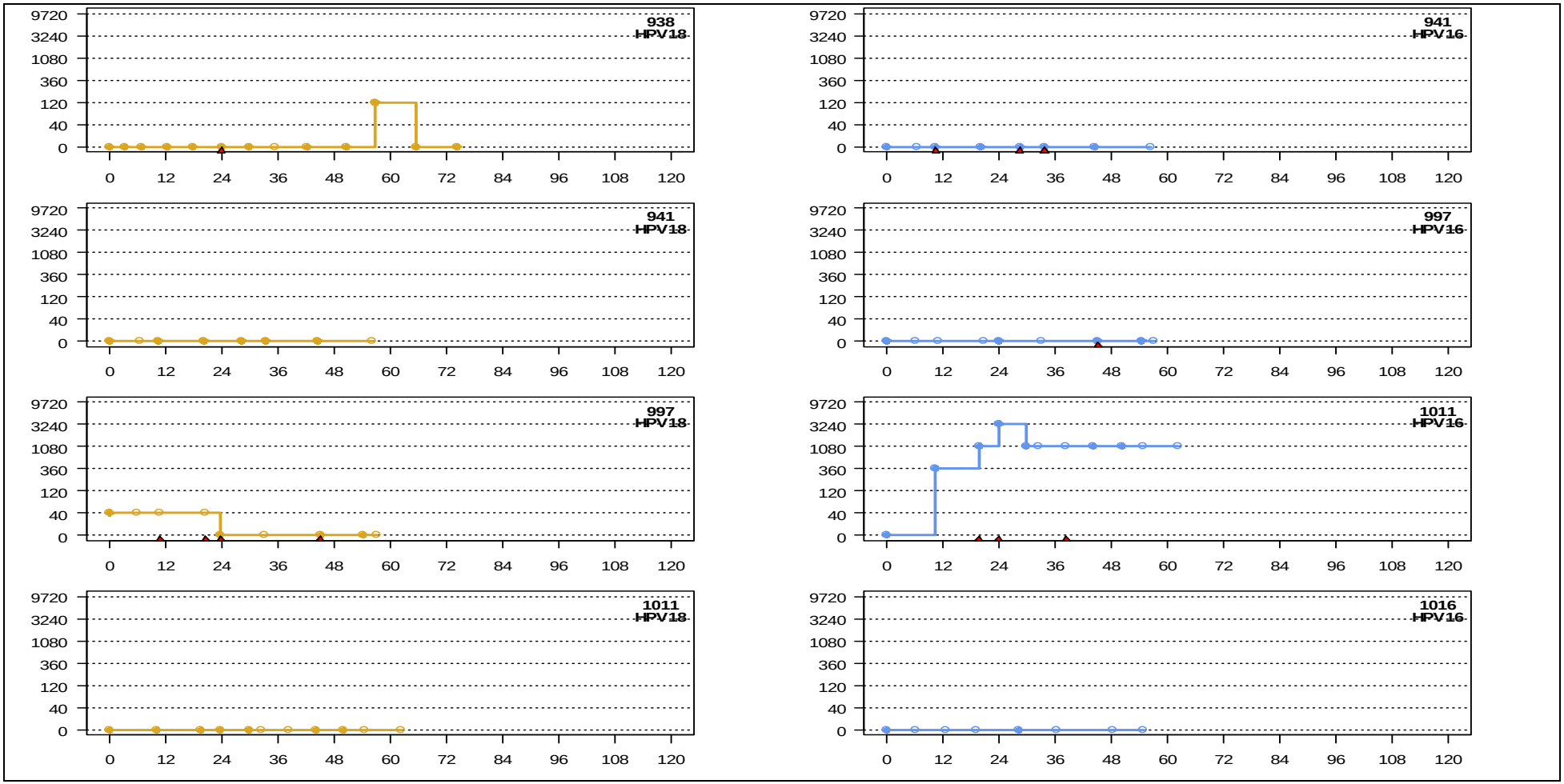


Figure 18.8 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time. The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months) and vertical scale (inverse of the 50% neutralizing antibody titre). A closed (open) plotting symbol indicates that a sample was taken (was not taken, or was unavailable): the 50% neutralizing antibody titre retains it's last known value until updated. Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

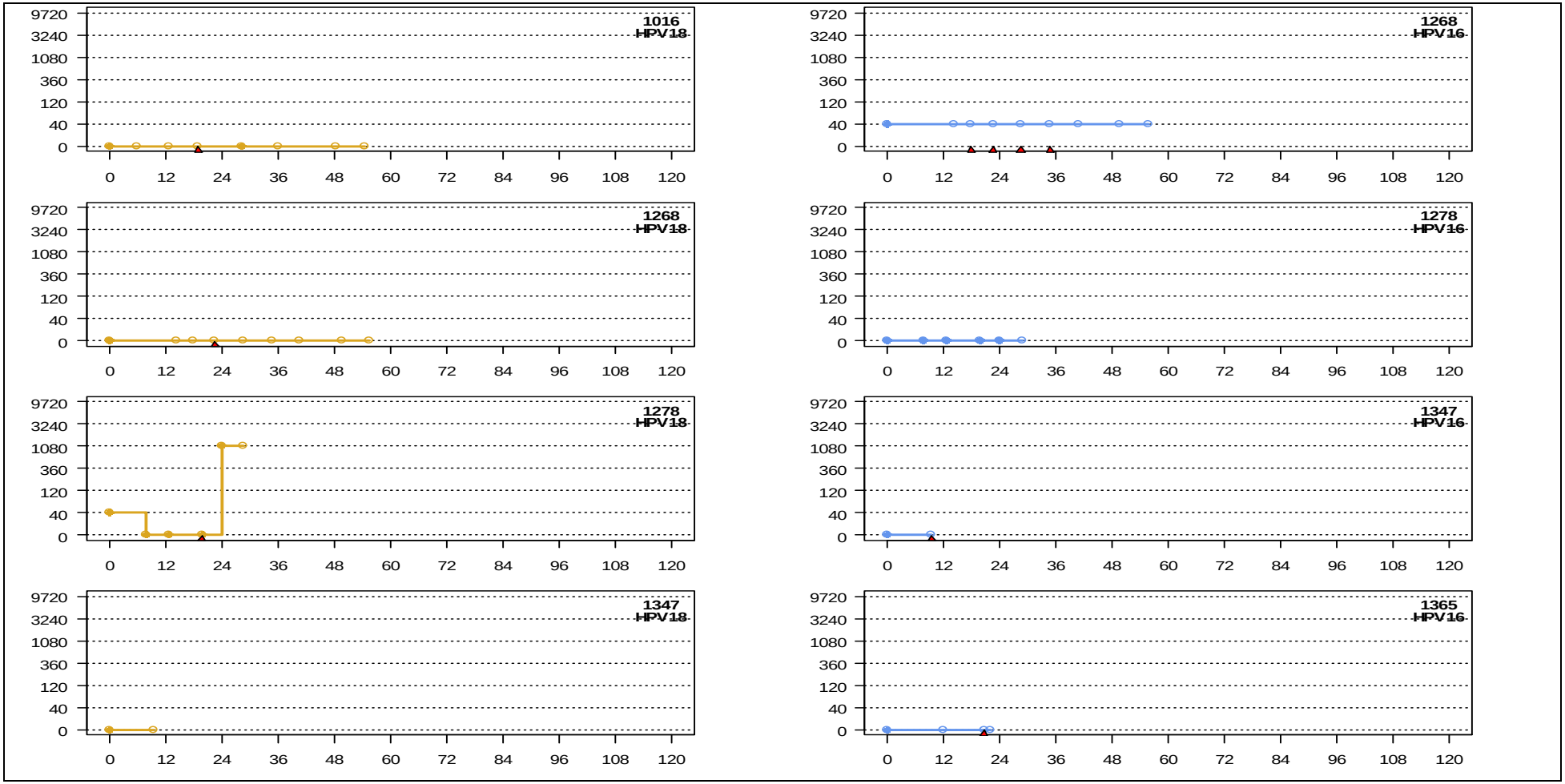


Figure 18.8 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time. The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months) and vertical scale (inverse of the 50% neutralizing antibody titre). A closed (open) plotting symbol indicates that a sample was taken (was not taken, or was unavailable): the 50% neutralizing antibody titre retains it's last known value until updated. Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

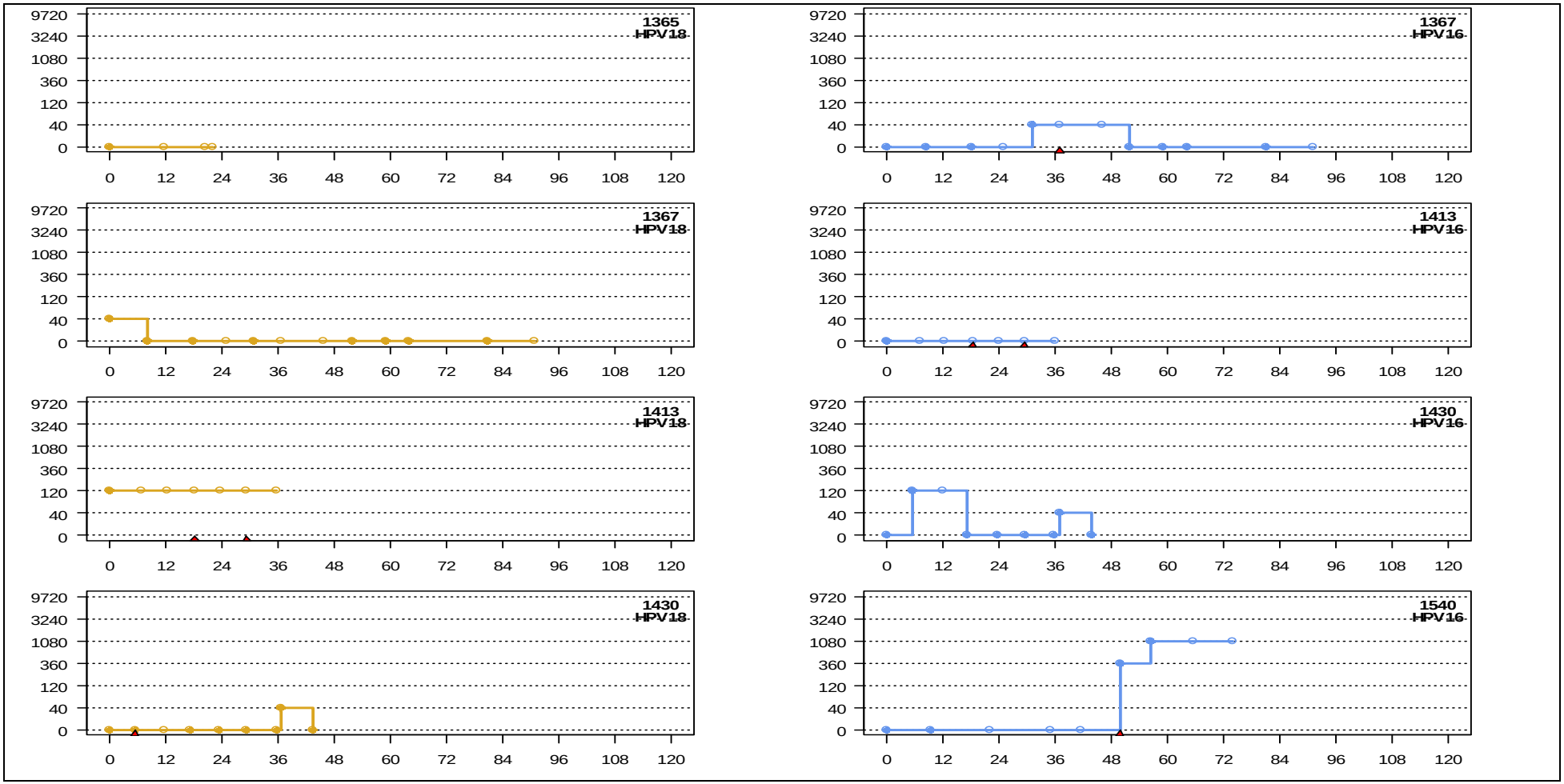


Figure 18.8 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time. The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months) and vertical scale (inverse of the 50% neutralizing antibody titre). A closed (open) plotting symbol indicates that a sample was taken (was not taken, or was unavailable): the 50% neutralizing antibody titre retains it's last known value until updated. Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

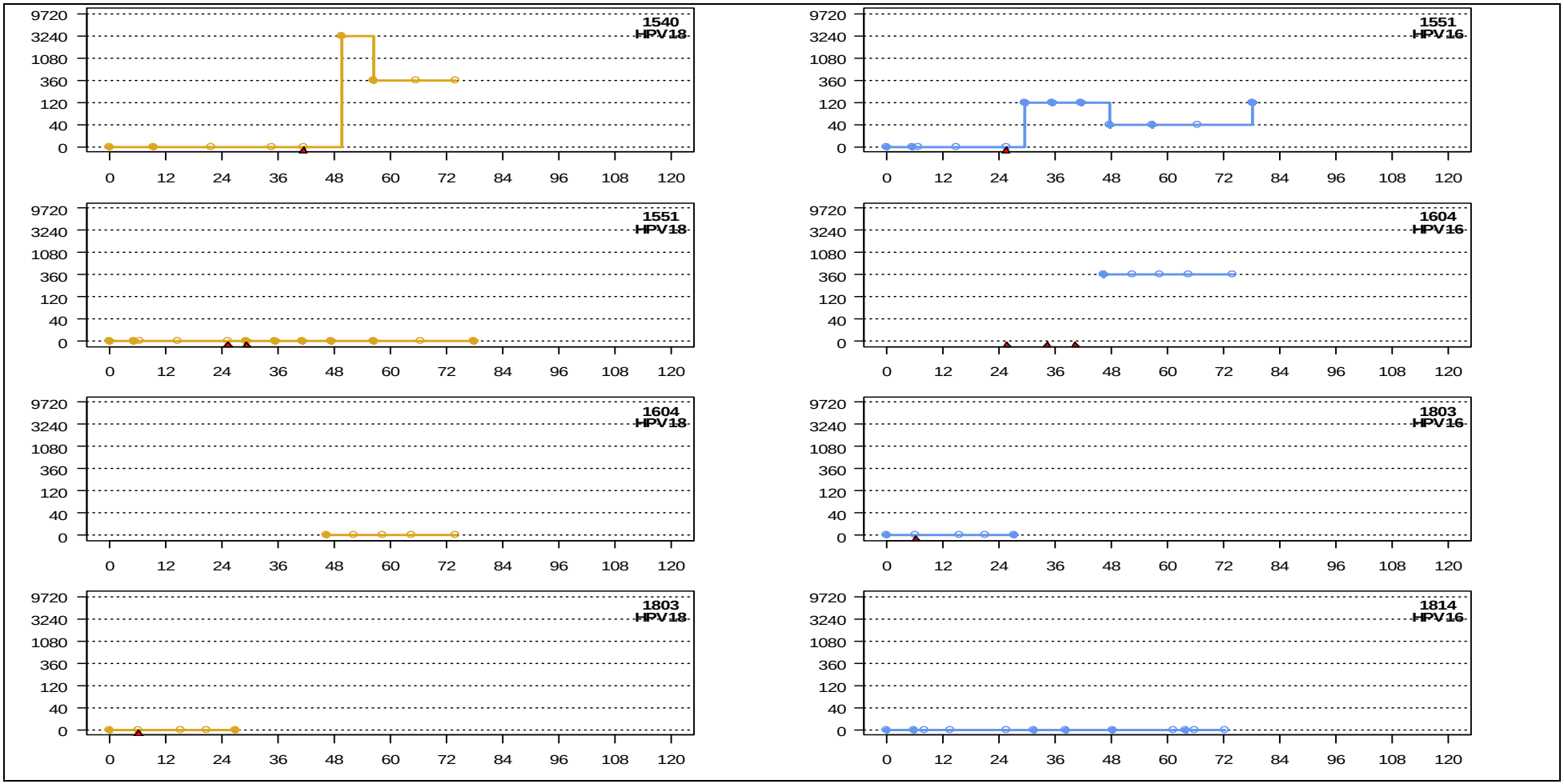
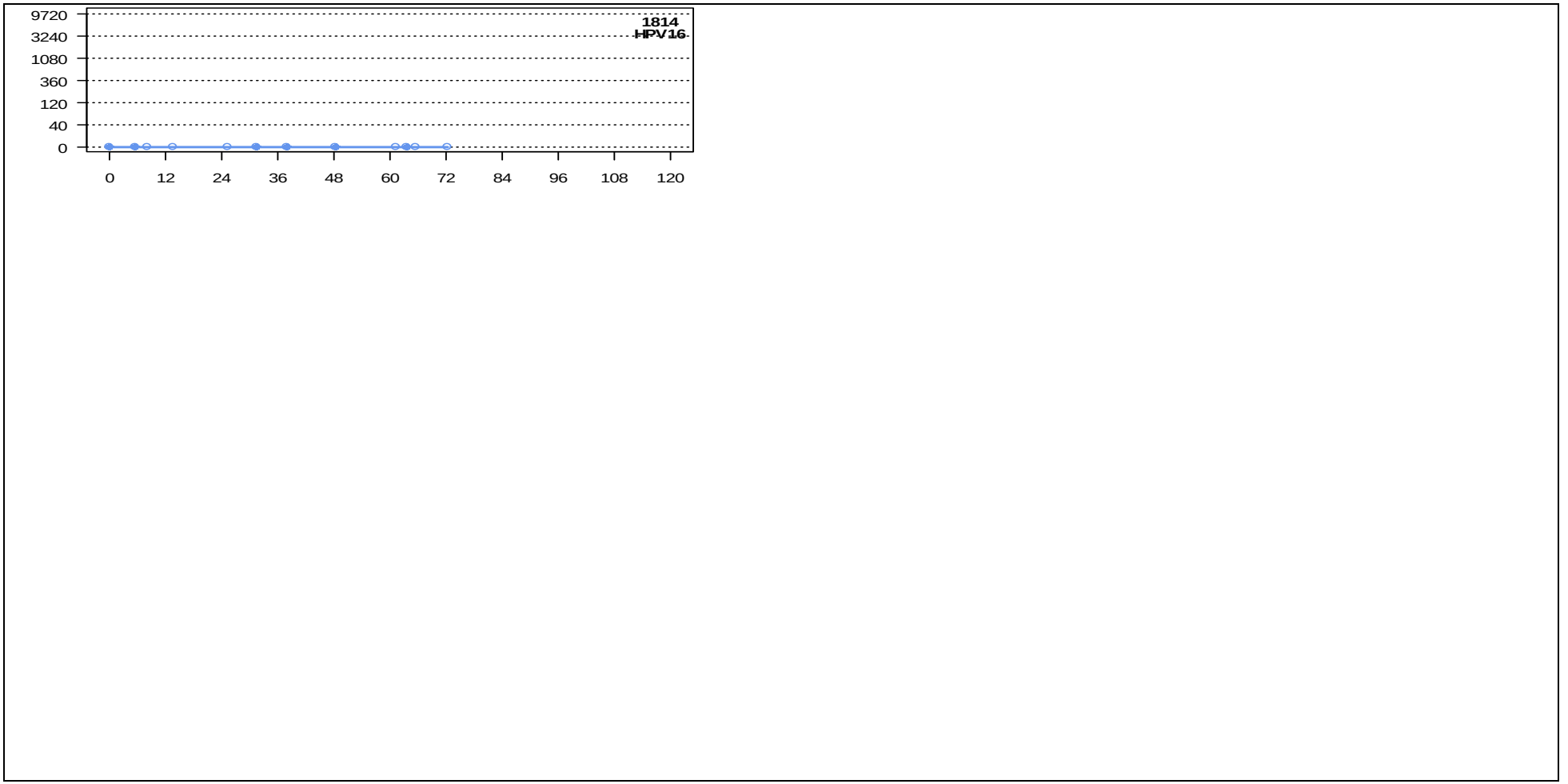


Figure 18.8 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time. The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months) and vertical scale (inverse of the 50% neutralizing antibody titre). A closed (open) plotting symbol indicates that a sample was taken (was not taken, or was unavailable): the 50% neutralizing antibody titre retains it's last known value until updated. Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.



When estimating the time between the date of the first acquisition of cervical infection with HPV16 or HPV18 DNA and seroconversion to that type, the Kaplan-Meier method and the approach adopted by Carter *et al*. (Carter 2000) was used. This was to enable a direct comparison to be made between the results from the cohort study described in this thesis those of the “Carter cohort”. With this approach, time to seroconversion was measured from the date of the visit corresponding to the sample in which HPV16 or HPV18 DNA was first detected, until the date of the visit corresponding to the sample in which women first seroconverted to that type, with censoring on the date of the last available serum sample. Both analyses were limited to women who acquired HPV16 or HPV18 DNA during follow-up, who were seronegative at all visits prior to the detection of HPV DNA, who had serology results available after the visit at which they were first HPV DNA-positive, and who were not missing a serology sample from the visit which corresponded to the first visit at which HPV DNA was detected, unless the woman was seronegative at the subsequent visit. Clearly, this approach is less than ideal, but is a pragmatic approach to dealing with the substantial amount of missing data.

18.8.1 The time to seroconversion to HPV16 in women who tested positive for HPV16 DNA using GP5+/GP6+

Thirty eight women were seronegative, and seven were seropositive, to HPV16 in their first evaluable serum sample. Twenty two of the women who were HPV16-seronegative in their first evaluable sample first tested positive for HPV16 DNA at a follow-up visit, as determined using GP5+/GP6+ general-primer-mediated PCR. Of

these, 17 had one or more evaluable sera samples at or subsequent to this visit, and 10 seroconverted to HPV16: two were first found to be seropositive to HPV16 at the same visit HPV16 DNA was first detected; four at a subsequent visit; and four before HPV16 DNA was first detected. Of those who first tested positive for HPV16 DNA and who subsequently seroconverted to HPV16 during follow-up, one each had a greatest titre of 1-in-40, 1-in-120, 1-in-360, and 1-in-9720. One woman (study number 1604) was seropositive at a visit subsequent to the first detection of HPV16 DNA, but had only a single serum sample.

In women with an incident cervical HPV16 infection, the median time from the first detection of HPV16 DNA in a cervical sample to seroconversion to HPV16 was 10.6 months.

18.8.2 The time to seroconversion to HPV18 in women who tested positive for HPV18 DNA using GP5+/GP6+

Thirty seven women were seronegative, and six were seropositive, to HPV18 in their first evaluable serum sample. Twenty three of the women who were HPV18-seronegative in their first evaluable sample first tested positive for HPV18 DNA at a follow-up visit, as determined using GP5+/GP6+ general-primer-mediated PCR. Of these, 21 had one or more evaluable sera samples at or subsequent to this visit, and 10 seroconverted to HPV18: two were first found to be seropositive to HPV18 at the same visit as HPV18 DNA was first detected; seven at a subsequent visit; and one before HPV18 DNA was first detected. Of those who first tested positive for HPV18

DNA and who subsequently seroconverted to HPV18 during follow-up, two each had a greatest titre of 1-in-40, 1-in-120, and 1-in-3240, and one had a greatest titre of 1-in-1080.

In women with an incident cervical HPV18 infection, the median time from the first detection of HPV18 DNA to seroconversion to HPV18 was 31.3 months.

18.8.3 The time to seroconversion to HPV16 and/or HPV18 in women who tested positive for HPV16 DNA and HPV18 DNA using GP5+/GP6+

The study population for the analysis of the time to seroconversion to HPV16 and/or HPV18 in women who tested positive for HPV16 DNA and HPV18 DNA using GP5+/GP6+ is restricted to the 42 women who were eligible for inclusion in both the HPV16 and HPV18 analyses. No woman was seropositive for both HPV16 and HPV18 in her first evaluable serum sample. Nine women who were seronegative for both HPV16 and HPV18 in their first evaluable serum sample, first tested positive for both of these HPV types during follow-up; four seroconverted to HPV16 alone, and two seroconverted to both HPV16 and HPV18.

Summary

Twenty two of the 38 women who were HPV16-seronegative in their first evaluable sample first tested positive for HPV16 DNA at a follow-up visit. Of these, 17 had one or more evaluable sera samples at or subsequent to this visit, and 10 seroconverted to HPV16. In women with an incident cervical HPV16 infection, the median time from the first detection of HPV16 DNA in a cervical sample to seroconversion to HPV16 was 10.6 months. Twenty three of the 37 women who were HPV18-seronegative in their first evaluable sample first tested positive for HPV18 DNA at a follow-up visit. Of these, 21 had one or more evaluable sera samples at or subsequent to this visit, and 10 seroconverted to HPV18. In women with an incident cervical HPV18 infection, the median time from the first detection of HPV18 DNA to seroconversion to HPV18 was 31.3 months. Nine women who were seronegative for both HPV16 and HPV18 in their first evaluable serum sample, first tested positive for both of these HPV types during follow-up; four seroconverted to HPV16 alone, and two seroconverted to both HPV16 and HPV18.

18.9 SEROCONVERSION TO HPV16 AND/OR HPV18 IN WOMEN WHO DID NOT TEST POSITIVE FOR HPV16 OR HPV18 DNA

18.9.1 Prior to Seroconversion

Four women who were HPV16-seronegative in their first evaluable serum sample, who first tested positive for HPV16 DNA at a follow-up visit, as determined using GP5+/GP6+ general-primer-mediated PCR, and who had one or more evaluable sera samples at, or subsequent to, this visit, seroconverted to HPV16 before HPV16 DNA was first detected; one had a greatest neutralizing antibody titre prior to the first detection of HPV16 DNA in a cervical sample of 1-in-40, two 1-in-360, and one a greatest titre of 1-in-1080. Four women were HPV16-seropositive in their first evaluable sample and first tested positive for HPV16 DNA at a follow-up visit; two

had a greatest neutralizing antibody titre prior to the first detection of HPV16 DNA in a cervical sample of 1-in-40, one 1-in-360, and one a greatest titre of 1-in-3240.

One woman who was HPV18-seronegative in her first evaluable serum sample, who first tested positive for HPV18 DNA at a follow-up visit, as determined using GP5+/GP6+ general-primer-mediated PCR, and who had one or more evaluable sera samples at, or subsequent to, this visit, seroconverted to HPV18 before HPV18 DNA was first detected: her greatest neutralizing antibody titre prior to the first detection of HPV18 DNA in a cervical sample was 1-in-3240. Five women were HPV18-seropositive in their first evaluable sample and first tested positive for HPV18 DNA at a follow-up visit; four had a greatest neutralizing antibody titre prior to the first detection of HPV18 DNA in a cervical sample of 1-in-40, and one a greatest titre of 1-in-1080.

18.9.2 At any time during follow-up

When HPV status was determined using GP5+/GP6+ general-primer-mediated PCR, seven women were found to have seroconverted to HPV16 without ever having tested positive for HPV16 DNA during the study period; two of these women were seropositive to HPV16 at study entry. The greatest titre observed in these seven women was 1-in-40 for one woman, 1-in-120 for four women, and the maximum possible titre of 1-in-9720 for two women.

Similarly, three women were found to have seroconverted to HPV18 without ever testing positive for HPV18 DNA, including one woman who was seropositive to HPV18 at study entry and who had no further serological samples. The greatest titre found during follow-up was 1-in-40, 1-in-120 and 1-in-1080 for one each of these three women.

Summary

Four women who were HPV16-seronegative in their first evaluable serum sample, who first tested positive for HPV16 DNA at a follow-up visit, and who had one or more evaluable sera samples at, or subsequent to, this visit, seroconverted to HPV16 before HPV16 DNA was first detected; for HPV18, the corresponding number was one woman. Seven women seroconverted to HPV16 without ever having tested positive for HPV16 DNA during the study period, including two who were seropositive to HPV16 at study entry; for HPV18, the corresponding numbers were three and one.

18.10 DURATION OF THE NEUTRALIZING ANTIBODY RESPONSE

18.10.1 HPV16

Of the 22 women who were seropositive to HPV16 at some time during follow-up, 15 remained seropositive to HPV16 until the end of follow-up. The median time interval between the dates of the first and last HPV16-seropositive samples was 859.5 days (range 102 to 2,058); the median number of consecutive HPV16-seropositive samples was 5.5 (range 2 to 10), excluding two women who were HPV16-

seropositive on only one occasion. Seven of the 22 women became seronegative to HPV16, of whom five had a subsequent sample which was seropositive to HPV16.

18.10.2 HPV18

Of the 17 women who were seropositive to HPV18 sometime during follow-up, nine remained seropositive to HPV18 until the end of follow-up. The median time interval between the dates of the first and last HPV18-seropositive samples was 1,106 days (range 203 to 2,836); the median number of consecutive HPV18-seropositive samples was 6 (range 2 to 12), excluding three women who were HPV18-seropositive on only one occasion. Eight of the 17 women became seronegative to HPV18, of whom five had a subsequent sample which was seropositive to HPV18.

Summary

Of the 22 women who were seropositive to HPV16 at some time during follow-up, 15 remained seropositive to HPV16 until the end of follow-up; for HPV18, the corresponding numbers were 17 and nine. The median time between the dates of the first and last seropositive samples was 859.5 days (range 102 to 2,058) for HPV16 and 1,106 days (203 to 2,836) for HPV18; the median number of consecutive seropositive samples was 5.5 (2 to 10) for HPV16, and 6 (2 to 12) for HPV18. Seven (8) of the 22 (17) women became seronegative to HPV16 (HPV18), of whom five had a subsequent sample which was seropositive to HPV16 (HPV18).

Chapter 19

THE KINETICS OF THE HUMORAL IMMUNE RESPONSE TO INCIDENT CERVICAL HPV16 AND HPV18 INFECTIONS

DISCUSSION

We set out to adapt the neutralization assay first developed by Pastrana *et al.*, 2004 (Pastrana 2004) for the purposes of testing sera samples from the cohort study described in this thesis. This assay uses HPV16 and HPV18 pseudovirions (PsV) carrying a secretory embryonic alkaline phosphatase (SEAP) reporter gene, and measures functionally-relevant HPV type-specific neutralizing antibodies.

19.1 RELIABILITY OF THE NEUTRALIZING ANTIBODY ASSAY

My first objective was to determine the reliability of the neutralizing antibody assay. This was an interactive, and iterative, process involving the candidate and the laboratory scientists collaborating on the study described in this section of the thesis. A laboratory analysis would be followed by a statistical analysis undertaken to measure the reliability of the assay. The results of this analysis would then be interpreted by, and a consensus reached between, the candidate and the laboratory scientists regarding the reliability of the assay, following which further assays either would, or would not, be undertaken. Results of statistical analyses were not the only criteria for judging success: however, some laboratory analyses undertaken early in the developmental stage of the assay appeared to produce *statistically reliable* results, but nevertheless this was insufficient to convince the laboratory scientists

that the results of the assay were in fact reliable. Similarly, only the results of a statistical analysis could undermine the confidence of the laboratory scientists in the results of some laboratory analyses of which they were otherwise uncritical.

The laboratory scientists were ultimately successful in adapting the assay to the purpose for which it was needed. The reproducibility of the HPV16 assay was excellent: the values of the weighted-kappa statistics measuring intra- and inter-assay reliability were 0.95 and 0.89, respectively, compared to a maximum possible value of 1.00. Discrepancies between titres from repeated measurements were at most one titre category for the final laboratory analyses undertaken during the developmental phase of the assay. The HPV18 assay was clearly more variable, and therefore less reliable, than the HPV16 assay: however, the values of the weighted-kappa statistics measuring intra- and inter-assay reliability were still impressive, both being 0.90. Although most discrepancies between replicates were of only one titre category, there were many such discrepancies, and two analyses each included one set of replicates for which there was a discrepancy of two titre categories between replicates. The greater variability of this assay almost certainly reflects the greater instability of the HPV18 PsV. HPV18 VLP are also more unstable than HPV16 VLP, which may have a biological basis, i.e. this may also be true of “real” viruses. A HPV18 VLP ELISA is therefore more likely than a HPV16 VLP-based ELISA to detect both cross-reactive and non-neutralizing antibodies. However, unlike for VLP-based assays, the impact of PsV degradation on the assay is at least immediately apparent: only an intact PsV will be able to “infect” a target virus and generate a detectable signal.

Titres against HPV were analysed on a categorical scale in analyses of reliability, although clearly the 50% neutralizing antibody titre is, strictly speaking, a continuous quantity. When the “exact” 50% neutralizing antibody titre was obtained by interpolation from the categorical titres, and an intra-cluster correlation coefficient used to estimate reliability, results were, in my opinion, deceptively impressive. This was true even for some of the early developmental assays, which were clearly sub-optimal. I believe the weighted-kappa statistic derived using categorical titres provides a more accurate summary of the performance of this assay: ultimately, impressive estimates of this statistic were obtained. Given the limited scope of the analyses subsequently performed, and the invariance of the results to the chosen scale of measurement, interpolated “exact” titres were not considered further.

Neutralization assays, such as have been described in this thesis, are reported to have a higher specificity than VLP-based ELISA (Buck 2005, Fleury 2008). Direct comparisons between the two suggest that the HPV16 neutralizing antibody assay is possibly more sensitive than the HPV16 VLP-based ELISA, whereas sensitivity is similar for the HPV18 version of this assay (Pastrana 2004). Therefore, although the HPV16 neutralizing antibody assay is superior to a VLP-based ELISA because it detects functionally relevant antibodies, i.e. those antibodies which prevent infection of a host cell by virus, the HPV16 VLP-based ELISA may have comparable performance. The HPV18 neutralizing antibody assay is superior to a HPV18 VLP-based ELISA on two counts, since not only is it more sensitive, it is also highly specific.

19.2 THE KINETICS OF THE NEUTRALIZING ANTIBODY RESPONSE

My second objective was to describe the kinetics of the neutralizing antibody response to incident cervical HPV16 and HPV18 infections in a cohort of young women who were recruited soon after first sexual intercourse, and who first tested positive for HPV16 DNA, or HPV18 DNA, or both, in cervical cytological samples during follow-up. Only limited inferences can be drawn from this small series. However, the findings reported in this thesis are broadly consistent with those of two other cohort studies which have measured the humoral immune response following an incident HPV infection using the technically simpler VLP-based ELISA.

In women with an incident cervical HPV16 infection in the cohort described in this thesis, the median time from the date of the first detection of HPV16 DNA in a cervical sample to the date of seroconversion to HPV16 was 10.6 months, greater than the 8.3 months reported by Ho *et al*. (Ho 2004), but similar to the 11.8 months reported by Carter *et al.* (Carter 2000). Carter *et al.* also report that 18 months after the first detection of HPV16 DNA, 59.1% of women with an incident HPV16 infection had seroconverted; the corresponding seroconversion rate for women in the cohort described in this thesis was 61.1%. Ho *et al*. report a 12 month cumulative seroconversion rate following an incident HPV16 infection of 56.7%; the corresponding rate for women in the cohort described in this thesis was 51.4%. In all three studies, most women who seroconverted remained seropositive until the end of follow-up.

In women with an incident cervical HPV18 infection in the cohort described in this thesis, the median time from the date of the first detection of HPV18 DNA in a cervical sample to the date of seroconversion to HPV18 was 31.3 months, substantially in excess of the 12.8 months reported by Carter *et al.*, although the numbers in both series are small. Carter *et al.* also report that 18 months after the first detection of HPV DNA, 54.8% of women with an incident HPV18 infection had seroconverted; the corresponding seroconversion rate for women in the cohort described in this thesis was 36.2%. Ho *et al* do not report seroconversion rates following an incident HPV18 infection.

Although in this thesis a potentially more sensitive and more type-specific neutralizing antibody assay than VLP-based ELISA was used, it was found, as others who have used a VLP-based ELISA have found, that some women failed to seroconvert. For example, among those women who failed to seroconvert to HPV16, one had five opportunities, i.e. measurements, to do so, and the other six; during which time these women tested positive for HPV16 using GP5+/GP6+ general-primer-mediated PCR on three and four occasions, respectively. We also found that seroconversion could be transient, and seropositivity intermittent, although this may merely reflect antibody levels which fluctuate around the threshold of detection. Although we recruited women soon after first sexual intercourse, some already appeared to have cleared a primary infection at study entry. For example, of 12 women who were seropositive for either HPV16 or HPV18 at study entry, and who tested negative for HPV DNA at this time, four had been sexually active for less than a year. Of course, it is also possible that vulval, or vaginal, infections, which are not

detected in cervical samples, may also generate an immune response (in particular, a neutralizing antibody response) against HPV (Winer 2003). The observations in this thesis, although not novel, again illustrate how DNA testing, or serology, may fail to identify those sexually active women who have previously been exposed to HPV. This would be an important consideration were the admission of such women to a prophylactic immunization program dependent upon adducing evidence of no previous exposure to vaccine types.

19.3 CONCLUSION

The final study population for this analysis differed from the intended study population. This was unintentional, as has been described earlier. The neutralization assay of Pastrana *et al.* proved to be complex to conduct in practice, and, although use of the assay was ultimately successful, as with all cell-based assays, reproducibility (intra- and inter-assay reliability) was not easily achieved or maintained. Given the expense of the assay, in terms of both monetary costs and laboratory staff time, the decision was made no to carry out any further testing, but instead to preserve the study sera samples for future analyses. In light of the experience gained with the use of the neutralizing antibody assay in a non-automated setting, the work described in this thesis cannot be used in support of the claim that this neutralizing antibody assay is a high-throughput assay suitable for use in large-scale epidemiological enquiries.

19.4 PUBLICATION

A paper based on the work described in chapters 16, 17, 18 and 19 of this thesis was published in 2008 (Steele 2008: see appendix 6). The definitive results of HPV viral load testing became available only *after* the completion of the analyses presented in the part of the thesis relating to the neutralizing antibody response to cervical HPV infections. An analysis of the association between HPV viral load and the humoral immune response may be the subject of a future report (see chapter 24).

Chapter 20

THE NATURAL HISTORY OF CERVICAL HPV16 AND HPV18 INFECTIONS – VIRAL LOAD

INTRODUCTION

Orientation

In this chapter I:

- define HPV viral load
- provide an overview of the evidence regarding the determinants of HPV viral load, its role in the kinetics of cervical HPV infection and its' association with the acquisition of cervical disease
- describe the measurement of HPV viral load

20.1 DEFINITION OF HPV VIRAL LOAD

The HPV *viral load* (or HPV *copy number*) of a given sample is the number of HPV virions, or the number of copies of the HPV genome, present in that sample.

20.2 BACKGROUND

The identification of high-risk HPV types as a *necessary* cause of cervical cancer offers not only the prospect of effective primary-prevention, but also the possibility of improving the efficiency of cervical screening programmes. Among women who test

positive for high-risk HPV types, cytological abnormality has been reported to be more common in those with a high- than in those with a low-viral load (Hall 1996, Nindl 1997, Ho 1998, Swan 1999, Heard 2000, Lillo 2005). The apparent consistency of this finding has persuaded many that the inclusion of a measurement of HPV viral load could improve the effectiveness of HPV-based cervical screening and triage strategies (Joseffson 2000). However, for this opportunity to be realised, a more complete understanding than we currently possess of the natural history of HPV infection, and its relationship to the acquisition of epithelial abnormalities of the cervix, is required.

20.3 LITERATURE OVERVIEW

In this section, I provide an overview of the evidence of how cervical HPV viral load varies according to virus- and disease-related factors, study design, and the method used to measure cervical HPV viral load. First, I describe studies which have investigated factors related to changes in HPV viral load itself.

20.3.1 The determinants of HPV viral load

20.3.1.1 Life cycle

Two studies investigating the association between HPV viral load and the risk of transmission of HPV in sexual couples have reported that high viral load is associated with an increased risk of transmission (Wickenden 1988, Bleeker 2005).

Cellular differentiation is vital to the life-cycle of HPV, and the physical state of the virus is a potentially important factor contributing to malignant transformation and disease progression (see section 3.1.5). Four studies have investigated the association between HPV viral load and cellular differentiation (Schneider 1987, Bedell 1991, Spink 2005, Wilson 2005): three were *in vitro* studies of HPV31, and one was a cross-sectional analysis in human subjects. The three *in vitro* studies demonstrated that viral load increased with increasing cellular differentiation, consistent with the postulated model for the life-cycle of cervical HPV infections (see section 3.1.2.3). Two studies, both performing cross-sectional analyses, found that viral load was lower in tumours containing integrated forms than in those which did not (Berumen 1995, Jeon 1995, Spartz 2005, Cricca 2006).

20.3.1.2 HPV type

Ten studies have investigated the association between HPV viral load and HPV type, including one study which has reported this association in two separate reports (Swan 1999, Zerbini 2001, Weissenborn 2003, Moberg 2004, Ho 2005, Moberg 2005, Carcopino 2006, Flores 2006, Lai 2006, Snijders 2006). Two of these studies were longitudinal, one of which nevertheless conducted a cross-sectional analysis, and the remainder were cross-sectional. All found that HPV viral load varied by HPV type, and that this was true in different grades (severities) of cytological abnormality, but the pattern was not consistent, either within or between HPV types.

20.3.1.3 Infections with more than one HPV type

Five cross-sectional studies have investigated the association between HPV viral load and the presence of multiple HPV types (Sherman 2003a, Weissenborn 2003, Giuliano 2004, Levi 2004, Flores 2006). Results were not consistent, either between studies, or within the same population. One report states that "as expected, the HPV viral load increases when more types are present", which, while accurately reflecting naïve *a priori* expectation, is too simplistic (Levi 2004). This expectation is by no means met in all of the remaining studies.

20.3.1.4 Duration of HPV infection

Eight studies have investigated the association between HPV viral load and the duration, or “persistence”, of HPV infection: all were necessarily longitudinal epidemiological studies (Clavel 2000, De Marco 2001, Van Duin 2002, Dalstein 2003, Molano 2003, Syrjanen 2005a, Cricca 2006, Song 2006). Results were inconsistent. Four studies found that HPV viral load was not associated with duration of infection, whereas the other four found that a higher viral load was associated with an increased duration of, or risk of persistent, HPV infection.

20.3.2 Association of HPV viral load with disease endpoints

Many studies have examined the association between HPV viral load and various aspects of the natural history of cervical disease (Woodman 2007). The conclusions of these studies are summarised in table 20.3.2.

Table 20.3.2. Association between high HPV viral load and the natural history of cervical neoplasia: number of studies reporting a positive-association.

Study design	Outcome variable	Positive-association	
		Yes	No
Cross-sectional	Presence of disease	29	1
Cross-sectional	Severity of disease	17	25
	HSIL>LSIL[bc]	*14*[a]	-
	LSIL>HSIL[bc]	*6*[a]	-
Longitudinal	Acquisition of disease	8	4
Longitudinal	Duration of disease	3	2
Longitudinal	Progression of disease	3	2

[a]in two studies, exposure-disease relationship varied with HPV type; [b]direction of association; [c]H(L)SIL=high(low)-grade squamous intraepithelial lesion.

Cross-sectional studies investigating the association between increasing HPV viral load and the presence of cervical disease of any severity have been mainly consistent in reporting a positive association. Cross-sectional studies have been less consistent in their reports of the association between HPV viral load and severity of disease: some have found an association, others none; and those which have found an association are in disagreement as to the direction of that association, i.e. whether increasing viral load is associated with increasing severity of disease, or *vice versa*.

Fewer longitudinal than cross-sectional studies have investigated the association between HPV viral load and the acquisition of cervical disease, or its impact on the duration or progression of disease. The majority have reported a positive-association, but findings from these more appropriate study designs have been less consistent than those from cross-sectional studies.

Factors relating to HPV viral load and characteristics of disease have also been investigated. Two cross-sectional studies have determined the association between HPV viral load and lesion size (Sun 2001, Sherman 2003b). Results were not consistent, with one reporting an increase in viral load with increasing lesion size (Sun 2001) and the other (Sherman 2003b) that viral load was associated with lesion size for CIN1 but not for CIN3. Of course, lesions are not necessarily homogeneous entities. The latter study also reported that cases of CIN3, or worse, for which the primary lesion was also surrounded by extensive CIN1, had the highest HPV viral loads, irrespective of the number of HPV types detected.

It should be noted that many of the results in table 20.3.2 are based on multiple reports from the same study. For example, the 12 reports investigating the association between HPV viral load and the risk of acquisition of cervical cytological disease are from six studies, including one study which has addressed this issue in four separate reports, one in three, and one in two: all were necessarily longitudinal (Clavel 2000, Josefsson 2000, Ylitalo 2000, Bory 2002, Castle 2002a, Lorincz 2002, Van Duin 2002, Schlecht 2003a, Moberg 2004, Castle 2005, Moberg 2005, Monnier-Benoit 2005). Multiple reports from the same study population were internally

consistent, except for those from one study population. One report found that HPV viral load in the baseline sample was not associated with the risk of acquiring CIN3, or worse (Lorincz 2002), in women who were positive for any oncogenic HPV type at baseline; the other found that viral load of HPV16 in the baseline sample *was* associated with the risk of acquiring CIN3, or worse (Castle 2005).

20.3.3 Critique of high-profile cohort studies describing an association between HPV viral load and the risk of the acquisition of cervical disease

In this section I critique reports from two particular longitudinal studies which have been extensively cited in support of an association between HPV viral load and the risk of acquiring high-grade CIN, and illustrate methodological shortcomings in two others. The names given to these studies are for convenience of reference only, and are not necessarily those with which the principal investigators endowed them.

20.3.3.1 The Uppsala cohort study

The Uppsala cohort study was a retrospective cohort study of women attending routine cervical cytological screening in Uppsala County, Sweden. The study population comprised women aged 25 to 49 years who were cytologically normal in their baseline smear and who were followed-up at intervals of three to four years (the routine screening interval). The HPV viral load of all smears taken from cases and controls prior to diagnosis of the case was measured using a modified quantitative PCR technique. Analyses were conducted using a nested case-control design.

Results from this cohort study have now been described in several publications (Joseffsson 1999, Ylitalo 1999, Josefsson 2000, Ylitalo 2000a, Ylitalo 2000b, Moberg 2003, Moberg 2004, Moberg 2005). The study design and the assumptions underpinning the analyses of HPV viral load are now discussed.

20.3.3.1.2 The Uppsala cohort study: Ascertainment of outcomes in cases and controls

In this study, cases and controls were matched only on age and the result of the baseline smear (Ylitalo 1999a, Josefsson 2000, Ylitalo 2000a, Ylitalo 2000b). The decision not to match on total number of smears was justified on the basis that “this variable is likely to be associated with sexual practices (e.g. the need for contraceptives, or treatment for sexually transmitted diseases) and hence with the probability of HPV infection”. This argument is not persuasive. The validity of inferences made from a cross-sectional case-control study depends on the assumption that controls selected for inclusion in the study were eligible to have become cases if they had developed disease. Similarly, in a *nested* case-control study, controls must remain eligible to become cases during follow-up. Cervical carcinoma *in situ* is an asymptomatic condition, the presence of which can only be suspected following an abnormal cervical smear test (see section 2.2.1.1). Therefore a control who has had only a baseline smear can never become a case during follow-up; over a quarter of the 608 controls included in one analysis fall into this category (Ylitalo 2000a).

Another requirement for the validity of statistical inferences with the case-control design is that the control is in fact free of disease i.e. that cases are not, in fact, being compared with cases at the point in time at which the control is matched to the case. In the absence of a smear reported as “normal” taken close to the date of diagnosis of the corresponding case, there is no certainty that a “control” is free of disease at the time the case was diagnosed. The median number of smears provided by cases was 4 (range 1 to 17), and by controls 2 (range 1 to 14). The authors attribute this mismatch in large part to an excess of “confirmatory” smears taken from cases in the year before diagnosis, and smears taken during this time were excluded from some analyses. However, although the excess is greatest during this interval, a substantial excess is also seen in the preceding years (Ylitalo 2000b): nearly half of the controls either had only one smear, or did not have a smear in the year preceding the date of diagnosis of the corresponding case, or both (Ylitalo 2000b).

In this instance, this study clearly violates two of the basic requirements of the case-control design.

20.3.3.1.3 The Uppsala cohort study: Estimation of HPV viral load

In the first paper from this cohort study describing the methods used to estimate HPV viral load, it is stated that, “after normalizing the samples for genomic DNA content, the mean number of copies per cell could be calculated”. However, when reporting the results of the cohort study, threshold cycle numbers (Ct values - see

section 21.1.2) and not the mean number of copies per cell are reported. Although in the final model, comparisons are adjusted for “small” and “large” quantities of DNA, such a crude comparison is in itself surprising, given the comment in their earlier methods paper that “Ct values for β-actin can vary with more than six cycles between individual samples” (Joseffsson 1999). In fairness, in two later reports are copy numbers normalized for cellular DNA content. In the study described in this thesis, all measurements of HPV viral load were normalized for cellular content.

20.3.3.2 The Amsterdam cohort study

The Amsterdam cohort study was a prospective cohort study describing the development, persistence, and progression, of CIN lesions in relation to the natural history of HPV infection, including the humoral immune response. The cohort consisted of 353 women aged 18 to 55 years who were routinely referred to a single colposcopy clinic in Amsterdam, The Netherlands. All women had mild to moderate or severe dyskaryosis at study entry, and were followed at intervals of three to four months until they *clinically* progressed to CIN3; a secondary endpoint was histologically-verified CIN3 at the end of follow-up. Biopsies were not taken during follow-up, but *all* women were biopsied at their last visit, including women with a normal colposcopy, for whom random biopsies were taken. PCR was used to test for 14 high-risk HPV types.

Results from this cohort study have now appeared in several publications (Remmink 1995, De Gruijl 1996a, Rozendaal 1996, De Gruiijl 1996b, De Gruijl 1997, Bontkes

1998a, Bontkes 1998b, Nobbenhuis 1999, Bontkes 1999a, Bontkes 1999b, De Gruijl 1999, Bontkes 2000, Rozendaal 2000, VanDuin 2002). The study design, and the assumptions underpinning the analyses of viral load, are now discussed.

20.3.3.2.1 The Amsterdam cohort study: Primary outcome

In a cohort study, subjects are followed-up until the outcome of interest occurs, thereby revealing the *temporal* aspects of any associations observed. A critical assumption in a cohort study, especially one in which the main interest is in the natural history of the exposure-disease relationship, is therefore that subjects are free of the outcome of interest at study entry.

Given that the primary outcome in this study was determined *clinically*, it should have been possible to confidently classify women as being free of the primary outcome at baseline. However, either two or four women apparently had CIN3 at baseline (Nobbenhuis 1999). And despite the use of a panel of experts to determine the primary outcome there was non-classification of colposcopic impression for 17 women; women for whom there was discordance between the “cytomorphological”, i.e. cytological, result and the colposcopic impression, were also considered to have failed to complete follow-up (Nobbenhuis 1999).

One report states that “women with normal Pap smears containing high-risk HPV genotypes were 116 times more at risk of developing CIN3, in contrast to women without high-risk HPV” (Rozendaal 1996). The characterisation of these women as

“cytologically normal” is questionable, since at baseline 21% of the study population had a smear reported as containing “very mild squamous dyskaryosis, including atypical squamous cells of unknown significance [ASCUS]”. Furthermore, cytological review revealed that, in fact, two of the seven women subsequently found to have CIN3 had moderate dyskaryosis in their baseline smear, and in two more a baseline smear which was initially graded as normal was subsequently re-graded as mildly abnormal. The end-point used also presents a difficulty, since women were not histologically sampled at baseline. Of the 353 women in the cohort, 133 were colposcopically adjudged to have at least CIN2 at study entry, i.e. these women were not free of disease at study entry. The study described in this thesis used a study population which comprised women who were negative for HPV DNA of any type, according to the GP5+/GP6+ system, and free of cervical cytological abnormalities, and undoubtedly free of high-grade CIN, at study entry.

20.3.3.2.2 The Amsterdam cohort study: Viral load

One report includes observations on members of two cohorts, one of which was the Amsterdam cohort described in the previous subsection (Van Duin 2002); the other study also recruited a cohort of women from the Amsterdam area of The Netherlands. Women with normal cytology were followed-up, 125 with a single HPV16 infection at baseline were identified, and a nested, matched, case-control study was undertaken. Cases comprised twelve women who progressed to high-grade CIN; each case was matched with four controls who had CIN1, or less, at the end of follow-up. However, the matching criteria are not stated in this report (the

other reports on this study describe cohort analyses); no reassurance is provided, for example, that cases and matched-controls were followed-up for the same period of time. It is also unclear from the description of the statistical methods whether the analysis was appropriate to the matched design.

In one report, mean copy numbers per sample are reported for HPV and β-globin, but no adjustment is made for cellular content (Van Duin 2002) (see section 20.4.1.1). It is stated that “a logistic regression analysis with adjustment for β-globin levels was used to study viral load in relation to the development of CIN”, but no further details of this sub-optimal analysis are provided. In the study described in this thesis, all measurements of HPV viral load were normalized for cellular content.

20.3.3.2.3 The Amsterdam cohort study: Statistical considerations

When estimating the risk of developing high-grade CIN associated with HPV viral load, the investigators dichotomized viral load into high- and low-viral load, around the median value of the observed distribution of copy numbers (Van Duin 2002). This is poor practice from a statistical standpoint since it is “well recognized in the methodological literature that dichotomization of continuous variables introduces major problems in the analysis and interpretation of models derived in a data-dependent fashion” (Altman 2006, Royston 2006). And not only does high-viral load defined in this way have no external validity, in this case it has no internal validity either. For example, in this paper (Van Duin 2002), high-viral load is defined as 240,000 copies or more per smear when estimating the risk of disease acquisition,

whereas later in the paper the authors choose greater than 4,300,000 copies as being a high-viral load (neither normalized for cellular DNA content).

Median HPV viral loads in baseline smears were found to be significantly lower in women who cleared their infection and in those with spontaneous regression of cytological abnormality. However: methods appropriate to the analysis of time-to-event data were not used; viral loads were not normalized for cellular content; the distribution of β-globin values was not reported, so even the suboptimal analysis performed for the previous study could not be reproduced; and, again, measures of association were based on a dichotomization of viral load around the median value.

20.3.3.3 The Ludwig-McGill cohort study

The Ludwig-McGill cohort study was a prospective cohort study of “persistent” HPV infection, and its role in the natural history of cervical disease. A systematic sample of 2,746 women aged 18 to 60 years was recruited from the family medicine, gynaecology, and family planning clinics at a single municipal hospital in Sao Paolo, Brazil, a high-risk area for cervical cancer. Follow-up was every four months during the first year, and every six months thereafter. PCR was used to test for HPV DNA.

Results from this cohort study have now appeared in several reports (Franco 1999a, Franco 1999b, Rousseau 2000, Villa 2000, Rousseau 2001, Schlecht 2001, Maciag 2002, Giuliano 2003, Rousseau 2003a, Rousseau 2003b, Schlecht 2003a, Schlecht

2003b, Schlecht 2003c). There are various unusual methodological approaches used in the analysis of certain of these reports.

In one report, the authors state that "we calculated average [*sic*] viral loads for each HPV-positive subject, using results from the first two follow-up visits to reduce the level of misclassification." This approach assumes that HPV viral load is a static quantity, rather than a dynamic quantity which is capable of changing rapidly over relatively short periods, which surely prejudges the natural history of HPV infections.

In at least two reports (Schlecht 2001, Schlecht 2003c), the investigators adopted an unusual approach to dealing with “missing” values. They state that “delays in returning for a given appointment were allowed, with information and specimens collected during any post-due visit being assigned to the delayed follow-up return, which precluded the occurrence of missing interval tests”. This statement is disingenuous: this merely describes the procedure selected by the investigators for dealing with missing values, the existence of missing values was still an issue which had to be dealt with. It is only made clear in one paper that time-to-event analyses were based on the date of the visit.

In one report, when estimating measures of association between outcomes and HPV, the investigators allocated to the “exposed” category any woman who tested positive for HPV at either of her first two study visits. However, time to occurrence of an event was measured from enrolment. Thus a woman who was found to have a squamous intraepithelial lesion (SIL) at her second visit, and who first tested positive

for HPV at this time, would be considered to have had an incident SIL following HPV exposure, although both events occurred at the same visit; one report suggests that a number of incident SIL were diagnosed at the second visit. In a later report (Schlecht 2003b), only events which occurred after exposure had been defined were included in time-to-event analyses. Similarly, “all incident cases of SIL were compared with subjects who had no detected lesions or ASCUS [atypical squamous cells of undetermined significance] during the entire period of follow-up”. Compared with women with normal smears, those with ASCUS had an increased risk of SIL. The number of women with ASCUS is not stated, and it is not clear why such a heterogeneous, and therefore suboptimal, reference category was chosen.

20.3.3.4 The Kaiser Permanente cohort study

The Kaiser Permanente cohort study was a prospective cohort study investigating the natural history of HPV infection. In total, 23,702 women from the Kaiser-Permanente prepaid health plan in Oregon, USA, were followed, passively, at approximately yearly intervals as part of standard cytological screening. Hybrid capture II (probe B) was used to test for HPV, and as a semiquantitative measurement of viral load.

Results from this cohort study have now appeared in several publications (Hildesheim 1998, Liaw 2001, Castle 2002a, Castle 2002b, Lorincz 2002, Castle 2003, Sherman 2003c, Khan 2005, Castle 2005). The only real criticism of this study is the use of the hybrid capture II assay, a semiquantitative assay, to measure viral

load: for reasons described below, use of this assay meant an opportunity to make a substantial contribution to clarifying issues surrounding HPV viral load was missed. The study described in this thesis used a quantitative assay to measure HPV viral load and so *was* able to make a substantial contribution to the understanding of HPV viral load.

20.4 THE MEASUREMENT OF HPV VIRAL LOAD

There are several possible explanations for inconsistent findings between studies, including differences in study design, investigation of different disease or infection outcomes, and methodological flaws. However, one of the greatest potential sources of inconsistent findings is the assay used to measure HPV viral load. Comparisons between results based on laboratory analyses from different studies must account for differences in the performance of the assays used.

20.4.1 Criteria for an adequate assay of viral load

Since testing for HPV DNA began, several methods have been used for measuring HPV viral load, most of which were used in at least one of the reports described in the previous sections. Table 20.4.1 presents the characteristics of an ideal assay of viral load, many of which are interdependent.

Table 20.4.1. Criteria for an adequate assay of HPV viral load.

	Criteria
1	Uses stable reagents
2	Uses appropriate controls
3	High-throughput
4	Inexpensive
5	Uses minimal sample material
6	High sensitivity
7	High specificity
8	High reproducibility
9	Adequate linear dynamic range
10	Quantitative
11	Type-specific
12	Enables normalisation for sample cellular content

The first two criteria are requirements of any laboratory assay. High-throughput, inexpensive, and use of minimal sample material are important considerations in any study which calls for the analysis of large numbers of samples, and for which samples may be tested on several occasions. High sensitivity, specificity and reliability are requirements for many methods of measurement.

The linear dynamic range is the working range of the assay, the range of true viral loads for which the assay is capable of producing accurate measurements. It is the range of viral loads over which a calibration curve remains linear: for example the range within which a doubling of viral load in a target sample yields a doubling in measured viral load (see section 21.2.7 for an example from this study). Again, this requirement is common to many measurement techniques which rely upon calibration. If this range is too narrow, either it will not be possible to measure all samples, or the quantity of DNA required for the assay may be prohibitive.

Methods for measuring viral load can be quantitative, yielding an “exact” count of the number of copies of a virus present in a sample, or semiquantitative, yielding measurements on a relative scale, such as high- or low-viral load. Categorisation of a continuous quantity should be avoided until it can be demonstrated that such a transformation does not lead to an important loss of information: a quantitative method is therefore preferred initially.

Measurements of HPV viral load which combine all HPV types present in a sample into a single measurement are of limited use in natural history studies, since there are now known to be critical differences between the behaviour and pathogenic potential of the individual HPV types. If cervical cancer is a clonal disease, and 90% of cancers are associated with the detection of a single virus type (Bosch 1995), then a measure of viral load accumulated across a number of HPV types may distort, or conceal, underlying type-specific associations. A type-specific assay is therefore a requirement for all but low-resolution analyses.

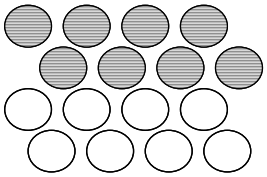
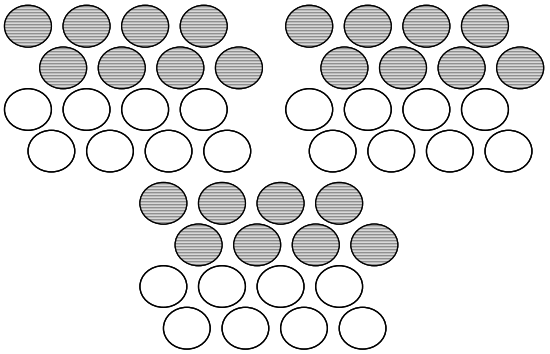
20.4.1.1 Criteria for an adequate assay of viral load: Normalisation for cellular content

To compare two measurements, it must be possible to express them using the same units of measurement. In the context of measuring viral load, this means that the measurement must account for the amount of cellular DNA present in a sample (a process often referred to as “normalisation for cellular content”). Unless an assay allows measurements of viral load to be normalised for cellular content,

interpretation of these measurements will be impossible. For example, if one sample has a viral load which is three times that of another, we cannot interpret this difference at face value, unless we know that the same number of cells are present in both samples.

Figure 20.4.1.1 illustrates the concept of normalisation of HPV viral load for cellular content. In this crude example, both sample A and sample B contain the same proportion of HPV-infected and uninfected cells, and each HPV-infected cell in each sample contains exactly the same number of copies of the HPV virus. However, since sample B contains three times as many cells as sample A, the HPV crude viral load for sample B is three times the crude viral load of sample A. Sample B would therefore be reported as containing a higher HPV viral load than sample A, simply because of a difference in cellular content. If crude viral loads are adjusted for cellular content, by dividing the crude viral load by the number of cells in each sample, the samples have identical *normalised* viral loads. Note that normalisation as described here implies that a reliable technique for counting the number of cells present in the sample is a prerequisite. Although cervical smears and cervical biopsies are taken using standardised instruments according to standard protocols, the cellular content of these samples will nevertheless vary greatly. Therefore, if it is the intention to make comparisons between viral load measurements in these samples, measurements *must* be normalised for cellular content.

Figure 20.4.1.1. Normalisation of viral load for cellular content.

	Sample A	Sample B
Uninfected cell *Cell infected by 10 virions*		
Crude viral load	*80*	*240*
No. of cells	*16*	*48*
Viral load after normalisation	*5 copies per cell*	*5 copies per cell*

20.4.2 Assays available for the measurement of HPV viral load

Early methods for measuring viral load, which now appear to have been superseded, at least in studies of HPV, include end-product PCR, PCR enzyme immunoassay, dot-blot and *in situ* hybridisation. All but *in situ* hybridisation are semiquantitative methods, and all fail to meet several of the criteria in table 20.4.1, in particular high-reproducibility (see section 21.1.2 for further discussion of the limitations of end-product PCR). *In situ* hybridisation in its most basic form involves denaturing sample DNA (i.e. separating the DNA strands so that they become “accessible”) so that the HPV DNA can be directly labelled with a probe and the labelled virus can then be counted directly. Whilst it is a quantitative method, this is a low-throughput,

insensitive, technique, with poor discrimination between samples with high-viral loads, and it requires large quantities of purified DNA (Hubbard 2003).

Methods for measuring HPV viral load which are still currently in use include hybrid capture, or rather it's successor hybrid capture II (HC2), low-stringency PCR, and real-time PCR. HC2 is a high-throughput semiquantitative method. However, although there are type-specific versions (Castle 2005), in general it is *not* type-specific, has a poor dynamic range, has low sensitivity, and does not enable normalisation for cellular content (Fontaine 2005). Low-stringency PCR, on the other hand, is a quantitative method, which is reproducible and has an adequate dynamic linear range. However, this method is labour-intensive, and although it allows for normalisation for cellular content without requiring an accurate quantification of the amount of DNA present in the initial sample, it does so by relying on measuring the ratio between the area and net intensity of a band on a gel and an arbitrary reference human band (Caballero 1995).

Real-time PCR is a quantitative method for measuring viral load, that meets all of the criteria adumbrated in table 20.4.1. Quantitative PCR (qPCR), as it is also known, and as it will be referred to throughout the remainder of this thesis, enables monitoring of the PCR, as it occurs rather than on completion of the assay as with conventional (end-point) PCR; it also allows for normalisation for cellular content. This assay is currently considered the "gold standard" for determining viral load, and was used to measure HPV viral load in the cohort study described in this thesis (see

section 21.2 for a full description of the qPCR technique used in testing study samples).

20.5 COMMENT

It is now clear that the relationship between HPV viral load and cervical disease is more complex than was previously thought. Whereas many cross-sectional studies have reported an increase in HPV viral load in cervical samples with increasing severity of cervical disease, others have found either no association, or a higher viral load in women with low-grade cervical lesions than in those with high-grade lesions. Longitudinal studies have also failed to find a consistent association between a baseline measurement of HPV viral load and: the duration of infection; the risk of clearance of disease; and the subsequent risk of acquisition of, or progression of, cervical disease. Many studies have had flaws in their design or in the way in which they were analysed, or have used a sub-optimal technique for measuring viral load.

Only well-designed, well-conducted, and correctly analysed longitudinal studies which use an appropriate quantitative assay to measure HPV viral load, are capable of providing reliable evidence regarding how cervical HPV viral load changes over time, and the role of HPV viral load in the natural history of cervical disease. I describe the analysis of measurements of the viral load of HPV16 and HPV18 in serial cervical samples taken during the follow-up of a cohort of young women recruited soon after they first had sexual intercourse and who first tested positive for HPV16 DNA, or HPV18 DNA, or both, in cervical samples, during follow-up.

Chapter 21

THE NATURAL HISTORY OF CERVICAL HPV16 AND HPV18 INFECTIONS – VIRAL LOAD

METHODS

Orientation

In this chapter I:

- describe the laboratory methods and procedures used to establish and conduct the assay used to measure HPV16 and HPV18 viral load in this thesis;
- describe the way in which the assay used to measure HPV16 and HPV18 viral load was carried out in practice;
- provide a numerical example to illustrate the data produced by the viral load assay and the method used to calculate HPV viral load;
- describe the statistical methods used in the assessment of the reliability and reproducibility of the viral load assay, and for the analysis of HPV16 and HPV18 viral load.

21.1 THE MEASUREMENT OF HPV VIRAL LOAD

21.1.1 The polymerase chain reaction

The polymerase chain reaction (PCR) is an *in vitro* method for primer-directed enzymatic amplification of specific target DNA sequences (Young 1996). An initial small quantity of target DNA sequence present in a biological specimen is selectively

amplified in a series of identical biochemical reactions to produce a final product containing an *exponentially-increased* number of copies of the target DNA sequence (Iftner 2003).

Reagents required for a PCR include: the sample potentially containing the target DNA to be amplified; a thermostable polymerase, an enzyme which mediates the synthesis of DNA; molecules corresponding to the four nucleotides of which DNA is composed; and primers (see below). The last two components are present in large quantities since they are the building-blocks for synthesising copies of the target DNA sequence: a PCR will continue to synthesise copies of the target sequence at a varying rate until these components have been exhausted (Rapley 2004).

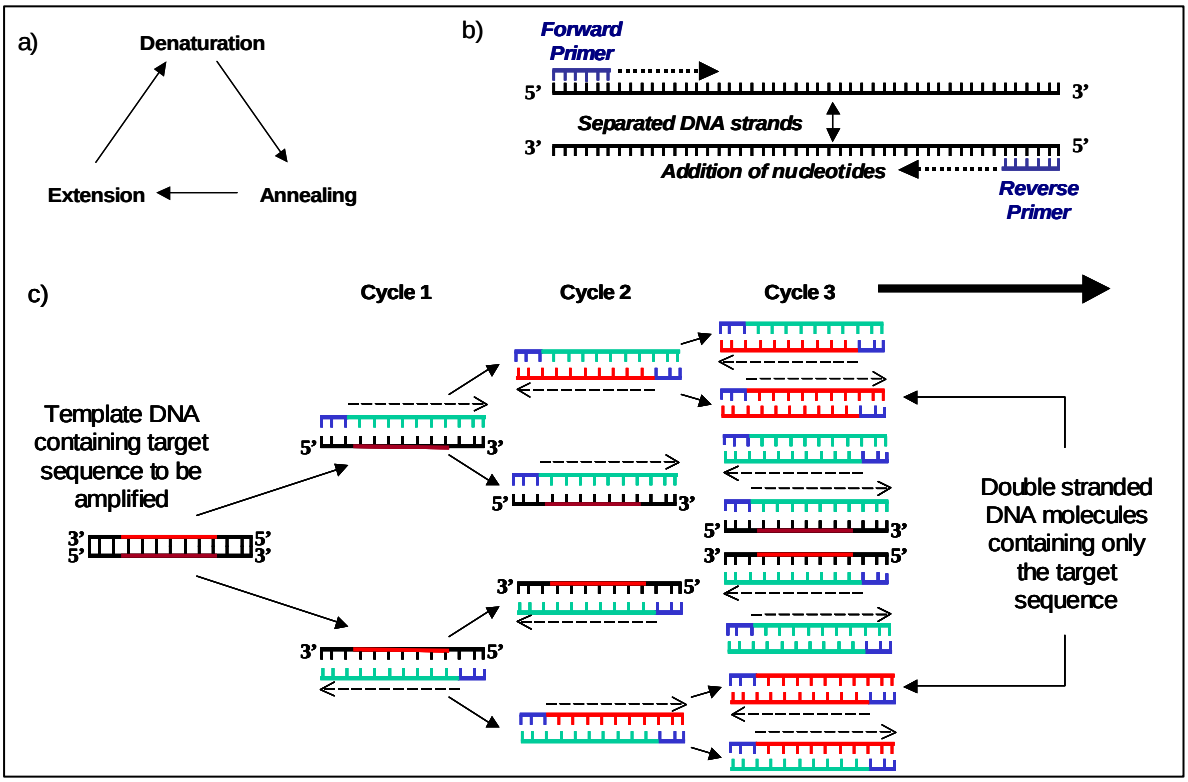
The PCR technique requires some knowledge of the DNA sequence which flanks the fragment of DNA to be amplified. Two oligonucleotide primers must be designed, one for each of the two strands of DNA, which are complementary to a sequence of DNA downstream of the target DNA. Primers are typically around 20 nucleotides in length (hence “oligonucleotide”). The most widely used PCR protocols, i.e. the set of instructions for conducting the PCR, employ consensus primers which are targeted at a highly-conserved region of the HPV L1 gene, and are therefore potentially capable of detecting all mucosal HPV types; among these are the consensus primers GP5+/GP6+ which, with further distinction of individual HPV types obtained using type-specific primers in an additional PCR, were used to test study samples from the cohort described in this thesis (see section 5.5.4.1).

A conventional (non-quantitative) PCR proceeds in a series of cycles, typically between 30 and 60 in number, each consisting of three steps carried out in order: denaturation, annealing and extension (figure 21.1.1 panel a). In the first step, the components of the reaction are heated to above 90^{o}C to denature, or separate, the DNA strands and make them accessible to the primers. In the second step, the temperature of the reaction is cooled to the annealing temperature of the primers, typically between 40^{o}C and 60^{o}C, to allow the primers to bind to their complementary sites. In the third step, *new* DNA is synthesised: both primers are extended by polymerisation until the new strands extend beyond the target DNA to be amplified; these new strands will now therefore contain a region which is complementary to the other primer, i.e. the primer which did not bind to this particular DNA strand (figure 21.1.1 panel b). This means that during the next cycle (cycle *n*, say), both the original and the new strands (strands synthesised during cycle *n-1*) will be available as templates for DNA-synthesis. As the reaction proceeds through successive cycles, all the strands present in the reaction will act as templates, and there will be an exponential increase in the amount of DNA produced (figure 21.1.1 panel c). After each cycle, there will be two new copies of the target sequence for every copy present in the previous cycle (a “single” copy comprises the pair of complementary sequences present in double-stranded DNA). Therefore after *n* cycles of a PCR, 2^{n} copies of the target sequence will have been synthesised i.e. 20 cycles is sufficient to amplify by a million (2^{20}=1,048,576) a single target sequence present in the starting sample. The vast majority of the copies present in the amplified product will have a precise length delimited exactly by the two regions complementary to the primers (the first double stranded molecules of exactly the right length appear in the

third cycle). On completion of the PCR, it can therefore be verified that the correct target sequence has been amplified, by detecting the product after electrophoresis on an agarose gel: a “band” will be produced on the gel which can be compared against controls of known sizes to ensure that the product of the PCR is of the correct size (Rapley 2004).

The sensitivity and specificity of PCR varies according to: the primers used; the size of the amplified product; reaction conditions; and the performance of the DNA polymerase. Particular care must be taken to avoid false-positive results derived from cross-contaminated specimens, or reagents.

Figure 21.1.1. The polymerase chain reaction.



21.1.2 Quantitative PCR

It may be thought reasonable to assume that the amount of target DNA present in the final product of a conventional PCR, as described in the previous section, has a deterministic relationship with the amount of target DNA present in the starting sample. For example, it might be anticipated that the brightness, or vertical extent, of a band on an agarose gel following electrophoresis would be directly-proportional to the amount of target DNA initially present in the target sample, thus providing at least a semiquantitative measurement of the viral load of that sample. However, over a large range of starting quantity of target DNA, whatever the starting quantity, the final quantity of target DNA amplified by the PCR will be the same (this is determined largely by the amount of primers added to the reaction). Conventional PCR, which is an endpoint assay only, i.e. the result is only obtained after the assay has been completed, is thus unable to distinguish even large differences in viral load. The result of a conventional PCR test is thus essentially only interpretable as “positive” or “negative”.

However, the addition of a fluorescent probe to a PCR enables the detection of the accumulation of copies of the target DNA sequence in a PCR in *real time* i.e. whilst the assay is underway, and thereby enables the quantification of the amount of target DNA present in the initial sample. A PCR undertaken in this way is known as a quantitative PCR (qPCR) (it is also known as “real-time” PCR), and provides both a positive or negative status, and an estimate of the number of copies of target DNA

present in a sample. If the target DNA sequence is a unique sequence within a HPV virus genome, qPCR can thus be used to determine HPV viral load.

The principle behind the use of a fluorescent probe, as used in the study described in this thesis, is illustrated in figure 21.1.2. A DNA-probe is designed which is complementary to a sequence within the region amplified by the primers. The probe has a fluorescent "reporter" at one end, and a "quencher" at the other. When the reporter and the quencher are in close proximity, as they are initially before the reaction begins (at which point the probe is intact), the quencher prevents detection of the fluorescence produced by the reporter. If the reporter is moved away from the quencher, the influence of the quencher is attenuated, and the reporters' fluorescence becomes detectable.

A qPCR assay is conducted in essentially the same way as a conventional PCR, with the probe as an additional component of the reaction. As a new DNA strand is synthesised from the primers by polymerisation during the PCR, the new strand increases in length until eventually it reaches the position of the probe. The polymerase used in the PCR has two actions: as exploited in a conventional PCR, it adds additional nucleotides to an existing primer (hence “polymerase”); and it also removes any nucleotides in the region of the site of activity (called its “exonuclease activity”). Hence, once the newly produced DNA strand reaches the probe, the polymerase starts to excise (remove) the nucleotides forming the probe so that they can be replaced by new nucleotides. Since the fluorescent reporter is attached to the first part of the probe reached by the polymerase, it is excised first. Thereafter, it

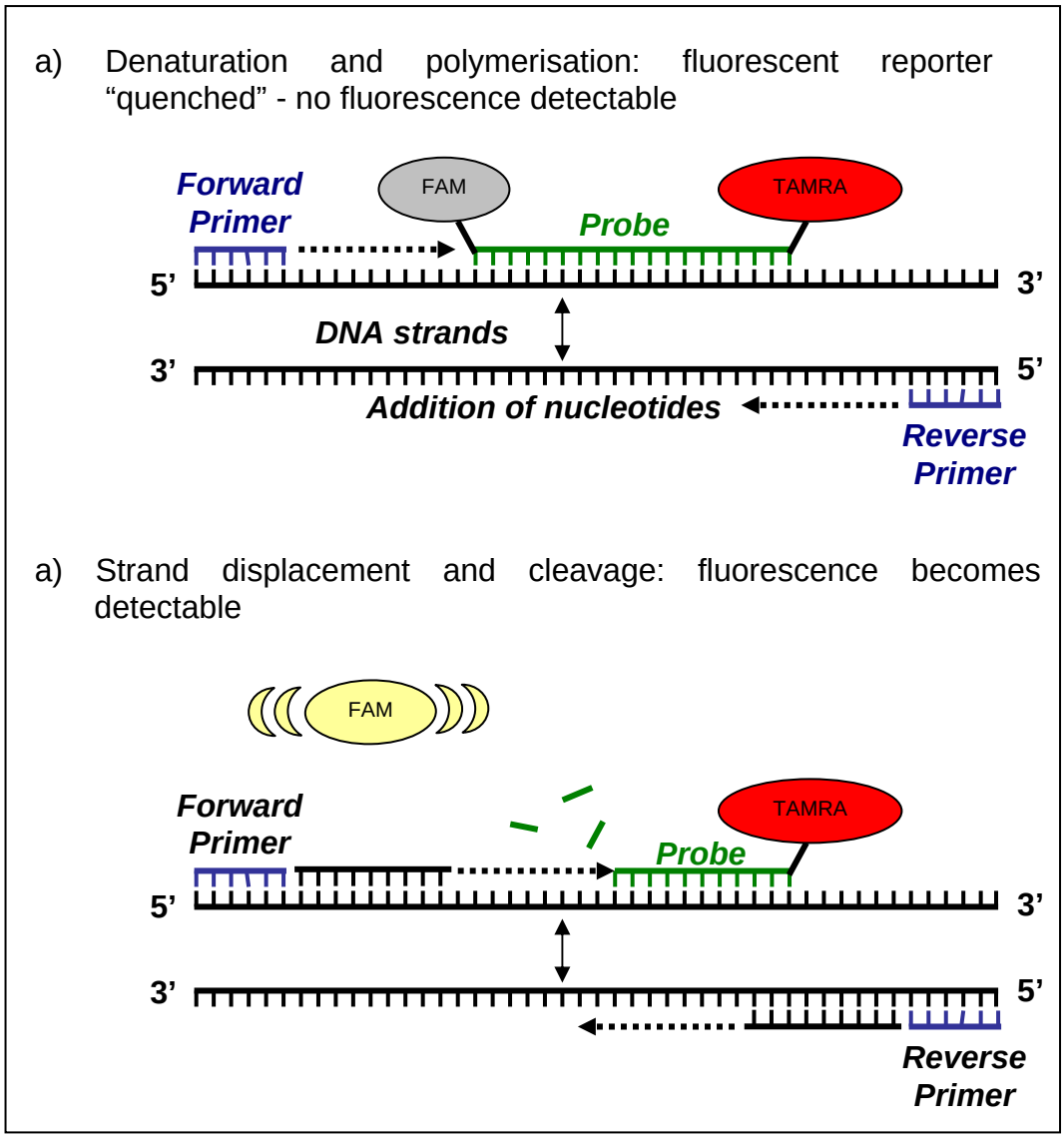
moves away from the quencher, and begins to fluoresce unhindered by the quencher. In this way, an increase in the number of copies of target DNA during each PCR cycle causes a proportional increase in fluorescence, which can be detected and quantified. The number of cycles at which fluorescence can first be detected, the cycle threshold, or “Ct”, value, is proportional to the initial target DNA concentration. If a “standard” sample, i.e. one with known amounts of target DNA, is included in the reaction together with a sample containing an unknown amount, it is then possible to quantify the number of copies of target DNA present in that unknown sample.

Note that the use of the fluorescent DNA probe not only enables quantification of the target DNA sequence, it also increases the specificity of the qPCR assay compared to conventional PCR. No assay is perfect, and one of the main threats to a PCR is non-specific binding, i.e. one or more of the primers bind to the wrong part of the DNA template, and hence amplify the incorrect DNA sequence. However, the probability of non-specific binding by both the primers *and* the probe in a qPCR is extremely low, and certainly less than the probability of non-specific binding by the primers alone in a conventional PCR.

It is possible to quantify several target DNA sequences simultaneously in the same PCR: such an assay is known as a *multiplex* assay. However, in the study described in this thesis, only one HPV type at a time was assayed in each qPCR i.e. a *singleplex* assay. However, the number of copies of two target DNA sequences present in the same patient sample were quantified in separate qPCR conducted on

the same well-plate, as will be described below. This was to enable measurements of HPV viral load to be adjusted (normalised) for the number of cells present in the sample, and thus to be expressed as copy number per 1,000 cells.

Figure 21.1.2. The principle behind the use of a fluorescent probe in a quantitative PCR assay.



21.2 THE MEASUREMENT OF HPV16 AND HPV18 VIRAL LOAD IN STUDY SAMPLES

Having described the principle behind the qPCR assay, I will now describe how it was conducted in practice for the analysis of patient samples from the cohort study described in this thesis; a formal technical description appears in appendix 7.

21.2.1 Sample Preparation

DNA extraction

A PCR assay cannot be performed on a cell which is structurally intact: the DNA content must first be separated from the remainder of the cell. This was done using a standard procedure, although in this instance the laboratory scientist refined the technique to ensure the maximum yield of DNA from the samples. Prior to testing study samples for HPV viral load, all of the DNA in that sample was extracted and stored under appropriate conditions.

Quality of the samples: DNA yield

Ideally, when a sample of cells is taken for subsequent DNA analysis, the cells should be stored immediately at -80°C, or ideally in liquid nitrogen, until the DNA can be extracted. Unfortunately, this was not the case for the cervical cytological samples taken in this study. All samples were initially stored as intact cellular

material at -20°C. Although storage at this temperature retards the activity of enzymes present in the sample (e.g. DNAses, RNAses, proteases, etc.), and therefore slows down the degradation of the samples, storage at -20°C does not arrest degradation completely. In addition, every time a sample is thawed (a necessary step prior to taking a subsample for a specific laboratory test), the activity of these enzymes increases. And samples sometimes arrived in the laboratory in poor condition, a result of being stored in apparently suboptimal conditions prior to being sent to the laboratory for final storage. Therefore, over time the integrity of the samples was deteriorating even when apparently lying dormant, and this is likely to have had an impact on the DNA yield from these samples.

21.2.2 Primer design

Initially, the primers and probes used for measuring HPV16 viral load were those designed by Gravitt *et al.* (Gravitt 2003). However, it was found that these primers were not specific, and also amplified a product [including the target DNA sequence] exceeding the length recommended by the manufacturers of the qPCR-kit. Therefore, for the analysis of samples from the cohort described in this thesis, new (optimised) primers and probes were designed for both HPV16 and HPV18. As well as being designed to amplify only the target DNA sequence of interest, primers for use in a PCR, quantitative or otherwise, must be designed within certain technical constraints; similarly for the fluorescent probe, which is specific to *q*PCR. In this study, primers and probes were designed with the constraints specified by the manufacturer as a guide only: primer parameters are presented in table A4a of

appendix 4. The HPV16 primers were designed to amplify the HPV16 E6 gene; the HPV18 primers were designed to amplify the HPV18 E7 gene.

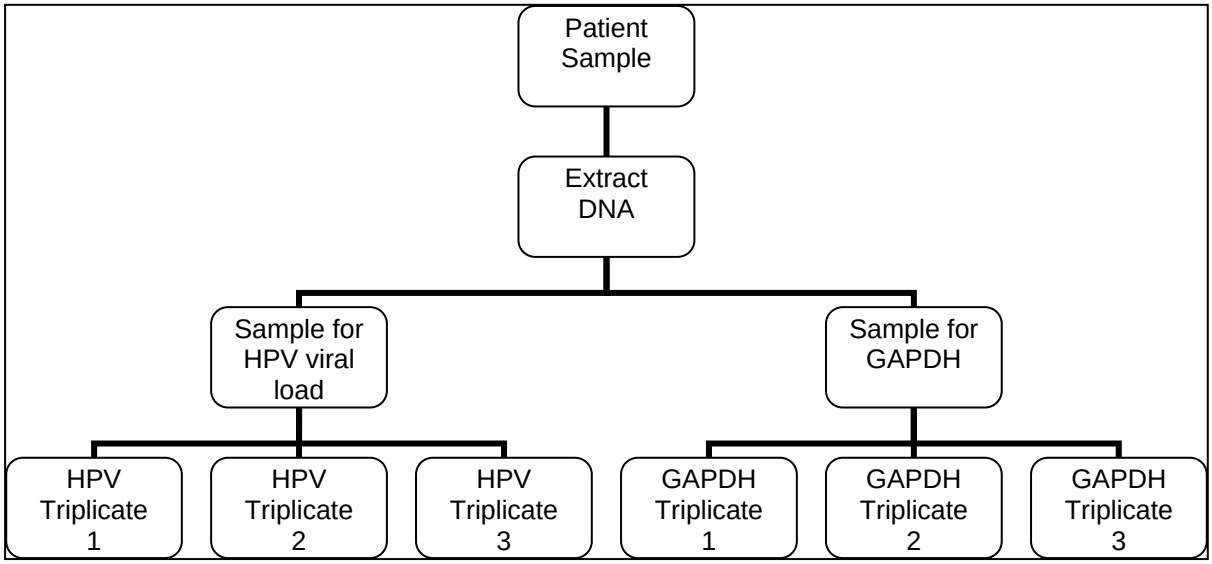
21.2.3 Generation of standard curves

Plasmids were used to generate standard curves for HPV16, HPV18, and the GAPDH housekeeping gene (the use of which is described subsequently). A plasmid is a small extra-chromosomal DNA molecule which is capable of independent replication. In brief, the gene of interest is incorporated into the plasmid, which is then inserted into bacterial “vectors”. Bacteria multiply at an extremely high rate, i.e. they undergo rapid cell-division, during which all DNA in the “parent” cell is copied and carried into “daughter” cells. Therefore, very large numbers of copies of a gene can be produced relatively easily. Commercially-available vectors were used for producing plasmids containing the HPV16 and HPV18 genomes. However, the production of the GAPDH plasmid was undertaken in-house, using a cell-free method (see table A4b of appendix 4 for PCR primer parameters).

21.2.4 Well-plate configuration

In this study only one HPV type at a time was assayed in each qPCR. However, each sample standard and control was tested three times in the same qPCR assay, with the result from each test subsequently being combined. Each patient sample was thus tested six times in the same qPCR assay. The procedure for testing a patient sample, for example, is illustrated in figure 21.2.4a.

Figure 21.2.4a. Description of the method for obtaining triplicate patient samples for GAPDH and HPV in the qPCR assay.



The qPCR assay was conducted using a 96-well well-plate, called a TaqMan® plate. Each well-plate can be considered to comprise a grid of eight rows and 12 columns of wells. A typical well-plate configuration for conducting the HPV16 assay is shown in table 21.2.4; the well-plate configuration for the HPV18 assay was similar. The wells can be divided into seven groups, each with a distinct purpose.

1) HPV type-specific plasmids

These wells are used for generating a standard curve for estimating the number of copies of HPV which were present. They hold serial ten-fold dilutions of plasmids containing a genome which incorporates the relevant portion of the HPV16, or HPV18, genome, as appropriate: therefore, within limits, these wells contain known numbers of copies of the HPV genome of the relevant type.

2) GAPDH plasmids

These wells are used for generating a standard curve for estimating the number of cells which are present in a sample. They hold serial ten-fold dilutions of plasmids containing a genome which incorporates the GAPDH housekeeping gene: therefore, within limits, these wells contain known numbers of copies of the GAPDH gene.

3) Patient samples: HPV

These wells are used for estimating the number of copies of HPV which are present in patient samples; they hold genomic DNA from patient samples. The genomic DNA will be a mixture of DNA from normal cells, DNA from cells with hyperploidy (i.e. an abnormal increase in the number of chromosomes usually found in a cell), and HPV DNA (if the sample was taken from an HPV-infected woman).

4) Patient samples: GAPDH

These wells are used for estimating the number of copies of the GAPDH housekeeping gene which are present in patient samples; they hold genomic DNA from patient samples.

5) Positive-control: HPV

These wells are used for estimating the number of copies of HPV which are present in the positive-control, with the objective of determining whether the assay has worked correctly, and therefore whether the results of the assay can be accepted. They contain DNA from a cervical carcinoma cell line with a known range of copy numbers per cell of the HPV type of interest. In assays used to measure HPV16 viral load, DNA from the HPV16-positive cervical cancer cell line SiHa was used as a positive internal control; for HPV18, the HPV18-positive cervical cancer cell line HeLa was used. SiHa has been reported to contain between one and 10 copies of the HPV16 genome per cell; HeLa between one and 50 copies of the HPV18 genome per cell. If the measured viral load per cell for a positive-control was not consistent with the previously reported range, the results of the assay were rejected.

6) Positive-control: GAPDH

These wells are used for estimating the number of copies of the GAPDH housekeeping gene which are present in patient samples; they hold DNA from a cell line containing known copy numbers of the HPV type of interest. Note that, although the number of copies of the virus per cell present in the sample is assumed known (i.e. within a known range), the number of cells in the sample used to measure that number is still unknown.

7) Negative-control: water

These wells are used to determine whether the assay has worked correctly, and therefore whether the results of the assay can be accepted. When preparing all wells on the well-plate, the same supply of DNase/RNase-free water was used to dilute samples to the necessary concentration. Therefore, any contamination of this water source, or of the equipment used to dilute the samples, would affect *all* samples. These wells are used to detect any such contamination. There are six of these wells in table 21.2.4, labelled in two groups: however, these wells essentially provide six identical negative control measurements.

Table 21.2.4. Well-plate configuration for conducting the HPV16 qPCR assay using a 96-well well-plate. Scientific notation is used, i.e. 10e*N* is equivalent to 10^N.

HPV16 Plasmid 10e8 copies	HPV16 Plasmid 10e8 copies	HPV16 Plasmid 10e8 copies	HPV16 Plasmid 10e7 copies	HPV16 Plasmid 10e7 copies	HPV16 Plasmid 10e7 copies	HPV16 Plasmid 10e6 copies	HPV16 Plasmid 10e6 copies	HPV16 Plasmid 10e6 copies	HPV16 Plasmid 10e5 copies	HPV16 Plasmid 10e5 copies	HPV16 Plasmid 10e5 copies
HPV16 Plasmid 10e4 copies	HPV16 Plasmid 10e4 copies	HPV16 Plasmid 10e4 copies	HPV16 Plasmid 10e3 copies	HPV16 Plasmid 10e3 copies	HPV16 Plasmid 10e3 copies	HPV16 Plasmid 10e2 copies	HPV16 Plasmid 10e2 copies	HPV16 Plasmid 10e2 copies	HPV16 Water	HPV16 Water	HPV16 Water
HPV16 Patient Sample 1 50 ng gDNA[a]	HPV16 Patient Sample 1 50 ng gDNA	HPV16 Patient Sample 1 50 ng gDNA	HPV16 Patient Sample 2 50 ng gDNA	HPV16 Patient Sample 2 50 ng gDNA	HPV16 Patient Sample 2 50 ng gDNA	HPV16 Patient Sample 3 50 ng gDNA	HPV16 Patient Sample 3 50 ng gDNA	HPV16 Patient Sample 3 50 ng gDNA	HPV16 Patient Sample 4 50 ng gDNA	HPV16 Patient Sample 4 50 ng gDNA	HPV16 Patient Sample 4 50 ng gDNA
HPV16 Patient Sample 5 50 ng gDNA	HPV16 Patient Sample 5 50 ng gDNA	HPV16 Patient Sample 5 50 ng gDNA	HPV16 Patient Sample 6 50 ng gDNA	HPV16 Patient Sample 6 50 ng gDNA	HPV16 Patient Sample 6 50 ng gDNA	HPV16 Patient Sample 7 50 ng gDNA	HPV16 Patient Sample 7 50 ng gDNA	HPV16 Patient Sample 7 50 ng gDNA	HPV16 SiHa 50 ng gDNA	HPV16 SiHa 50 ng gDNA	HPV16 SiHa 50 ng gDNA
GAPDH Plasmid 10e8 copies	GAPDH Plasmid 10e8 copies	GAPDH Plasmid 10e8 copies	GAPDH Plasmid 10e7 copies	GAPDH Plasmid 10e7 copies	GAPDH Plasmid 10e7 copies	GAPDH Plasmid 10e6 copies	GAPDH Plasmid 10e6 copies	GAPDH Plasmid 10e6 copies	GAPDH Plasmid 10e5 copies	GAPDH Plasmid 10e5 copies	GAPDH Plasmid 10e5 copies
GAPDH Plasmid 10e4 copies	GAPDH Plasmid 10e4 copies	GAPDH Plasmid 10e4 copies	GAPDH Plasmid 10e3 copies	GAPDH Plasmid 10e3 copies	GAPDH Plasmid 10e3 copies	GAPDH Plasmid 10e2 copies	GAPDH Plasmid 10e2 copies	GAPDH Plasmid 10e2 copies	GAPDH Water	GAPDH Water	GAPDH Water
GAPDH Patient Sample 1 50 ng gDNA	GAPDH Patient Sample 1 50 ng gDNA	GAPDH Patient Sample 1 50 ng gDNA	GAPDH Patient Sample 2 50 ng gDNA	GAPDH Patient Sample 2 50 ng gDNA	GAPDH Patient Sample 2 50 ng gDNA	GAPDH Patient Sample 3 50 ng gDNA	GAPDH Patient Sample 3 50 ng gDNA	GAPDH Patient Sample 3 50 ng gDNA	GAPDH Patient Sample 4 50 ng gDNA	GAPDH Patient Sample 4 50 ng gDNA	GAPDH Patient Sample 4 50 ng gDNA
GAPDH Patient Sample 5 50 ng gDNA	GAPDH Patient Sample 5 50 ng gDNA	GAPDH Patient Sample 5 50 ng gDNA	GAPDH Patient Sample 6 50 ng gDNA	GAPDH Patient Sample 6 50 ng gDNA	GAPDH Patient Sample 6 50 ng gDNA	GAPDH Patient Sample 7 50 ng gDNA	GAPDH Patient Sample 7 50 ng gDNA	GAPDH Patient Sample 7 50 ng gDNA	GAPDH SiHa 50 ng gDNA	GAPDH SiHa 50 ng gDNA	GAPDH SiHa 50 ng gDNA

[a]gDNa=genomic DNA.

21.2.5 qPCR conditions

Optimal qPCR conditions were determined and pre-programmed into a machine called a thermal cycler. The thermal cycler conducts the qPCR automatically, monitors accumulating fluorescence, and outputs the data to a computer in a format suitable for analysis by specialised software.

During the initial cycles in a qPCR, there is no detectable increase in fluorescence: the cycle number at which fluorescence can first be detected (the cycle threshold, or Ct, value) is determined automatically.

21.2.5.1 The Ct value and a qPCR-negative result

The number of cycles in the qPCR assay used in this thesis was arbitrarily chosen to be 50 (typically, qPCRs use between 30 and 70 cycles). This means that any sample whose Ct value wass 50 would be measured as having a viral load of zero. This may have been because the sample was in fact HPV-negative for the specific type assayed; or because the number of copies of HPV present in the sample was very low, and the number of cycles undertaken was insufficient to produce detectable fluorescence. Thus “qPCR-negative” (or, equivalently, a viral load of zero) refers to an arbitrary cut-off value. However, this is no different from a “negative” in a conventional PCR which also uses an arbitrary, and finite, number of cycles.

21.2.6 Determination of virus copy numbers and cell copy numbers

Inverse regression was used to obtain estimates of the number of copies of HPV (the HPV viral load) which were present in cervical samples taken from patients in the study samples. A *standard curve* for the specific virus type assayed was obtained by plotting the logarithm of the observed Ct values from the triplicate measurements made on the standard virus plasmid controls, against the “known” input concentration (serial ten-fold dilutions containing “known” copy numbers of the

relevant HPV-type were used). A line of best-fit was then estimated using ordinary least-squares regression. To quantify viral load in a target sample, the mean of the Ct values from each of the accepted triplicate wells was referred to the fitted regression line, and the corresponding concentration for that Ct value was determined. The measurement so obtained was then used as an estimate of the total number of copies of the virus genome present in the target sample.

To “normalise” virus copy number for specimen cellular content, the number of copies of the GAPDH gene present in the target sample was estimated in the same way, but using the standard GAPDH plasmid controls. GAPDH is a so-called “housekeeping” gene which is assumed to be present in the genome of all human cells, with one copy of this gene per DNA strand: each host-cell of the target sample was therefore assumed to contain two copies of the GAPDH gene. Therefore, since each normal cell contains two strands of DNA, and each strand of DNA contains a single copy of the GAPDH gene, the number of cells present in the target sample was estimated by dividing the estimate of the number of copies of the GAPDH gene which were present by two. The viral load of the target sample, normalised or “corrected” for cellular content, was then obtained by dividing the estimate of the number of copies of the virus which were present in the target sample, by the estimate of the number of cells which were present in the target sample.

21.2.6.1 GAPDH failures

The GAPDH gene sometimes failed to amplify for a particular sample: this could indicate that the sample contained no DNA, or that a problem had occurred with the PCR-amplification of that sample. On repeating the assay, sometimes the problem could be resolved, but not always. If GAPDH failed to amplify for a sample, the sample could not be evaluated for viral load since it could not be normalised for cellular DNA content, unless the result for the virus indicated that it was negative.

21.2.7 Numerical example

A numerical example is the best way to demonstrate the procedure for obtaining estimates of HPV viral load.

In this example, numbers are approximate: more accurate values were used in the actual calculations for study samples. Examples of the two results tables from the qPCR assay, as supplied to the candidate by the laboratory scientist following preliminary analysis, are presented in tables 21.2.7a and 21.2.7b. Five patient samples were tested in this assay (study ID's 247.5, 247.6, 247.7, 247.8 and 247.9).

1) The first thing to notice is that all three of the negative-control wells for GAPDH are negative, as are all three of the negative-control wells for HPV16 (the row labelled "water": Ct value=50, in all cases). Thus the results of this assay are not rejected on the grounds of contamination.

2) Each well-plate yields three Ct values for each of the samples on that well-plate. Ultimately, a mean Ct value across the replicates for each sample was used as the observation for that sample. However, only Ct values which did not display excessive variability amongst replicates were retained in the analysis. “Excessive variability” was defined as follows. If the standard deviation among the three Ct values was 1.5 or more, the most extreme Ct value would be excluded; the remaining two Ct values would then be used to form the mean Ct value for that sample. Apparently, this is a common practice in other laboratory assays where triplicates are used. However, the use of a standard deviation of 1.5 is clearly arbitrary, although more stringent than others would normally use within the laboratory where the qPCR assay described in this thesis was undertaken. That being said, this procedure typically led to the exclusion of a clear outlier. The same rule was then applied to the remaining two Ct values: if the standard deviation among the remaining two Ct values was 1.5 or more, the results for this sample were excluded, and the sample was tested again. Where the standard deviation among the three Ct values exceeded 1.5, and there was more than one choice of pairs of Ct values for which the standard deviation among the remaining pair of Ct values did not exceed 1.5 (i.e. where it is was not clear-cut which Ct value should be excluded), the pair of Ct values yielding the smallest standard deviation was used. If the Ct values measuring the GAPDH copy number suggested the result should be excluded, but the Ct values measuring the virus copy number indicated a zero viral load (i.e. that a sample which was negative for the specific HPV type), the zero viral load result was accepted. In addition, Ct values of 50 required special handling. A Ct value of 50 was not included in calculations of standard deviations since these

represented either a specific value of viral load (zero), or a censored value i.e. a viral load below the limits of detection. If two of the triplicates had Ct values of 50, the third (non-50) Ct value was excluded. If two triplicates had non-50 Ct values, the Ct value of 50 was excluded. These rules were applied for all types of sample on the well-plate, but excluding Ct values from samples used to fit the standard curves had a greater impact on estimates of viral load than did excluding Ct values which only affected the results of an individual patient or control sample.

In this example, none of the values from the GAPDH analysis require further manipulation. In the HPV16 analysis (table 21.2.7b), Ct value 2 for patient sample 247.6 was excluded, making this sample HPV16-qPCR-negative; and Ct value 1 for patient sample 247.9, indicating a HPV-negative result, was also excluded.

Table 21.2.7a. Example of the Ct values output from the qPCR assay for plate VL061: GAPDH. Scientific notation is used, i.e. 10e*N* is equivalent to 10^N.

Sample	Ct Value 1	Ct Value 2	Ct Value 3	Mean Ct value	Ct value SD[a]
Plasmid: 10e8 copies[b]	14.96	14.94	14.80	14.90	0.09
Plasmid: 10e7 copies	17.09	16.74	17.06	16.96	0.19
Plasmid: 10e6 copies	19.92	20.36	20.52	20.27	0.31
Plasmid: 10e5 copies	23.18	23.11	23.26	23.18	0.08
Plasmid: 10e4 copies	25.47	26.22	25.95	25.88	0.38
Plasmid: 10e3 copies	28.30	28.51	28.41	28.41	0.11
Plasmid: 10e2 copies	33.39	33.77	33.91	33.69	0.27
Water	50.00	50.00	50.00	50.00	NA
SiHa	23.71	23.72	23.75	23.73	0.02
Patient sample 247.5	31.09	29.78	29.62	30.16	0.81
Patient sample 247.6	42.45	40.10	40.91	41.15	1.19
Patient sample 247.7	29.11	29.24	29.06	29.14	0.09
Patient sample 247.8	25.74	25.46	25.37	25.52	0.19
Patient sample 247.9	24.71	24.71	24.47	24.63	0.14

[a]Standard deviation; [b]Scientic notation used: 10e8 = 10 x 10^8

Table 21.2.7b. Example of the Ct values output from the qPCR assay for plate VL061: HPV16. Scientific notation is used, i.e. 10e*N* is equivalent to 10^N.

Sample	Ct Value 1	Ct Value 2	Ct Value 3	Mean CT value	Ct value SD[a]
Plasmid: 10e8 copies[b]	14.28	14.19	14.23	14.23	0.05
Plasmid: 10e7 copies	18.30	18.28	18.20	18.26	0.05
Plasmid: 10e6 copies	21.76	21.69	21.67	21.71	0.05
Plasmid: 10e5 copies	27.20	26.56	27.23	27.00	0.38
Plasmid: 10e4 copies	29.86	29.49	29.62	29.66	0.19
Plasmid: 10e3 copies	32.77	33.24	33.20	33.07	0.26
Plasmid: 10e2 copies	36.25	36.11	36.87	36.00	0.40
Water	50.00	50.00	50.00	50.00	NA
SiHa	23.76	23.68	23.76	23.73	0.05
Patient sample 247.5	21.80	21.58	21.33	21.57	0.24
Patient sample 247.6	50.00	46.71	50.00	50.00	NA[c]
Patient sample 247.7	31.07	31.23	31.11	31.14	0.08
Patient sample 247.8	33.02	33.22	32.20	32.81	0.54
Patient sample 247.9	50.00	39.56	39.11	39.34	0.32[c]

[a]Standard deviation; [b]Scientic notation used: 10e8 = 10 x 10^8; [c]Excluding the outlier Ct value.

3) Next two standard curves were generated based on the observed Ct values from the wells containing known copy numbers of plasmids, one for GAPDH (figure 21.2.7a) and one for the virus (figure 21.2.7b). Each (accepted) Ct value was plotted against the natural logarithm of the known copy number, and a straight line was then fitted using ordinary least squares regression.

Figure 21.2.7a. Example standard curve from the qPCR assay for plate VL061: GAPDH.

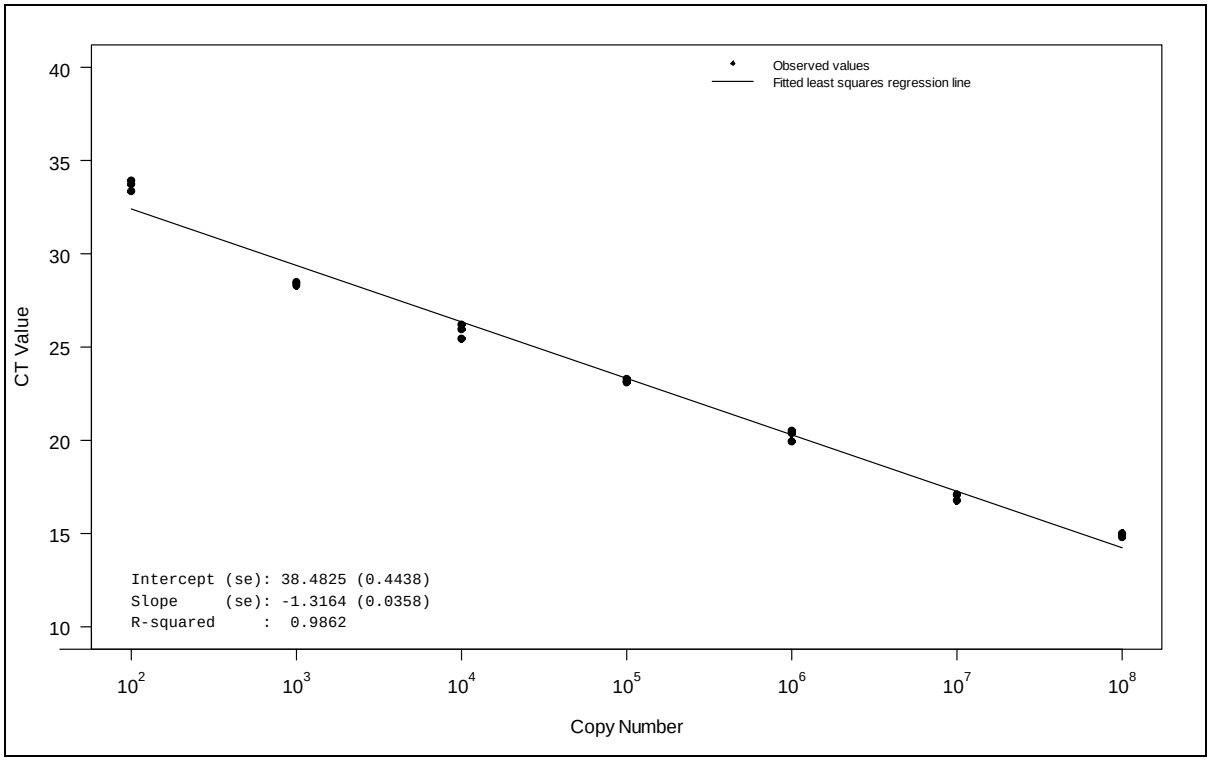
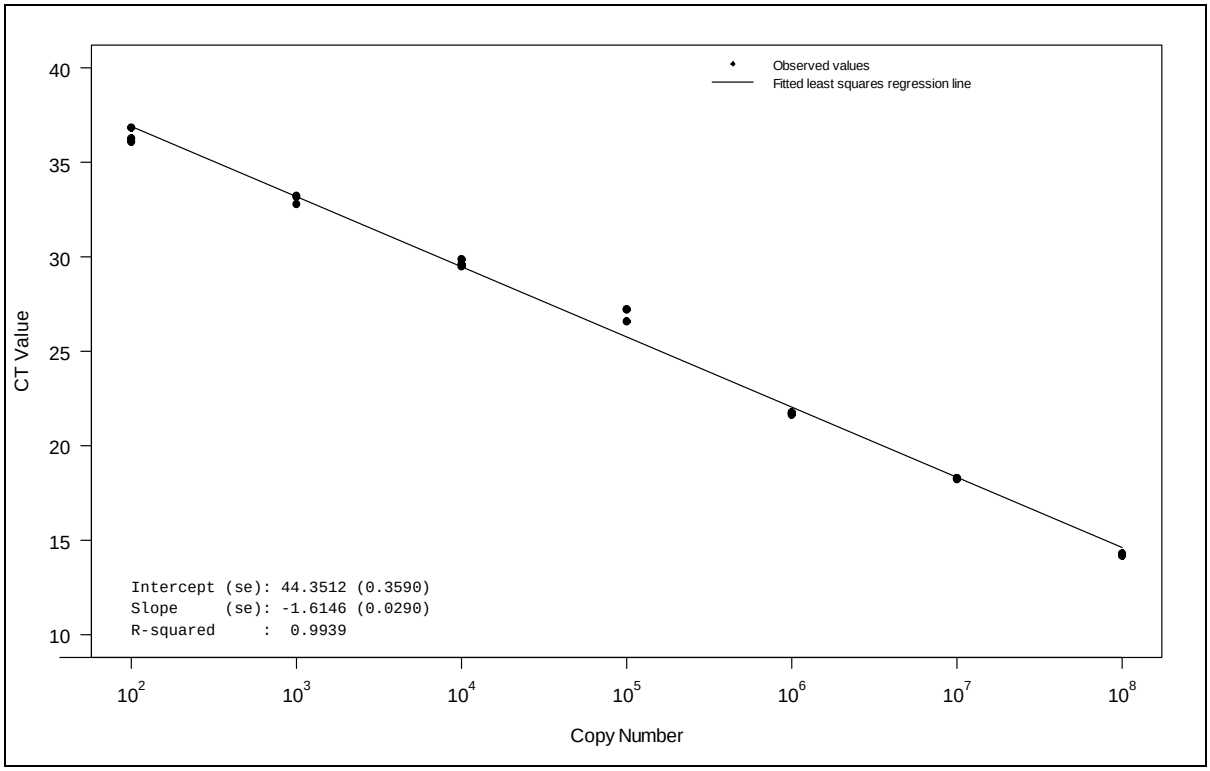


Figure 21.2.7b. Example standard curve from the qPCR assay for plate VL061: HPV16.



Slopes and intercepts of the regression curves were estimated using the internal functions in Microsoft Excel. These were subsequently verified using the R statistical package, with no differences being found. The slope and intercept were taken from the equation displayed on an Excel chart, which are reported to four decimal places, and were used in calculations.

In this case the equation of the GAPDH standard curve is

$$\text{Ct Value}_{GAPDH} = 38.4250 - (1.3164 \times \ln(\text{Copy Number}_{GAPDH}))$$

Hence:

$$\text{Copy Number}_{GAPDH} = \exp((38.4250 - \text{Ct Value}_{GAPDH})/1.3164) \qquad (1)$$

Similarly, for HPV16:

$$\text{Copy Number}_{HPV16} = \exp((44.3512 - \text{Ct Value}_{HPV16})/1.6146) \qquad (2)$$

Given a Ct value from a sample containing an unknown number of copies of HPV16, or of GAPDH, we are now in a position to estimate this unknown quantity.

4) The first unknown copy number which must be calculated is that for the positive-control. If the measured viral load per cell for the positive-control is not consistent

with the acceptable range for the cell line, the results of the assay must be rejected and further calculations are unnecessary.

The mean Ct value for HPV16 for the positive-control is 23.73. Hence the number of copies of virus present in this sample is 352,113.26 (from equation 2 above). By itself, this value is of little use: we need to know the number of cells present in that sample. The mean Ct value for GAPDH for the positive-control is also 23.73 (unusual, but by no means a unique phenomenon). Hence the number of copies of GAPDH present in this sample is 70,475.50 (from equation 1 above). If it is assumed that each cell has two copies of the GAPDH gene (one copy per DNA strand), then there are 35,248 cells (70,475.50/2) present in this sample. Hence, this sample of the SiHa cell line contains 10 copies (352,113.26/35,248) of HPV16 per cell, consistent with the previously reported range of one to 10. Hence the results of this assay are not rejected due to failure of the positive-control, and the results of the assay can be accepted.

5) The number of copies of HPV16 present in the sample can now be estimated, and further expressed as copy number per cell. Figure 21.2.7c illustrates this process for GAPDH, and figure 21.2.7d for HPV16 for the patient samples tested in this qPCR assay. Clearly, for sample 247.6 and GAPDH, it was necessary to extrapolate beyond the range of copy numbers used to generate the standard curve; similarly for sample 247.9 and HPV16, but the extrapolation is not as extreme in this case.

Figure 21.2.7c. Illustration of the calculation of the GAPDH copy number from the standard curve for patient samples on qPCR well-plate VL061. The horizontal numbers in green are patient sample IDs; the vertical numbers in green are estimated ln(copy number) for those samples.

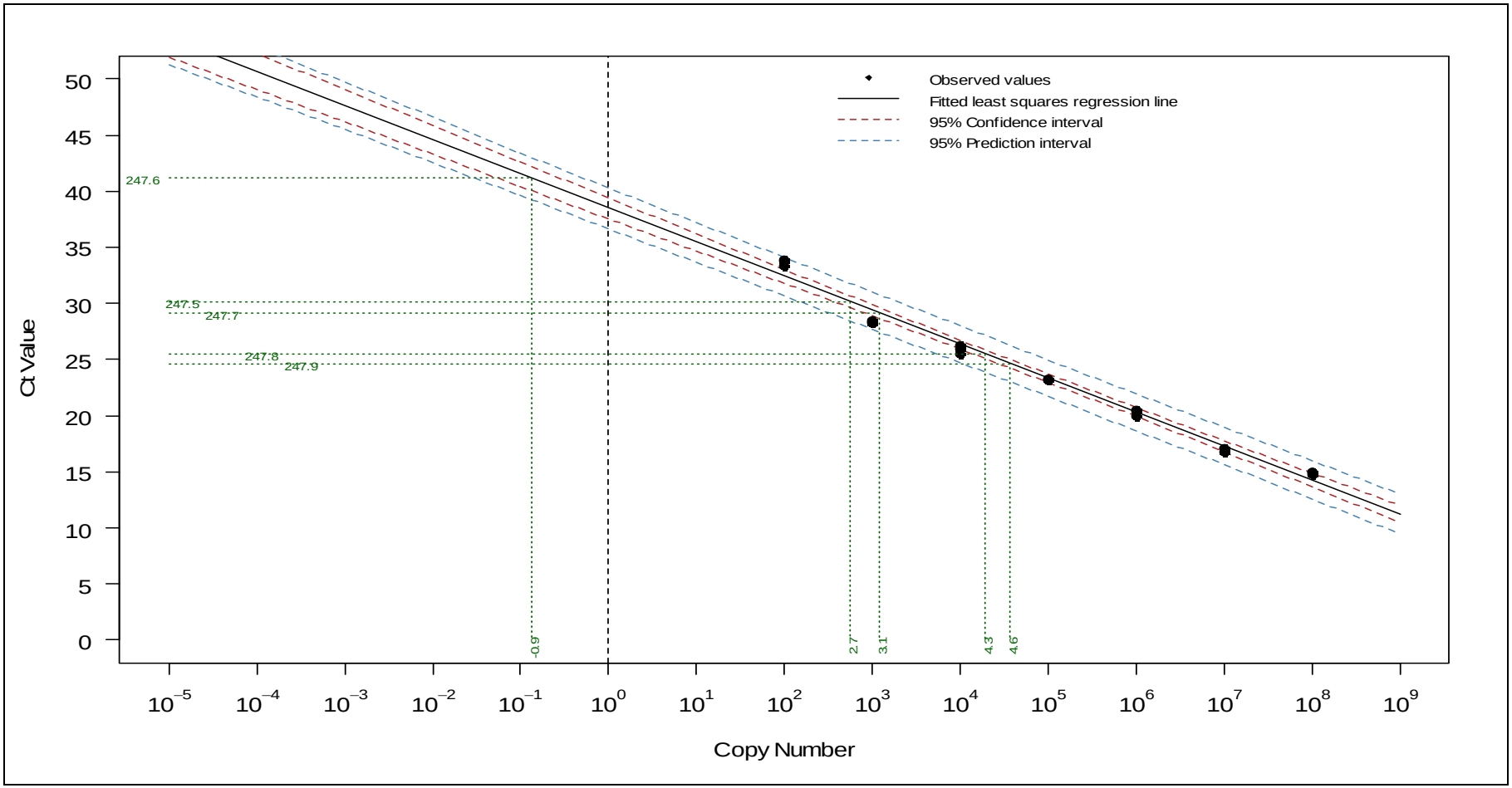
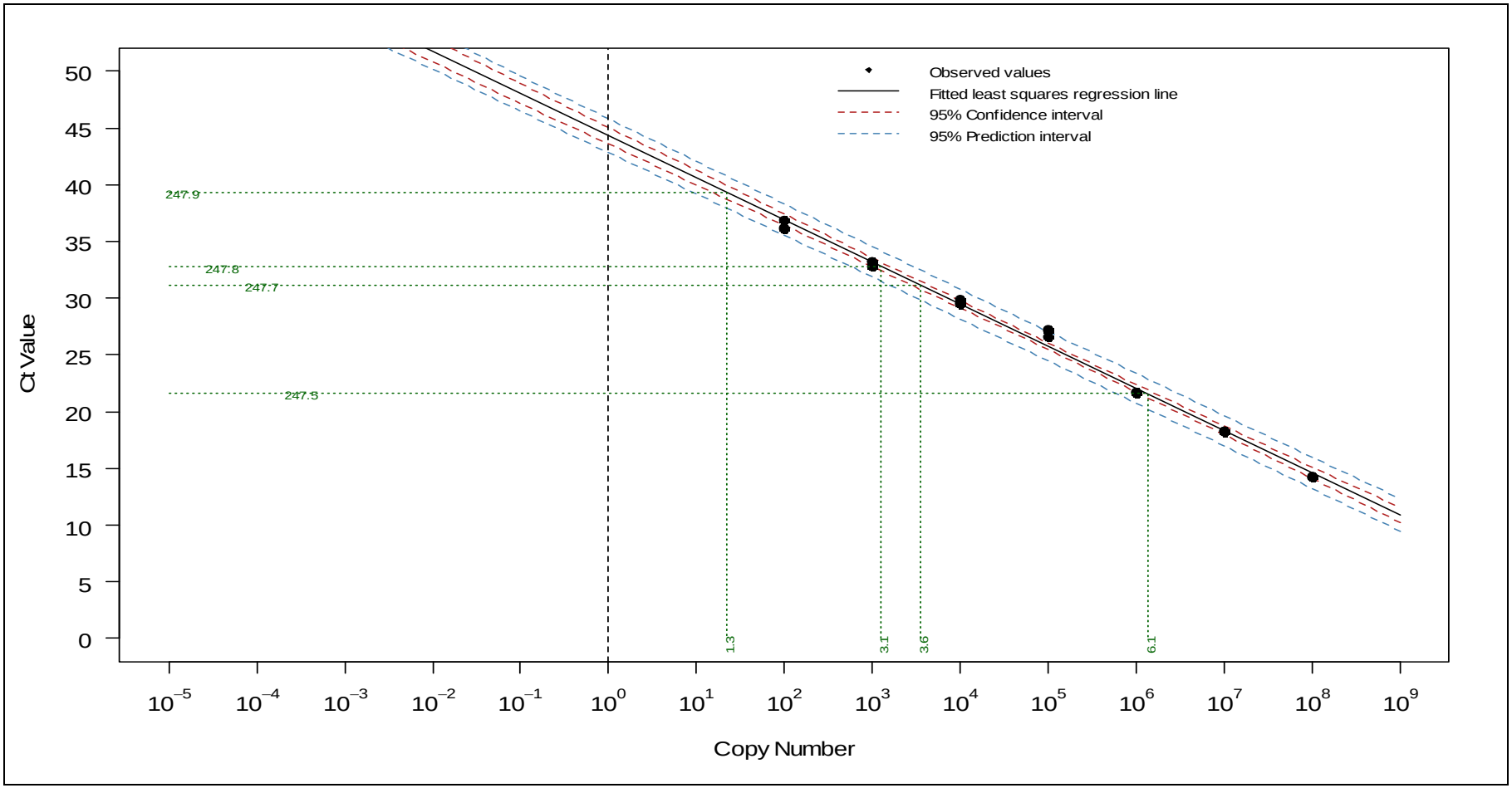


Figure 21.2.7d. Illustration of the calculation of the HPV16 copy number from the standard curve for patient samples on qPCR well-plate VL061. The horizontal numbers in green are patient sample IDs; the vertical numbers in green are estimated ln(copy number) for those samples.



Although the estimation of copy numbers (viral load) using the method described above is clearly an exercise in regression, specifically *inverse regression*, it is not treated as a statistical procedure in standard laboratory practice (in the laboratory testing the samples used in this thesis, or any other apparently). Since statistical techniques are used to estimate the standard curve, there is uncertainty in the estimates produced using that curve. Figures 21.2.7e and 21.2.7f also include a 95% confidence interval, and a 95% prediction interval, for the standard curve. Clearly, in this particular assay, measured copy numbers have a substantial range of values which are consistent with the point estimate of the observed copy number (tables 21.2.7c and 21.2.7d illustrate this for GAPDH and HPV16, respectively).

Table 21.2.7c. Point estimates of GAPDH copy numbers and 95% confidence limits for patient samples assayed on qPCR plate VL061.

Sample ID	Mean Ct Value	GAPDH Copy Number	Lower 95% Confidence Limit	Upper 95% Confidence Limit
247.5	30.16	555.43	367.57	811.45
247.6	41.15	0.13	0.06	0.28
247.7	29.14	1,211.56	830.07	1,718.39
247.8	25.52	18,855.62	14,184.90	24,794.10
247.9	24.63	37,167.74	28,325.02	48,458.45

Table 21.2.7d. Point estimates of HPV16 copy numbers and 95% confidence limits for patient samples assayed on qPCR plate VL061.

Sample ID	Mean Ct Value	HPV16 Copy Number	Lower 95% Confidence Limit	Upper 95% Confidence Limit
247.5	21.57	1,341,237.24	1,103,697.01	1,641,883.58
247.6	50.00	0.00	0.00	0.00
247.7	31.14	3,583.76	2,881.32	4,415.79
247.8	32.81	1,268.71	993.41	1,600.46
247.9	39.34	22.35	15.40	31.66

The uncertainty in these estimates could easily be reduced. It should be noted that standard “curve” is a misnomer since whatever Ct values are observed following the assay, a straight line will always be fitted to the data rather than any other function, e.g. a quadratic curve. There are theoretical reasons to justify why a straight-line relationship on a logarithmic scale is expected rather than any other, but conventionally no diagnostics are performed in the modelling of the standard curve. Given that this is the case, it would be better to fit the line using only two points with known copy numbers, with a greater number of observations concentrated at the extremes of the range of interest, rather than spreading them over the entire range (Draper 1998). This would increase precision within the range of copy numbers covered by the standard curve and reduce uncertainty in estimates.

In standard laboratory practice, the uncertainty in estimates is ignored: estimated values are treated as being known precisely. In particular, when calculating viral load normalised for cellular content, a ratio of two correlated values estimated with uncertainty from two regression curves is formed and used thereafter without regard for any uncertainty in the estimates. A crude estimate of the range of values potentially consistent with the observed normalised copy numbers for the patient samples in this assay is presented in table 21.2.7e. The lower limit was calculated by dividing the lower 95% confidence limit for the HPV16 copy number by the upper 95% confidence limit for the GAPDH copy number; the upper limit by dividing the upper 95% confidence limit for the HPV16, copy number by the lower 95% confidence limit for the GAPDH copy number.

Table 21.2.7e. Observed and corrected HPV16 copy numbers and range of plausible values for normalised copy number for patient samples assayed on qPCR plate VL061.

Sample ID	HPV16 Copy Number	HPV16 Normalised Copy Number	Lower Limit Of Plausible Values	Upper Limit Of Plausible Values	Ratio Of Upper To Lower Limit
247.5	1,341,237.241	4,829.549	2,720.299	8,933.672	3.3
247.6	0.000	0.000	0.000	0.000	-
247.7	3,583.759	5.916	3.354	10.640	3.2
247.8	1,268.709	0.135	0.080	0.226	2.8
247.9	22.347	0.001	0.001	0.002	3.5

21.3 qPCR ASSAY PERFORMANCE

Dynamic range

When a standard curve is generated, a straight line is fitted to the observed Ct values by ordinary least squares regression. This curve will only be valid if it is reasonable to assume that a straight line is the appropriate model. To determine the linear dynamic range of the assay i.e. the range of copy numbers over which a straight line is a valid model for the data, serial ten-fold dilutions of plasmids were assayed for viral load. For this assay, the dynamic range exceeded 10^9 to one copies.

Type specificity

Primer cross-reactivity was assessed by testing DNA from the HPV18-positive cervical cancer cell line HeLa using the HPV16 primers; and DNA from the HPV16-positive cervical cancer cell line CaSki using the HPV18 primers. No cross-reactivity was found i.e. the HPV16 primers did not amplify HPV18, and *vice versa*. As

desired, both sets of primers also failed to amplify when the HPV-negative cervical cancer cell line C33a was tested.

Inhibition and PCR artefacts

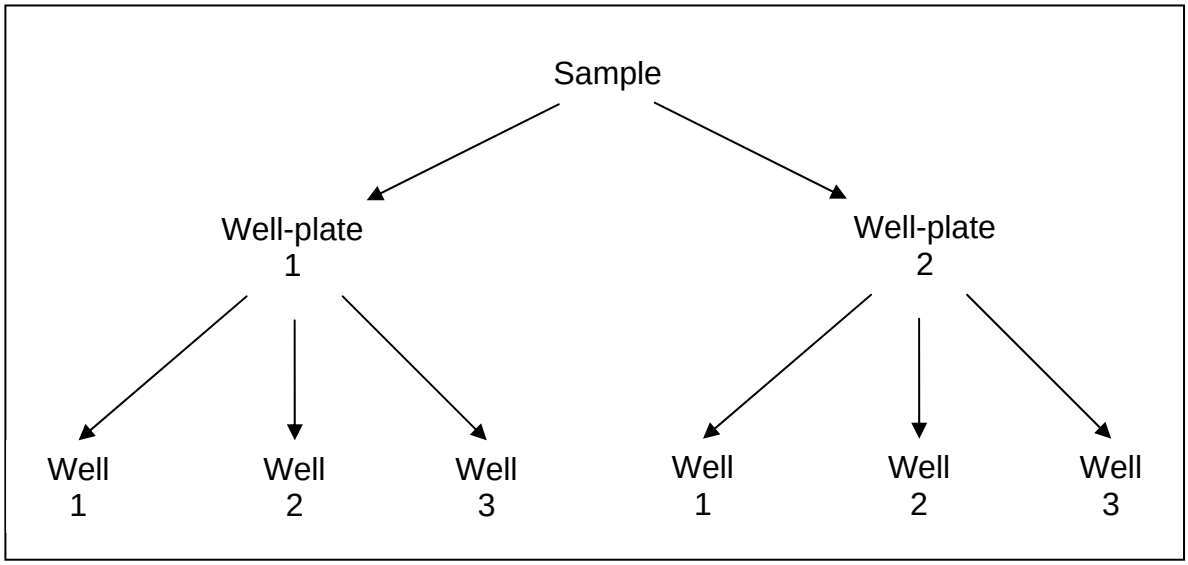
Samples containing varying concentrations of DNA from the HPV16-positive cell lines SiHa and CaSki, or the HPV18-positive cell line HeLa (1000 ng; 500 ng; 100 ng; 50 ng; 10 ng and 5 ng) were used to identify the presence of inhibitors of PCR or concentration-dependent PCR artefacts. Results were as expected for all concentrations, with no evidence of interference from inhibitors, or of PCR artefacts: the estimated number of copies of the virus genome per cell were approximately the same in each assay. Competitive inhibition of the qPCR assay for measuring HPV16 and HPV18 viral load was assessed by mixing SiHa and HeLa DNA together in varying concentrations, and no competitive inhibition was found, i.e. the presence of HPV16 did not have an effect on the performance of the HPV18 qPCR, and *vice versa*.

21.4 STATISTICAL METHODS

21.4.1 Measuring assay reliability

The data structure for the experiments used to measure intra- and inter-assay reliability is shown in figure 21.4.1. Each sample was measured in three wells on each of two well-plates.

Figure 21.4.1. Data structure for experiments used to measure intra- and inter-assay reliability.



21.4.1.1 Intra-assay reliability

An intra-cluster correlation coefficient (ICC) was used to measure intra-assay and inter-assay reliability. The ICC takes values of one for perfect reliability (perfect agreement), and values close to zero for little reliability.

Both a random effects model and a closed formula were used to estimate the ICC, with no material differences between the results of the two methods: only the results derived from the random effects model are presented (Pinheiro 2002, Rousson 2002).

In this context, intra-assay reliability refers to within-plate agreement between the three Ct values whose mean value was ultimately used to determine viral load for a

single sample: the same sample was not tested more than once on the same plate. Further, given the procedure used for selecting valid Ct values (see section 21.2.7 point 2), intra-assay reliability was effectively constrained to be almost perfect: only samples with Ct values within strict limits were retained in the analysis. ICC were calculated for the observed Ct values used for measuring both viral DNA copy number and GAPDH gene copy number.

The calculation of the ICC assumes that the data arise from a normal, or at least a continuous, distribution. In this case, observed Ct values range from 1 to 50, with a value of 50 being “HPV-negative” for the type tested for. Therefore, *a priori*, these data cannot be considered to be a sample from a normal distribution. Any transformation to a different scale, e.g. using a complementary log-log transformation, would still present problems in dealing with numerical zeroes, as well as with assay-specific “zeroes”, i.e. Ct=50. When appropriate, analyses were carried out both including and excluding these values. Given that in each case where a Ct value of 50 was observed for any replicate, all other replicate Ct values were also 50, excluding such perfectly replicated results is likely to have yielded a more conservative estimate of reliability.

If the mean Ct value from the replicated GAPDH wells was used to estimate the copy number of the GAPDH gene in cytological samples, then each experiment could also be considered as yielding three estimates of viral copy number. ICC were therefore also calculated for observed virus copy numbers per cell.

21.4.1.2 Inter-assay reliability

An intra-cluster correlation coefficient (ICC) was used to estimate inter-assay reliability, calculated using a random effects model.

In this context, inter-assay reliability refers to the agreement between the virus copy number per cell of a sample measured on each of two separate plates. A single measurement of virus copy number per cell was obtained for each sample from each well-plate. These well-plates were then repeated, yielding two measurements of virus copy number per cell for each sample. Inter-assay reliability, unlike intra-assay reliability, was not artificially constrained within tight limits. In this test-retest reliability situation, there was no “learning” effect: therefore the ICC is the usual product-moment correlation coefficient.

21.4.2 The analysis of HPV viral load

21.4.2.1 Measurement of HPV viral load as a time-varying covariate

When repeated measurements over time are made on both exposure and outcome, determining how exposure influences the risk of that outcome presents a complex challenge (see section 6.6). Analysing HPV viral load as a time-fixed covariate, by including in the analysis only HPV status at baseline for example, is one possibility. However, estimates of association may be unreliable if *changes* in HPV viral load are also important. In an analysis incorporating a time-varying measurement of HPV

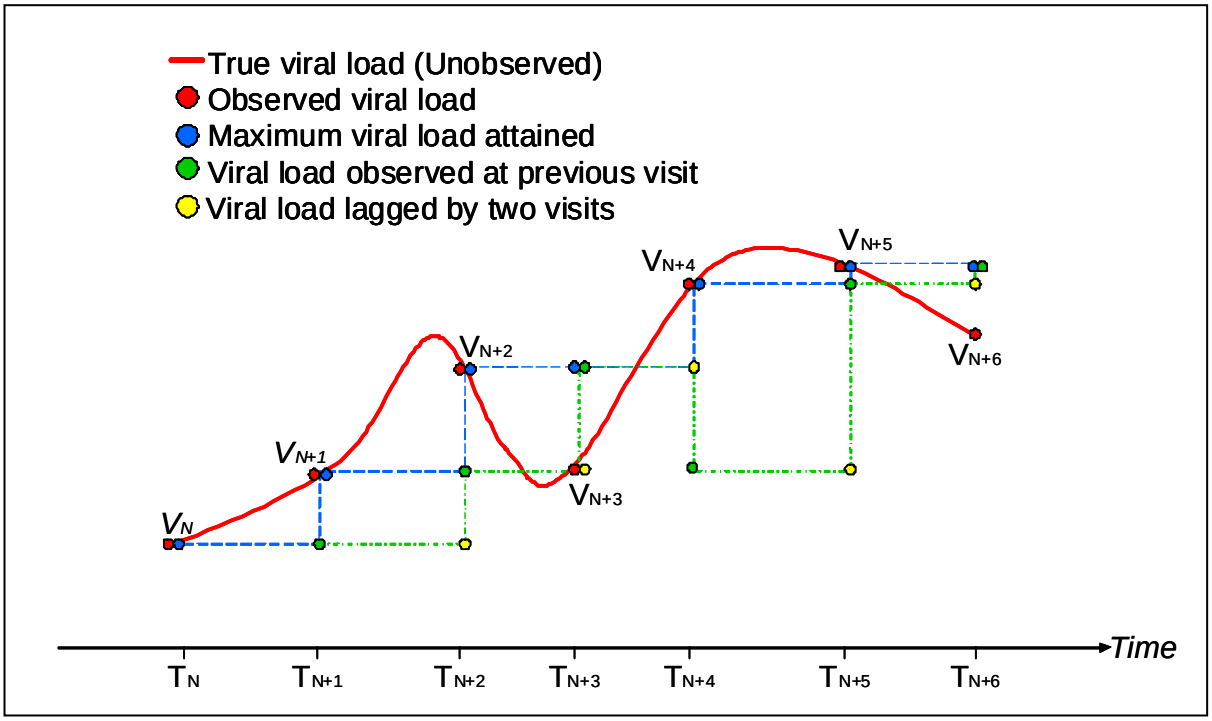
viral load, several possibilities exist for how to measure viral load. For example, at a given point in time the most recent measurement of viral load can be used; or the measurement of viral load taken at an earlier time, such as at the previous visit, or the measurement from two visits earlier. Alternatively, it may be preferable to use a variable which summarises how viral load has changed over time, or one which captures important features of that change ,e.g. mean viral load, area under curve, maximum viral load attained, etc. But again, if this approach is adopted, the variable can take its current value, or its value from an earlier point in time.

For example, figure 21.4.2.1 illustrates how HPV viral load changes with time for one hypothetical woman. The true form of the viral load curve is unknown; instead we observe the discrete set of viral load measurements V_N to V_{N+6} taken at measurement times T_N to T_{N+6}, respectively. For each of the measurement times, the viral load observed at the previous time is shown in green: this is undefined at time T_N. Similarly, the viral load observed two measurement occasions in the past is shown in yellow: this is undefined at times T_N and T_{N+1}. The maximum viral load attained at each of the measurement times is shown in blue: this can only increase as time increases, i.e. as more measurements are made.

In the cohort study described in this thesis measurements of viral load and outcome, for example cervical cytological status, were made at the same time (see figures 5.2.3b and 6.1). Therefore outcome status and the potential risk factor for that outcome are measured contemporaneously. Using measurements of risk factor variables taken contemporaneously with measurements on outcome status

introduces difficulties of interpretation (see section 6.6.3). For example, the presence of cervical cytological abnormality may lead to an increase in viral production and therefore viral load: in this case current viral load status would not be a “cause” but rather a “consequence” of the presence of cytological abnormality. If the analysis is to be able to establish temporality, it is preferable to relate current outcomes to previous values of risk factor variables (note that it still cannot be ruled out that a cervical abnormality at a stage at which it cannot be detected by cytology still leads to an increase in HPV viral load which can be detected).

Figure 21.4.2.1. Measurement of HPV viral load as a time-varying covariate



21.4.2.2 Time to the acquisition of incident cervical cytological abnormality

Analyses of the time to the acquisition of incident cervical cytological abnormality were undertaken using methods appropriate to interval-censored time-to-event data.

Time to the acquisition of cervical cytological abnormality was measured from the date of the first qPCR-evaluable sample until the interval between the date of the first detection of cervical cytological abnormality and the date of the immediately preceding cytologically normal smear; censoring occurred on the earliest of the date of the diagnosis of high-grade CIN, the date of the last qPCR-evaluable sample, and the date of the last smear.

Estimates of hazards ratios were obtained using a semi-parametric method for modelling interval-censored time-to-event data as a generalized linear model (see section 6.5.2), with various estimates of viral load treated as time-varying covariates. The optimal analysis of the association between viral load and the incidence of cervical cytological abnormality is not immediately apparent. Consequently, five exposure variables were investigated in analyses:

- a time-fixed binary variable indicating HPV viral load status at baseline (positive or negative);
- a time-fixed continuous variable measuring HPV viral load at baseline;
- a time-varying binary variable indicating “ever HPV viral load positive” up-to and including the previous visit (yes or no);
- a time-varying continuous variable measuring HPV viral load at previous visit;
- a time-varying continuous variable measuring the maximum HPV viral load attained up-to and including the previous visit.

For example, for the woman illustrated in figure 21.4.2.1, at time T_{N+4} these variables have the values “positive”, V_N, “yes”, V_{N+3}, and V_{N+2}, respectively.

95% confidence intervals were constructed from parameter estimates and their standard errors; and statistical tests of hypotheses were undertaken using likelihood ratio tests.

21.4.2.3 Time to clearance of an incident cervical HPV infection

Analyses of the duration of incident cervical HPV infections were undertaken using methods appropriate to interval-censored time-to-event data.

An incident qPCR HPV infection was defined as an infection in a woman who was qPCR-negative for that type in her first qPCR-evaluable sample, but who subsequently became qPCR-positive during follow-up. An *episode* of qPCR infection was then defined as being the interval between this first qPCR-positive sample and the last subsequent qPCR positive sample, during which interval of time the woman was continuously qPCR-positive. Duration of an incident cervical HPV16 or HPV18 qPCR infection was measured from the date of the first qPCR-positive sample until the interval between the date of the first subsequent qPCR-negative sample and the date of the immediately preceding qPCR-evaluable sample; censoring occurred on the earliest of: the date of the diagnosis of high-grade CIN; and the date of the last qPCR-evaluable sample.

Estimates of the cumulative risk of the clearance of a qPCR infection were obtained using a non-parametric maximum likelihood estimator (Turnbull 1976) (see section 6.4.1). Estimates of hazards ratios were obtained using a semi-parametric method for modelling interval-censored time-to-event data as a generalized linear model (Smith 1997), with various estimates of HPV viral load treated as time-varying covariates. Three exposure variables were investigated in analyses:

- a time-fixed continuous variable measuring HPV viral load at the start of the episode;
- a time-varying continuous variable measuring HPV viral load at the previous visit;
- and a time-varying continuous variable measuring maximum HPV viral load attained by the previous visit

21.4.2.4 The risk of being free of cervical HPV infection in the last qPCR-evaluable sample

An additional exploratory analysis, which was related to the time to clearance of a cervical HPV qPCR-infection, was used to determine the association between the maximum HPV16 and HPV18 viral load observed during follow-up and the probability of being clear of HPV16 or HPV18 infection in the last evaluable sample, after controlling for the length of follow-up. Estimates of odds ratios measuring the risk of being free of cervical HPV infection in the last qPCR-evaluable sample were obtained using binary logistic regression.

21.4.3 Estimation and hypothesis testing

All tests of statistical significance were conducted at the 5% two-sided significance level.

In multivariate analyses, all numerical values of HPV viral load were log-transformed; 95% confidence intervals were constructed from parameter estimates and their standard errors; and tests of hypotheses were undertaken using likelihood ratio tests.

Chapter 22

THE NATURAL HISTORY OF CERVICAL HPV16 AND HPV18 INFECTIONS – VIRAL LOAD

RESULTS

Orientation

In this chapter I:

- determine the reliability of an assay used for the measurement of the viral load of HPV16 and HPV18 in cervical samples
- describe how the distribution of HPV viral load varies with infecting HPV type, for HPV16 and HPV18
- describe how HPV viral load changes over time in women with incident cervical HPV16 and/or HPV18 infections
- describe the relationship between HPV16 and HPV18 viral load in cervical samples and the detection of epithelial abnormalities of the cervix

22.1 RESTATEMENT OF OBJECTIVES

To investigate how the natural history of cervical HPV infection varies with viral load.

- To evaluate the performance of an assay for the measurement of the viral load of cervical infections with HPV16 and HPV18.

- To describe changes in HPV16 and HPV18 viral load in cervical samples taken from a cohort of young women who were recruited soon after first sexual intercourse, and who first tested positive for HPV16 DNA, or HPV18 DNA, or both, during follow-up.

- To describe the association between the acquisition of epithelial abnormalities of the cervix and changes in HPV16 and HPV18 viral load in cervical samples taken from a cohort of young women who were recruited soon after first sexual intercourse, and who first tested positive for HPV16 DNA, or HPV18 DNA, or both, during follow-up.

22.2 ASSESSMENT OF TEST PERFORMANCE FOR THE MEASUREMENT OF VIRAL LOAD BY QPCR

The development and implementation of the qPCR assay used for measuring HPV viral load occurred in three stages, with some overlap in practice. In the first stage, the experienced laboratory scientist who would be carrying out the analysis of study samples: designed HPV type-specific primers for use in the assay; and established the procedures for conducting the assay (section 21.2). The second stage was to determine whether the assay was valid, i.e. yielded accurate measurements of HPV viral load, and *reliable*, i.e. gave the “same” result in repeated measurements of the same sample. The third stage was the testing of study samples: this stage could only begin when development reached a point at which the technique could be used with confidence.

22.2.1 Selection of study samples for use in validation of the qPCR assay

The samples from the study described in this thesis are a limited resource, and so must be used sparingly with future uses kept in mind. The samples from the study cohort were therefore chosen with the larger aims of the study in view. Most women in the analyses used to develop this assay contributed only a single sample because they only ever had one evaluable cervical cytological sample, or because they were from the "prevalent cohorts" i.e. women who were HPV DNA-positive, or had a cytologically abnormal cervical smear, at study entry, or both. Such women are already diseased, or infected with HPV, at study entry, and are therefore less valuable in an analysis of incident cervical HPV infections or cervical disease.

Samples were selected from those provided by women who were either cytologically abnormal in their cervical smear, or HPV-positive at study entry, or both, and tended to be from women with limited, or no, further follow-up. Initially, samples were selected neither at random nor in a systematic way, but rather to give a mixture of samples which were cytologically normal and cytologically abnormal in the corresponding cervical smear, and which were positive for the HPV type of interest according to the GP5+/GP6+ system (see section 5.5.4.1).

No formal sample size calculation was performed. It was decided that, initially, three qPCR experiments including a minimum of 15 samples should be performed. It was hoped that this would be sufficient to establish that the qPCR assay was *reliable*,

and therefore that only one qPCR per sample would be needed when analysing patient samples from the study population of interest.

22.2.1.1 Study samples and analyses undertaken for the validation of the HPV16 qPCR assay

To determine whether the qPCR assay for measuring HPV16 viral load was *reliable*, 10 analyses were undertaken (table 22.2.1.1).

Table 22.2.1.1. Analyses undertaken for determining the reliability of the qPCR assay measuring HPV16 viral load.

Well-plate ID	Analysis		Primers	Comments
	Date	Time		
VL005	12/10/05	12pm	Gravitt	
VL006	12/10/05	5pm	Gravitt	Repeat of VL005
VL007	13/10/05	12pm	Gravitt	
VL008	13/10/05	230pm	Gravitt	Repeat of VL007
VL021	20/01/06	12pm	Optimized	Repeat of VL007
VL022	20/01/06	230pm	Optimized	Repeat of VL021
VL023	20/01/06	5pm	Optimized	
VL024	20/01/06	7pm	Optimized	Repeat of VL023
VL025	23/01/06	230pm	Optimized	
VL026	23/01/06	5pm	Optimized	Repeat of VL025

Note: apparently missing plate ID's relate to interim analyses undertaken for assay-refinement purposes.

Strictly speaking, only four analyses were required to estimate intra-plate and inter-plate reliability. However, following the first set of experiments, the primers were redesigned (optimized): this decision was based on technical reasons and was *not* influenced by the results of the reliability analyses using the original primers (see section 21.2.2); the results of these first four analyses are not considered further. Further opportunistic analyses were also undertaken.

The study population for assessing the reliability of the HPV16 qPCR assay comprised 21 samples from 19 women. Well-plates VL021 to VL024 inclusive were the primary analyses undertaken to establish reliability. These analyses tested 16 samples, with 12 women each providing one sample, and two providing two samples (one sample from each of two separate visits: one woman had one sample tested on well-plates VL021 and VL022, and another on well-plates VL023 and VL024, and one had *both* samples tested on well-plates VL023 and VL024). Well-plates VL025 and VL026 were additional opportunistic analyses which also contributed to establishing the reliability of the qPCR assay. These analyses tested seven samples, with seven women each providing one sample. None of these samples had been tested before, and none of the women included had previously had other samples tested.

22.2.1.2 Study samples and analyses undertaken for the validation of the HPV18 qPCR assay

To determine whether the qPCR assay for measuring HPV18 viral load was reliable, four analyses were undertaken (table 22.2.1.2).

Table 22.2.1.2. Analyses undertaken for determining the reliability of the qPCR assay used for measuring HPV18 viral load.

Well-plate ID	Analysis		Primers	Comments
	Date	Time		
VL012	13/12/05	12pm	Optimized	
VL013	13/12/05	230pm	Optimized	Repeat of VL012
VL014	14/12/05	12pm	Optimized	
VL015	14/12/05	5pm	Optimized	Repeat of VL014

Well-plates VL012 to VL015 inclusive, used nine samples; on each well-plate, nine women each provided one sample, but one woman contributed two samples to the analysis. In all of these analyses, redesigned (optimized) primers were used. The study population for assessing the reliability of the HPV18 qPCR assay thus comprised 18 samples from 17 women.

22.2.2 Intra-assay reliability

22.2.2.1 Intra-assay reliability of the qPCR assay measuring HPV16 viral load

Estimates of intraclass correlation coefficients used to measure intra-assay (within-assay) reliability are presented in table 22.2.2.1. Using either observed Ct values, or observed virus copy numbers, estimates of intraclass correlation coefficients used to measure intra-assay reliability exceeded 0.99: this was the case whether the analysis included all six well-plates used in the reliability analyses, or was restricted to the four “definitive” well-plates; and whether assay “zeroes” were included or excluded. Similar results were obtained when GAPDH Ct values were analyzed. Most of the variability was clearly due to different samples being tested, rather than different wells being used to test those samples. As discussed earlier (see section 21.4.1.1), this is not a particularly informative finding but rather a consequence of the approach adopted to accepting Ct values as valid.

Table 22.2.2.1. Estimates of the intraclass correlation coefficient used to measure intra-assay reliability of the qPCR assay measuring the viral load of HPV16.

Analyses	Samples	Negative Values	Quantity	Number of Well-plates	Number of Measurements	Proportion of Variability			ICC[a]
						Plate	Sample	Well	
All	Viral DNA	Included	Ct Value	6	41	5.1%	94.6%	0.3%	0.997
All	Viral DNA	Excluded	Ct Value	6	38	12.5%	87.0%	0.5%	0.994
Definitive	Viral DNA	Included	Ct Value	4	27	9.8%	90.0%	0.2%	0.998
Definitive	Viral DNA	Excluded	Ct Value	4	24	23.9%	75.6%	0.4%	0.994
Opportunistic	Viral DNA	Included	Ct Value	2	14	0.0%	99.3%	0.7%	0.993
Opportunistic	Viral DNA	Excluded	Ct Value	2	14	0.0%	99.3%	0.7%	0.993
All	GAPDH	Included	Ct Value	6	42	8.5%	91.1%	0.4%	0.995
Definitive	GAPDH	Included	Ct Value	4	28	0.0%	99.8%	0.2%	0.998
Opportunistic	GAPDH	Included	Ct Value	2	14	0.0%	99.3%	0.7%	0.993
All	Viral DNA	Included	Copy Number	6	41	1.3%	97.9%	0.8%	0.992
All	Viral DNA	Excluded	Copy Number	6	38	0.0%	99.2%	0.8%	0.992
Definitive	Viral DNA	Included	Copy Number	4	27	7.4%	91.9%	0.7%	0.993
Definitive	Viral DNA	Excluded	Copy Number	4	24	6.8%	92.6%	0.7%	0.993
Opportunistic	Viral DNA	Included	Copy Number	2	14	0.0%	99.1%	0.9%	0.991
Opportunistic	Viral DNA	Excluded	Copy Number	2	14	0.0%	99.1%	0.9%	0.991

[a]intraclass correlation coefficient, calculated essentially as Sample/(Sample+Well)

22.2.2.2 Intra-assay reliability of the qPCR assay measuring HPV18 viral load

Estimates of intraclass correlation coefficients used to measure intra-assay reliability of the HPV18 qPCR assay are presented in table 22.2.2.2. In these analyses, all samples tested were qPCR-positive for HPV18. Using either observed Ct values or observed virus copy numbers, estimates of intraclass correlation coefficients used to measure intra-assay reliability exceeded 0.97: this was the case whether the analysis included all four well-plates used in the reliability analyses, or was restricted

to subsets of these well-plates; similar results were obtained when GAPDH Ct values were analyzed. This is not a particularly informative finding, but rather a consequence of the approach adopted to accepting Ct values as valid.

Table 22.2.2.2. Estimates of the intraclass correlation coefficient used to measure the intra-assay reliability of the qPCR assay for measuring HPV18 viral load.

Analyses	Samples	Quantity	Number of Plates	Number of Measurements	Proportion of Variability			ICC[a]
					Plate	Sample	Well	
All	Viral DNA	Ct Value	4	36	0.0%	99.6%	0.4%	0.996
VL012 and VL013	Viral DNA	Ct Value	2	18	0.0%	99.5%	0.5%	0.995
VL014 and VL015	Viral DNA	Ct Value	2	18	0.0%	99.6%	0.4%	0.996
All	GAPDH	Ct Value	4	36	0.0%	99.3%	0.7%	0.993
VL012 and VL013	GAPDH	Ct Value	2	18	0.0%	99.5%	0.5%	0.995
VL014 and VL015	GAPDH	Ct Value	2	18	0.0%	99.1%	0.9%	0.991
All	Viral DNA	Copy Number	4	36	1.1%	96.8%	2.2%	0.978
VL012 and VL013	Viral DNA	Copy Number	2	18	0.0%	98.4%	1.6%	0.984
VL014 and VL015	Viral DNA	Copy Number	2	18	0.0%	97.7%	2.3%	0.977

[a]intraclass correlation coefficient, calculated essentially as Sample/(Sample+Well)

22.2.3 Inter-assay reliability

22.2.3.1 Inter-assay reliability of the qPCR assay measuring HPV16 viral load

The HPV16 viral load for each sample, measured on each of two occasions, is presented in table 22.2.3.1a. The median viral load in the first measurement was 2,198 copies per 1,000 cells (range 0 to 249,628), compared with 1,849 copies per 1,000 cells (range 0 to 172,481) in the second measurement, a decrease of 16%.

Sample 553.10 was excluded from analyses of inter-assay reliability since only one observation was available. For all but one sample (ID 138.3), there were notable differences between the first and second measurement of viral load. However, the relative differences tended to be small. The exception was the sample with ID 1191.3, which was HPV16-qPCR-negative in the first measurement, but had a viral load of 935 copies per 1,000 cells in the second. Recall that a Ct value of 50 means "negative" for the qPCR assay for the specific type as it was used in these analyses (see section 21.2.5.1). The three virus Ct values for this sample were 50, 38.11, and 50 i.e. two qPCR-negative values and one qPCR-positive value. Given the system adopted for accepting Ct values as valid, for this sample the qPCR-positive value was excluded, meaning that it was classified as HPV16-qPCR-negative. On more than one occasion, the Ct value which was excluded was a value of 50 in the presence of two non-50 values, i.e. something appeared to have gone awry with a sample which was in fact qPCR-positive. Of course, if it is possible for one Ct value to be aberrant, it is also possible for two. If the qPCR-positive value had been accepted for sample 1191.3, and the qPCR-negative values excluded, this sample would have been scored as having 804 copies per 1,000 cells, a ratio of 0.86 relative to the first measurement. Nevertheless, the discrepancy in values for this sample have to be accepted in this analysis.

Excluding the samples with IDs 1191.3 and 553.10, two samples had a different rank order for measurement two compared to measurement one, and *vice versa*; in each case, there was a difference of only one rank (see section 23.2 for a discussion of the disparity between the GP5+/GP6+ result and the qPCR result).

Table 22.2.3.1a. Raw data for the measurement of inter-assay reliability of the viral load assay for HPV16.

ID[a]	Visit	HPV status (GP5+/GP6+)[b]	Cytology[c]	Test 1		Test 2		Ratio
				Analysis	*Viral Load*[d]	*Analysis*	*Viral Load*[d]	T2:T1[e]
83	3	Negative	Normal	VL023	16.86	VL024	10.41	0.62
138	3	16;31	Moderate	VL021	49,407.41	VL022	49,518.06	1.00
168	4	Negative	US	VL023	47.31	VL024	60.28	1.27
553	10	Negative	Normal	VL023	Virus Ct[f]	VL024	0.28	-
702	2	16	Moderate	VL021	2,408.02	VL022	1,988.66	0.83
878	1	16	US	VL021	30,470.92	VL022	28,438.01	0.93
937	3	16;18	Moderate	VL021	713.70	VL022	679.63	0.95
1191	3	Type X	Normal	VL021	0	VL022	934.61	-
1191	7	Negative	Moderate	VL023	0	VL024	0	-
1466	2	16	Mild	VL021	75.40	VL022	60.59	0.80
1621	3	16	Mild	VL023	2,459.88	VL024	2,710.01	1.10
1621	4	Negative	Normal	VL023	0.50	VL024	0.30	0.59
1784	4	6/11;16	Normal	VL021	2,447.32	VL022	2,855.84	1.17
1832	5	Negative	Normal	VL023	130.65	VL024	109.78	0.84
126	2	Negative	Normal	VL025	1.15	VL026	0.63	0.55
294	2	16	Normal	VL025	3,125.84	VL026	4,912.92	1.57
426	1	16	Mild	VL025	249,627.98	VL026	172,481.45	0.69
869	2	16	Moderate	VL025	6,743.14	VL026	5,067.09	0.75
1400	1	16;33	Mild	VL025	64,752.71	VL026	59,610.61	0.92
1463	1	16	Mild	VL025	2,093.30	VL026	1,848.79	0.88
2091	1	16	Normal	VL025	2,303.03	VL026	2,662.67	1.16

[a]Unique identifier for a woman; [b]HPV DNA status according to the GP5+/GP6+ system; [c]Severity of dyskaryosis; [d]Copy number per 1,000 cells; [e]Ratio of the viral load from the second test compared to that from the first test; [f]Result excluded due to excessive variability in observed Ct values for HPV16.

Estimates of intraclass correlation coefficients (ICC) used to measure inter-assay test-retest reliability are presented in table 22.2.3.1b. In the definitive analysis of inter-assay reliability, the estimate of ICC exceeded 0.99; combining results from all analyses yielded an ICC in excess of 0.93. In either case, inter-assay reliability is excellent.

Table 22.2.3.1b. Estimates of the intraclass correlation coefficient (ICC) used to measure inter-assay test-retest reliability of the qPCR assay measuring viral load for HPV16.

Analyses	Number of Well-plates	Number of Samples	ICC
All	6	20	0.939
Definitive	4	13	0.999
Opportunistic	2	7	0.934

22.2.3.2 Inter-assay reliability of the qPCR assay measuring HPV18 viral load

The HPV18 viral load for each sample, measured on each of the two occasions, is presented in table 22.2.3.2a. In these analyses, all samples tested were HPV18 qPCR-positive. The median viral load in the first measurement was 4,079 copies per 1,000 cells (range 0.23 to 886,139), compared to 4,809 copies per 1,000 cells (range 0.39 to 336,070) in the second measurement, an increase of 18%. For all but one sample (sample ID 138.3), there were notable differences between the first and second measurements of viral load. However, the relative differences tended to be small.

The exception was the sample with ID 1708.2, which had a viral load of 886,139 copies per 1,000 cells in the first measurement compared with 336,070 in the second; although both measurements clearly indicate a high-viral load, the second measurement indicates a viral load which is 62% less than that indicated by the first. There is no indication from the assay results as to why this discrepancy may have occurred.

Compared with their rank order in measurement two, three samples had a different rank order in measurement one; there was a difference of one rank for two of these samples, and two ranks for the other. Compared with their rank order in measurement one, four samples had a different rank order in measurement two; in each case, there was a difference of only one rank.

Table 22.2.3.2a. Raw data for the measurement of inter-assay reliability of the qPCR assay for measuring HPV18 viral load.

ID[a]	Visit	HPV status (GP5+/GP6+)[b]	Cytology[c]	Test 1		Test 2		Ratio
				Analysis	*Viral Load*[d]	*Analysis*	*Viral Load*[d]	T2:T1[e]
107	1	18	Normal	VL014	350.25	VL015	304.34	0.87
178	1	18	Mild	VL012	8.74	VL013	9.74	1.11
459	2	18	Normal	VL012	9,249.61	VL013	13,635.06	1.47
582	2	18	Moderate	VL014	68,541.00	VL015	97,660.62	1.42
652	2	18	US	VL014	0.23	VL015	0.39	1.75
698	1	18	Normal	VL012	138.18	VL013	125.98	0.91
968	1	18	Normal	VL012	73.07	VL013	78.70	1.08
974	1	18	No result	VL014	5,022.00	VL015	6,304.19	1.26
1008	2	16;18	Normal	VL014	74,237.68	VL015	62,381.39	0.84
1171	3	18	Normal	VL014	27,739.75	VL015	25,840.30	0.93
1625	1	18	Normal	VL012	90.54	VL013	130.33	1.44
1625	3	Negative	Normal	VL014	4,999.32	VL015	7,641.18	1.53
1688	1	18	Normal	VL012	2,370.26	VL013	2,760.10	1.16
1695	2	18	Moderate	VL012	33,528.59	VL013	36,444.69	1.09
1708	2	16;18	Moderate	VL014	886,138.92	VL015	336,069.68	0.38
1739	2	18	Normal	VL012	3,158.04	VL013	3,314.30	1.05
1864	1	18	Normal	VL012	8,128.04	VL013	7,487.90	0.92
2000	1	18;52	Normal	VL014	214.27	VL015	203.22	0.95

[a]Unique identifier for a woman; [b]HPV DNA status according to the GP5+/GP6+ system; [c]Severity of dyskaryosis; [d]Copy number per 1,000 cells; [e]Ratio of the viral load from the second test compared to that from the first test.

Estimates of ICC used to estimate test-retest reliability for the HPV18 qPCR assay are presented in table 18.4.2b. In the definitive analysis of inter-assay reliability, the estimate of ICC was only 0.66. This poor reliability was due to the effect of the sample with ID 1708.2; when this sample is excluded, all estimates of ICC exceed

0.94, indicating excellent inter-assay reliability. However, the result for sample 1708.2 clearly remains a concern.

Table 22.2.3.2b. Estimates of the intraclass correlation coefficient (ICC) used to measure inter-assay test-retest reliability of the qPCR assay for measuring the viral load of HPV18.

Analyses	Number of Well-plates	Number of Samples	ICC
All	4	18	0.655
VL012 vs VL013	2	9	0.989
VL014 vs VL015	2	9	0.644
All excluding 1708.2	4	17	0.952
VL014 vs VL015 excluding 1708.2	2	8	0.941

22.2.4 Viral load in replicate samples and samples tested more than once

Although this section is an aside to some extent, and relies on analyses described in subsequent sections, it is related to the issue of repeated testing; it is also important in that it establishes which measurement of viral load is reported and analysed for the remainder of the chapter.

Some women in the study population had more than one sample taken at the same study visit, which I refer to as replicates; and some had samples from the same study visit which were tested more than once. In the analysis of both HPV16 and HPV18, the majority of tests on replicates were concordant with respect to HPV status, with HPV16 results being slightly more concordant than those for HPV18. In the concordant *positive* pairs of replicates, the ratio of the largest copy number to the smallest ranged from 1 to 9,388 for HPV16, and 2 to 14,058 for HPV18. For both

types, in the replicates with discrepant results, copy number in the positive sample was numerically low in the majority of cases.

One woman had three of her samples tested for HPV16 viral load on two occasions: all samples were found to have zero HPV16 viral load in all tests. Three women had a total of four samples which were tested for HPV18 viral load on two occasions. Results were in close agreement, with no discrepant results in terms of positive or negative status.

For the remainder of this chapter, where a woman had several replicates available for a given visit and/or more than one test result for the same sample, the qPCR status and viral load at that visit was determined by the mean viral load across all replicates and all tests.

Summary

The primers used in the qPCR assays for measuring both HPV16 and HPV18 viral load were shown to be sensitive, with a linear dynamic range exceeding 10^9 to one copies; and specific, with no cross-reactivity when tested in cell lines. Using either observed Ct values, or observed virus copy numbers, estimates of intraclass correlation coefficients (ICC) used to measure intra-assay reliability of the qPCR assay exceeded 0.99 and 0.97 for the HPV16 and HPV18 assay, respectively; similar results were obtained when GAPDH Ct values were analyzed: the intra-assay reliability of both assays was therefore excellent. For both assays, the ICC measuring inter-assay reliability exceeded 0.94: the inter-assay reliability of both assays was therefore also excellent. The assay was now considered to be suitable for use in testing samples from the study population.

22.3 SELECTION AND IDENTIFICATION OF STUDY POPULATIONS

In this section I, describe the identification of the study population for use in subsequent analyses.

22.3.1 Intended study population for the analysis of HPV16 and HPV18 viral load

The study population for the analysis of incident cervical HPV16 infection comprises all women who were HPV DNA-negative and cytologically normal in their first evaluable sample, and who were subsequently found to be HPV16 DNA-positive prior to the end of follow-up; HPV status was determined using the GP5+/GP6+ system, with type-specific primers for HPV16.

The study population for the analysis of incident cervical HPV18 infection comprises all women who were HPV DNA-negative and cytologically normal in their first evaluable sample, and who were subsequently found to be HPV18 DNA-positive prior to the end of follow-up; HPV status was determined using the GP5+/GP6+ system, with type-specific primers for HPV18.

One hundred and fifty five women fulfilled these criteria: 91 women acquired an incident cervical HPV16 infection, 45 acquired an incident cervical HPV18 infection, and 19 acquired both, at some point prior to the end of follow-up, either contemporaneously or otherwise. When appropriate, these three cohorts will be referred to as the “HPV16-only”, “HPV18-only” and “HPV16+HPV18” cohorts,

respectively. When combined by type, the 110 women who acquired incident cervical HPV16 infection during follow-up will be referred to as the “incident HPV16 cohort”; similarly, the 64 women who acquired incident cervical HPV18 infection during follow-up will be referred to as the “incident HPV18 cohort”.

22.3.2 Study population tested

It was the intention to test all women in the incident HPV16 cohort for HPV16 viral load, and all women in the incident HPV18 cohort for HPV18 viral load. To accomplish this, for practical reasons rather than any other, women in the study population were divided into subsets for testing.

1) First, women with a “complete” set of cervical cytological samples available would be tested for HPV16 and/or HPV18 viral load. These were women for whom samples from consecutive visits 1 to n (n>1) were available, with no missing samples within the sequence.

2) Second, women with an “incomplete” set of cervical cytological samples available would be tested for HPV viral load.

Given that all women would eventually be tested, this artificial division should have been of little consequence. As described in the methods section, prior to testing a study sample for viral load, all DNA in that sample was extracted and stored in appropriate conditions (see section 21.2.1). In the first instance, DNA was extracted

from samples from women in subset 1, which were then tested for viral load; and then DNA was extracted from women in subset 2, and these were then tested for viral load

The finding that some samples which were negative for HPV DNA of a specific type (or even of any type) using the GP5+/GP6+ system were nevertheless positive for that type according to the qPCR assay, was unexpected (see section 23.2). This implied that the study cohorts were potentially not as well-defined as anticipated; and that samples from women in the HPV16 cohort would also have to be tested for HPV18 viral load, and *vice versa*. At a stroke, this doubled the amount of work which it had been anticipated would be required for this assay. Although the qPCR assay was not difficult to perform *per se*, and comparatively inexpensive, the unanticipated requirement to test all samples for both HPV types was not budgeted for. The assay was also time consuming, and so had substantial opportunity costs associated with it. Due to competing priorities within the study, the decision was therefore made to stop the testing of samples when women from the “complete samples” group had been tested. At that point, sufficient data was available to draw conclusions about the utility of a measurement of HPV viral load.

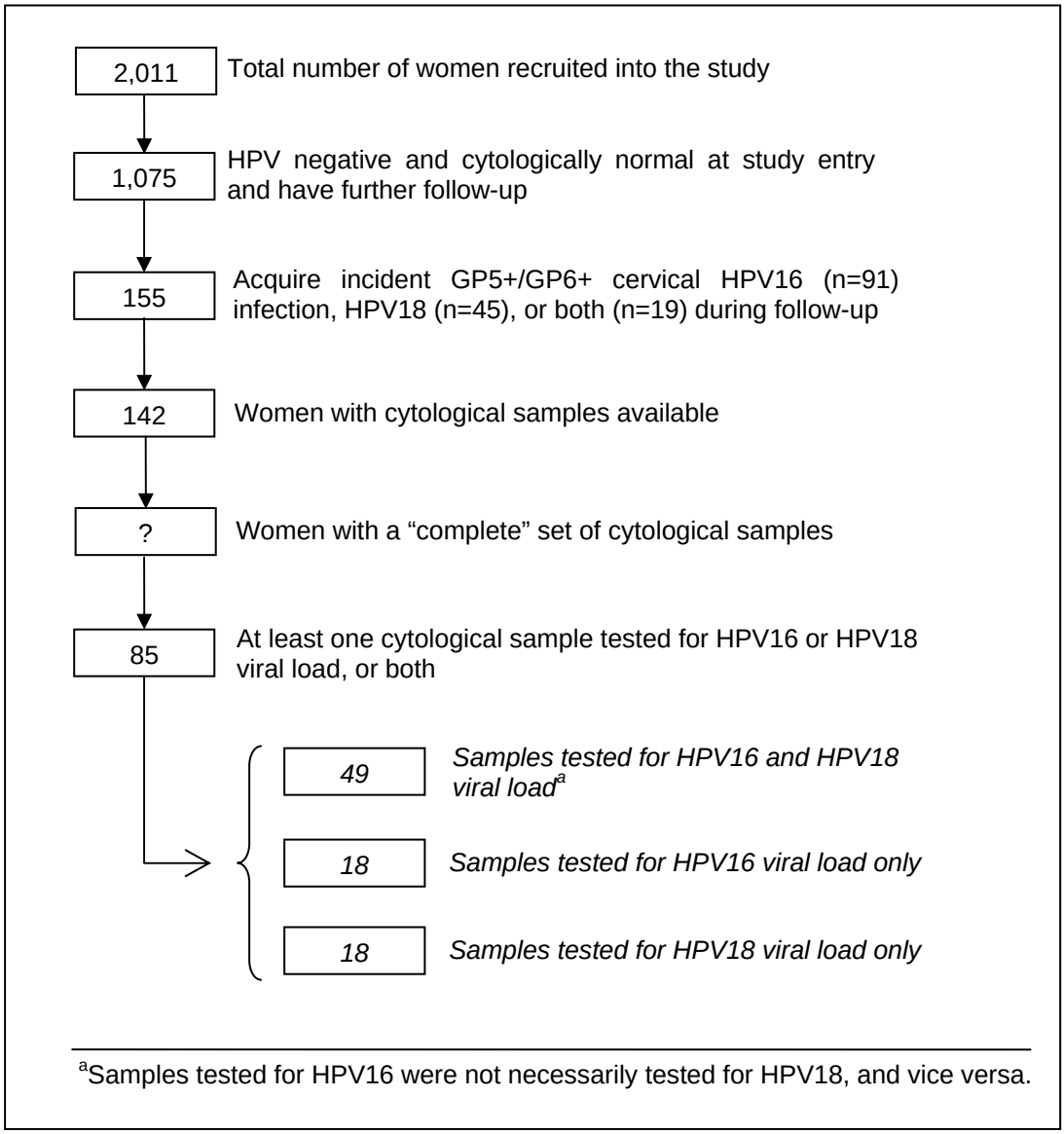
Table 22.3.2 describes the relationship between the intended study population and the population actually tested. All but one of the women in the HPV16+HPV18 cohort who could have been tested, were tested, compared with 86% of the women in the HPV18-only cohort, and 43% of women in the HPV16-only cohort.

Table 22.3.2. Relationship between the study population it was intended to test and the population actually tested for HPV16 and/or HPV18 viral load in the incident HPV16 or HPV18 cohorts. Numbers in the table are numbers of *women.*

Population	Cohort			Total
	HPV16-only	HPV18-only	HPV16+HPV18	
Intended study population	91	45	19	155
Number with cytological samples available	90	36	16	142
Population actually tested	39	31	15	85

The derivation of the final study population from the full cohort of 2,011 women recruited is shown in figure 22.3.2. The final study population for the analysis of HPV16 and HPV18 viral load comprised 85 women: 67 could contribute to the analysis of HPV16 viral load; (a different) 67 could contribute to the analysis of HPV18; and the 51 women common to both groups could contribute to the analysis of both types. For ease of reference, in this subsection, the 67 women will be referred to as the HPV16 cohort, the other 67 women the HPV18 cohort, and the 51 women as the HPV16+HPV18 cohort.

Figure 22.3.2. Derivation of the final study population tested for HPV16 and/or HPV18 viral load from the cohort of all 2,011 women recruited to the study. Numbers in the figure are numbers of *women.*



22.3.2.1 Follow-up of the study population tested

The availability and qPCR testing history of cervical cytological samples from each study visit for each woman in the study population are illustrated in table 22.3.2.1a.

For convenience of presentation, this table has been separated according to cohort. Some examples to illustrate interpretation may be helpful.

Note the deliberate choice of colours to aid interpretation: HPV16 testing is indicated by a blue cell, HPV18 testing by a yellow cell. When blue and yellow are mixed the result is green: hence a sample tested for both HPV16 and HPV18 is indicated by a green cell.

Example 1. The woman with study number (ID) 511 belongs to the HPV16-only cohort, and so appears in part one of the table. She made nine visits to the study: the maximum number of visits made by any woman was 16, hence this woman has her last seven cells “blacked out” (16-9=7). For eight of her visits, at least one cytological sample was available for testing; for the remaining visit (her last), no sample was ever taken, to the best of our knowledge. For her third visit, there were two samples available. This woman was unusual in that samples from visits 1, 2, 3, 5 and 6 were all tested for both HPV16 and HPV18 viral load; however, the sample from visit 5 was tested for HPV16 only, whereas the samples from visits 7 and 8 were tested for HPV18 viral load only.

Example 2. The woman with study number (ID) 1413 belongs to the HPV16+HPV18 cohort, and so appears in part three of the table. She made seven visits to the study and had a single cytological sample available for testing at all of her visits. All of her samples were tested for both HPV16 and HPV18 viral load. However, her samples from visits 2, 3 and 4, were each tested for HPV16 twice, but only once for HPV18.

Example 3. The woman with study number (ID) 253 belongs to the HPV16-only cohort and made 11 visits to the study. She had a single cytological sample available for testing from nine of her visits, but two samples were available for visit 8; the sample from visit 7 was missing. The samples from visits 1 to 6 inclusive were tested for HPV16 viral load only; all of her remaining samples were tested for both HPV16 and HPV18 viral load. Although she had two samples available at visit 8, testing was "discrepant" for the two types: one sample was tested for both types, but the other was tested for HPV16 only. The same occurred for study number (ID) 1011 at visit 7. In contrast, for study number (ID) 1814 at visit 10, both samples were tested for both HPV types, but one was tested twice for HPV18 but only once for HPV16; the single sample at visit 8 for this woman was also tested twice for HPV18, but only once for HPV16.

Example 4. The woman with study number (ID) 27 belongs to the HPV18-only cohort and made 14 visits to the study. She had a single cytological sample available for testing from 13 of her visits, but no sample was taken at visit 2. The samples from visits 6 to 14 inclusive were tested for HPV18 viral load only. However, although samples from visits 1, 3, 4 and 5 were available, they were not tested since there was insufficient DNA remaining in these samples.

Note: It is not always possible to distinguish a visit for which the cytological sample was missing, i.e. had been lost, from a visit at which no sample was ever taken. If a GP5+/GP6+ result was available, but the sample could not be found when an attempt was made to test it for viral load, then this sample was unequivocally

missing. However, at some visits no attempt was made to collect a cytological sample for HPV DNA testing. Typically, these were visits at the end of follow-up, after the diagnosis of high-grade CIN for example, or during the final follow-up phase of women with cytological abnormalities who were in the process of being "discharged" from the study. But this practice may also have occurred at interim visits. In a few instances, a sample for which no GP5+/GP6+ result was available was, however, located and tested for viral load (e.g. 69.14, 69.16, 128.1, 1011.5).

The median number of study visits was 8 (range 2 to 16). The median number of study visits for which at least one cytological sample was available was 7 (range 2 to 16), with four women providing only two samples and 17 providing at least 10. The median number of study visits for which no sample was available was 0 (range 0 to 6): 47 women had a "complete" set of samples available for testing. Fifty four women had a single sample tested from each of their visits, whereas 26 women had two samples tested from one of their visits, and five had two samples tested from two visits.

Table 22.3.2.1b summarises the viral load testing status for all women in the study population. Forty-nine women had at least one sample tested for both HPV16 and HPV18 viral load, compared with 18 whose samples were only tested for HPV16, and 18 whose samples were only tested for HPV18.

Figure 22.3.2.1a (Part 1-The HPV16-only cohort). Availability and qPCR testing status of cervical cytological samples from each study visit for each woman in the study population. ID=study number; C=Cohort; V=Number of study visits; N=number of visits with no sample; S=Number of visits for which at least one cytological sample is available; R=Number of visits for which two cytological samples are available; T=Total number of times the sample was tested; X=Number of samples tested twice for 16 and 18, separated by a dash; Blue=Tested for HPV16 viral load only; Yellow=Tested for HPV18 viral load only; Green=Tested for HPV16 and HPV18 viral load; Red=Not tested; Grey=Sample missing; Diagonal lines on white background=No sample taken; Black=End of full follow-up (a diagonal line in a coloured cell indicates a sample taken after first diagnosis of high-grade CIN and therefore after the formal end of follow-up); 2S=Two samples available for this visit; Tn/m=sample tested for HPV16[HPV18] n[m] times; E=Excised i.e. treated for high-grade CIN.

ID	C	V	N	S	R	T	X	Visit Number																
								1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	
18	16	13	0	13	0	13	0																	
69	16	16	0	16	0	16	0																	
73	16	3	0	3	0	3	0																	
150	16	12	2	10	0	10	0											E						
244	16	9	1	8	0	8	0																	
253	16	11	1	10	1	11	0								2S/TD									
287	16	4	0	4	0	4	0																	
305	16	2	0	2	0	2	0																	
320	16	4	0	4	0	4	0																	
330	16	5	1	4	0	4	0				E													
345	16	7	0	7	0	7	0																	
355	16	4	0	4	1	5	0				2S													
408	16	10	1	9	1	10	0						2S											
433	16	6	0	6	0	6	0																	
508	16	5	0	5	0	5	0																	
511	16	9	1	8	1	9	0			2S														
515	16	6	0	6	0	6	0																	
519	16	16	2	14	1	15	0																	2S
545	16	4	0	4	0	4	0																	
556	16	5	0	5	0	5	0																	
590	16	8	0	8	0	8	0																	
622	16	5	0	5	1	6	0				2S													
761	16	2	0	2	0	2	0																	
820	16	9	4	5	0	5	0																	
910	16	4	0	4	0	4	0																	
914	16	2	0	2	0	2	0																	
928	16	12	0	12	0	12	0																	
932	16	7	0	7	0	7	0																	
941	16	8	1	7	1	8	0			2S														
981	16	2	0	2	0	2	0																	
1005	16	7	3	4	0	4	0					E												
1011	16	11	2	9	1	10	0							2S/TD										
1066	16	4	0	4	0	4	0																	
1115	16	4	0	4	0	4	0																	
1181	16	14	2	12	1	13	0											2S						
1205	16	5	0	5	0	5	0																	
1231	16	8	0	8	0	8	0																	
1365	16	4	1	3	0	3	0																	
1367	16	12	1	11	2	13	0										2S	2S						

Figure 22.3.2.1a (Part 2-The HPV18-only cohort). Availability and qPCR testing status of cervical cytological samples from each study visit for each woman in the study population. ID=study number; C=Cohort; V=Number of study visits; N=number of visits with no sample; S=Number of visits for which at least one cytological sample is available; R=Number of visits for which two cytological samples are available; T=Total number of times the sample was tested; X=Number of samples tested twice for 16 and 18, separated by a dash; Blue=Tested for HPV16 viral load only; Yellow=Tested for HPV18 viral load only; Green=Tested for HPV16 and HPV18 viral load; Red=Not tested; Grey=Sample missing; Diagonal lines on white background=No sample taken; Black=End of full follow-up (a diagonal line in a coloured cell indicates a sample taken after first diagnosis of high-grade CIN and therefore after the formal end of follow-up); 2S=Two samples available for this visit; Tn/m=sample tested for HPV16[HPV18] n[m] times; E=Excised i.e. treated for high-grade CIN.

ID	C	V	N	S	R	T	X	Visit Number															
								1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*
27	18	14	5	9	0	9	0																
29	18	12	1	11	0	11	0																
43	18	10	0	10	0	10	0																
49	18	8	2	6	1	7	0			2S													
292	18	16	2	14	1	15	0																2S
381	18	12	1	11	1	12	0			2S													
406	18	6	0	6	0	6	0																
446	18	8	1	7	1	8	0							2S									
490	18	9	1	8	2	10	0			2S				2S									
524	18	8	0	8	1	9	0				2S												
561	18	3	0	3	1	4	0		2S														
577	18	4	0	4	0	4	0																
692	18	8	0	8	0	8	0																
739	18	5	0	5	0	5	0																
744	18	8	1	7	1	8	0			2S													
825	18	4	0	4	1	5	0		2S														
905	18	8	1	7	0	7	0																
938	18	13	6	7	1	8	0			2S													
1016	18	8	0	8	0	8	0																
1067	18	7	1	6	1	7	0			2S													
1086	18	7	1	6	0	6	0																
1276	18	4	0	4	0	4	0																
1278	18	6	0	6	0	6	0																
1296	18	9	0	9	0	9	0																
1430	18	9	1	8	1	9	0			2S													
1559	18	5	0	5	0	5	0																
1613	18	5	0	5	0	5	0																
1629	18	9	1	8	1	9	0								2S								
1665	18	6	0	6	0	6	0																
1732	18	3	0	3	0	3	0																
1814	18	12	3	9	1	10	0-2								T1/2		2S/TD	E					

Figure 22.3.2.1a (Part 3-The HPV16+HPV18 cohort). Availability and qPCR testing status of cervical cytological samples from each study visit for each woman in the study population. ID=study number; C=Cohort; V=Number of study visits; N=number of visits with no sample; S=Number of visits for which at least one cytological sample is available; R=Number of visits for which two cytological samples are available; T=Total number of times the sample was tested; X=Number of samples tested twice for 16 and 18, separated by a dash; Blue=Tested for HPV16 viral load only; Yellow=Tested for HPV18 viral load only; Green=Tested for HPV16 and HPV18 viral load; Red=Not tested; Grey=Sample missing; Diagonal lines on white background=No sample taken; Black=End of full follow-up (a diagonal line in a coloured cell indicates a sample taken after first diagnosis of high-grade CIN and therefore after the formal end of follow-up); 2S=Two samples available for this visit; Tn/m=sample tested for HPV16[HPV18] n[m] times; E=Excised i.e. treated for high-grade CIN.

ID	C	V	N	S	R	T	X	Visit Number															
								1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*
2	16+18	10	2	8	0	8	0																
92	16+18	5	1	4	0	4	0																
118	16+18	9	1	8	1	9	0	2S															
128	16+18	16	3	13	0	13	0																
147	16+18	11	0	11	1	12	0-1			2S					T1/2								
240	16+18	7	1	6	2	8	0		2S			2S											
247	16+18	9	0	9	0	9	0																
350	16+18	15	4	11	1	12	0												2S				
393	16+18	7	2	5	0	5	0					E											
1268	16+18	9	0	9	1	10	0	2S															
1413	16+18	7	0	7	0	7	3-0		T2/1	T2/1	T2/1												
1512	16+18	12	2	10	2	12	0									2S			2S				
1551	16+18	12	1	11	2	13	0										2S		2S				
1803	16+18	5	0	5	0	5	0																
1913	16+18	5	0	5	0	5	0																

Table 22.3.2.1b. Summary of the qPCR testing status for all women in the study population according to GP5+/GP6+ cohort.

	HPV16-only	HPV18-only	HPV16+ HPV18	Total
Number tested	39	31	15	85
Number of women with all samples tested for:				
HPV16 and HPV18	15	10	15	40
HPV16 only	17	0	0	17
HPV18 only	1	17[a]	0	18
Number of women with some samples tested for both HPV16 and HPV18, with others tested for:				
HPV16 only	4	2	0	6
HPV18 only	0	2	0	2
HPV16 and HPV18 but not both	2	0	0	2

[a]Includes woman with ID 27 for whom not all available samples were tested

Table 22.3.2.1c summarises the qPCR testing history of women with respect to cytological samples taken at each visit. During the course of making a total of 664 visits, the 85 women in the study population provided at least one sample for viral load testing on 600 of these visits, with samples from four of these visits (all from the woman with ID 27) not being tested for viral load.

Table 22.3.2.1c. Summary of the qPCR testing history of cytological samples taken at each visit according to GP5+/GP6+ cohort.

Cohort	Total number of visits	Number of visits with at least one sample tested for viral load:				Number of visits with:		
		HPV16 and HPV18	*HPV16 only*	*HPV18 only*	*Total*	Samples not tested	All samples missing	No sample taken
HPV16-only	279	136	114	6	256	0	3	20
HPV18-only	246	98	12	108	218	4	12	12
HPV16+HPV18	139	122	0	0	122	0	5	12
Total	664	356	126	114	596	4	20	44

Summary

Eighty-five women who first tested positive for HPV16, or HPV18, or both, using GP5+/GP6+ primers contributed samples to this analysis. Of these 39 women had tested positive during follow-up for HPV16 but not HPV18, 31 for HPV18 but not for HPV16, and 15 for both HPV16 and HPV18, either in the same sample or at different follow-up visits. Sixty-seven and 67 women had at least one sample tested for HPV16 or HPV18 viral load, respectively, with 51 women common to both groups. During the course of making a total of 664 visits, the 85 women in the study population provided at least one sample for viral load testing on 600 of these visits.

22.4 HPV TYPE AND THE DISTRIBUTION OF HPV VIRAL LOAD

22.4.1 HPV type and the distribution of HPV viral load in cervical samples which tested positive for HPV16 by qPCR

Sixty seven women had at least one cervical sample tested for HPV16 viral load; these samples were obtained during the course of making a total of 482 study visits. The median number of visits for which a woman had at least one sample tested for HPV16 viral load was 7 (range 2 to 16).

It is possible for a sample to be positive by qPCR, but for the viral load in that sample to be unevaluable i.e. it is not possible to obtain a numerical measurement of copy number. Results from six and two visits were non-evaluable for HPV16 viral load due to excessive variability in virus and GAPDH Ct values, respectively; and results from 16 visits were non-evaluable due to GAPDH failures (see section 21.2.6.1 and section 21.2.7 point 2). Among the results for the remaining 458 study

visits, 251 (52%) from 64 women had a non-zero HPV16 viral load; and 207 (48%) from 54 women had a HPV16 viral load of zero.

The median number of visits per woman for which an evaluable measurement of HPV16 viral load was available was 7 (range 2 to 15). The median number of visits per woman for which a non-zero measurement of HPV16 viral load was obtained was 3 (range 0 to 10). The distribution of HPV16 viral load in study samples which tested qPCR-positive for HPV16, and which had evaluable results, is shown in tables 22.4.1a and 22.4.1b.

In samples which were found to test qPCR-positive for HPV16, the median HPV16 viral load was 41 copies per 1,000 cells. Eight (3%) of the qPCR HPV16-positive samples were found to have a HPV16 viral load in excess of 1 million copies per 1,000 cells; four women had one such sample, and one woman had three taken at consecutive visits with intervals of 10 and three months between visits. Forty seven (19%) of the HPV16-qPCR-positive samples were found to have a HPV16 viral load of less than 1 copy per 1,000 cells, with the smallest viral load observed being 7 copies per million cells. Twenty eight women had at least one such “weak-positive” result: 15 had one, seven had two, and six had three; for four women, all of their positive results were weak-positives.

Table 22.4.1a. The distribution of HPV16 viral load (expressed as copy number per 1,000 cells) in all cervical cytological samples with evaluable HPV16 viral loads and which were HPV16-qPCR-positive.

HPV16 Viral Load[a]	Number of samples	
	n	%
(0,1)	47	18.7
[1,100)	92	36.7
[100,1000)	34	13.5
[1000,1000000)	71	28.3
1,000,000+	7	2.8
Total	251	100.0

[a][x, y) is a mathematical convention indicating that viral load is greater than or equal to x, but less than y

Table 22.4.1b. The distribution of HPV16 viral load (expressed as copy number per 1,000 cells) in cervical cytological samples with evaluable HPV16 viral loads according to the qPCR system. N is the number of evaluable samples; W is the number of "weak positive" samples, defined as samples with a non-zero viral load of less than 1 copy per 1,000 cells.

Group	Number of Women	HPV16 Viral Load				
		N	*W*	*Median*	*Min*	*Max*
HPV16-qPCR-positive	64	251	47	41	0.01	420,953,588

22.4.2 HPV type and the distribution of HPV viral load in cervical samples which tested positive for HPV18 by qPCR

Sixty-seven women had at least one sample tested for HPV18 viral load; these samples were obtained during the course of making a total of 470 study visits. The median number of visits for which a woman had at least one sample tested for HPV18 viral load was 7 (range 2 to 14).

Results from nine and one visits were non-evaluable for HPV18 viral load due to excessive variability in virus and GAPDH Ct values, respectively; and results from 16 visits were non-evaluable due to GAPDH failures. Among the results for the remaining 444 study visits, 235 (53%) from 60 women had a non-zero HPV18 viral load; and 209 (47%) from 55 women tested qPCR-negative for HPV18.

The median number of visits per woman for which an evaluable measurement of HPV18 viral load was available was 6 (range 2 to 14). The median number of visits per woman at which a non-zero measurement of HPV18 viral load was obtained was 1 (range 0 to 14). The distribution of HPV18 viral load in study samples which tested qPCR-positive for HPV18 and which had evaluable results, is shown in tables 22.4.2a and 22.4.2b.

In samples which were found to test qPCR-positive for HPV18, the median HPV18 viral load was 4 copies per 1,000 cells. Two (0.9%) of the HPV18-qPCR-positive samples were found to have a HPV18 viral load in excess of 1 million copies per 1,000 cells; two women had one such sample each. Seventy-three (31%) of the qPCR HPV18-positive samples were found to have a HPV18 viral load of less than 1 copy per 1,000 cells, with the smallest viral load observed being 0.1 copies per million cells. Forty women had at least one such “weak-positive” result: 20 had one, eight had two, 11 had three, and one had four; for 14 women, all of their positive results were weak-positives.

Table 22.4.2a. The distribution of HPV18 viral load (expressed as copy number per 1,000 cells) in all cervical cytological samples with evaluable HPV18 viral loads and which were HPV18-qPCR-positive.

HPV18 Viral Load[a]	Number of samples	
	n	%
(0,1)	73	31.1
[1,100)	97	41.3
[100,1000)	29	12.3
[1000,1000000)	34	14.5
1,000,000+	2	0.9
Total	235	100.0

[a][x, y) is a mathematical convention indicating that viral load is greater than or equal to x, but less than y

Table 22.4.2b. The distribution of HPV18 viral load (expressed as copy number per 1,000 cells) in all cervical cytological samples with evaluable HPV18 viral loads according to the qPCR system. N is the number of evaluable samples; W is the number of "weak positive" samples, defined as samples with a non-zero viral load of less than 1 copy per 1,000 cells.

Group	Number of Women	HPV18 Viral Load				
		N	*W*	*Median*	*Min*	*Max*
HPV18-qPCR-positive	60	235	73	4	0.0001	25,564,280

22.4.3 HPV type and the distribution of HPV viral load in cervical samples according to both HPV16 and HPV18 status

The study population for this analysis is restricted to samples which were qPCR-evaluable for both HPV16 and HPV18. This is to ensure that all results relate to the same study population, allowing direct comparisons to be made.

Forty-nine women had at least one sample tested for both HPV16 and HPV18 viral load; these samples were obtained from a total of 356 study visits. The median

number of visits per woman for which an evaluable qPCR result was available for both HPV16 and HPV18 was 7 (range 2 to 14). Results from five and one visits were non-evaluable for HPV16 viral load only due to excessive variability in virus and GAPDH Ct values, respectively; the corresponding numbers for the HPV18 analysis are six and zero. The median number of visits per woman for which an evaluable measurement of both HPV16 and HPV18 viral load was available was 7 (range 2 to 14). Forty-nine women had 328 study visits for which an evaluable measurement of both HPV16 and HPV18 viral load was available.

The distribution of HPV16 and HPV18 viral load in all cervical cytological samples for which both an evaluable HPV16 and HPV18 viral load was available is shown in table 22.4.3. HPV16 viral load in HPV16-qPCR-positive samples which were also HPV18-qPCR-positive was slightly lower than in HPV16-qPCR-positive samples which were HPV18-qPCR-negative (45 compared to 69 copies per 1,000), but the difference was slight. Differences in HPV18 viral load according to HPV16 qPCR status were trivial, but the median viral loads were also very small.

Table 22.4.3. The distribution of HPV16 and HPV18 viral load (expressed as copies per 1,000 cells) in all cervical cytological samples for which both evaluable HPV16 and HPV18 viral loads were available. Rows in the table are not mutually exclusive. For example, the 166 samples which were HPV16 qPCR-positive (labelled "HPV16+ve") combines two groups of samples: the 60 which were also HPV18 qPCR-positive, and the 106 which were HPV18 qPCR-negative.

Sample Classification[a]	Number of women	Number of samples	Viral Load (Copies per 1,000 cells)					
			HPV16			HPV18		
			Median	*Min*	*Max*	*Median*	*Min*	*Max*
HPV16+ve; HPV18-ve	32	106	69	0.01	4,803,266	0	0	0
HPV16-ve; HPV18+ve	22	77	0	0	0	3	0.0001	25,564,280
HPV16+ve; HPV18+ve	34	60	45	0.02	420,953,588	2	0.0001	236,315
HPV16+ve	45	166	52	<1	420,953,588	0	0	236,315
HPV18+ve	39	137	0	0	420,953,588	3	<1	25,564,280
HPV16-ve	40	162	0	0	0	0	0	25,564,280
HPV18-ve	49	191	0.27	0	4,803,266	0	0	0
Total	49	328	0.03	0	420,953,588	0	0	25,564,280

[a]According to qPCR

Summary

In qPCR-positive samples, the median HPV16 and HPV18 viral loads were numerically low, but were greater for HPV16 than for HPV18. In samples for which both an evaluable HPV16 and HPV18 viral load was available, viral load for one type did not differ greatly according to the presence or absence of the other type.

22.5 CERVICAL CYTOLOGICAL STATUS AND THE DISTRIBUTION OF HPV VIRAL LOAD

The study population for this analysis comprises all visits with evaluable viral load results for both HPV16 and HPV18, and for which an evaluable cytological smear result was available.

Forty-nine women had at least one sample tested for both HPV16 and HPV18 viral load, obtained from a total of 356 study visits. Results from 45 visits were excluded:

27 due to GAPDH failures, or excessive variability in the Ct values for either GAPDH or HPV, in the HPV16 and/or the HPV18 qPCR assays; 17 due to the cytological smear result being missing, or the smear being unreadable; and one due to both reasons. The final study population for this analysis thus comprises evaluable viral load and cytological results from 311 study visits made by 49 women.

22.5.1 Cytologically normal cervical smears

Forty-eight women had a total of 224 cytologically normal cervical smears with evaluable qPCR results for both HPV16 and HPV18 in the corresponding virological sample. The median number of normal cervical smears per woman was 5 (range 0 to 10). One woman contributed no cytologically normal cervical smears to the analysis since qPCR results were not evaluable in the corresponding virological sample.

The distribution of HPV16 and HPV18 viral load in all cytologically normal cervical smears for which both evaluable HPV16 and HPV18 viral loads were available, is shown in table 22.5.1. The median viral load in cytologically normal cervical smears was zero, for both HPV16 and HPV18. However, clearly some women had substantial HPV16 or HPV18 viral loads at the same time as they had no *detectable* cervical cytological abnormality. Viral load for one of the HPV types did not vary according to the presence or absence of the other type.

Table 22.5.1. The distribution of HPV16 and HPV18 viral load (expressed as copies per 1,000 cells) in all cytologically normal cervical smears for which both evaluable HPV16 and HPV18 viral loads were available. Rows in the table are not mutually exclusive. For example, the 89 samples which were HPV18 qPCR-positive comprise two groups of samples: the 35 which were also HPV16 qPCR-positive, and the 54 which were HPV16 qPCR-negative.

Sample Classification[a]	Women	Samples	Viral Load (Copies per 1,000 cells)					
			HPV16			HPV18		
			Median	*Min*	*Max*	*Median*	*Min*	*Max*
HPV16+ve; HPV18-ve	27	66	5	0.01	1,341,411	0	0	0
HPV16-ve; HPV18+ve	19	54	0	0	0	1	0.0001	25,564,280
HPV16+ve; HPV18+ve	26	35	4	0.02	5,482	1	0.0006	85,024
HPV16+ve	41	101	5	0.01	1,341,411	0	0	85,024
HPV18+ve	34	89	0	0	5,482	1	0.0001	25,564,280
HPV16-ve	38	123	0	0	0	0	0	25,564,280
HPV18-ve	40	135	0	0	1,341,411	0	0	0
Total	48	224	0	0	1,341,411	0	0	25,564,280

[a]According to qPCR

22.5.2 Cytologically abnormal cervical smears

Thirty women had a total of 87 cytologically abnormal cervical smears with evaluable qPCR results for both HPV16 and HPV18 in the corresponding virological sample. The median number of abnormal cervical smears per woman was 3 (range 1 to 7). The number of cervical smears reported as containing borderline nuclear abnormalities (BNA) only, mild dyskaryosis, moderate dyskaryosis, or severe dyskaryosis, were 36 (41%), 47 (54%), 3 (3%) and 1 (1%), respectively.

The distribution of HPV16 and HPV18 viral load in all cytologically abnormal cervical smears for which both evaluable HPV16 and HPV18 viral loads were available, by severity of abnormality, is shown in table 22.5.2.

The median HPV16 viral load in all cytologically abnormal cervical smears was 74 copies per 1,000 cells, compared to 0 for cytologically normal cervical smears. In smears with a non-zero HPV16 viral load, the median HPV16 viral load was substantially higher when cytological abnormalities were present (2,134 copies per 1,000 cells) than when they were not (5 copies per 1,000 cells). HPV16 viral load increased with increasing severity of cervical cytological abnormality, within the limited range available in this analysis. In all abnormal cervical smears combined, HPV16 viral load did not vary according to the presence or absence of HPV18 in these smears; however, in cervical smears containing at worst BNA, HPV16 viral load was higher in the absence of HPV18 than when it was present; the reverse was true in cervical smears containing at worst mild dyskaryosis. In both cases, the number of samples in the analysis was small.

The median HPV18 viral load in all cytologically abnormal cervical smears was zero, as it was in cytologically normal smears. However, the greatest HPV18 viral load observed in cytologically *abnormal* smears was considerably less than that seen in cytologically *normal* smears. In smears with a non-zero HPV18 viral load, the median HPV18 viral load was substantially higher when cytological abnormalities were present (376 copies per 1,000 cells) than when they were not (1 copy per 1,000 cells), a much smaller difference than for HPV16. In contrast to HPV16, HPV18 viral load decreased with increasing severity of cervical cytological abnormality, although the numbers of samples available in these analyses were clearly small. In all abnormal cervical smears combined, and in smears containing at

worst BNA, HPV18 viral load was greater in the absence of HPV16 than in its presence; the reverse was true in smears containing at worst mild dyskaryosis.

Table 22.5.2. The distribution of HPV16 and HPV18 viral load in all cytologically abnormal cervical smears for which both evaluable HPV16 and HPV18 viral loads were available, by severity of abnormality. Note that in this table, for example, the 42 cytologically abnormal cervical smears which were HPV18 qPCR-positive combine two groups of smears: the 22 which were also HPV16 qPCR-positive, and the 20 which were HPV16 qPCR-negative.

Sample Classification[a]	Number of women	Number of samples	Viral Load (Copies per 1,000 cells)					
			HPV16			HPV18		
			Median	*Min*	*Max*	*Median*	*Min*	*Max*
Cytology abnormal								
HPV16+ve; HPV18-ve	18	38	2,134	0.6	4,803,266	0	0	0
HPV16-ve; HPV18+ve	9	20	0	0	0	599	0.9	332,323
HPV16+ve; HPV18+ve	19	22	1,533	0.2	420,953,588	39	0.02	236,315
HPV16-ve; HPV18-ve	5	7	0	0	0	0	0	0
HPV16+ve	27	60	2,134	0.2	420,953,588	0	0	236,315
HPV18+ve	20	42	2	0	420,953,588	376	0.02	332,323
HPV16-ve	13	27	0	0	0	17	0	332,323
HPV18-ve	20	45	748	0	4,803,266	0	0	0
Total	30	87	74	0	420,953,588	0	0	332,323
Cytology BNA								
HPV16+ve; HPV18-ve	10	12	304	0.6	78,795	0	0	0
HPV16-ve; HPV18+ve	7	12	0	0	0	837	0.9	144,188
HPV16+ve; HPV18+ve	8	9	47	5	48,806	11	0.03	236,315
HPV16+ve	15	21	123	0.6	78,795	0	0	236,315
HPV18+ve	13	21	0	0	48,806	693	0.03	236,315
HPV16-ve	10	15	0	0	0	505	0	144,188
HPV18-ve	11	15	166	0	78,795	0	0	0
Total	20	36	8	0	78,795	0.7	0	236,315
Cytology Mild Dyskaryosis								
HPV16+ve; HPV18-ve	13	22	7,142	1	4,803,266	0	0	0
HPV16-ve; HPV18+ve	4	8	0	0	0	12	2	332,323
HPV16+ve; HPV18+ve	11	13	9,530	0.2	420,953,588	42	0.02	31,265
HPV16+ve	21	35	9,530	0.2	420,953,588	0	0	31,265
HPV18+ve	12	21	391	0	420,953,588	37	0.02	332,323
HPV16-ve	6	12	0	0	0	5	0	332,323
HPV18-ve	14	26	1,668	0	4,803,266	0	0	0
Total	23	47	748	0	420,953,588	0	0	332,323

[a]According to qPCR

Summary

The median viral load in cytologically normal cervical smears was zero, for both HPV16 and HPV18, but with some substantial viral loads measured for both types. In abnormal cervical smears, the median HPV16 viral load was 74 copies per 1,000 cells, compared to zero for HPV18. For both HPV types, cervical smears with non-zero viral loads had substantially higher viral loads when cytological abnormalities were present than when they were not. HPV16, but not HPV18, viral load increased with increasing severity of cervical cytological abnormality. In normal cervical smears, viral load for one HPV type (HPV16 or HPV18) did not vary according to the presence or absence of the other type. In cervical smears containing at worst BNA, viral load of either HPV type was higher in the absence of the other type than when it was present; the reverse was true in smears containing at worst mild dyskaryosis.

22.6 CHANGES IN HPV VIRAL LOAD OVER TIME

Data is available from 85 women with incident cervical (GP5+/GP6+) HPV16 or HPV18 cervical infections with which to describe how viral load changes over time. Typically, the first stage in a longitudinal analysis is to model and summarise these changes. The aim of statistical modelling in this situation, as in others, is to capture the important features of the underlying process, without being unduly influenced by irrelevant fluctuations in measurements. If this can be done, the data will be reduced from a large number of individual viral load profiles, to a manageable set of parameters describing key aspects of how viral load changes over time.

I present changes over time in HPV16 and HPV18 viral load for all women tested for HPV16 or HPV18 viral load, or both in figure 22.6. For ease of comparison, all figures are plotted on the same time scale (x-axis) and viral load (y-axis) scales. In each case, time is measured from the date of the first evaluable smear. For aesthetic

reasons, and to avoid giving prominence to very small viral loads, all samples with a viral load of 0.1 copies per 1,000 cells (i.e. 1 copy per 10,000 cells) or less, including samples with a zero viral load, were given a viral load value of 0.1; when plotting this figure, the label for 10^{-1} was changed to “0”. Again for aesthetic reasons, if a sample from a visit was not available, not tested, or not evaluable, the previously measured value of viral load was plotted (observed values are plotted with a closed plotting symbol, backwards-interpolated values with an open symbol) to form a left-continuous curve. This is a pragmatic approach, and matches how missing viral load would presumably be handled in practice, in a screening programme for example: a woman’s viral load status will be assumed to be that of her last measurement, until updated with a new measurement. Unique identifiers for women (study numbers) appear in bold in the top right-hand corner of each of the graphs which appear on a page. A vertical coloured line indicates the date of the first detection of cervical cytological abnormality: a blue line represents a smear containing at worst borderline nuclear abnormalities; a green line, mild dyskaryosis (other severities of abnormality were not detected).

I have been unable to find a suitable statistical method for modelling changes in viral load over time. Any model must be able to simultaneously account for large numbers of zero viral loads, small viral loads, and extreme viral loads. Within the range of statistical methods available, in particular for repeated measurements, I have been unable to find a model which meets these criteria. In addition, it seems unlikely that a *single* statistical model will be able to adequately describe the changes in HPV18 viral load seen for study numbers 292, 1367, 1512 and 1814 (figure 22.6). It appears

Figure 22.6. Changes in HPV16 (—) and HPV18 (—) viral load over time. All graphs are plotted on the same time scale (months) and vertical scale (HPV copy number per 1,000 cells). A closed (open) plotting symbol indicates that a sample was taken (was not taken, or was unavailable): viral load retains its last known value until updated. A coloured vertical line indicates date of first detection of cytological abnormality (blue=borderline nuclear abnormalities; green=mild dyskaryosis). The woman's study number appears in the top right-hand corner.

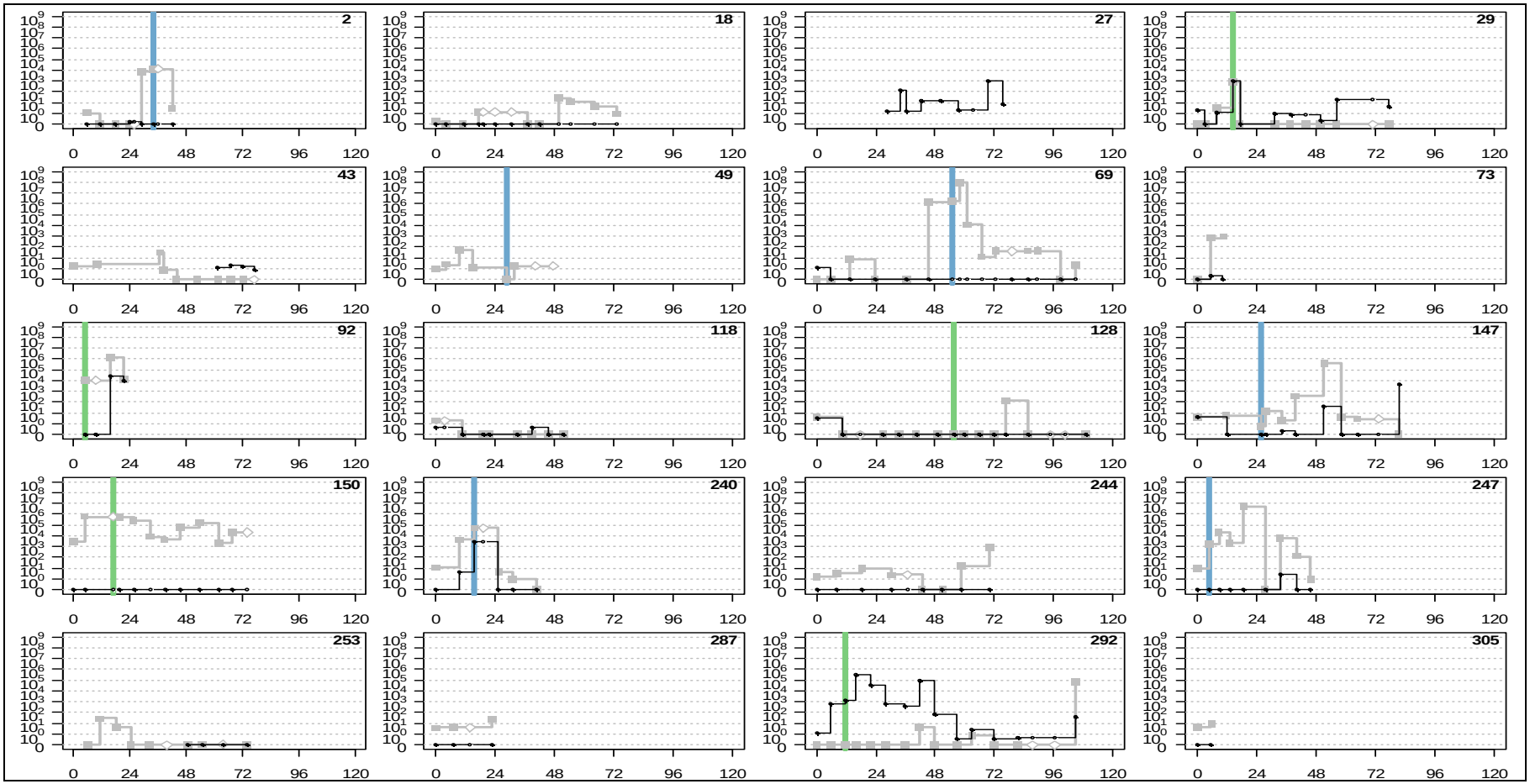


Figure 22.6 (continued). Changes in HPV16 (—) and HPV18 (—) viral load over time. All graphs are plotted on the same time scale (months) and vertical scale (HPV copy number per 1,000 cells). A closed (open) plotting symbol indicates that a sample was taken (was not taken, or was unavailable): viral load retains its last known value until updated. A coloured vertical line indicates date of first detection of cytological abnormality (blue=borderline nuclear abnormalities; green=mild dyskaryosis). The woman's study number appears in the top right-hand corner.

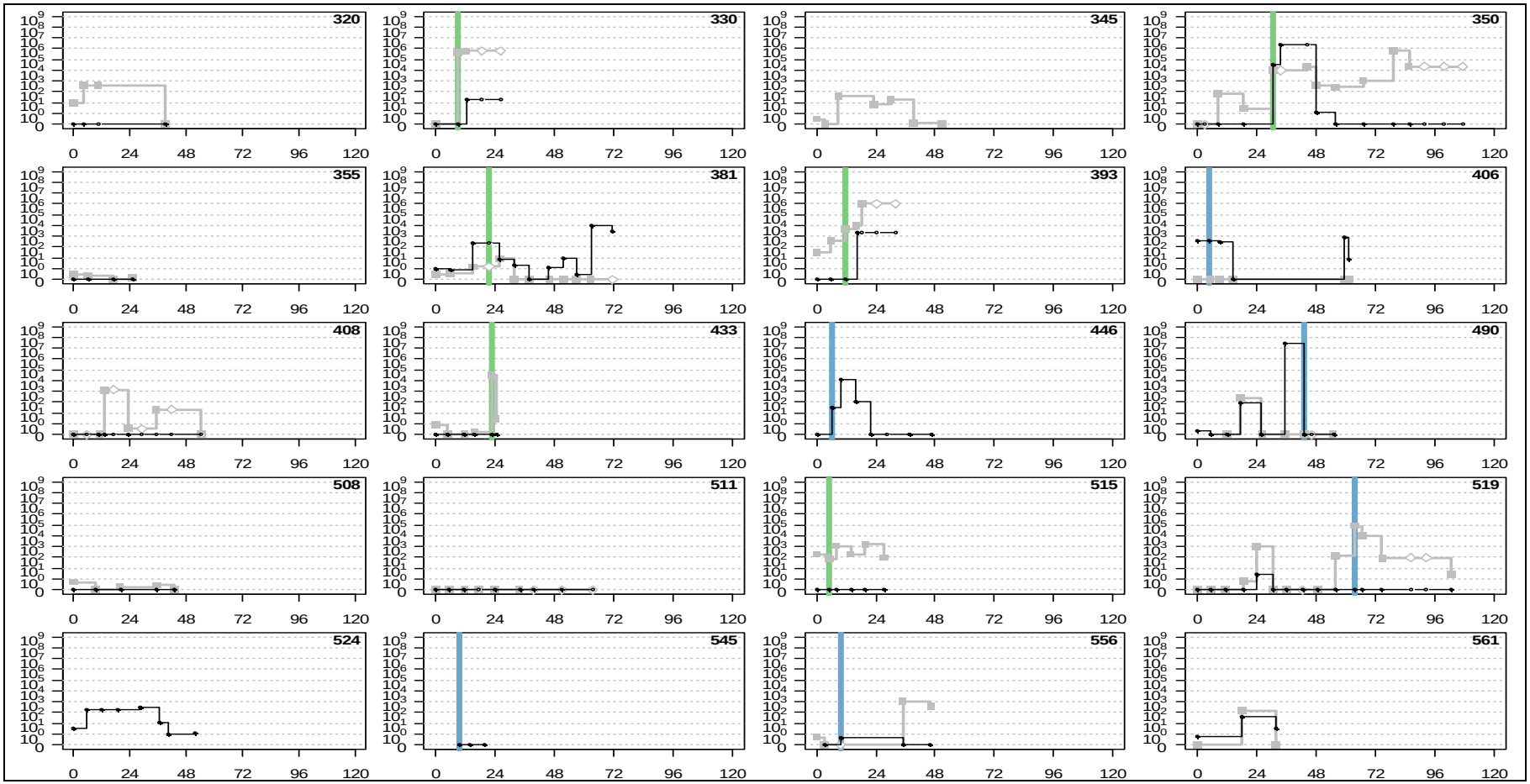


Figure 22.6 (continued). Changes in HPV16 (—) and HPV18 (—) viral load over time. All graphs are plotted on the same time scale (months) and vertical scale (HPV copy number per 1,000 cells). A closed (open) plotting symbol indicates that a sample was taken (was not taken, or was unavailable): viral load retains its last known value until updated. A coloured vertical line indicates date of first detection of cytological abnormality (blue=borderline nuclear abnormalities; green=mild dyskaryosis). The woman's study number appears in the top right-hand corner.

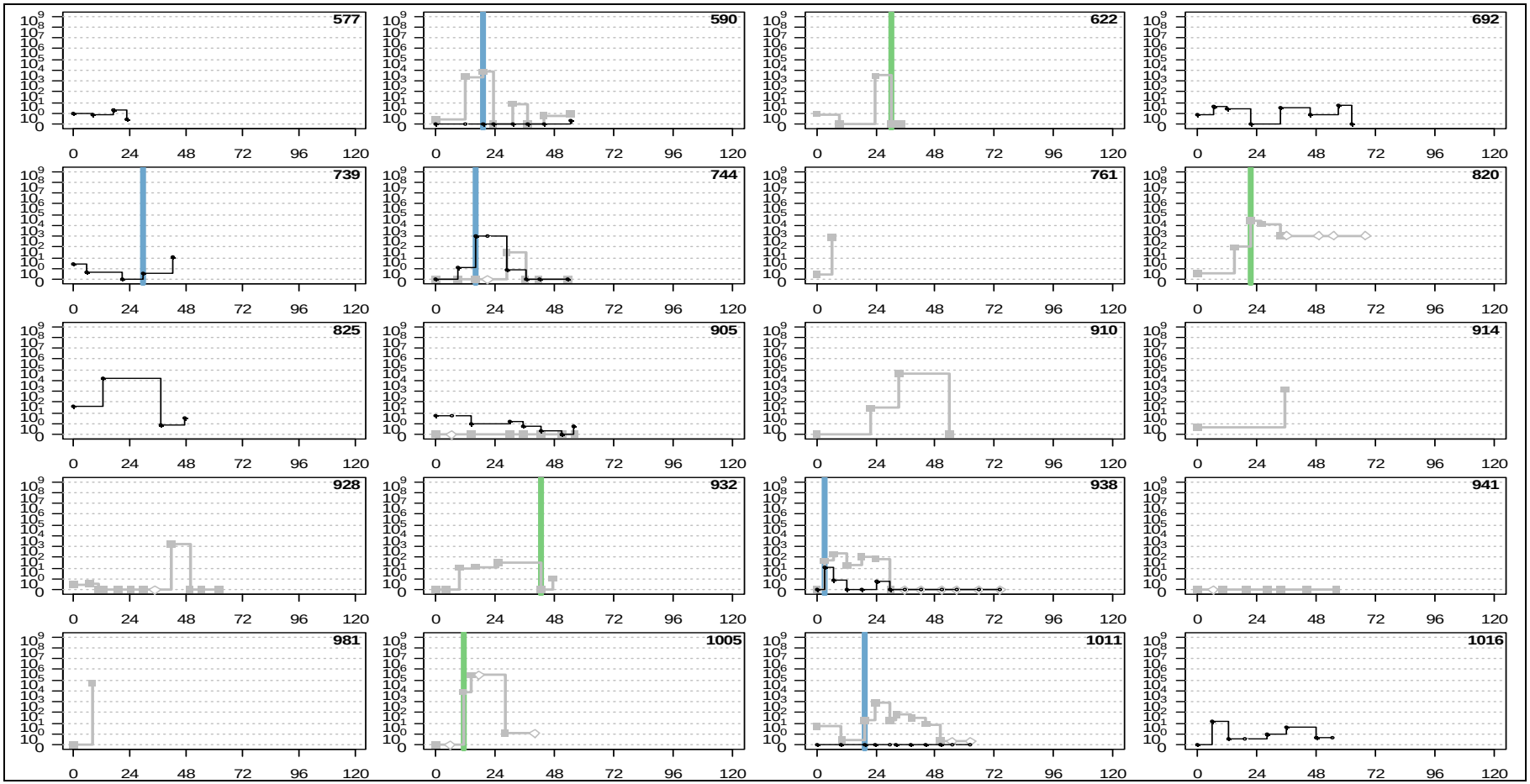


Figure 22.6 (continued). Changes in HPV16 (—) and HPV18 (—) viral load over time. All graphs are plotted on the same time scale (months) and vertical scale (HPV copy number per 1,000 cells). A closed (open) plotting symbol indicates that a sample was taken (was not taken, or was unavailable): viral load retains its last known value until updated. A coloured vertical line indicates date of first detection of cytological abnormality (blue=borderline nuclear abnormalities; green=mild dyskaryosis). The woman's study number appears in the top right-hand corner.

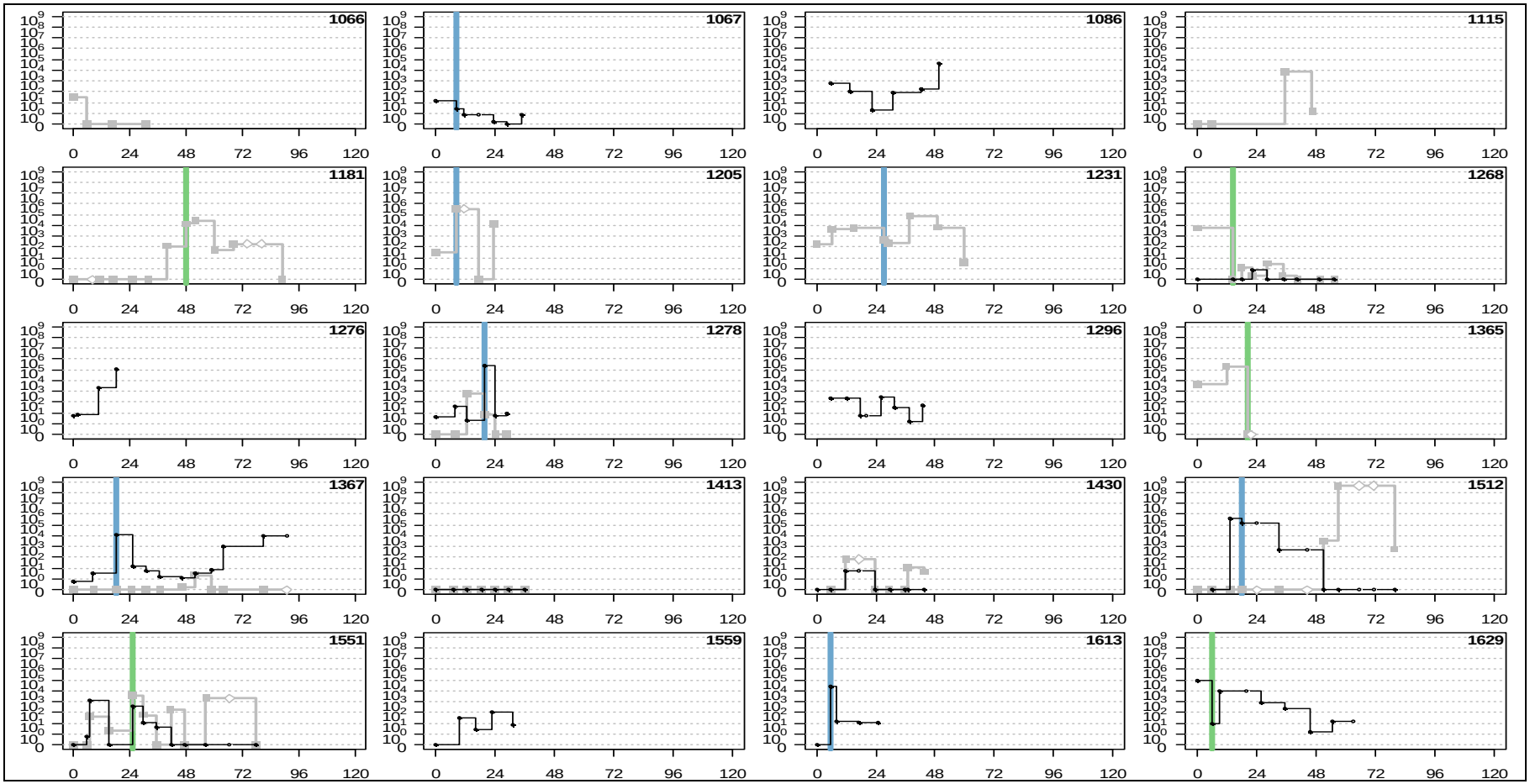
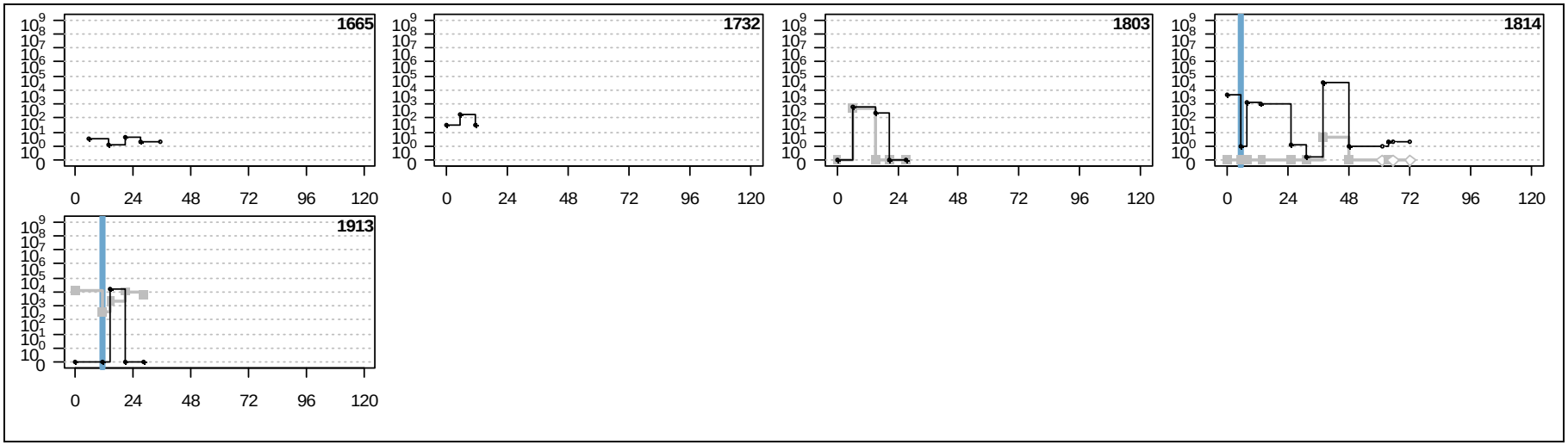


Figure 22.6 (continued). Changes in HPV16 (—) and HPV18 (—) viral load over time. All graphs are plotted on the same time scale (months) and vertical scale (HPV copy number per 1,000 cells). A closed (open) plotting symbol indicates that a sample was taken (was not taken, or was unavailable): viral load retains its last known value until updated. A coloured vertical line indicates date of first detection of cytological abnormality (blue=borderline nuclear abnormalities; green=mild dyskaryosis). The woman's study number appears in the top right-hand corner.



that each woman requires a specific model tailored to her own profile, which defeats the aim of modelling as set out at the start of this subsection. Although these profiles may fall into a smaller set of apparently common groups, these would be artificial and for convenience only, and would not necessarily shed any light on the evolution of HPV viral load over time in these women.

22.7 HPV VIRAL LOAD AND THE DURATION OF INCIDENT CERVICAL HPV INFECTIONS

The association between viral load and the duration of an incident cervical qPCR HPV infection was investigated in a time-to-event analysis. Recall that in a time-to-event analysis of duration, the event is “clearing” the infection (see section 21.4.2.3); and therefore, a hazards ratio less than one indicates a *decreased risk* of clearing an infection compared to the reference group, or alternatively an infection of longer duration.

Three variables measuring HPV viral load were considered for analysis: a time-fixed variable measuring viral load at the start of the episode; a time-varying covariate measuring viral load at the previous visit; and a time-varying covariate measuring maximum viral load attained by the previous visit. All numerical values of viral load were log-transformed.

22.7.1 HPV16

The study population for this analysis comprised 27 women who had incident cervical HPV16 qPCR infections, and who had further follow-up after their first HPV16 qPCR-positive sample.

Twenty women cleared their incident cervical HPV16 infections; the median duration of an incident cervical HPV16 infection was 11.0 months (inter-quartile range (IQR) 6.3 to 28.8).

22.7.1.1 HPV16 viral load in the first qPCR-positive sample and the duration of incident cervical HPV16 qPCR infections

A multivariate analysis for interval censored time-to-event data was used to determine the association between the HPV16 viral load observed at the start of an episode of infection, and the probability of clearing that infection. The results are shown in table 22.7.1.1.

For women who cleared their incident cervical HPV16 qPCR infections, median (IQR) [range] HPV16 viral load in the first sample of the episode was 29 copies per 1,000 cells (3 to 81) [0.03 to 676]; in women who did not clear their HPV16 infections, the median (IQR) [range] viral load was 3,270 copies per 1,000 cells (381 to 8,170) [0.01 to 397,966]. The duration of an incident cervical HPV16 qPCR infection increased as the viral load in the first qPCR-positive sample increased;

however, the association was not statistically significant. The risk of clearing an infection decreased by 9% (1.00-0.91) for every ten-fold increase in viral load in the first sample of an episode of infection, i.e. the hazards ratio was 0.91 (1-0.91=9%) for every increase of one in log_{10} of viral load, which on this scale corresponds to a ten-fold increase in viral load.

Table 22.7.1.1. The association between HPV16 viral load in the first sample of an incident episode of infection and the duration of incident cervical qPCR HPV16 infections.

Variable	Frequency	Events	Hazards ratio	95% CI	LR test[a]
Log_{10} of the HPV16 viral load in the first sample	27	20	0.91	0.70 to 1.18	0.58;1;0.45

[a]Likelihood-ratio test: Chi-squared statistic; degrees of freedom; p-value.

22.7.1.2 Viral load in the sample taken at the previous visit and the duration of incident cervical qPCR HPV16 infections

A multivariate analysis for interval censored time-to-event data was used to determine the association between the HPV16 viral load in the sample taken at the previous visit as a time-varying covariate, and the probability of clearing that infection. The results are shown in table 22.7.1.2.

The duration of an incident cervical qPCR HPV16 infection increased as the viral load in the sample taken at the previous visit increased; however, the association was not statistically significant. The risk of clearing an infection decreased by 19%

(1.00-0.81) for every ten-fold increase in viral load in the sample taken at the previous visit.

Table 22.7.1.2. The association between HPV16 viral load in the sample taken at the previous visit and the duration of an incident cervical qPCR HPV16 infection.

Variable	Frequency	Events	Hazards ratio	95% CI	LR test[a]
Log_{10} of the HPV16 viral load at the previous visit	27	20	0.81	0.63 to 1.05	3.23;1;0.07

[a]Likelihood-ratio test: Chi-squared statistic; degrees of freedom; p-value.

22.7.1.3 Greatest viral load observed and the duration of incident cervical qPCR HPV16 infections

A multivariate analysis for interval censored time-to-event data was used to determine the association between the greatest HPV16 viral load observed during the episode of infection, measured as a time-varying covariate updated at each visit, and the probability of clearing that infection. The results are shown in table 22.7.1.3.

Each woman had a maximum viral load observed during her episode of infection: when the median of these viral loads is taken we obtain the “median maximum” viral load. For women who cleared their HPV16 infections, the median (IQR) [range] maximum HPV16 viral load observed during their infections was 106 copies per 1,000 cells (6 to 694) [0.03 to 45,774]; in women who did not clear their HPV16 infections, the median (IQR) [range] maximum HPV16 viral load was 282,186 copies

per 1,000 cells (4,713 to 631,595) [884 to 420,953,588]. The duration of an incident HPV16 infection increased as the maximum viral load attained during the episode increased; the association was of borderline statistical significance (I dislike this phrase, but use it because of the disagreement between the 95% confidence interval and the p-value from the likelihood ratio test). The risk of clearing an infection decreased by 25% (1-0.75) for every ten-fold increase in the maximum viral load observed during the episode of infection.

Table 22.7.1.3. The association between the maximum HPV16 viral load observed during an episode of infection and the duration of an incident cervical qPCR HPV16 infection.

Variable	Frequency	Events	Hazards ratio	95% CI	LR test[a]
Log_{10} of the greatest HPV16 viral load observed to date	27	20	0.75	0.55 to 1.02	4.24;1;0.04

[a]Likelihood-ratio test: Chi-squared statistic; degrees of freedom; p-value.

Summary

An increase in viral load was associated with a reduced probability of clearing an incident HPV16-qPCR infection, but the association was not statistically significant. This was the case whether the analysis included a variable measuring viral load in the first positive sample of the episode, a variable measuring viral load in the previous evaluable sample, or a variable measuring the greatest viral load attained during the episode. There was very little difference between the estimates of association. The magnitude of the association was least for the variable measuring viral load in the first positive sample, and greatest for that measuring greatest viral load attained to date, which was of borderline significance. Numbers are small in these analyses, so results must be treated with caution.

22.7.2 HPV18

The study population for this analysis comprises 21 women who had incident cervical qPCR HPV18 infections, and who had further follow-up after their first HPV18 qPCR-positive sample.

Sixteen women cleared their incident cervical HPV18 infections; the median duration of an incident cervical HPV18 infection was 6.0 months (IQR 0.7 to 21.6).

22.7.2.1 HPV18 viral load in the first qPCR-positive sample and the duration of incident cervical qPCR HPV18 infections

A multivariate analysis for interval censored time-to-event data was used to determine the association between the HPV18 viral load observed at the start of an episode of infection, and the probability of clearing that infection. The results are shown in table 22.7.2.1.

For women who cleared their incident cervical HPV18 infections, the median (IQR) [range] HPV18 viral load in the first sample of the episode was 2 copies per 1,000 cells (0.4 to 190) [0.001 to 413,356]; in women who did not clear their HPV18 infections, the median (IQR) [range] viral load was 35 copies per 1,000 cells (0.004 to 26,054) [0.004 to 28,588]. The duration of an incident HPV18 infection increased as the viral load in the first qPCR-positive sample increased; however, the association was not statistically significant. The risk of clearing an infection

decreased by 15% (1-0.85) for every ten-fold increase in viral load in the first sample of an episode of infection.

Table 22.7.2.1. The association between HPV18 viral load in the first sample of an incident episode of infection and the duration of an incident cervical qPCR HPV18 infection.

Variable	Frequency	Events	Hazards ratio	95% CI	LR test[a]
Log_{10} of the HPV18 viral load in first sample	21	16	0.85	0.68 to 1.07	2.08;1;0.15

[a]Likelihood-ratio test: Chi-squared statistic; degrees of freedom; p-value.

22.7.2.2 Viral load in the sample taken at the previous visit and the duration of incident cervical qPCR HPV18 infections

A multivariate analysis for interval censored time-to-event data was used to determine the association between the HPV18 viral load in the sample taken at the previous visit, measured as a time-varying covariate, and the probability of clearing that infection. The results are shown in table 22.7.2.2.

The duration of an incident cervical HPV18 qPCR infection increased as the viral load increased in the sample taken at the previous visit increased; however, the association was not statistically significant. The risk of clearing an infection decreased by 20% (1.00-0.80) for every ten-fold increase in viral load in the sample taken at the previous visit.

Table 22.7.2.2. The association between HPV18 viral load in the sample taken at the previous visit and the duration of an incident qPCR HPV18 infection.

Variable	Frequency	Events	Hazards ratio	95% CI	LR test[a]
Log_{10} of the HPV18 viral load at the previous visit	21	16	0.80	0.63 to 1.03	3.58;1;0.06

[a]Likelihood-ratio test: Chi-squared statistic; degrees of freedom; p-value.

22.7.2.3 Greatest viral load observed and the duration of incident qPCR HPV18 infections

A multivariate analysis for interval censored time-to-event data was used to determine the association between the maximum HPV18 viral load observed during the episode of infection, measured as a time-varying covariate, and the probability of clearing that infection. The results are shown in table 22.7.2.3.

For women who cleared their HPV18 infections, the median (IQR) [range] maximum HPV18 viral load observed during their infection was 8 copies per 1,000 cells (0.4 to 4,313) [0.002 to 2,325,781]; in women who did not clear their HPV18 infections, the median (IQR) [range] maximum HPV18 viral load was 106 copies per 1,000 cells (0.03 to 26,054) [0.01 to 28,588]. The duration of an incident cervical HPV18 qPCR infection increased as the maximum viral load attained during the episode increased; the association was not statistically significant. The risk of clearing an infection decreased by 12% (1-0.88) for every ten-fold increase in maximum observed viral load during the episode of infection.

Table 22.7.2.3. The association between greatest HPV18 viral load observed and the duration of an incident qPCR HPV18 infection.

Variable	Frequency	Events	Hazards ratio	95% CI	LR test[a]
Log_{10} of greatest HPV18 viral load to date	21	16	0.88	0.71 to 1.09	1.57;1;0.21

[a]Likelihood-ratio test: Chi-squared statistic; degrees of freedom; p-value.

Summary

An increase in viral load was associated with a reduced probability of clearing an incident HPV18-qPCR positive episode, but the association was not statistically significant. This was the case whether the analysis included a variable measuring viral load in the first positive sample of the episode, a variable measuring viral load in the previous evaluable sample, or a variable measuring the greatest viral load attained during the episode. There was very little difference between the estimates of association. The magnitude of the association was least for the variable measuring greatest viral load attained, and greatest for that measuring viral load at the previous visit. Numbers were small in these analyses, so results must be treated with caution.

22.8 HPV VIRAL LOAD AND HPV STATUS AT THE END OF FOLLOW-UP

The association between the maximum viral load observed during follow-up and the probability of being clear of infection in the last evaluable qPCR sample was also investigated. This was a cross-sectional analysis based on the qPCR status of the last sample only. Nevertheless, there are elements of a longitudinal study since the viral load exposure variable was updated during a known period of follow-up, and length of follow-up was controlled for in the analysis.

22.8.1 HPV16

The study population for this analysis comprised 64 women who had two or more samples with evaluable HPV16 viral load during follow-up, and who were HPV16-qPCR-positive on at least one occasion.

Logistic regression analysis was used to determine the association between the maximum HPV16 viral load observed during follow-up and the probability of being clear of HPV16 infection in the last evaluable sample, after controlling for length of follow-up. The outcome in this analysis was being positive in the last sample. The results of this analysis are shown in table 22.8.1.

Thirty-seven women were HPV16-qPCR-positive in their last qPCR-evaluable sample. For women who were free of HPV16 infection in their last evaluable sample, the median (IQR) [range] maximum HPV16 viral load observed during follow-up was 209 copies per 1,000 cells (18 to 3,295) [0.03 to 393,824]; in women who were HPV16-positive in their last evaluable sample, the median (IQR) [range] maximum HPV16 viral load was 8,076 copies per 1,000 cells (742 to 282,186) [0.3 to 420,953,588]. Women were significantly less likely to be free of HPV16 infection in their last evaluable sample as the maximum observed viral load during follow-up increased, after controlling for the length of follow-up. The odds of being HPV16-positive in the last evaluable sample increased by 63% (1.63-1) for every ten-fold increase in the maximum observed HPV16 viral load during follow-up. There was no evidence of an interaction between time and HPV16 viral load.

Table 22.8.1. Logistic regression analysis of the association between the maximum HPV16 viral load observed during follow-up and the probability of being clear of HPV16 infection in the last evaluable sample.

Variable	Frequency	Events[c]	Odds ratio	95% CI	LR test[ab]
Log_{10} of the maximum HPV16 viral load	64	37	1.63	1.19 to 2.23	12.09;1;<0.001

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [b]controlling for length of follow-up; [c]the event modelled was: being positive in the last evaluable sample.

22.8.2 HPV18

The study population for this analysis comprised 59 women who had two or more samples with evaluable HPV18 viral load during follow-up, and who were HPV18-qPCR-positive on at least one occasion.

Logistic regression analysis was used to determine the association between the maximum HPV18 viral load observed during follow-up and the probability of being clear of HPV18 infection in the last evaluable sample, after controlling for length of follow-up. The results of this analysis are shown in table 22.8.2.

Thirty-seven women were HPV18-qPCR-positive in their last qPCR-evaluable sample (this number is a coincidence with the HPV16 analysis). For women who were free of HPV18 infection in their last evaluable sample, median (IQR) [range] maximum HPV18 viral load observed during follow-up was 3 copies per 1,000 cells (0.3 to 1,237) [0.0006 to 25,564,280]; in women who were HPV18-positive in their last evaluable sample, the median (IQR) [range] maximum HPV18 viral load was

312 copies per 1,000 cells (5 to 15,681) [0.002 to 413,356]. Women were less likely to be free of HPV18 infection in their last evaluable sample as the maximum observed viral load during follow-up increased, after controlling for the length of follow-up, but the association was not statistically significant. The odds of being HPV18-positive in the last evaluable sample increased by 22% (1.22-1) for every ten-fold increase in the maximum observed HPV18 viral load during follow-up. There was no evidence of an interaction between time and HPV18 viral load.

Table 22.8.2. Logistic regression analysis of the association between the maximum HPV18 viral load observed during follow-up and the probability of being clear of HPV18 infection in the last evaluable sample.

Variable	Frequency	Events[c]	Odds ratio	95% CI	LR test[ab]
Log_{10} of the maximum HPV18 viral load	59	37	1.22	0.97 to 1.53	3.05;1;0.08

[a]Likelihood ratio test: chi-squared statistic; degrees of freedom; p-value; [b]controlling for length of follow-up; [c]the event modelled was: being positive in the last evaluable sample.

Summary

There was a statistically significant association between the maximum viral load observed during follow-up and the probability of being free of infection in the last evaluable sample for HPV16, but not for HPV18. The odds of being positive in the last evaluable sample increased by 63% for every ten-fold increase in the maximum viral load observed during follow-up for HPV16, but by only 22% for HPV18.

22.9 HPV VIRAL LOAD AND THE ACQUISITION OF CERVICAL CYTOLOGICAL ABNORMALITY

The association between the HPV viral load of cervical HPV infections and the risk of acquiring an incident cervical cytological abnormality was investigated in a time-to-event analysis.

Five viral load exposure variables were considered for analysis: a time-fixed binary variable indicating viral load status at baseline (positive or negative); a time-fixed continuous variable measuring viral load at baseline; a time-varying binary variable indicating “ever viral load positive” status at the previous visit (yes or no); a time-varying continuous covariate measuring viral load at the previous visit; and a time-varying continuous covariate measuring maximum viral load attained by the previous visit.

22.9.1 HPV16

The study population for this analysis was confined to 62 women who had an evaluable HPV16 viral load at study entry, and who had further follow-up after this time.

The median HPV16 viral load in the cervical sample taken at study entry was 0.3 copies per 1,000 cells (IQR 0 to 4.1; range 0 to 11,930.7); excluding the 28 women

who had a HPV16 viral load of zero at study entry, the median HPV16 viral load was 2.9 copies per 1,000 cells (IQR 0.5 to 25.2; range 0.2 to 11,930.7).

Thirty-seven women acquired an incident cervical cytological abnormality during follow-up: 19 had smears reported as containing at worst borderline nuclear abnormalities at first detection; and 18 had mild dyskaryosis.

22.9.1.1 The association between HPV16 viral load at study entry and the acquisition of cervical cytological abnormality

The results of a univariate analysis of the association between HPV16 viral load at baseline and the risk of acquiring an incident cervical cytological abnormality, are shown in table 22.9.1.1. The association between HPV16 qPCR status at study entry and the risk of a subsequent cervical cytological abnormality was not statistically significant, when this variable was analysed as a binary variable. However, when this variable was included in the model as a (transformed) continuous covariate quantifying viral load, the association became highly significant, with an increase in risk of acquiring cervical cytological abnormality of 75% (1.75-1.00) for every ten-fold increase in HPV16 viral load.

Table 22.9.1.1. Univariate analysis of the association between HPV16 viral load at baseline and the risk of acquiring an incident cervical cytological abnormality.

Variable	Frequency	Events	Hazards ratio	95% CI	LR test[a]
HPV16-qPCR-positive at baseline					
No	28	17	1.00	Reference	0.01;1;0.93
Yes	34	20	1.03	0.53 to 1.98	
Log_{10} of HPV16 viral load at baseline[b]	62	37	1.75	1.28 to 2.41	10.82;1;0.001

[a]Likelihood-ratio test: Chi-squared statistic; degrees of freedom; p-value; [b]Model includes a binary variable indicating HPV16-qPCR positivity status at baseline.

22.9.1.2 The association between HPV16 viral load in the sample taken at the previous visit and the risk of acquiring an incident cervical cytological abnormality

The results of a univariate analysis of the association between HPV16 viral load in the sample taken at the previous visit and the risk of acquiring an incident cervical cytological abnormality are shown in table 22.9.1.2. The association between HPV16 qPCR status in the sample taken at the previous visit and the risk of a subsequent cervical cytological abnormality was not statistically significant, when this variable was analysed as a binary variable. However, when this variable was included in the model as a (transformed) continuous covariate quantifying viral load, the association became highly significant, with an increase in risk of 54% (1.54-1.00) for every ten-fold increase in HPV16 viral load.

Table 22.9.1.2. Univariate analysis of the association between HPV16 viral load in the sample taken at the previous visit and the risk of acquiring an incident cervical cytological abnormality.

Variable	Frequency	Events	Hazards ratio	95% CI	LR test[a]
Ever HPV16-qPCR-positive					
No	14	10	1.00	Reference	0.64;1;0.42
Yes	48	27	1.35	0.63 to 2.86	
Log_{10} of HPV16 viral load in sample taken at previous visit[b]	62	37	1.54	1.22 to 1.93	13.33;1;<0.001

[a]Likelihood-ratio test: Chi-squared statistic; degrees of freedom; p-value; [b]Model including a binary variable indicating HPV16-qPCR positivity status in previous sample.

22.9.1.3 The association between the greatest HPV16 viral load during follow-up and the risk of acquiring cervical cytological abnormality

For the 37 women who acquired a cervical cytological abnormality, the median of the maximum HPV16 viral loads observed during follow-up was 30.7 copies per 1,000 cells (IQR 0 to 1,050.4; range 0 to 1,341,410.6); in the 25 women who did not acquire a cervical cytological abnormality, the median maximum HPV16 viral load was 15.8 copies per 1,000 cells (IQR 0.3 to 414.9; range 0 to 45,774.2)) [Wilcoxon rank-sum test with continuity correction: W=460; p=0.60]. In the sample taken immediately prior to the date of the first detection of cervical cytological abnormality, HPV16 viral load: was unevaluable for one woman; was zero for 10 women, nine of whom had zero viral loads in all evaluable samples taken prior to and including this sample, with the remaining woman having at least one viral-load positive sample; was less than the maximum observed viral load in all samples taken prior to detection for six women; attained its greatest observed value for 20 women.

The results of a univariate analysis of the association between the greatest HPV16 viral load observed prior to and including the sample taken at the previous visit and the risk of acquiring an incident cervical cytological abnormality are shown in table 22.9.1.3. The association between a log-transformed continuous covariate measuring a woman's greatest observed HPV16 viral load to date and the risk of acquiring an incident cervical cytological abnormality was highly statistically significant, with an increase in risk of 76% (1.76-1.00) for every ten-fold increase in HPV16 viral load.

Table 22.9.1.3. Univariate analysis of the association between the greatest HPV16 viral load and the risk of acquiring an incident cervical cytological abnormality.

Variable	Frequency	Events	Hazards ratio	95% CI	LR test[a]
Log_{10} of greatest HPV16 viral load[b]	62	37	1.76	1.38 to 2.25	21.77;1;<0.001

[a]Likelihood-ratio test: Chi-squared statistic; degrees of freedom; p-value; [b]model includes a binary variable indicating "ever HPV16-qPCR-positive".

Summary

Each of the analyses in this subsection suggests that it is not the fact of being HPV16-qPCR-positive which is important, but the observed magnitude of viral load. All three continuous covariates measuring HPV16 viral load: viral load at baseline; viral load in the previous sample; and the greatest viral load at the time of the previous sample, were significantly associated with the risk of acquiring cervical cytological abnormality. The magnitude of the increase in risk associated with a ten-fold increase in viral load was virtually identical for all three variables.

22.9.2 HPV18

The study population for this analysis was confined to 56 women who had an evaluable HPV18 viral load at study entry, and who had further follow-up after this time.

The median HPV18 viral load in samples taken at study entry was 0.0004 copies per 1,000 cells (IQR 0 to 1.5; range 0 to 83,292.6); excluding the 27 women who had zero HPV18 viral load at study entry, the median HPV18 viral load was 1.3 copies per 1,000 cells (IQR 0.5 to 5.0; range 0.0001 to 83,292.6).

Thirty-two women acquired an incident cervical cytological abnormality during follow-up: 19 had smears reported as containing at worst borderline nuclear abnormalities at first detection; and 13 had mild dyskaryosis.

22.9.2.1 The association between HPV18 viral load at study entry and the risk of acquiring cervical cytological abnormality

The results of a univariate analysis of the association between HPV18 viral load at baseline and the risk of acquiring an incident cervical cytological abnormality are shown in table 22.9.2.1. The association between HPV18-qPCR positivity status at study entry and the risk of a subsequent cervical cytological abnormality was not statistically significant, when this variable was analysed as a binary variable. However, when this variable was included in a model as a (transformed) continuous

covariate quantifying viral load, the association became highly statistically significant, with an increase in risk of 67% (1.67-1.00) for every ten-fold increase in HPV18 viral load.

Table 22.9.2.1. Univariate analysis of the association between HPV18 viral load at baseline and the acquisition of an incident cervical cytological abnormality.

	Frequency	Events	Hazards ratio	95% CI	LR test[a]
HPV18-qPCR-positive at baseline					
No	27	16	1.00	Reference	0.06;1;0.81
Yes	29	16	0.92	0.46 to 1.84	
Log_{10} of HPV18 viral load at baseline[b]	56	32	1.67	1.15 to 2.43	6.26;1;0.01

[a]Likelihood-ratio test: Chi-squared statistic; degrees of freedom; p-value; [b]model includes a binary variable indicating HPV18-qPCR positivity status at baseline.

22.9.2.2 The association between HPV18 viral load in the sample taken at the previous visit and the risk of acquiring cervical cytological abnormality

The results of a univariate analysis of the association between HPV18 viral load in the sample taken at the previous visit and the risk of acquiring an incident cervical cytological abnormality, are shown in table 22.9.2.2. The association between HPV18 qPCR status in the sample taken at the previous visit and the risk of a subsequent cervical cytological abnormality was not statistically significant, when this variable was analysed as a binary variable. However, when this variable was included in a model as a (transformed) continuous covariate quantifying viral load,

the association became highly statistically significant, with an increase in risk of 77% (1.77-1.00) for every ten-fold increase in HPV18 viral load.

Table 22.9.2.2. Univariate analysis of the association between HPV18 viral load in the sample taken at the previous visit and the risk of acquiring an incident cervical cytological abnormality.

Variable	Frequency	Events	Hazards ratio	95% CI	LR test[a]
Ever HPV18-qPCR-positive					
No	14	10	1.00	Reference	0.001;1;0.98
Yes	42	22	1.01	0.48 to 2.14	
Log_{10} of HPV18 viral load in sample taken at previous visit[b]	56	32	1.77	1.29 to 2.43	12.20;1;<0.001

[a]Likelihood-ratio test: Chi-squared statistic; degrees of freedom; p-value; [b]model includes a binary variable indicating HPV18-qPCR positivity status in the previous sample was eliminated.

22.9.2.3 The association between the greatest HPV18 viral load and the risk of acquiring an incident cervical cytological abnormality

For the 32 women who acquired cervical cytological abnormality, the median of the maximum HPV18 viral loads observed during follow-up was 1.2 copies per 1,000 cells (IQR 0 to 21.6; range 0 to 25,564,280); in the 24 women who did not acquire cervical cytological abnormality, the median maximum HPV18 viral load was 1.4 copies per 1,000 cells (IQR 0.002 to 54.7; range 0 to 15,680.6)) [Wilcoxon rank-sum test with continuity correction: W=432; p=0.42]. In the sample taken immediately

prior to the date of the first detection of cervical cytological abnormality, HPV18 viral load: was unevaluable for one woman; was zero for 18 women, 13 of whom had zero viral loads in all evaluable samples taken prior to and including this sample, with the remaining five women having at least one HPV18-qPCR-positive sample; was less than the maximum observed viral load in all samples taken prior to detection for two women; and attained its greatest observed value for 11 women.

The results of a univariate analysis of the association between greatest HPV18 viral load observed prior to and including the sample taken at the previous visit and the risk of acquiring an incident cervical cytological abnormality are shown in table 22.9.2.3. The association between a log-transformed continuous covariate measuring a woman's greatest observed HPV18 viral load and the risk of acquiring an incident cervical cytological abnormality was highly significant, with an increase in risk of 59% (1.59-1.00) for every ten-fold increase in HPV18 viral load.

Table 22.9.2.3. Association between the greatest HPV18 viral load and the risk of acquiring an incident cervical cytological abnormality.

	Frequency	Events	Hazards ratio	95% CI	LR test[a]
Log_{10} of greatest HPV18 viral load to date[b]	56	32	1.59	1.25 to 2.03	12.14;1;<0.001

[a]Likelihood-ratio test: Chi-squared statistic; degrees of freedom; p-value; [b]Model includes a binary variable indicating "ever HPV18-qPCR-positive" at baseline.

Summary

Each of the analyses in this subsection suggests that it is not the fact of being HPV18-qPCR-positive which is important, but the observed magnitude of viral load. All three continuous time-varying covariates measuring HPV18 viral load: viral load at baseline; viral load in the previous sample; and the greatest viral load at the time of the previous sample, were significantly associated with the risk of acquiring cervical cytological abnormality. An increase in viral load was associated with an increased risk for all three variables: the magnitude of the association was least for the variable measuring greatest viral load attained, and greatest for that measuring viral load in the first evaluable sample, but the differences were not great.

22.10 HPV VIRAL LOAD AND THE ACQUISITION OF HIGH-GRADE CIN

The previous section examined the association between HPV viral load and the risk of acquiring cervical cytological abnormality. A more appropriate outcome would be the incidence of high-grade CIN, the precursor of cervical cancer. The data to address this question within the cohort described in this thesis is severely limited by number of events, and the age of the women. Nevertheless, given its performance, some descriptive results will be presented.

Within the cohort of 1,075 women, the number who were diagnosed with high-grade CIN during follow-up was 28, 17 of whom were in the intended study population for viral load testing of 155 women: 10, two and five in the HPV16-only, HPV18-only and HPV16+HPV18 cohorts, respectively. Of these, only eight were tested for HPV viral load: four, one, and three in each of the respective cohorts, with two of the women in the HPV16-only cohort being tested for HPV16 viral load only, but with all other

women tested for both HPV16 and HPV18 viral load. The results of cytological, histological, virological, and serological testing of all women in the study population who were diagnosed with high-grade CIN and who were tested for viral load prior to diagnosis, is shown in table 22.10.

All of these women had HPV16-associated high-grade CIN when HPV status was determined using qPCR, including the woman from the HPV18-only cohort who was positive on only one occasion with a viral load of 5 copies per 1,000 cells. Three women were infected with HPV16 at study entry, with the remaining four having incident HPV16 qPCR infections. Seven women who subsequently acquired HPV16-associated high-grade CIN had numerically low viral loads in their first evaluable sample, including the four women who were qPCR HPV16-negative. One woman was qPCR HPV16-negative on five separate occasions over the course of three years, before subsequently having a sample containing over 400 million copies per 1,000 cells. Four women had monotonically-increasing HPV16 viral loads prior to the diagnosis of high-grade CIN, i.e. the HPV16 viral load started low and increased in all subsequent samples, one had an increase in viral load which was maintained until diagnosis, one had a viral load which fluctuated, and two had increases followed by a decrease.

Five women had HPV18-associated high-grade CIN when HPV status was determined using qPCR, one did not, and two women were not tested for HPV18 viral load. One woman was infected with HPV18 at study entry, one was non-evaluable, and the remaining three had incident HPV18 qPCR infections. Four

women who subsequently acquired HPV18-associated high-grade CIN had numerically low viral loads in their first evaluable sample, including the three women who were qPCR HPV18-negative. One woman was qPCR HPV18-negative on six separate occasions over the course of nearly seven years, including the date of diagnosis of high-grade CIN, but also had a sample containing over 2 million copies per 1,000 cells. Two women were only found to be HPV18-qPCR-positive at the date of diagnosis of high-grade CIN, and three women had viral loads that increased from numerically low values to a peak high-value, before returning to a numerically-low value.

Summary

The number of women with incident high-grade CIN in this analysis was small, so observations are suggestive rather than conclusive. All women with incident high-grade CIN who were tested for viral load in this cohort were qPCR-positive for HPV16 at some point during follow-up; many women only ever had numerically-low viral loads, and others had zero viral loads for extended periods. Similar results were seen for HPV18.

Table 22.10 (Part 1). The results of cytological, histological, and virological, testing of all women in the study population who were diagnosed with high-grade CIN and who had cytological samples tested for viral load prior to diagnosis. Data presented is restricted to visits prior to and including the date of diagnosis of high-grade CIN. Where multiple samples were available, or the same sample was tested more than once, the mean viral load across all samples and tests is presented.

ID	Visit	Time since study entry (days)	Cytology[a]	Histology[b]	HPV DNA (GP5+/6+)[c]	HPV16 VL[d]	HPV18 VL[d]
150	1	0	Normal	—	Negative	2,904	0
	2	140	Normal	—	Negative	583,227	0
	3	511	Mild dyskaryosis	—	16	Missing	Missing
	4	601	Mild dyskaryosis	CIN2	6/11; 16	547,175	0
330	1	0	Normal	—	Negative	0	0
	2	268	Mild dyskaryosis	—	16	397,966	0
	3	370	Mild dyskaryosis	CIN3	16	583,861	18
350	1	0	Result missing	—	Negative	0	0
	2	98	Normal		No sample	NA	NA
	3	259	Normal	—	X	72	0
	4	575	Normal	—	Negative	3	0
	5	941	Normal	—	18	9,530	31,265
	6	1035	Mild dyskaryosis	—	X	Ct-G[e]	2,325,781
	7	1360	Mild dyskaryosis	—	Negative	21,579	Ct-V[f]
	8	1462	Mild dyskaryosis	CIN1	X	391	1
	9	1707	Mild dyskaryosis	HPV	16	269	0
	10	2043	Mild dyskaryosis	—	Negative	1,153	0
	11	2414	Mild dyskaryosis	CIN2	X	679,329	0
393	1	0	Normal	—	Negative	34	0
	2	181	Normal	—	6/11	373	0
	3	360	Mild dyskaryosis	—	6/11	4,402	0
	4	493	BNA	CIN3	16;18	9,579	2,419
820	1	0	Normal	—	Negative	<1	Not Tested
	2	455	Normal	—	16	100	Not Tested
	3	664	Mild dyskaryosis	—	16	25,559	Not Tested
	4	789	Mild dyskaryosis	—	16	13,733	Not Tested
	5	1027	Mild dyskaryosis	CIN2	16	1,100	Not Tested
1005	1	0	Normal	—	Negative	0	Not Tested
	2	184	Result missing	—	No sample	NA	Not Tested
	3	329	Mild dyskaryosis	—	16	8,265	Not Tested
	4	625	Mild dyskaryosis	CIN3	16	282,186	Not Tested

[a]BNA=borderline nuclear abnormality (the least severe grade of cervical cytological abnormality), US=unreadable smear; [b]a dash means NA, not missing; [c]X means HPV-positive with an unknown type which however was known not to be 6/11, 16, 18, 31, 33, 52 or 58; [d]Viral load expressed as copy number per 1,000 cells; [e]sample had excess variability in GAPDH Ct values, but is nevertheless known to be positive; [f]sample had excess variability in virus Ct values, but is nevertheless known to be positive; [g]Failure in GAPDH assay.

Table 22.10 (Part 2). The results of cytological, histological, and virological, testing of all women in the study population who were diagnosed with high-grade CIN and who had cytological samples tested for viral load prior to diagnosis. Data presented is restricted to visits prior to and including the date of diagnosis of high-grade CIN. Where multiple samples were available, or the same sample was tested more than once, the mean viral load across all samples and tests is presented.

ID	Visit	Time since study entry (days)	Cytology[a]	Histology[b]	HPV DNA (GP5+/6+)[c]	HPV16 VL[d]	HPV18 VL[d]
1512	1	0	Normal	—	Negative	0	GF[g]
	2	180	Normal	—	X	0	0
	3	405	Normal	—	18	0	413,356
	4	546	BNA	—	18	0	144,188
	5	739	BNA	HPV	Negative	GF[g]	GF[g]
	6	1005	BNA	HPV	X	0	505
	7	1355	Mild dyskaryosis	—	33	GF[g]	GF[g]
	8	1551	Normal	—	Negative	3,270	0
	9	1733	Mild dyskaryosis	CIN1	6/11;16	420,953,588	<1
	10	1995	Mild dyskaryosis	—	No sample	NA	NA
	11	2170	BNA	CIN2	No sample	NA	NA
1814	1	0	Normal	—	Negative	0	4,849
	2	171	BNA	—	Negative	0	<1
	3	244	BNA	HPV	Negative	0	1,464
	4	412	BNA	—	X	0	982
	5	776	Normal	CIN1	Negative	0	1
	6	958	Normal	—	Negative	0	<1
	7	1161	BNA	HPV	18	5	30,417
	8	1469	Normal	—	Positive	0	<1
	9	1864	Normal	CIN3	No sample	NA	NA

[a]BNA=borderline nuclear abnormality (the least severe grade of cervical cytological abnormality), US=unreadable smear; [b]a dash means NA not missing; [c]X means HPV-positive of unknown type which however was known not to be 6/11, 16, 18, 31, 33, 52 or 58; [d]Viral load expressed as copy number per 1,000 cells; [e]sample had excess variability in GAPDH Ct values, but is nevertheless known to be positive; [f]sample had excess variability in virus Ct values, but is nevertheless known to be positive; [g]Failure in GAPDH assay.

Chapter 23

THE NATURAL HISTORY OF CERVICAL HPV16 AND HPV18 INFECTIONS – VIRAL LOAD

DISCUSSION

I have assessed the test performance of an assay measuring HPV16 and HPV18 viral load with respect to reliability, sensitivity, specificity, and type-specificity, described how the distribution of HPV viral load varies with infecting HPV type for types 16 and 18, and described the relationship between HPV16 and HPV18 viral load and the detection of epithelial abnormalities of the cervix.

23.1 ASSAY PERFORMANCE AND INTERPRETATION

The assay used had an excellent performance: the linear dynamic range exceeded 10^9 to one copies, and the primers used were shown to be sensitive and specific, with no cross-reactivity when tested in cervical cancer cell lines known to contain specific high-risk HPV types. The assay was found to be highly reproducible for measuring both HPV16 and HPV18 viral load. Estimates of intraclass correlation coefficients used to estimate within-assay reliability, using either observed CT values or observed virus copy numbers, corresponded to “almost-perfect” reliability on the Landis scale for both the HPV16 and HPV18 assays. Although this was, in part, a consequence of the approach adopted to accepting measurements as valid, this is primarily a reflection of the excellent reproducibility of these assays. Similarly, estimates of intraclass correlation coefficients used to measure inter-assay test-

retest reliability, while smaller in magnitude than those measuring intra-assay reliability, still corresponded to “almost-perfect” reliability for both the HPV16 and the HPV18 assay. There is a concern that the HPV18 assay produced an aberrant result for one sample, but this can never be ruled out entirely with *any* assay

Another concern is that some of the viral load results were obtained by extrapolation from the standard curves, although the number of results involved was very small. To reduce the number of wells used to generate the standard curves, the range of these curves was restricted to 10^2 to 10^8 copies for HPV16, and 10^3 to 10^7 for HPV18. However, the dynamic range exceeded 10^9 to 1 copies, and the vast majority of extrapolation occurred towards the lower end of this range, i.e. within the known dynamic range.

It is not clear how to interpret some of the numerically very-low viral loads observed in the cohort described in this thesis, such as 1 copy per 100,000 cells, or the very-high viral loads, such as 100,000 copies per *cell*. There is no benchmark against which to compare the observed values of viral load, and no means by which to determine whether a viral load is “low” or “high” in some sense. One study using qPCR, with GAPDH to normalise for cellular content, reported that the 25th percentile of HPV18 viral loads in cytologically normal smears was 2 copies per 10,000 cells, indicating that even lower viral loads were also found (Gravitt 2003b). One study using qPCR to estimate HPV viral load in cervico-vaginal cytological specimens, with albumin to normalise for cellular content, suggests that 110,000 copies per 1,000 cells is a high-viral load for HPV16, or 760,000 copies per 1,000

cells for HPV18 (Carcopino 2006). Another study in HIV-infected and uninfected women, using qPCR, with β-globin to normalise for cellular content, reports viral loads for various HPV types which range from less than 100 copies per 1,000 cells to greater than 4.7 million copies per 1,000 cells (Weissenborn 2003); another study using the same technique reported a viral load of 5.8 million copies per 1,000 cells (Snijders 2006). Clearly, very-high HPV viral loads are possible, and given that the qPCR assay is, in theory, capable of detecting a single copy of the virus, very-low viral loads are also plausible.

23.2 IDENTIFICATION OF THE STUDY POPULATION

Prior to undertaking this analysis, the intention was to measure HPV viral load in all women in the study cohort who had incident HPV16 and/or HPV18 infections. Women with incident cervical HPV infections were defined on the basis of the results of the first HPV testing system we had used, the GP5+/GP6+ system. It was believed that this would yield a well-defined study population for use in the analysis of both HPV viral load and the humoral immune response to incident cervical HPV infection, since it was anticipated that only samples which were HPV DNA positive according to the original *nominal* system would be positive using the *quantitative* system.

In practice, agreement between the two systems was substantial, although not perfect. Not all samples which tested positive for HPV16 or HPV18 using GP5+/GP6+ also tested positive for the corresponding type using qPCR; and many

which tested negative using GP5+/GP6+, tested positive using qPCR. In retrospect, this could perhaps have been anticipated: others have also found a discrepancy between the GP5+/GP6+ system and the more sensitive qPCR system (Van Duin 2002, Gravitt 2003b). It is clear that within the study population of the cohort study described in this thesis, more women are likely to be HPV16 and/or HPV18 DNA-positive than were identified on the basis of the GP5+/GP6+ system. Ideally analyses previously undertaken using measurements of HPV exposure defined on the basis of the GP5+/GP6+ system would be repeated using qPCR measurements, to see if previous conclusions still hold: of course, this is true for all studies which have used techniques less sensitive than qPCR to detect HPV. However, given that only a very limited subset of the cohort described in this thesis was tested using qPCR, this is not possible in this study population. Funding for such a study is now unlikely ever to be obtained.

23.3 DISTRIBUTION OF VIRAL LOAD ACCORDING TO HPV TYPE AND PRESENCE OF DISEASE

HPV viral load varied according to HPV type and cervical cytological abnormality status. For both HPV types which were the focus of analyses in this thesis, smears with non-zero viral loads had substantially higher viral loads when cytological abnormalities were present than when they were not, and HPV16 but not HPV18 viral load increased with increasing severity of cytological abnormality. In normal smears, HPV viral load for one type did not vary according to the presence or absence of the other type. However, in smears containing at worst BNA, viral load of

either type was higher in the absence of the other than when it was present; the reverse was true in smears containing at worst mild dyskaryosis.

Note that these results relate to all samples taken from all women. Most other studies which have investigated the association between HPV viral load and HPV type have included only one sample from each woman, typically that taken at diagnosis (Swan 1999, Zerbini 2001, Sherman 2003a, Weissenborn 2003, Giuliano 2004, Levi 2004, Moberg 2004, Ho 2005, Moberg 2005, Carcopino 2006, Flores 2006, Lai 2006, Snijders 2006). Most have also found that HPV viral load varies by HPV type, with several reporting an increase in HPV16 viral load with increasing disease severity (Swan 1999, Zerbini 2001, Gravitt 2003b, Weissenborn 2003, Carcopino 2006, Lai 2006). One report from a study which found no significant difference between the viral load of single and multiple infections (Giuliano 2004) was later contradicted by a report on the same population, which found that viral load was higher in samples containing any single HPV type compared to samples with multiple infections (Flores 2005). Another found that HPV16 viral load is higher in high-grade squamous intraepithelial lesions when another type is also present (Weissenborn 2003).

23.4 DURATION OF EPISODE OF HPV INFECTION IN RELATION TO VIRAL LOAD

An increase in viral load was associated with an increase in the duration of incident cervical HPV16 and HPV18 infections, but the association was not statistically

significant. This was the case whether the analysis included a variable measuring viral load in the first positive sample of the episode, a variable measuring viral load in the previous evaluable sample, or a variable measuring the greatest viral load attained during the episode. There was very little difference between the magnitudes of the estimates of association for either HPV type. For HPV16, the magnitude of the association was least for the variable measuring viral load in the first positive sample, and greatest for that measuring greatest viral load attained, which was of borderline statistical significance. For HPV18, the magnitude of the association was least for the variable measuring the greatest viral load attained, and greatest for that measuring viral load at the previous visit. Numbers were small in these analyses, so results must be treated with caution. Similarly, there was a statistically significant association between the maximum viral load observed during follow-up and the probability of being free of infection in the last evaluable sample for HPV16, but not for HPV18.

Other studies which have investigated the association between HPV viral load and the duration, or “persistence” of, infection, disagree in their conclusions (Clavel 2000, De Marco 2001, Van Duin 2002, Dalstein 2003, Molano 2003, Syrjanen 2005a, Cricca 2006, Song 2006). Four studies found that HPV viral load was not significantly associated with duration of infection, although risk estimates suggest a positive association, whereas the other four found that a higher viral load was associated with an increased duration of, or risk of, persistent infection. Unfortunately, both studies which have reported measures of duration or association have used the semiquantitative hybrid capture II system to measure viral load and so

results are not comparable with the study described in this thesis (Dalstein 2003, Syrjanen 2005b).

23.5 VIRAL LOAD AND CYTOLOGICAL ABNORMALITY

An increase in viral load was associated with an increased risk of acquiring an incident cervical cytological abnormality. This was the case for both HPV types, and for all three continuous variables used to measure viral load. However, for HPV16, the magnitude of the increase in risk associated with a ten-fold increase in viral load was identical, whether that ten-fold change related to the viral load in the baseline sample or in the maximum observed viral load during follow-up, and whether the maximum used was based on the current, or the previous, sample (analysis not shown). It is not immediately obvious why baseline and greatest viral load are essentially interchangeable in a model. Of course on practical, if not purely statistical grounds, if a baseline covariate provides as much predictive power as a covariate measured at each visit, then the baseline covariate would be preferred. However, being HPV-negative in the baseline sample would not then imply that the woman is at low-risk of cytological abnormality, since if she subsequently becomes HPV-positive her risk still increases with increasing viral load, and increases to the same extent.

Fewer longitudinal than cross-sectional studies have investigated the association between HPV viral load and the risk of the acquisition of cervical disease. Eight reports from these studies have indicated a positive association, including some

from studies which have been criticised on methodological grounds earlier (section 20.3.3). Only one study has used a comparable disease outcome, but also used the semiquantitative hybrid capture II technique to measure viral load (Castle 2002a). In a cohort of 2,020 women who tested positive for HPV at study entry, the risk of a subsequent abnormal smear of any severity was significantly greater when the baseline sample contained a high-viral load than when it contained a low-viral load (odds ratio 2.7; 95% CI 1.7 to 4.1).

With respect to high-grade CIN, the ideal disease outcome for this cohort study, numbers are severely limited, so observations can only be suggestive. All women with high-grade CIN who were tested for viral load in this cohort were HPV16-qPCR-positive at some point during follow-up; many women only ever had numerically-low viral loads, and others had zero viral loads for extended periods. Similar results were seen for HPV18.

23.6 BIOLOGICAL FACTORS WHICH MAY CONFOUND THE HPV VIRAL LOAD-CERVICAL DISEASE RELATIONSHIP

Although normalised for cellular DNA content, comparisons made using the qPCR assay may still be confounded by differences in the proportion of infected and uninfected cells present in a sample. Sampling variation, or heterogeneity within cervical lesions, may also distort comparisons: for example, viral load is reported to vary with the endocervical cell content of the cytological sample (Depuydt 2006); and

to be higher in women with high-grade CIN when low-grade CIN is also present, than when it is absent (Sherman 2003a).

Temporal changes in viral load may reflect other aspects of viral infection which are in themselves associated with disease progression. For example, whereas the prevalence of integrated forms of HPV increases with increasing disease severity, integration is followed by a fall in viral load; however, in almost all cross-sectional and longitudinal surveys which have measured HPV viral load, HPV integration status is undefined and may, of course, change over time (Berumen 1995, Jeon 1995, Spartz 2005, Cricca 2006). Similarly, the acquisition of new HPV types is associated with both changes in viral load and with the development of new cervical lesions; therefore, measures of association may be unreliable in longitudinal studies which rely on a single measurement of exposure made at baseline (Woodman 2002, Crum 2004).

23.7 CLINICAL UTILITY OF A MEASUREMENT OF HPV VIRAL LOAD

Although the observations above are of interest when considering how the exposure-disease relationship varies with the infecting HPV type, their clinical utility is debateable.

At present, cervical screening programmes usually offer screening, by cervical cytology, to all women at regular intervals of between three and five years (see section 4.2). However, among women who test positive for high-risk HPV types,

cytological abnormality is more common in those with a high-viral load, and this has been found to be the case irrespective of the method used to measure viral load (Hall 1996, Nindl 1997, Ho 1998a, Swan 1999, Heard 2000, Lillo 2005). It has therefore been suggested that viral load status can be used to improve the efficiency of cervical cytology screening programmes by identifying high-risk (high-viral load) women who can then be screened more frequently; whereas low-risk (low-, or zero-viral load) women could be screened less frequently.

The first problem with this suggestion relates to the sampling of cervical HPV infection at indeterminate points during it's natural history. In figure 23.7a, subjects A, B and C all have HPV infections starting at different points in calendar time, and all are tested for HPV viral load on the same day. In each case, the natural history of the infection is identical: after a subject acquires the infection, viral load increases steadily to a maximum, and then declines to zero (or to below the threshold of detectability). In this scenario, whether subject B is identified as being at high-risk depends only on where the threshold is set for viral load to be considered "high". Subjects A and C, with identical (low) measures of viral load, would be categorised as low-risk. A single measurement of viral load is unable to determine that these three HPV infections are identical and therefore pose identical risks to all three subjects. In particular, a single test is unable to determine whether a low-viral load measurement is low but increasing towards a (high) maximum, or low following a fall from a peak (high) viral load. Additional testing would also not clarify matters. Depending on the time scale for these infections, the next test would certainly give a

zero-viral load for subject A, a low- or possibly zero-viral load for subject B, and either a high- or low-viral load for subject C.

Figure 23.7a. Changes in viral load over time for three hypothetical subjects with identical HPV infections tested for HPV viral load at a single point in time.

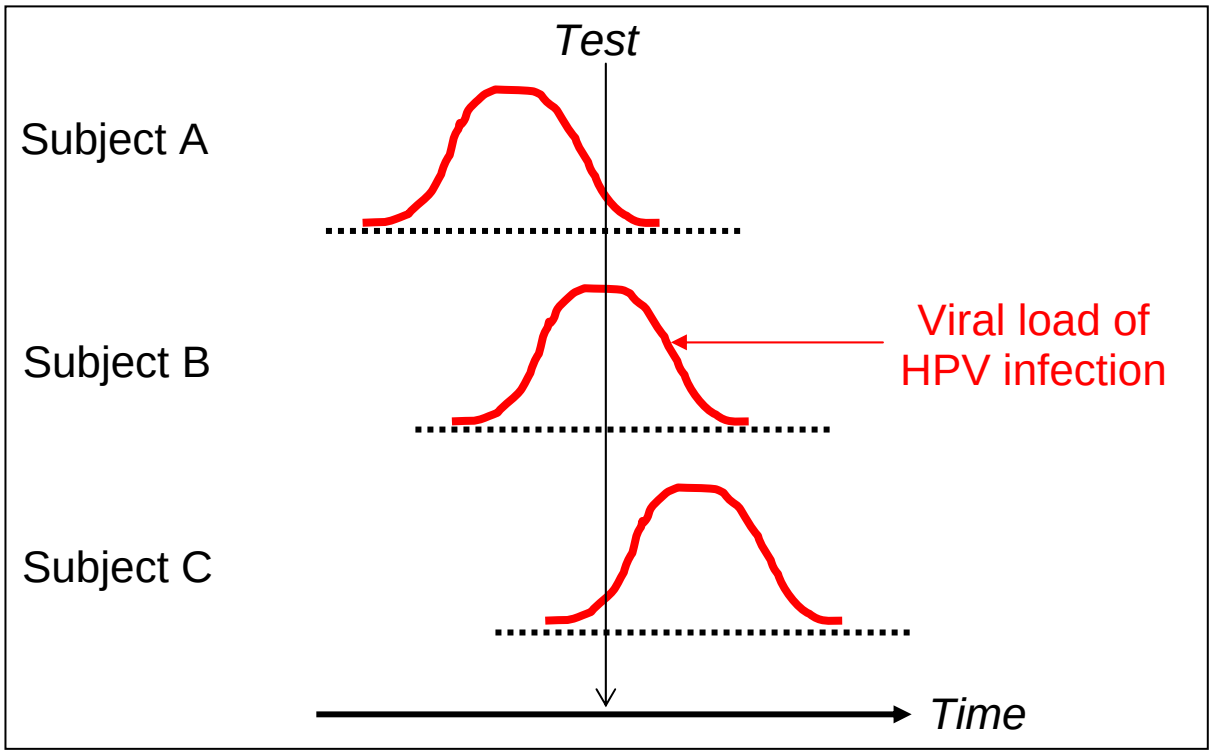


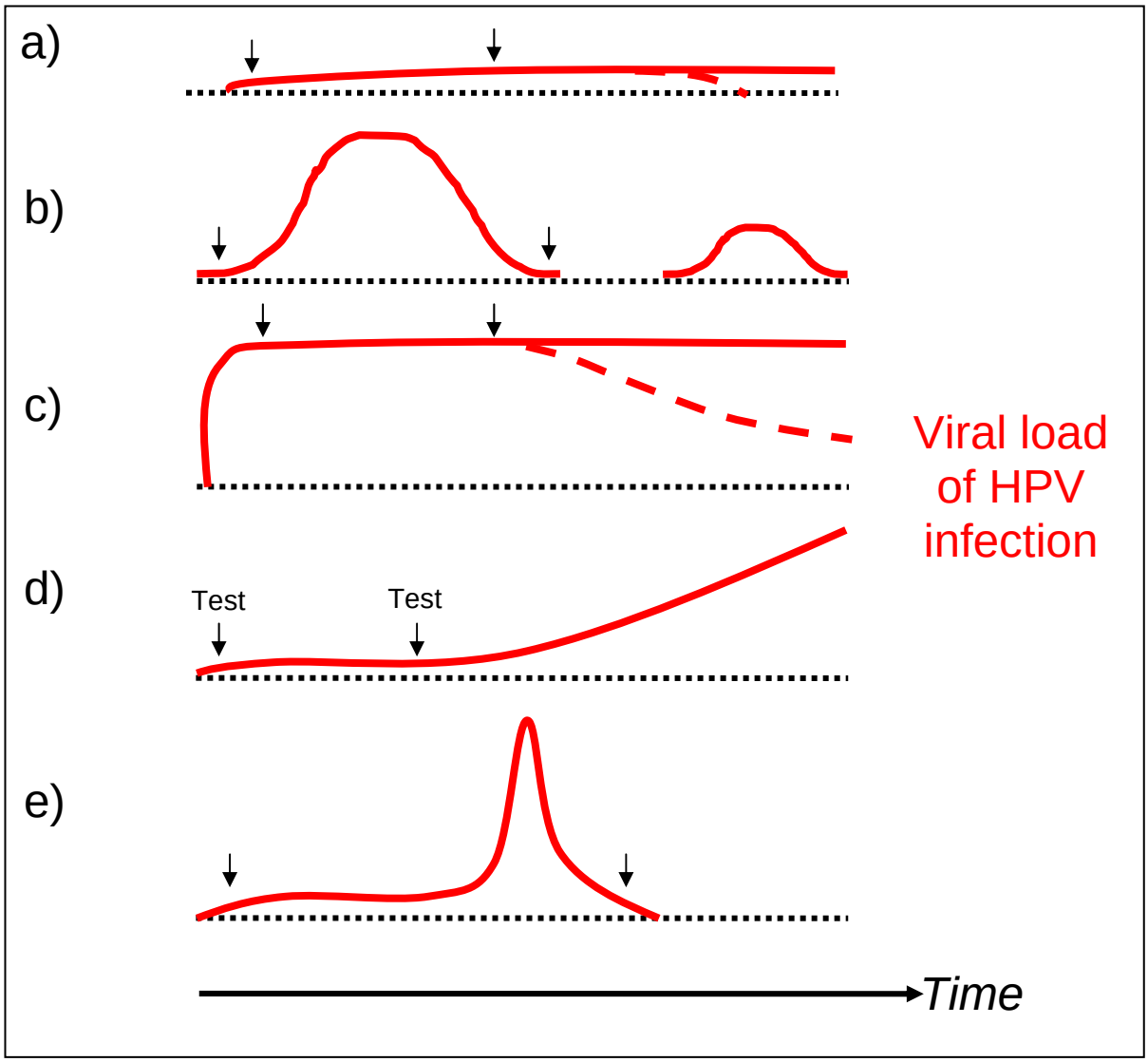
Figure 23.7a makes the assumption that the natural history of cervical HPV infection, in terms of changes in viral load, is essentially identical in different women, and that the manner in which viral load changes over time is smooth and symmetric. Of course, if this artificial assumption is confirmed in practice, it would render a measurement of viral load unnecessary, since it would then be the mere fact of having been infected with HPV which is relevant. However, if a single measurement of HPV viral load, or even a small number of measurements, are to tell us anything about the risk of disease, the assumption must be made that these measurements

are made at comparable points in the natural history, and that they capture an important aspect of the infection as a whole, not merely a transient feature.

Figure 23.7b illustrates some other possibilities for how HPV viral load may change over time. In example (a), the solid line illustrates a chronic low viral load infection: a single test will suggest that this woman has a low-viral load infection, and the second test will “confirm” this suggestion; similarly for example (c) and high-viral load. The women in examples (a) and (c) will thus be categorised as low- and high-risk, respectively. In these cases, although viral load is apparently measuring a consistent feature of the infection, it is still not certain, for example, that the risk-levels attributed to these women are valid. Why does the woman in example (a) have a long term infection? If the virus has become integrated, then viral low will typically be low; but an integrated virus may represent more of a risk than a high-viral load infection in which the virus remains episomal. Clearly, the two tests conducted on woman (a) cannot distinguish an infection which is cleared (example (a) dashed line), or an infection in which viral load subsequently increases (example (d)), or an infection in which an increase in viral load was shortly followed by a decrease, perhaps as a result of viral integration (examples (b) and (e)). If the aspect of an infection which is relevant to the risk of disease is *ever having had* a high-viral load, then of the four women in these examples with a high-viral load at some point in time, three will be incorrectly categorised as low-risk because testing was not undertaken frequently enough. If this is the case, to correctly categorise women, women will have to be tested at every visit until a high-viral load is observed, thus increasing the costs of the screening programme and undermining any supposed efficiency gains obtained

by introducing viral load testing. And if all HPV infections have a period during which they have a high-viral load, assuming testing occurs frequently enough to detect the infection when it is in this period, then there will be no opportunity to reduce the frequency of screening.

Figure 23.7b. Changes in HPV viral load over time for five hypothetical subjects with HPV infections tested for HPV viral load at two points in time.



Note: depending on how much of the curve is observed, and how long the time period, the curves in this figure are not mutually exclusive.

But in practice, does viral load change in the ways illustrated in figure 23.7b? Within limits, yes, certainly for the young women in the cohort described in this thesis.

Figure 22.6 presents changes over time in HPV16 and HPV18 viral load for all women tested for HPV16 or HPV18 viral load, or both, in the cohort described in this thesis. Examples of women with “type a” profiles are study number 244, who had low HPV16 viral load for a period of over five years, and study number 1296, who had a low, but non-zero, HPV18 viral load for a period of over two years. Study number 1367 has a “type b” profile for HPV18 since she can be considered to have two consecutive rises in HPV18 viral load, but without becoming HPV18-negative in between. Study number 150 has a “type c” profile for HPV16, since HPV16 viral load was within the range between 10^3 and 10^6 copies for over six years, but seemed to fluctuate around a constant value rather than demonstrating a particular pattern. From a value of zero at baseline, HPV16 viral load steadily increased over the following nine years for study number 350, making her HPV16 profile a “type d”; for HPV18 she has a “type e” profile. Study number 1512 has “type e” profiles for both HPV16 and HPV18: in both cases viral load rose precipitously from zero to a high-viral load, before returning equally precipitously to a zero, or at least a low, viral load. For other women, study numbers 1551 and 1814, for example, it is difficult to conceive of the smooth underlying process which corresponds to the observed changes in both HPV16 and HPV18 viral load.

However women are categorised according to their viral load profile, what is immediately apparent is that viral load tends not to be a constant, i.e. a

representative, feature of an infection: both HPV16 and HPV18 viral load appear to be able to undergo substantial changes over a short period of time (it is difficult to say what a “substantial change” is, but with the chosen scale a ten-fold change is, of course, easy to distinguish). This may or may not be important: it may be more relevant to know how often, or for how long, a high-viral load infection remains a high-viral load infection; and similarly for a low-viral load infection. The question may then be not whether the viral load of an infection fluctuates, but does it fluctuate within a range in which it remains a high-, or low-viral load infection?

As figure 22.6 illustrates, a woman’s viral load status changes over time in terms of both whether or not she is positive, and copy number; for the latter, changes can be quite dramatic from one sample to the next, i.e. over a period of approximately six months. It is therefore doubtful whether a *single* measurement of HPV viral load can ever be used to predict a woman’s risk of cervical disease acquisition, or progression. Even if the assumption is made that once a woman is infected with HPV16 she remains infected; and that any “negative” viral load measurement is in fact a positive measurement which lies below the threshold of detectability for the assay; then it is still readily apparent that a single measurement of HPV viral load is uninformative with respect to predicting risk status.

A final consideration with respect to the use of viral load as a clinical biomarker is that it must be remembered that viral replication is not necessary for maintaining the malignant phenotype (see section 3.1.2.1.2). *In situ* assays on cervical cancer cell lines show that not all HPV copies are transcriptionally active; and in pure tumour

samples taken from women with invasive disease, E6/E7 messenger RNA levels were not correlated with HPV DNA copy number, after normalisation for cellular content (Van Tine 2004, De Boer 2007).

Within the restricted resources of a health service, women can only be offered a limited number of screening tests. In the UK, cervical screening is currently offered to women aged 25 to 64 years at intervals of three to five years, depending on age (see section 4.2.3). In the cohort study described in this thesis, which had intervals of only six months between study visits, it was possible to describe the kinetics of cervical HPV infections at a resolution impossible to obtain when repeated tests are separated by such large time intervals. The suggestion that cervical HPV viral load can be used to increase the efficiency of cervical cytology screening programmes appears to be untenable in young women; whether it is tenable in older women has yet to be demonstrated one way or the other, but must be open to doubt.

23.8 PUBLICATION

A paper based on the work described in chapters 20, 21, 22 and 23 of this thesis was accepted for publication in August 2009 (Constandinou-Williams 2009). The definitive results of HPV viral load testing became available only *after* the completion of the analyses relating to the neutralizing antibody response to cervical HPV infections. An analysis of the association between HPV viral load and the humoral immune response may be the subject of a future report (see chapter 24).

Chapter 24

SUMMARY AND FURTHER ANALYSIS

24.1 What has been learned about the natural history of early cervical HPV infections?

The candidate is now in a position to construct a tentative biological model of the natural history of early cervical HPV infection. This is possible as a result of the work undertaken to complete this thesis, and as a result of other work related to the cohort study described in this thesis with which the candidate was also intimately involved.

Prior to the time at which the work described in this thesis was undertaken, cervical HPV infection was known to be a necessary but not sufficient cause of cervical cancer. It had been established that cervical HPV infection was a very common sexually transmitted infection, which was acquired soon after the onset of first sexual intercourse.

The longitudinal observations generated by the cohort described in this thesis, and which were analysed by the candidate, enabled the estimation, for the first time, of the incidence of cervical HPV infection, in a cohort of young women who had recently experienced penetrative vaginal sexual intercourse for the first time. This enabled me to show that cervical HPV infection was acquired soon after first sexual intercourse; that the sequential detection of different HPV types was very common;

that most incident cervical HPV infections were transient and of short duration; and that cervical infection with HPV was frequently associated with cervical cytological abnormalities, which were themselves transient.

This cohort also enabled an estimate to be obtained of the incidence of high-grade CIN in young women: this estimate underpinned the sample size calculations proposed by the federal drug administration in the USA when negotiating the need for trials of HPV vaccines with pharmaceutical companies.

The longitudinal observations generated by the cohort described in this thesis, enabled me to show that not only is cervical HPV infection very common, but that it is acquired as a result of low-risk sexual behaviour, i.e. even during the first sexual relationship. It was also possible to confirm the effect of age of sexual partner in the risk of the acquisition of cervical HPV infection.

Subsequent work, for which the candidate undertook all statistical analyses, demonstrated that: the cervical cytological abnormalities associated with HPV18 could be understated compared to those associated with HPV16, making the detection of HPV18-associated disease by cytological screening more difficult than the detection of HPV (this could undermine the effectiveness of such programmes); and that HPV16 and HPV18 both appeared to become integrated into the host-cell genome within a relatively short period of time following infection, although HPV18 appeared to become integrated more rapidly than HPV16. I was able to show, using

the qPCR assay described in this thesis, that integration appears to be followed by a fall in viral load.

Turning now specifically to the areas covered in depth in this thesis.

Aspects of sexual behaviour were known to be the most important risk factors for cervical HPV infection. However, as a result of the analyses undertaken to complete this thesis, it has proven possible to contradict the popular belief, resulting from the use of sub-optimal cross-sectional study designs, that the adolescent cervix is at inherently greater risk of infection. As the results presented in this thesis are derived from a study using a longitudinal design and a study population recruited during the relevant exposure period; I was able to demonstrate that the risk of an incident cervical HPV infection of any type *increases* as the time interval between menarche and first sexual intercourse increases: this risk does not appear to decrease. In an analysis of this kind is always difficult in distinguishing the effects of possibly correlated variables. However, while gynaecological age and calendar age at first sexual intercourse are positively-correlated, it has been shown that gynaecological age at first sexual intercourse does not appear to be merely a surrogate for calendar age at first sexual intercourse. The results from this thesis provide the most convincing epidemiological evidence to date that the adolescent cervix is not at greater risk of HPV infection: while many results suggest non-significant associations, all point estimates of association suggest an *increasing* risk of cervical HPV infection with an increasing gynaecological age at first sexual intercourse. Stronger evidence could perhaps be obtained by making more frequent observations

in younger women, using a more sensitive HPV detection method; however, it now seems unlikely that this will occur, in particular in countries in which vaccination of pre-adolescent girls is now planned.

Prior to undertaking this thesis, cervical HPV infection was known to be a necessary but not sufficient cause of cervical cancer: therefore it was clear that it was necessary to identify other cofactors which modulate the progression of cervical HPV infection to cervical cancer. Cigarette smoking was one potential cofactor, which was believed to act through any of several biological mechanisms, either alone or in combination with other factors, to increase the risk of acquiring cervical HPV infection and epithelial abnormalities of the cervix. In this thesis, cigarette smoking, typically considered an indicator of high-risk sexual behaviour, was found to be an independent (of cervical HPV infection) risk factor for the acquisition of high-grade cervical intraepithelial neoplasia; however, it was *not* found to be an independent risk factor for the incidence of cervical HPV infection, once sexual behaviour was taken into account. Smoking did not prolong the duration of cervical HPV infections. There is now a compelling case for investigating mechanisms which mediate the association between smoking and the acquisition of cervical epithelial abnormalities. For example, the cohort described in this thesis is being used to explore the extent to which smoking-induced promoter methylation of the tumour suppressor gene p16 (CDKN2A) might explain the association between smoking and the acquisition of epithelial abnormalities of the cervix.

The kinetics of the immune response to cervical HPV infections were not well characterised prior to undertaking this thesis. The sample size used in analyses was

less than planned, which was certainly disappointing, but this thesis did make a substantial contribution to our knowledge of the immune response to early cervical HPV infections. Many studies have investigated some aspect of the humoral immune response to HPV infections. However, the study described in this thesis was able to describe the functionally-relevant aspect of the humoral immune response: the neutralizing antibody response. It was demonstrated that women do generate an often substantial neutralizing antibody response to natural infections, and maintain this response for considerable periods of time, although seropositivity may also be intermittent. The relationship between the neutralizing antibody response to cervical HPV16 and HPV18 infections and the viral load of these infections will be the subject of further inquiries based on the cohort described in this thesis (see below). Not all women with sera samples which *could* have been tested *were* tested. There is no intention to complete the testing of these women using the neutralising antibody assay. However, it is hoped that in the future, all sera samples available will be tested using the technically simpler VLP-based ELISA: clearly, this will yield a much larger sample size.

Prior to the results of this thesis, it was believed that HPV viral load was of potential clinical utility, in that it could be used to distinguish between women who were at high- and low-risk of epithelial abnormalities of the cervix. This belief was based on the results of studies which were often contradictory. This thesis has clearly illustrated that the viral load of cervical HPV infections in young women is not a static quantity, but varies a great deal throughout the course of an infection. The risk of acquiring epithelial abnormalities of the cervix certainly increases with increasing

HPV16 and HPV18 viral load. However, the increase in risk is associated with increasing viral load in *any* sample taken from that woman, it is not confined to the first sample taken. This renders the clinical utility of a measurement of cervical HPV viral load questionable. Nevertheless, a sensitive measure of viral load which is able to detect low-copy numbers will be essential for continuing investigations of the natural history of cervical HPV infections. For example, although it is now clear that HPV can establish a “persistent” infection, it remains to be determined whether persistent infections are characterized by the continuing presence of low-copy numbers of HPV, or by a state of latency during which the virus remains undetectable, only to reappear later.

24.2 Further analyses

Clearly, in the study described in this thesis, it was not possible to exploit the full potential of the cohort for defining the kinetics of the humoral immune response to cervical HPV infection. It is hoped that it will be possible to do so in future, but this will require a suitable high-throughput neutralising antibody assay.

However, with the results currently available, I can now present preliminary results describing the association between HPV viral load and the neutralizing antibody response to HPV infection, together, for the subset of women who contribute to both analyses. Figure 24.2 presents changes over time in the neutralizing antibody titre against HPV16 and HPV18, with changes over time in viral load of the relevant HPV

type superimposed for all women potentially eligible for inclusion in this analysis (note that figure 24.2 is essentially an amalgamation of figures 18.8 and 22.6).

A few preliminary observations can be made regarding this data. One might consider that the detection of HPV of a given type, as measured using qPCR, might generate an immune response against that type, in particular a neutralizing antibody response. Study numbers 43 (for HPV16), 240 (18), 350 (16), 393 (16) and 561 (16) might be considered to be examples of this behaviour. Similarly, an immune response elicited by an “insult” due to the presence of a HPV infection of a given type might persist until that infection is “cleared”. Take study number 292 and HPV18, for example: HPV18 becomes detectable in large copy-numbers per 1,000 cells; this appears to cause a very large neutralizing antibody response, which persists for a considerable period of time (a titre of around 1-in-3240 is observed for approximately eight years); the presence of this large neutralizing antibody response appears to lead to the clearance of HPV18 (the black line measuring HPV18 viral load continues to increase initially, but subsequently falls to around zero). Study numbers 446 (for HPV18), 590 (16), 744 (18), 938 (16) and 1011 (16) might be considered to be further examples of this behaviour. However, sometimes a large viral load seems to have failed to elicit a neutralizing antibody response during the period when women were under observation in this study, study numbers 147 (for HPV16), 350 (18), 1367 (18) and 1814 (18) are good examples of this. There is also the question of which viral load “event” causes which neutralizing antibody response “event”, and what the time-lag between those events is. For example, for study numbers 2 (16), 247 (16) and 393 (16), an increase in viral load seems to be

followed shortly after by an increase in the neutralizing antibody response to the relevant HPV type; however, for study number 490 (HPV18) there may be a threshold effect, with the increase in the neutralizing antibody response being caused by the second increase in HPV18 viral load rather than the first, but this is likely to prove difficult to determine in practice. This is not the least of the statistical challenges posed by the data: these are repeat measurements; there is substantial missing data; and HPV viral load may act as both an outcome variable and as an explanatory factor. For example, consider study number 292 and HPV18 discussed earlier. In this case, HPV18 viral load is both an outcome variable, the values of which are of interest in themselves, and also an explanatory variable, since it potentially “causes” an immune response which then affects subsequent values of HPV18 viral load.

24.3 Conclusion

The natural history of *early* cervical HPV infections is perhaps more complex than was previously thought: until demonstrated otherwise, one can only assume that the natural history of cervical HPV infections in older women is similarly complex. It has been suggested that the association between cervical HPV infection and cervical neoplasia can be exploited to improve the efficiency and effectiveness of primary- and secondary-prevention programmes for cervical cancer. The results of this thesis suggest that achieving these improvements still requires a greater understanding of the natural history of early cervical HPV infection and its role in the acquisition of epithelial abnormalities of the cervix, than we currently possess.

Figure 24.2. Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time with changes over time in viral load of the relevant HPV type superimposed (—). The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months), left-vertical scale (inverse of the 50% neutralizing antibody titre), and right-vertical scale (HPV copy number per 1,000 cells). The measurements of the 50% neutralizing antibody titre and viral load retain their last known value until updated. A coloured vertical line indicates date of first detection of cytological abnormality (blue=borderline nuclear abnormalities; green=mild dyskaryosis). Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

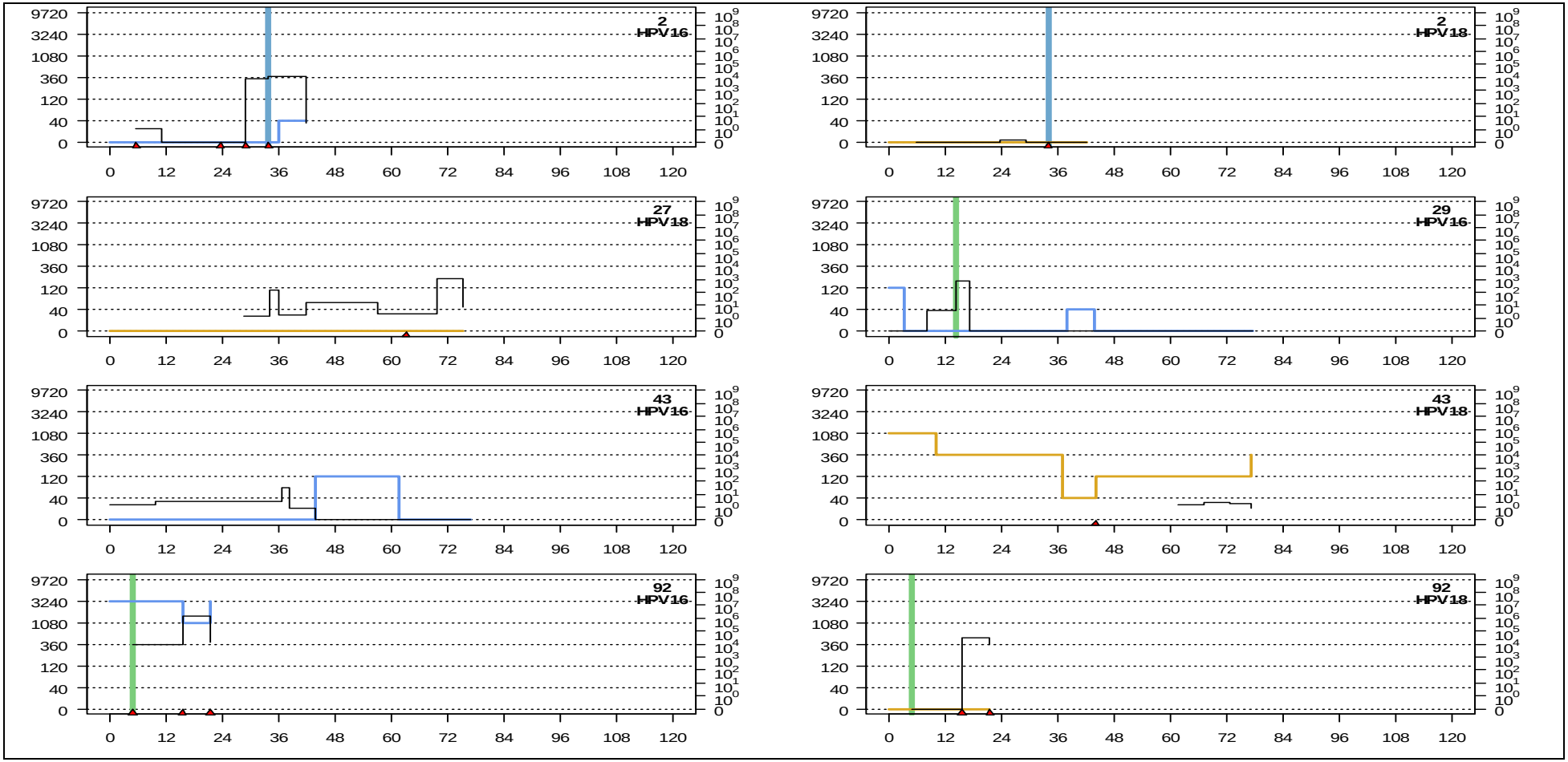


Figure 24.2 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time with changes over time in viral load of the relevant HPV type superimposed (—). The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months), left-vertical scale (inverse of the 50% neutralizing antibody titre), and right-vertical scale (HPV copy number per 1,000 cells). The measurements of the 50% neutralizing antibody titre and viral load retain their last known value until updated. A coloured vertical line indicates date of first detection of cytological abnormality (blue=borderline nuclear abnormalities; green=mild dyskaryosis). Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

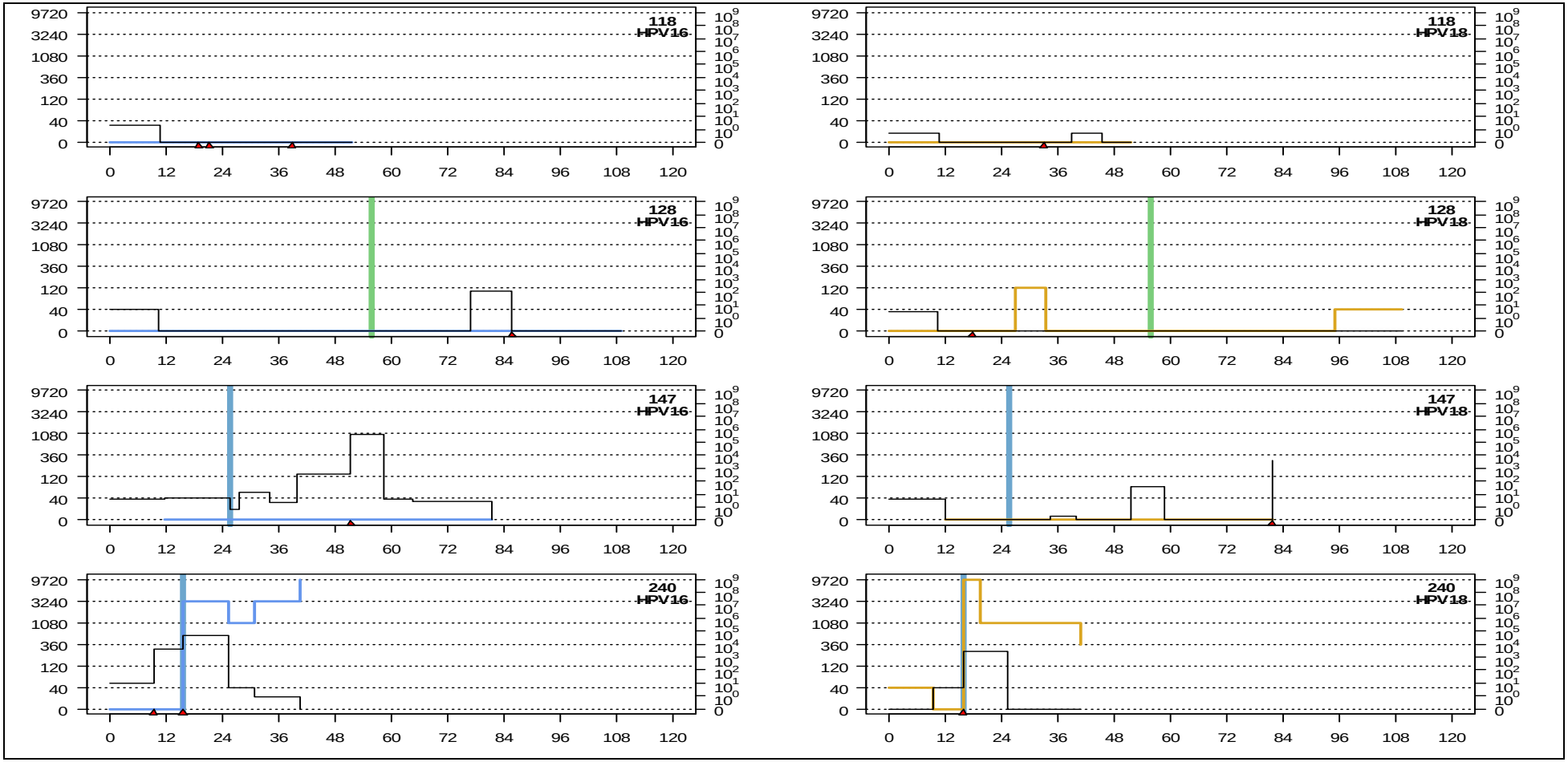


Figure 24.2 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time with changes over time in viral load of the relevant HPV type superimposed (—). The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months), left-vertical scale (inverse of the 50% neutralizing antibody titre), and right-vertical scale (HPV copy number per 1,000 cells). The measurements of the 50% neutralizing antibody titre and viral load retain their last known value until updated. A coloured vertical line indicates date of first detection of cytological abnormality (blue=borderline nuclear abnormalities; green=mild dyskaryosis). Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

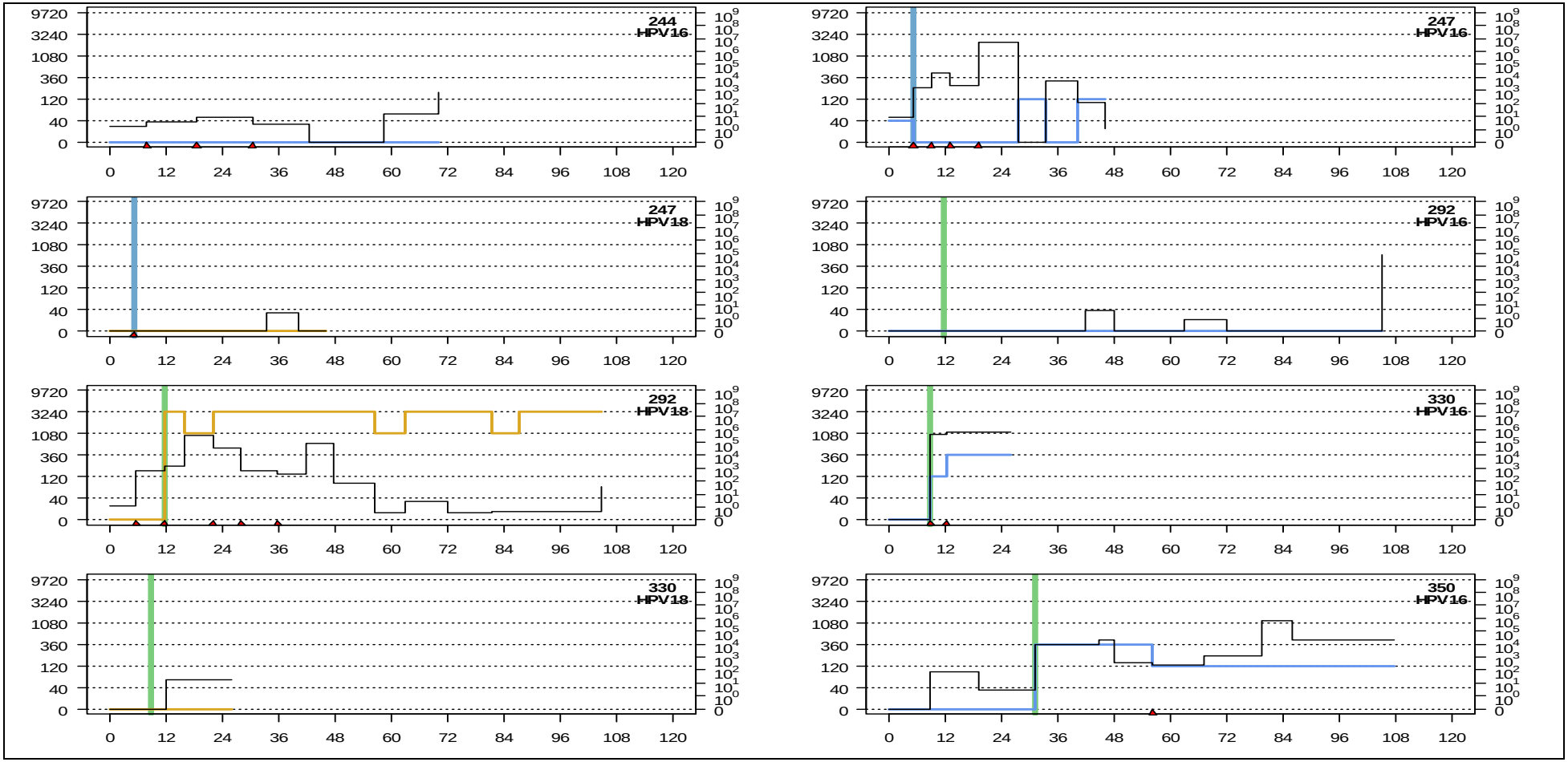


Figure 24.2 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time with changes over time in viral load of the relevant HPV type superimposed (—). The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months), left-vertical scale (inverse of the 50% neutralizing antibody titre), and right-vertical scale (HPV copy number per 1,000 cells). The measurements of the 50% neutralizing antibody titre and viral load retain their last known value until updated. A coloured vertical line indicates date of first detection of cytological abnormality (blue=borderline nuclear abnormalities; green=mild dyskaryosis). Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

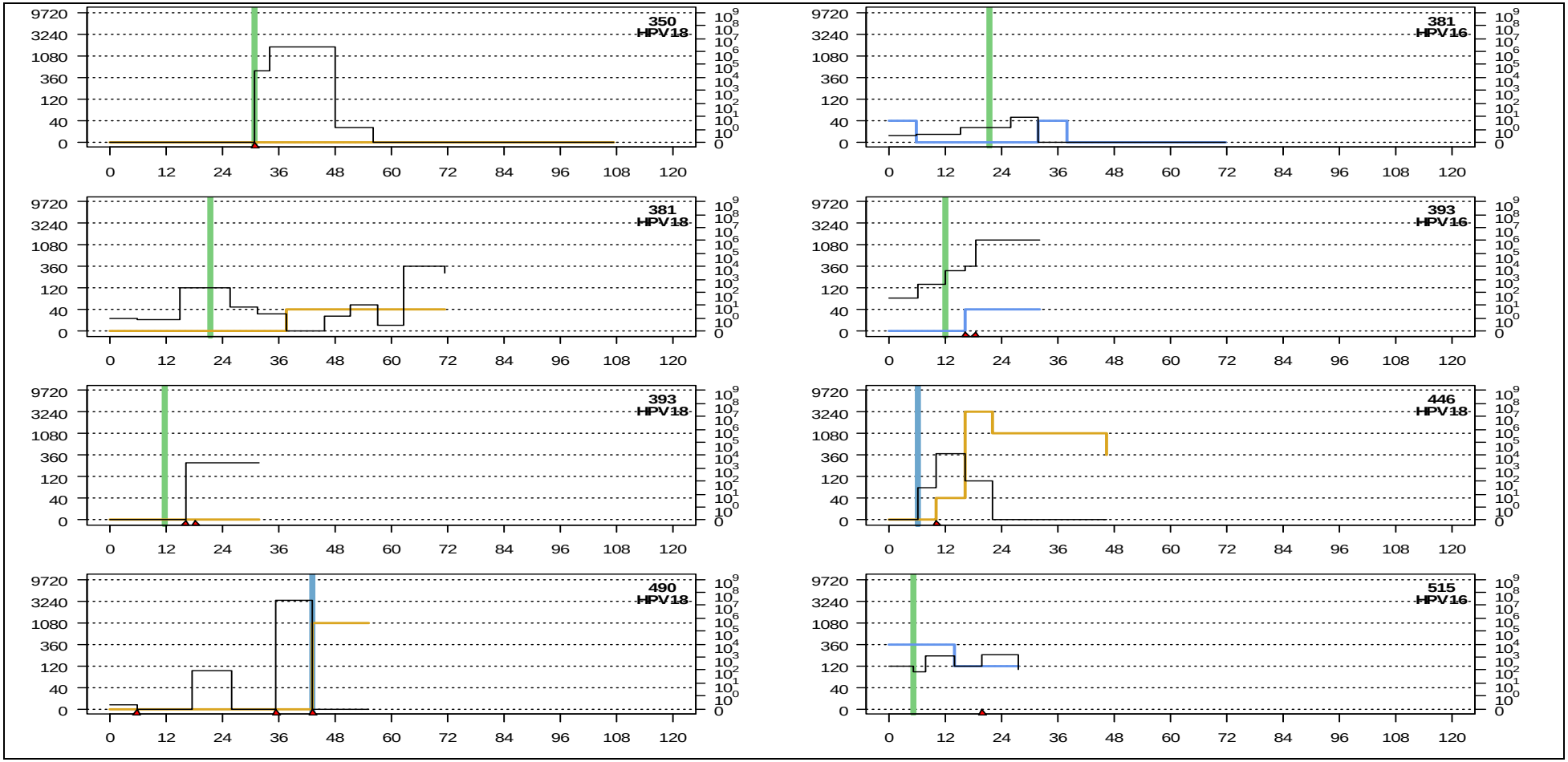


Figure 24.2 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time with changes over time in viral load of the relevant HPV type superimposed (—). The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months), left-vertical scale (inverse of the 50% neutralizing antibody titre), and right-vertical scale (HPV copy number per 1,000 cells). The measurements of the 50% neutralizing antibody titre and viral load retain their last known value until updated. A coloured vertical line indicates date of first detection of cytological abnormality (blue=borderline nuclear abnormalities; green=mild dyskaryosis). Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

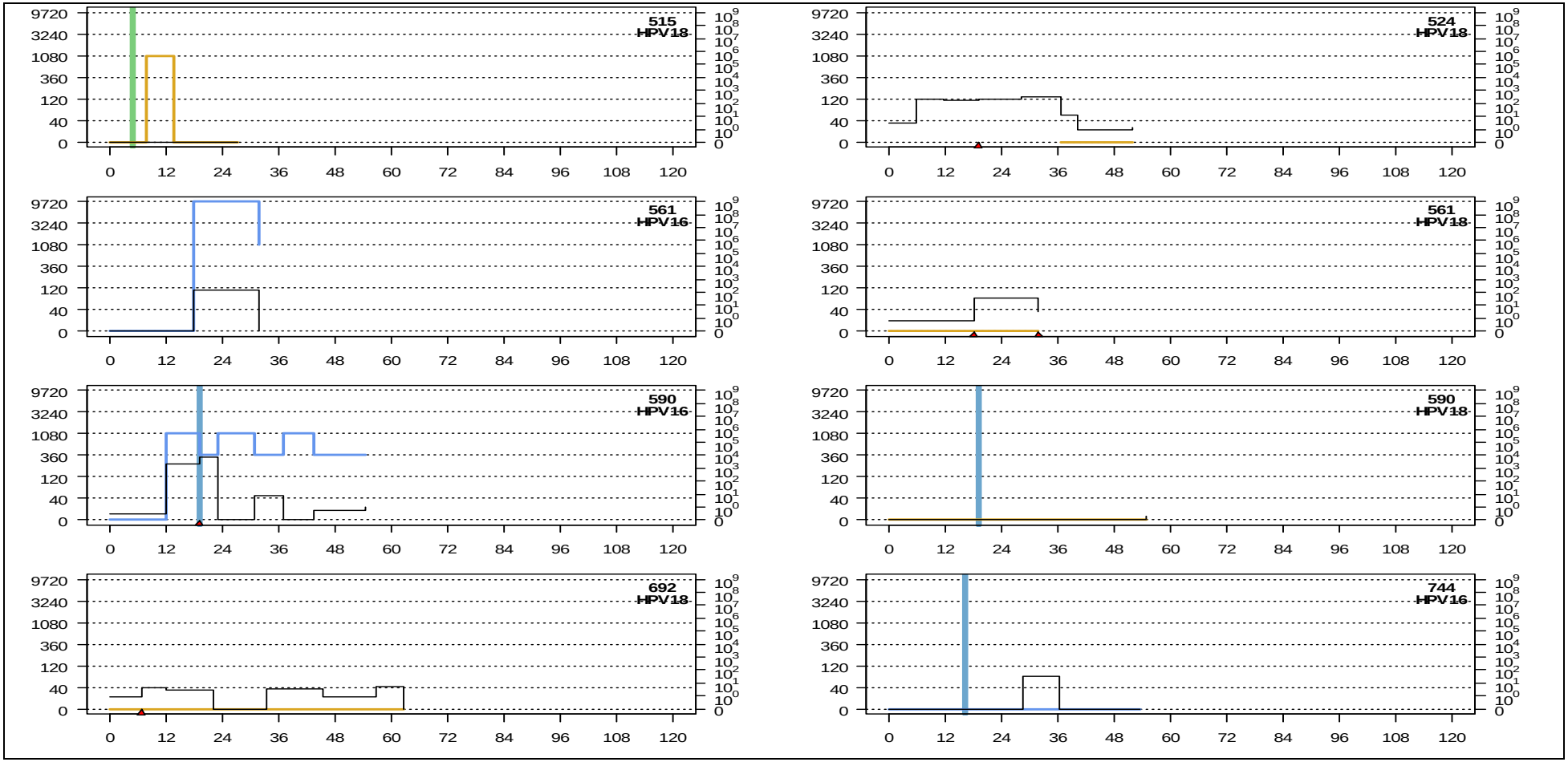


Figure 24.2 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time with changes over time in viral load of the relevant HPV type superimposed (—). The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months), left-vertical scale (inverse of the 50% neutralizing antibody titre), and right-vertical scale (HPV copy number per 1,000 cells). The measurements of the 50% neutralizing antibody titre and viral load retain their last known value until updated. A coloured vertical line indicates date of first detection of cytological abnormality (blue=borderline nuclear abnormalities; green=mild dyskaryosis). Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

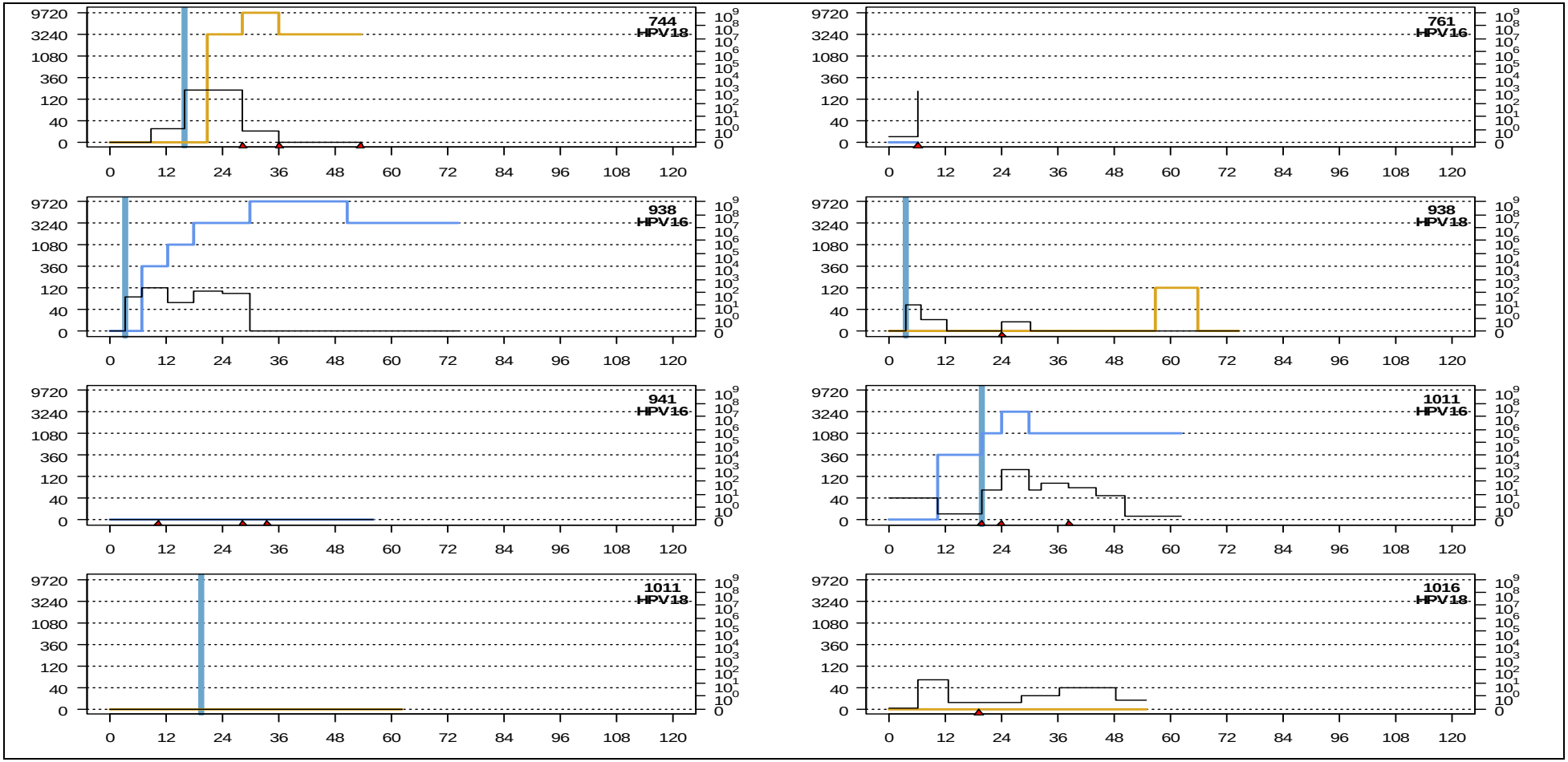


Figure 24.2 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time with changes over time in viral load of the relevant HPV type superimposed (—). The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months), left-vertical scale (inverse of the 50% neutralizing antibody titre), and right-vertical scale (HPV copy number per 1,000 cells). The measurements of the 50% neutralizing antibody titre and viral load retain their last known value until updated. A coloured vertical line indicates date of first detection of cytological abnormality (blue=borderline nuclear abnormalities; green=mild dyskaryosis). Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

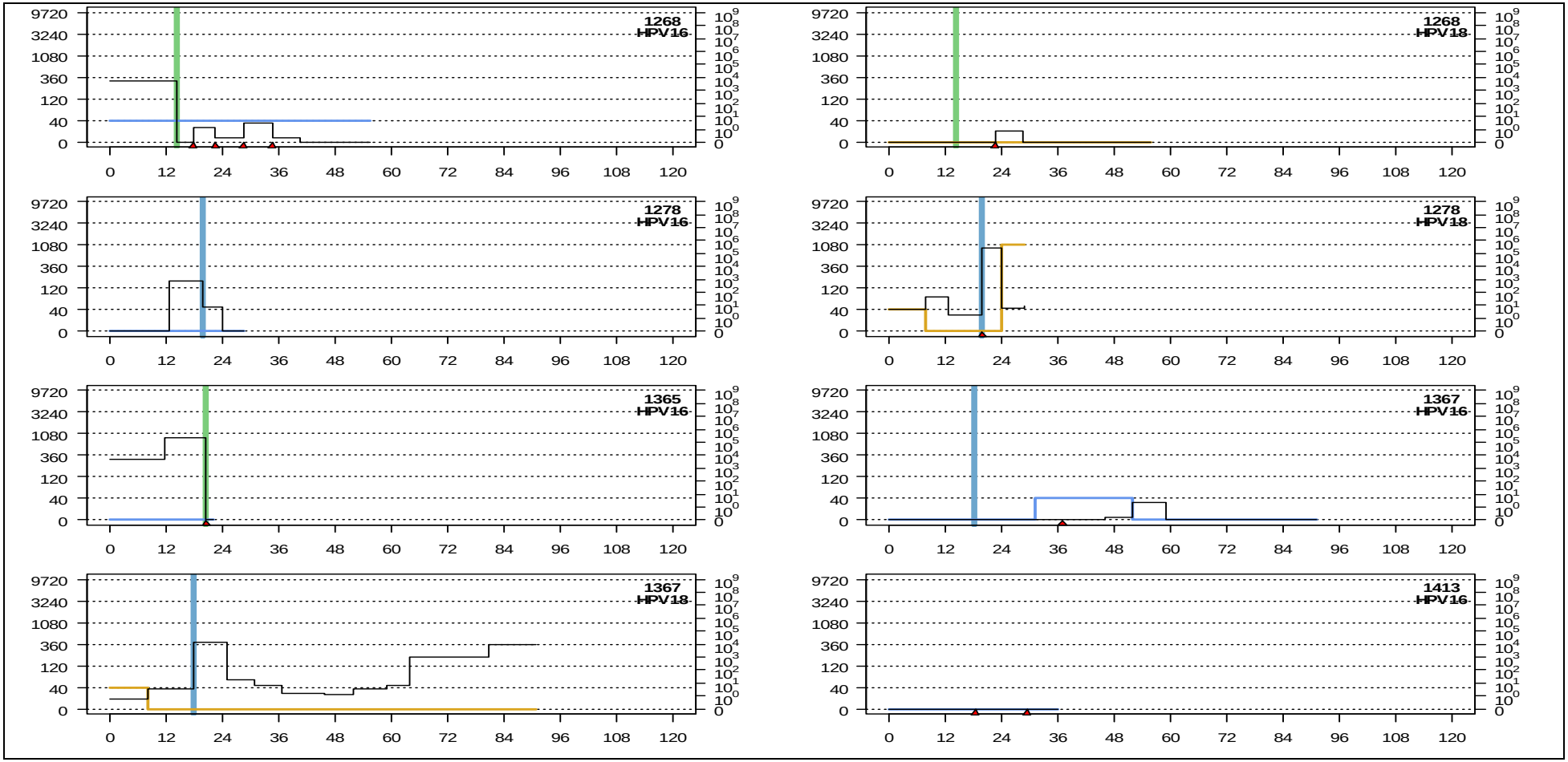


Figure 24.2 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time with changes over time in viral load of the relevant HPV type superimposed (—). The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months), left-vertical scale (inverse of the 50% neutralizing antibody titre), and right-vertical scale (HPV copy number per 1,000 cells). The measurements of the 50% neutralizing antibody titre and viral load retain their last known value until updated. A coloured vertical line indicates date of first detection of cytological abnormality (blue=borderline nuclear abnormalities; green=mild dyskaryosis). Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.

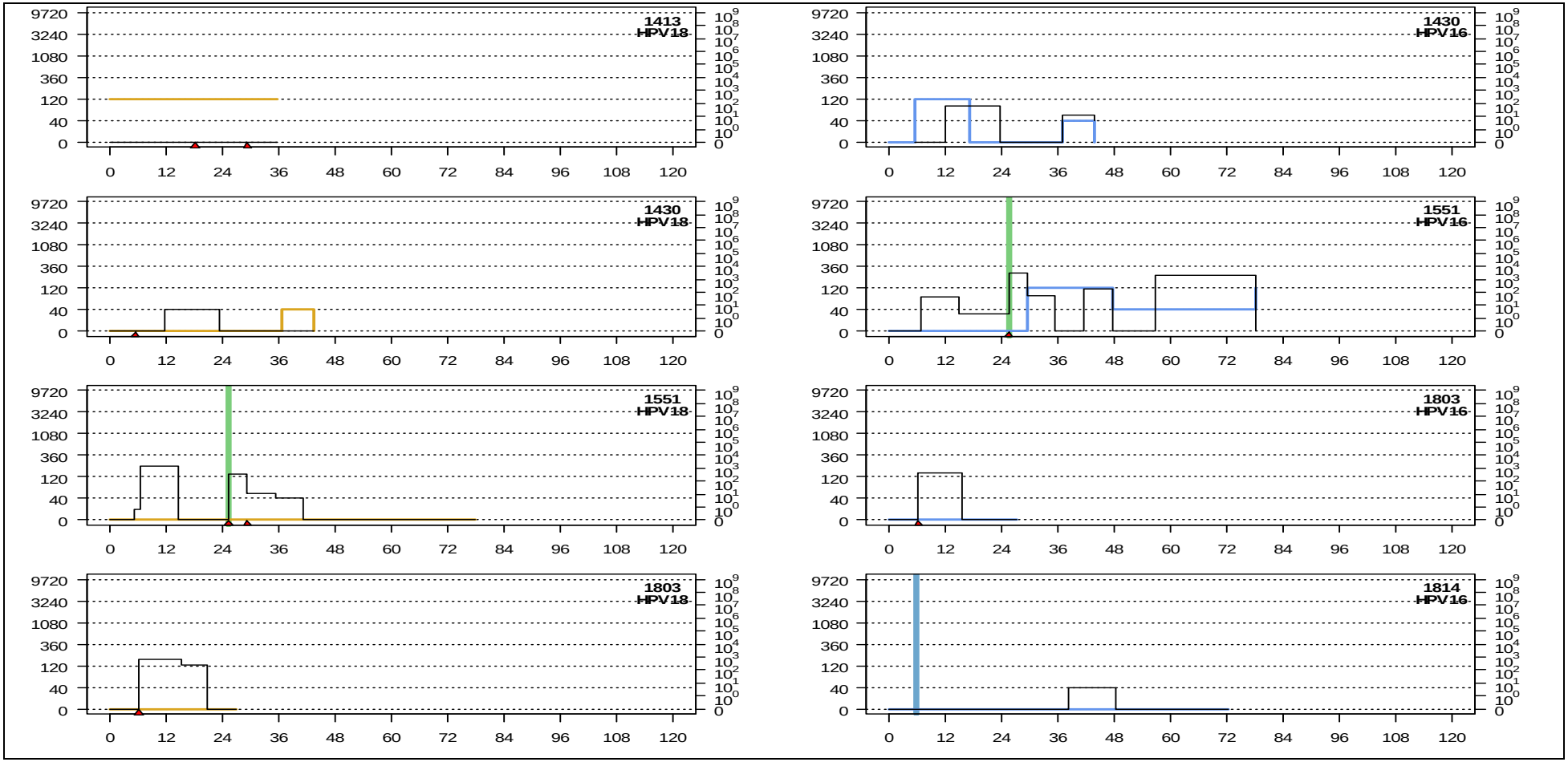
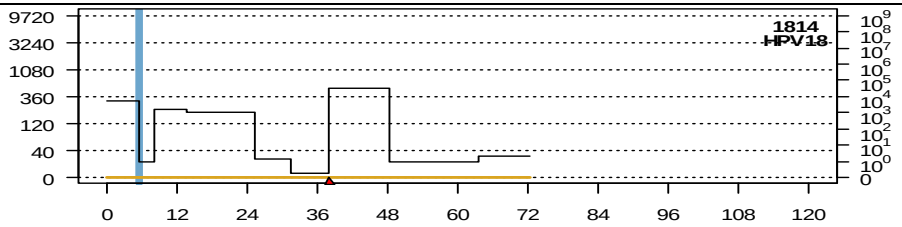


Figure 24.2 (continued). Changes in HPV16 (—) and HPV18 (—) 50% neutralizing antibody titre over time with changes over time in viral load of the relevant HPV type superimposed (—). The woman's study number and the HPV type being analysed appear in the top right-hand corner. All graphs are plotted on the same time scale (months), left-vertical scale (inverse of the 50% neutralizing antibody titre), and right-vertical scale (HPV copy number per 1,000 cells). The measurements of the 50% neutralizing antibody titre and viral load retain their last known value until updated. A coloured vertical line indicates date of first detection of cytological abnormality (blue=borderline nuclear abnormalities; green=mild dyskaryosis). Red triangles indicate a GP5+/GP6+ positive sample of the relevant type.



Appendix 1

PATIENT QUESTIONNAIRE

Retrospective

PATIENT DETAILS

Study Number |_|_|_|_|_|_|_|_|_|_| Brook Number |_|_|_|_|_|_|_|_|_|_|

Surname _____ Forenames _____

Address (Resident) _____

_____

_____

Post code _____ Correspondence _____

Family Address
(if different) _____

_____

_____

Post code _____ Correspondence _____

Post restante
(if necessary) _____

_____

_____

Post code _____ Correspondence _____

G P Address _____

_____

_____

Post code _____ Contact _____

1. Date of interview ⌊__⌋⌊__⌋ / ⌊__⌋⌊__⌋ / ⌊__⌋⌊__⌋

2. Date of birth ⌊__⌋⌊__⌋ / ⌊__⌋⌊__⌋ / ⌊__⌋⌊__⌋ Age ⌊__⌋⌊__⌋

3. Ethnic Origin ⌊__⌋

1 Caucasian 2 Afro Carribean 3 IndoPakistani
4 Oriental 5 Hispanic 6 Other _____

4. Marital status ⌊__⌋

1 Single 2 Married 3 Separated
4 Divorced 5 Widowed

5. Education Are you still at school (Y=Yes N=No) ⌊__⌋

Are you 1 Undertaking Further education ⌊__⌋
2 Undertaking Higher ed/Prof training
3 Finished Further education
4 Finished Higher ed
5 None

At what age did you leave school ⌊__⌋⌊__⌋

Do you intend to go on to 1 Further educ ⌊__⌋
2 Higher ed/Prof training
3 None

6. Occupation What is your occupation _____ ⌊__⌋⌊__⌋⌊__⌋

Social class ⌊__⌋

Your fathers occupation _____ ⌊__⌋⌊__⌋⌊__⌋

Social class ⌊__⌋

Your mothers occupation _____ ⌊__⌋⌊__⌋⌊__⌋

Social class ⌊__⌋

Your partners occupation _____ ⌊__⌋⌊__⌋⌊__⌋

Social class ⌊__⌋

Are your parents separated or divorced (Y=Yes N=No) ⌊__⌋

7. Housing — What type of house do you occupy

1 Boarding school
2 Tennant
3 Owner occupied
4 NFA
5 Student residence
6 Local Authority Care
7 Work residence

Who do you live with

1 Parents
2 Boyfriend
3 Husband
4 Friends
5 Alone
6 Institution
7 Relatives

8. Medical — Have you ever suffered with any of the following (Y=Yes N=No D=Dont know)

- Measles
- Chicken Pox
- Rubella
- Mumps
- Glandular Fever
- Ever had an appendicectomy

9. Smoking

- Have you ever regularly smoked cigarettes
- How old were you when you started
- Do you still smoke
- If you have given up, how old were you
- How many cigarettes per day

1 1-9
2 10-19
3 20-29
4 30-39
5 40+

10. Alcohol — What is your daily consumption of alcohol (a drink is equivalent to a short measure of spirits, a glass of wine, half a pint of beer).

1 Nil
2 1-3/month
3 1-5/week
4 1-2/day
5 3-4/day
6 5-6/day

└__┘

11. Gynacological History — Age at menarche └__┘

12. Obstetric History — Have you ever been pregnant └__┘

How old when first pregnant └__┘

Parity └__┘

13. Family History of Cancer - Has anyone in your close family ever had cancer

Who	Age	Site of cancer, treatment and conclusion
└__┘	└__┘	
└__┘	└__┘	
└__┘	└__┘	

Relative code
M Mother — F Father — B Brother
S Sister — G Grandmother — A Aunt
N Niece — C Cousin

Has anyone had cancer of the cervix in your family (code as above) └__┘

└__┘

14. Sexual History When did you first have intercourse ⌊__⌋/⌊__⌋

How old were you when you first had vaginal intercourse. ⌊__⌋

How many men have you had intecourse with ⌊__⌋

When did you start taking the pill ⌊__⌋/⌊__⌋

Would you mind answering some questions about the last 6 partners you had

	Month/Year first intercourse	Age partnr	Month/Year last intercourse	Frequency	Cont used	For how long	Barr used	For how long	Partnr virgin
Curr	⌊__⌋/⌊__⌋	⌊__⌋	⌊__⌋/⌊__⌋	⌊__⌋	⌊_⌋	⌊__⌋	⌊_⌋	⌊__⌋	⌊_⌋
2	⌊__⌋/⌊__⌋	⌊__⌋	⌊__⌋/⌊__⌋	⌊__⌋	⌊_⌋	⌊__⌋	⌊_⌋	⌊__⌋	⌊_⌋
3	⌊__⌋/⌊__⌋	⌊__⌋	⌊__⌋/⌊__⌋	⌊__⌋	⌊_⌋	⌊__⌋	⌊_⌋	⌊__⌋	⌊_⌋
4	⌊__⌋/⌊__⌋	⌊__⌋	⌊__⌋/⌊__⌋	⌊__⌋	⌊_⌋	⌊__⌋	⌊_⌋	⌊__⌋	⌊_⌋
5	⌊__⌋/⌊__⌋	⌊__⌋	⌊__⌋/⌊__⌋	⌊__⌋	⌊_⌋	⌊__⌋	⌊_⌋	⌊__⌋	⌊_⌋
6	⌊__⌋/⌊__⌋	⌊__⌋	⌊__⌋/⌊__⌋	⌊__⌋	⌊_⌋	⌊__⌋	⌊_⌋	⌊__⌋	⌊_⌋

Contraception codes 1 Combined oral contraceptive pill (CP) 2 Progesterone only pill (POP)
3 Sheath 4 Cap 5 Intrauterine device
6 Withdrawal 7 Sterilized (self) 8 Sterilized (partner)
9 None

If aged 20 or ~~over~~ can you please give details of your history for the previous 6 years

Year	How many new partners	Frequency	Contr used	Ever use a barrier method
1st coital year	⌊__⌋	⌊__⌋	⌊_⌋	⌊_⌋
2nd coital year	⌊__⌋	⌊__⌋	⌊_⌋	⌊_⌋
3rd coital year	⌊__⌋	⌊__⌋	⌊_⌋	⌊_⌋
4th coital year	⌊__⌋	⌊__⌋	⌊_⌋	⌊_⌋
5th coital year	⌊__⌋	⌊__⌋	⌊_⌋	⌊_⌋
6th coital year	⌊__⌋	⌊__⌋	⌊_⌋	⌊_⌋

15. History

Do you suffer from vulval warts ⌊__⌋

Have you ever had vulval warts in the past ⌊__⌋

Do you ever suffer from warts elsewhere (code where) ⌊__⌋ ⌊__⌋

Have you ever had warts elsewhere (code where) ⌊__⌋ ⌊__⌋

Does your boyfriend have penile warts ⌊__⌋

Has your boyfriend ever had penile warts ⌊__⌋

Have your previous partners ever had penile warts ⌊__⌋

Site of warts code

1 Hands	2 Feet
3 Legs	4 Arms
5 Perianal	6 Trunk
7 Head	

16.

Do you currently have genital herpes ⌊__⌋

Have you ever had genital herpes ⌊__⌋

Does your partner currently have genital herpes ⌊__⌋

Has your partner ever had genital herpes ⌊__⌋

Have your previous partners ever had genital herpes ⌊__⌋

17.

Do you currently have oral cold-sores ⌊__⌋

Have you ever had oral cold-sores ⌊__⌋

Does your partner currently have oral cold-sores ⌊__⌋

Has your partner ever had oral cold-sores ⌊__⌋

18. Have you ever been to a "Special Clinic" ⌊__⌋

Have you ever been treated at a "Special Clinic" ⌊__⌋

If so, do you know what for _____ ⌊__⌋

blank None
1 Syphilis 2 Gonorrhoea 3 Thrush
4 Trichonomas 5 Gadinella 6 Pelvic Inflamm disease
7 Warts 8 Non-spc Ureth 9 Chlamydia
A Genital herpes B Other C Not known

Has your present partner ever been treated at a "Special Clinic" ⌊__⌋

If so, do you know what for _____ ⌊__⌋
(code as above)

Have any of your previous partners ever been treated at a "Special Clinic" ⌊__⌋

If so, do you know what for _____ ⌊__⌋
(code as above)

19. Has your present partner ever had glandular fever ⌊__⌋

Have your previous partners ever had glandular fever ⌊__⌋

20. Examination Macroscpøic Warts Vulval ⌊__⌋

Cervix ⌊__⌋

Other abnormalities _____

Date of smear ⌊__⌋⌊__⌋ / ⌊__⌋⌊__⌋ / ⌊__⌋⌊__⌋

Result of smear ⌊__⌋⌊__⌋⌊__⌋

E ⌊__⌋

M ⌊__⌋

I ⌊__⌋

Date of previous smear ⌊__⌋⌊__⌋ / ⌊__⌋⌊__⌋ / ⌊__⌋⌊__⌋

Result of previous smear ⌊__⌋⌊__⌋⌊__⌋

Appendix 2

PATIENT QUESTIONNAIRE

Prospective

DETAILS - change only

Study Number |_|_|_|_|_|_|_| Brook Number |_|_|_|_|_|_|_|

Surname _____ Maiden name _____

Address (Resident) _____

_____

_____

Post code _____ Correspondence _____

Family Address
(if different) _____

_____

_____

Post code _____ Correspondence _____

Post restante
(if necessary) _____

_____

_____

Post code _____ Correspondence _____

G P Address _____

_____

_____

Post code _____ Contact _____

Previous post code _____

1. Date of interview - Visit 1 __ / __ / __

- Visit 2 __ / __ / __
- Visit 3 __ / __ / __
- Visit 4 __ / __ / __
- Visit 5 __ / __ / __
- Visit 6 __ / __ / __
- Visit 7 __ / __ / __
- Visit 8 __ / __ / __
- Visit 9 __ / __ / __
- Visit 10 __ / __ / __
- Visit 11 __ / __ / __
- Visit 12 __ / __ / __

2. Marital status

1	2	3	4	5	6	7	8	9	10	11	12

1 Single 2 Married 3 Separated
4 Divorced 5 Widowed

3. Job status

1	2	3	4	5	6	7	8	9	10	11	12

1 At school
2 At college of further education
3 Higher education, professional training
4 Unemployed
5 Housewife
6 Employed - same occupation
7 Employed - different or new occupation

If employment (code 7) has changed give new occupation

Visit 1 _____ Visit 2 _____

Visit 3 _____ Visit 4 _____

Visit 5 _____ Visit 6 _____

Visit 7 _____ Visit 8 _____

Visit 9 _____ Visit 10 _____

Visit 11 _____ Visit 12 _____

4. Social class

1	2	3	4	5	6	7	8	9	10	11	12

5. What type of housing do you occupy?

1	2	3	4	5	6	7	8	9	10	11	12

1 Work residence 2 Tennant
3 Owner occupied 4 NFA
5 Student residence 6 Local Authority residence

6. Who do you live with?

1	2	3	4	5	6	7	8	9	10	11	12

1 Parents 2 Boyfriend
3 Husband 4 Friends
5 Alone 6 Institution

7. Are you cigarette smoker?

1	2	3	4	5	6	7	8	9	10	11	12

8. Have you changed the no. you smoke since last visit?

1	2	3	4	5	6	7	8	9	10	11	12

1 Stopped 2 Same????? 3 Started 4 Cut down 5 Increased

9. If smoking, How many per day?

1 2 3 4 5 6 7 8 9 10 11 12

1 1-9 2 10-19 3 20-29 4 30-39 5 40+

10. Consumption of alcohol?

1 2 3 4 5 6 7 8 9 10 11 12

1 nil 2 1-3/month 3 1-5/week
4 1-2/day 5 3-4/day 6 5-6/day

11. Have you become pregnant since last visit

1 2 3 4 5 6 7 8 9 10 11 12

12. If Yes, what was the outcome?

1 2 3 4 5 6

7 8 9 10 11 12

13. Has anyone in your close family developed cancer since I last interviewed you?

Who	Age	Site of cancer, treatment and conclusion

Relative code M Mother F Father B Brother S Sister
G Grandmother A Aunt N Niece C Cousin

14. Sexual History

	Date of interview	Contin. partner	Date of start of relation	Age partner	Date of end of relation	Frequency	Contr used	How long	Barr used	How long	Soc cls partner	Occup	
14	__/__/__	__	__/__/__	__	__/__/__	____	__	____	__	____	__	____	
13	__/__/__	__	__/__/__	__	__/__/__	____	__	____	__	____	__	____	
12	__/__/__	__	__/__/__	__	__/__/__	____	__	____	__	____	__	____	
11	__/__/__	__	__/__/__	__	__/__/__	____	__	____	__	____	__	____	
10	__/__/__	__	__/__/__	__	__/__/__	____	__	____	__	____	__	____	
9	__/__/__	__	__/__/__	__	__/__/__	____	__	____	__	____	__	____	
8	__/__/__	__	__/__/__	__	__/__/__	____	__	____	__	____	__	____	
7	__/__/__	__	__/__/__	__	__/__/__	____	__	____	__	____	__	____	
6	__/__/__	__	__/__/__	__	__/__/__	____	__	____	__	____	__	____	
5	__/__/__	__	__/__/__	__	__/__/__	____	__	____	__	____	__	____	
4	__/__/__	__	__/__/__	__	__/__/__	____	__	____	__	____	__	____	
3	__/__/__	__	__/__/__	__	__/__/__	____	__	____	__	____	__	____	
2	__/__/__	__	__/__/__	__	__/__/__	____	__	____	__	____	__	____	
1	__/__/__	__	__/__/__	__	__/__/__	____	__	____	__	____	__	____	

Contaception codes 1 Combined oral contraceptive pill (CP) 2 Progesterone only pill (POP) 3 Sheath 4 Cap 5 Intrauterine device
6 Withdrawa 7 Sterilized (self) 8 Sterilized (partner) 9 None

How long codes D Day M Month Y Year X Times
ALL All the time MFC Main form of contraception

Appendix 3

15. Have you contracted venereal disease or vaginal infection since last visit?

1 2 3 4 5 6 7 8 9 10 11 12

blank None
1 Syphilis 2 Gonorrhoea 3 Thrush 4 Trichonomas
5 Gadinella 6 Pelvic Inflammatory disease 7 Warts 8 Non-spec Ureth
9 Chlamydia A Genital herpes B Other C Not known

16. Have your partner contracted venereal disease or vaginal infection since last visit?

1 2 3 4 5 6 7 8 9 10 11 12

Code as above

17. Have you had oral cold sores since last visit?

1 2 3 4 5 6 7 8 9 10 11 12

18. Have your partner had oral cold sores since last visit?

1 2 3 4 5 6 7 8 9 10 11 12

19. Have you had glandular fever since last visit?

1 2 3 4 5 6 7 8 9 10 11 12

20. Have your partner had glandular fever since last visit?

1 2 3 4 5 6 7 8 9 10 11 12

21. Smear Results

No	Date of smear	Result	E+	M+	I
1	__/__/__	____	__	__	__
2	__/__/__	____	__	__	__
3	__/__/__	____	__	__	__
4	__/__/__	____	__	__	__
5	__/__/__	____	__	__	__
6	__/__/__	____	__	__	__
7	__/__/__	____	__	__	__
8	__/__/__	____	__	__	__
9	__/__/__	____	__	__	__
10	__/__/__	____	__	__	__
11	__/__/__	____	__	__	__
12	__/__/__	____	__	__	__

PROXIMITY OF MENARCHE TO FIRST SEXUAL INTERCOURSE AND THE RISK OF CERVICAL HPV INFECTION

PUBLISHED PAPER

The work described in chapters 8, 9, 10 and 11 of this thesis has been published. This paper is reproduced here.

Appendix 4

SMOKING, CERVICAL HPV INFECTION, AND THE RISK OF CERVICAL NEOPLASIA

PUBLISHED PAPER

The work described in chapters 12, 13, 14 and 15 of this thesis has been published. This paper is reproduced here.

Appendix 5

THE KINETICS OF THE HUMORAL IMMUNE RESPONSE TO INCIDENT CERVICAL HPV16 AND HPV18 INFECTIONS

LABORATORY METHODS

The following is a formal and precise description of the laboratory techniques used to undertake the neutralization assay used to measure the neutralizing antibody response in the cohort described in this thesis, given in sufficient detail to enable it to be repeated. Where relevant these laboratory techniques have been described in chapter 17 in a more accessible manner.

The preparation of pseudovirions (PsV) for use in the neutralization assays was performed essentially as described by Pastrana et al, 2004 (Pastrana 2004). Briefly, pseudovirus stocks were prepared by transfection of 293TT cells with vectors (all kindly provided by Dr John Schiller, Division of Basic Sciences, National Cancer Institute, Bethesda, MD 20892-4040, USA) expressing HPV16 L1 (p16L1h), HPV16 L2 (p16L2h), HPV18 L1 (peL1fB), HPV18 L2 (peL2bhb), BPV1 L1/L2 (pSheLL) together with secreted alkaline phosphatase (pYSEAP). For the production of HPV16 and HPV18 pseudovirions, L1/L2 pairs were co-transfected with the reporter plasmid. Three days following transfection, cells were harvested and pseudovirus isolated over Optiprep gradients (Accurate Chemical, Westbury, NY). For each pseudovirus, fractions were titrated to determine the minimum amount of PsV required to give a robust signal in an assay for secreted alkaline phosphatase activity following infection of 293TT cells. This concentration of PsV was then used in subsequent neutralization assays. For the titration, 293TT cells (3 x 10^4/well) were plated in 96-well plates in neutralization assay medium (NAM; 100µl/well) comprising DMEM without phenol red (Gibco-Invitrogen), 10% heat-inactivated FCS, 1% MEM Non-Essential Amino acids (Gibco-Invitrogen), 1% Glutamax (Gibco-Invitrogen) and 1% Antibiotic-antimycotic (Gibco-Invitrogen) and incubated for 2 – 5 hours at 37^{o}C. Serial dilutions of the PsV (1:200 - 1:50,000) were prepared using siliconised tips in 96-well polystyrene plates (Corning Costar) in triplicate, with and without heparin treatment (Sigma; incubation for 1hr at 4^{o}C), added to the preplated cells (final volume 200µl/well) and incubated at 37^{o}C for 72hours. Cell supernatants were then assayed for alkaline phospatase activity by chemiluminescense.

The method for detection of SEAP in 293TT cell supernatants was again essentially as described in Pastrana et al, 2004 (Pastrana 2004) using the Great Escape SEAP detection kit (BD Bioscience). Dilution buffer from the kit was added directly to the wells of a black optiplate-96 assay plate (Perkin Elmer; 45µl/well). The 96-well plates containing 293TT cells and PsV were lightly shaken and 50µl of a homogeneous cell suspension from each well was

transferred to the corresponding wells of a fresh plate and centrifuged at 800g for 5min. Clarified supernatant from each well (15µl) was removed and added to wells containing dilution buffer on the optiplate, the plate covered with plastic and then incubated for 30min at 65^{o}C. After cooling on ice for 2 - 5min, assay buffer was added (60µl/well) and plates incubated at room temperature for 5min. Chemiluminescence substrate was added (50µl/well) and following a final incubation period of 20min at room temperature, the plates were read on a microplate reader set at Glow-Endpoint 0.2sec/wellRAW Data.

Neutralization assays were performed by first diluting various concentrations of the test serum in NAM in polystyrene 96-well plates and incubating with titrated HPV16, HPV18 and BPV1 pseudovirus stocks (in triplicate) for 1hr on ice in a total volume of 100µl. Triplicate wells of titrated pseudovirus alone were also set up to determine the maximum signal in the absence of neutralizing serum. The contents of each well was transferred to 96-well tissue culture plates previously seeded with 293TT cells (3 x 10^{4}/well) in 100µl NAM. Plates were incubated for 72hr at 37^{o}C and SEAP activity in cell supernatants performed using the chemiluminescent assay described above. The 50% neutralizing titre was determined for each serum i.e. the highest dilution at which the signal was less than 50% of the mean signal seen in the “no sera” control wells. A serum was considered to be positive in the HPV16 or HPV18 assay if it was neutralizing at a dilution at least 4-fold higher than the titre observed in the BPV1 assay.

Appendix 6

THE KINETICS OF THE HUMORAL IMMUNE RESPONSE TO INCIDENT CERVICAL HPV16 AND HPV18 INFECTIONS

PUBLISHED PAPER

The work described in chapters 16, 17, 18 and 19 of this thesis has been published. This paper is reproduced here.

Appendix 7

THE NATURAL HISTORY OF HPV16 AND HPV18 INFECTIONS – VIRAL LOAD

LABORATORY METHODS

The following is a formal and precise description of the laboratory techniques used to undertake the qPCR assay used to measure HPV16 and HPV18 viral load in the cohort described in this thesis, given in sufficient detail to enable it to be repeated. Where relevant these laboratory techniques have been described in chapter 21 in a more accessible manner.

HPV viral load was measured in a modified singleplex quantitative PCR (qPCR) assay, with standard curves used to correct measurements of viral load for the number of cells present in the sample.

DNA was isolated from study samples using Proteinase K digestion and phenol/chloroform extraction. Type-specific primers for the amplification of the HPV16 E6 and HPV18 E7 oncogenes, and the GAPDH gene, were designed using a web-based computer program called Primer3 (http://fokker.wi.mit.edu/primer3/input.htm). Standard curves were generated using serial 10-fold dilutions of plasmids, which contained known numbers of copies of HPV16, HPV18, or GAPDH. A HPV16/Bluescript II KS^+ plasmid (9,111 bp), which contains HPV16 up to the BAMHI site (6,150 bp), was used in assays of HPV16; and a HPV18/pGEM-2 plasmid (10,726 bp), which contains the complete HPV18 genome (7,857 bp), was used for HPV18 (both plasmids kindly provided by Sally Roberts). The GAPDH plasmid which was used was produced by amplification of a 2,039 bp PCR product from DNA obtained from the HPV16-positive cervical carcinoma cell line W12.

All assays were performed in singleplex, with target samples, standards and controls all tested in triplicate in each assay. Genomic DNA (50 ng), standards, positive controls and negative controls were loaded onto TaqMan® plates. Each reaction was amplified using 1x TaqMan® Universal PCR Master Mix (Applied Biosystems, Foster City, USA) with 0.4 μmol/L primer mix and 0.1 μmol/L of appropriate fluorogenic probe. TaqMan® plates were centrifuged at 2,750 RPM for 5 minutes at 4°C and placed into the ABI 7700 SDS thermal cycler. Cycle conditions for amplification of GAPDH, HPV16 E6 and HPV18 E7 were: 50°C for 2 min, 95°C for 12 min, followed by 50 cycles of 95°C for 15 s and 55°C for 30 s. Following amplification, the ABI 7700 SDS.1 detection software was used to set the cycle threshold (Ct) level within the linear phase of the exponential amplification in the growth curves.

Table A4a. Primer parameters for the qPCR assay.

Target	Accession Number	Nucleotide Sequence Location	Nucleotide Sequence	Type	Length (bases)	C	G	G+C Content (%)	Annealing Temperature (Tm)	Amplicon Length (bases)
GAPDH	AY340484	3701	*5'-GCTCAAGGGAGATAAAATTC-3'*	Forward primer	20	3	5	40%	55^{o}C	158
		3858	*5'-CGACCAAATCTAAGAGACAA-3'*	Reverse primer	20	5	3	40%		
		3783	*FAM-5'-CCTAGGGCTGCTCACATATT-3'-TAMRA*	Probe	20	6	4	50%		
HPV16 E6	NC_001526	368	*5'-GAACAGCAATACAACAAACC -3'*	Forward primer	20	6	2	40%	55^{o}C	161
		528	*5'-GATCTGCAACAAGACATACA -3'*	Reverse primer	20	5	3	40%		
		418	*FAM-5'- CTGTCAAAAGCCACTGTGTC-3'-TAMRA*	Probe	20	6	4	50%		
HPV18 E7	NC_001357	76	*5'-GTTGACCTTCTATGTCACGA -3'*	Forward primer	20	5	4	45%	55^{o}C	151
		226	*5'-CAATTCTGGCTTCACACTTA -3'*	Reverse primer	20	6	2	40%		
		97	*FAM-5'-CAATTAAGCGACTCAGAGGAA-3'-TAMRA*	Probe	21	4	5	43%		

Table A4b. Primer parameters for manufacturing the GAPDH plasmid.

Target	Accession Number	Nucleotide Sequence Location	Nucleotide Sequence	Type	Length (bases)	C	G	G+C Content (%)	Annealing Temperature (Tm)	Amplicon Length (bases)
GAPD Plasmid	AY340484	3233	5'-CCACACACATGCACTTACCT-3'	Forward primer	20	9	1	50%	57.3^{o}C	2039
		5271	5'-AAGTCAGAGGAGACCACCTG-3'	Reverse primer	20	5	6	55%	57.3^{o}C	

REFERENCES

Abbas AK, Lichtman AH (2003). Cellular and molecular immunology: fifth edition. Saunders, Philadelphia.

Alam S, Conway MJ, Chen HS and Meyers C (2008). The cigarette smoke carcinogen benzo[a]pyrene enhances human papillomavirus synthesis. *J Virol* **82:**1053-8.

Alani RM and Munger K (1998). Human papillomaviruses and associated malignancies. *J Clin Oncol* **16:**330-7.

Alioum A and Commenges D (1996). A proportional hazards model for arbitrarily censored and truncated data. *Biometrics* **52:**512-24.

Altman DG and Royston P (2006). The cost of dichotomising continuous variables. *Bmj* **332:**1080.

Andersen MH, Schrama D, Thor Straten P and Becker JC (2006). Cytotoxic T cells. *J Invest Dermatol* **126:**32-41.

Andersson-Ellstrom A, Hagmar BM, Johansson B, *et al.* (1996a). Human papillomavirus deoxyribonucleic acid in cervix only detected in girls after coitus. *Int J STD AIDS* **7:**333-6.

Andersson-Ellstrom A, Dillner J, Hagmar B, *et al.* (1996b). Comparison of development of serum antibodies to HPV16 and HPV33 and acquisition of cervical HPV DNA among sexually experienced and virginal young girls. A longitudinal cohort study. *Sex Transm Dis* **23:**234-8.

Arends MJ, Buckley CH and Wells M (1998). Aetiology, pathogenesis, and pathology of cervical neoplasia. *J Clin Pathol* **51:**96-103.

Bacchetti P and Quale C (2002). Generalized additive models with interval-censored data and time-varying covariates: application to human immunodeficiency virus infection in hemophiliacs. *Biometrics* **58:**443-7.

Baseman JG and Koutsky LA (2005). The epidemiology of human papillomavirus infections. *J Clin Virol* **32 Suppl 1:**S16-24.

Becker TM, Wheeler CM, McGough NS, *et al.* (1994). Cigarette smoking and other risk factors for cervical dysplasia in southwestern Hispanic and non-Hispanic white women. *Cancer Epidemiol Biomarkers Prev* **3:**113-9.

Bedell MA, Hudson JB, Golub TR, *et al.* (1991). Amplification of human papillomavirus genomes in vitro is dependent on epithelial differentiation. *J Virol* **65:**2254-60.

Berumen J, Unger ER, Casas L and Figueroa P (1995). Amplification of human papillomavirus types 16 and 18 in invasive cervical cancer. *Hum Pathol* **26:**676-81.

Bleeker MC, Hogewoning CJ, Berkhof J, *et al.* (2005). Concordance of specific human papillomavirus types in sex partners is more prevalent than would be expected by chance and is associated with increased viral loads. *Clin Infect Dis* **41:**612-20.

Blomfield PI, Lancashire RJ and Woodman CB (1998). Can women at risk of cervical abnormality be identified? *Br J Obstet Gynaecol* **105:**486-92.

Bontkes HJ, Walboomers JM, Meijer CJ, Helmerhorst TJ and Stern PL (1998a). Specific HLA class I down-regulation is an early event in cervical dysplasia associated with clinical progression. *Lancet* **351:**187-8.

Bontkes HJ, van Duin M, de Gruijl TD, *et al.* (1998b). HPV 16 infection and progression of cervical intra-epithelial neoplasia: analysis of HLA polymorphism and HPV 16 E6 sequence variants. *Int J Cancer* **78:**166-71.

Bontkes HJ, de Gruijl TD, Bijl A, *et al.* (1999a). Human papillomavirus type 16 E2-specific T-helper lymphocyte responses in patients with cervical intraepithelial neoplasia. *J Gen Virol* **80 (Pt 9):**2453-9.

Bontkes HJ, de Gruijl TD, Walboomers JM, *et al.* (1999b). Immune responses against human papillomavirus (HPV) type 16 virus-like particles in a cohort study of women with cervical intraepithelial neoplasia. II. Systemic but not local IgA responses correlate with clearance of HPV-16. *J Gen Virol* **80 (Pt 2):**409-17.

Bontkes HJ, de Gruijl TD, van den Muysenberg AJ, *et al.* (2000). Human papillomavirus type 16 E6/E7-specific cytotoxic T lymphocytes in women with cervical neoplasia. *Int J Cancer* **88:**92-8.

Bory JP, Cucherousset J, Lorenzato M, *et al.* (2002). Recurrent human papillomavirus infection detected with the hybrid capture II assay selects women with normal cervical smears at risk for developing high grade cervical lesions: a longitudinal study of 3,091 women. *Int J Cancer* **102:**519-25.

Bosch FX, Manos MM, Munoz N, *et al.* (1995). Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. *J Natl Cancer Inst* **87:**796-802.

Bosch FX and de Sanjose S (2003). Chapter 1: Human papillomavirus and cervical cancer--burden and assessment of causality. *J Natl Cancer Inst Monogr* **31:**3-13.

Bray F, Loos AH, McCarron P, *et al.* (2005a). Trends in cervical squamous cell carcinoma incidence in 13 European countries: changing risk and the effects of screening. *Cancer Epidemiol Biomarkers Prev* **14:**677-86.

Bray F, Carstensen B, Moller H, *et al.* (2005b). Incidence trends of adenocarcinoma of the cervix in 13 European countries. *Cancer Epidemiol Biomarkers Prev* **14:**2191-9.

Brisson J, Morin C, Fortier M, *et al.* (1994). Risk factors for cervical intraepithelial neoplasia: differences between low- and high-grade lesions. *Am J Epidemiol* **140:**700-10.

Brock KE, MacLennan R, Brinton LA, *et al.* (1989). Smoking and infectious agents and risk of in situ cervical cancer in Sydney, Australia. *Cancer Res* **49:**4925-8.

Buchinsky FJ, Carter JJ, Wipf GC, *et al.* (2006). Comparison of oral fluid and serum ELISAs in the determination of IgG response to natural human papillomavirus infection in university women. *J Clin Virol* **35:**450-3.

Buck CB, Pastrana DV, Lowy DR and Schiller JT (2005). Generation of HPV pseudovirions using transfection and their use in neutralization assays. *Methods Mol Med* **119:**445-62.

Bulkmans NW, Berkhof J, Rozendaal L, *et al.* (2007). Human papillomavirus DNA testing for the detection of cervical intraepithelial neoplasia grade 3 and cancer: 5-year follow-up of a randomised controlled implementation trial. *Lancet* **370:**1764-72.

Burk RD, Kelly P, Feldman J, *et al.* (1996a). Declining prevalence of cervicovaginal human papillomavirus infection with age is independent of other risk factors. *Sex Transm Dis* **23:**333-41.

Burk RD, Ho GY, Beardsley L, *et al.* (1996b). Sexual behavior and partner characteristics are the predominant risk factors for genital human papillomavirus infection in young women. *J Infect Dis* **174:**679-89.

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

Campaner AB, Piato S, Galvao MA, dos Santos RE and Nadais RF (2006). Langerhans cells in cervical intraepithelial neoplasia related to smoking habits. *J Low Genit Tract Dis* **10:**223-8.

Cann AJ (2001). Principles of molecular virology: third edition. Academic press, London.

Carcopino X, Henry M, Benmoura D, *et al.* (2006). Determination of HPV type 16 and 18 viral load in cervical smears of women referred to colposcopy. *J Med Virol* **78:**1131-40.

Carstensen B (1996). Regression models for interval censored survival data: application to HIV infection in Danish homosexual men. *Stat Med* **15:**2177-89.

Carter JJ, Koutsky LA, Wipf GC, *et al.* (1996). The natural history of human papillomavirus type 16 capsid antibodies among a cohort of university women. *J Infect Dis* **174:**927-36.

Carter JJ, Koutsky LA, Hughes JP, *et al.* (2000). Comparison of human papillomavirus types 16, 18, and 6 capsid antibody responses following incident infection. *J Infect Dis* **181:**1911-9.

Castellsague X and Munoz N (2003). Chapter 3: Cofactors in human papillomavirus carcinogenesis--role of parity, oral contraceptives, and tobacco smoking. *J Natl Cancer Inst Monogr* 20-8.

Castellsague X, Diaz M, de Sanjose S, *et al.* (2006). Worldwide human papillomavirus etiology of cervical adenocarcinoma and its cofactors: implications for screening and prevention. *J Natl Cancer Inst* **98:**303-15.

Castle PE, Wacholder S, Sherman ME, *et al.* (2002a). Absolute risk of a subsequent abnormal pap among oncogenic human papillomavirus DNA-positive, cytologically negative women. *Cancer* **95:**2145-51.

Castle PE, Wacholder S, Lorincz AT, *et al.* (2002b). A prospective study of high-grade cervical neoplasia risk among human papillomavirus-infected women. *J Natl Cancer Inst* **94:**1406-14.

Castle PE, Lorincz AT, Scott DR, *et al.* (2003). Comparison between prototype hybrid capture 3 and hybrid capture 2 human papillomavirus DNA assays for detection of high-grade cervical intraepithelial neoplasia and cancer. *J Clin Microbiol* **41:**4022-30.

Castle PE, Schiffman M, Bratti MC, *et al.* (2004). A population-based study of vaginal human papillomavirus infection in hysterectomized women. *J Infect Dis* **190:**458-67.

Castle PE, Schiffman M, Scott DR, *et al.* (2005a). Semiquantitative human papillomavirus type 16 viral load and the prospective risk of cervical precancer and cancer. *Cancer Epidemiol Biomarkers Prev* **14:**1311-4.

Castle PE, Schiffman M, Herrero R, *et al.* (2005b). A prospective study of age trends in cervical human papillomavirus acquisition and persistence in Guanacaste, Costa Rica. *J Infect Dis* **191:**1808-16.

Castle PE, Jeronimo J, Schiffman M, *et al.* (2006). Age-related changes of the cervix influence human papillomavirus type distribution. *Cancer Res* **66:**1218-24.

Clavel C, Masure M, Levert M, *et al.* (2000). Human papillomavirus detection by the hybrid capture II assay: a reliable test to select women with normal cervical smears at risk for developing cervical lesions. *Diagn Mol Pathol* **9:**145-50.

Collett D (2003). Modelling survival data in medical research: second edition. Chapman and Hall/CRC, London.

Collins S, Mazloomzadeh S, Winter H, *et al.* (2002). High incidence of cervical human papillomavirus infection in women during their first sexual relationship. *Br J Obstet Gynaecol* **109:**96-8.
Collins SI, Mazloomzadeh S, Winter H, *et al.* (2005). Proximity of first intercourse to menarche and the risk of human papillomavirus infection: a longitudinal study. *Int J Cancer* **114:**498-500.

Collins SI, Constandinou-Williams C, Wen K *et al.* (2009). Disruption of the E2 gene is a common and early event in the natural history of cervical human papillomavirus infection: a longitudinal cohort study. *Cancer Res* **69:**3828-32.

Collins S, Rollason TP, Young LS, Woodman CBJ (2010). Cigarette smoking is an independent risk factor for cervical intraepithelial neoplasia in young women after controlling for HPV infection: a longitudinal study. *European Journal Of Cancer* **46:**405-11.

Constandinou-Williams C, Collins SI, Roberts S, Young LS, Woodman CBJ, Murray PG (2009). Is human papillomavirus viral load a clinically useful predictive marker: a longitudinal study. *Cancer Epidemiology Biomarkers and Prevention* In press.

Coursaget P (2003). Serology for human papillomavirus. *Salud Publica Mex* **45 Suppl 3:**S361-6.

Cricca M, Venturoli S, Morselli-Labate AM, *et al.* (2006). HPV DNA patterns and disease implications in the follow-up of patients treated for HPV16 high-grade carcinoma in situ. *J Med Virol* **78:**494-500.

Crosbie EJ and Kitchener HC (2006). Human papillomavirus in cervical screening and vaccination. *Clin Sci (Lond)* **110:**543-52.

CRUK Website. Website of Cancer Research UK. http://www.cancerresearchuk.org/.

Crum CP, Beach KJ, Hedley ML, *et al.* (2004). Dynamics of human papillomavirus infection between biopsy and excision of cervical intraepithelial neoplasia: results from the ZYC101a protocol. *J Infect Dis* **189:**1348-54.

Cuzick J, Singer A, De Stavola BL and Chomet J (1990). Case-control study of risk factors for cervical intraepithelial neoplasia in young women. *Eur J Cancer* **26:**684-90.

Cuzick J, Sasieni P and Singer A (1996). Risk factors for invasive cervix cancer in young women. *Eur J Cancer* **32A:**836-41.

Dalstein V, Riethmuller D, Pretet JL, *et al.* (2003). Persistence and load of high-risk HPV are predictors for development of high-grade cervical lesions: a longitudinal French cohort study. *Int J Cancer* **106:**396-403.

Davison AC, Hinkley DV (1997). Bootstrap methods and their application. Cambridge University Press, Cambridge.

de Boer MA, Jordanova ES, Kenter GG, *et al.* (2007). High human papillomavirus oncogene mRNA expression and not viral DNA load is associated with poor prognosis in cervical cancer patients. *Clin Cancer Res* **13:**132-8.

de Gruijl TD, Bontkes HJ, Walboomers JM, *et al.* (1996a). Analysis of IgG reactivity against Human Papillomavirus type-16 E7 in patients with cervical intraepithelial neoplasia indicates an association with clearance of viral infection: results of a prospective study. *Int J Cancer* **68:**731-8.

de Gruijl TD, Bontkes HJ, Stukart MJ, *et al.* (1996b). T cell proliferative responses against human papillomavirus type 16 E7 oncoprotein are most prominent in cervical intraepithelial neoplasia patients with a persistent viral infection. *J Gen Virol* **77 (Pt 9):**2183-91.

de Gruijl TD, Bontkes HJ, Walboomers JM, *et al.* (1997). Immunoglobulin G responses against human papillomavirus type 16 virus-like particles in a prospective nonintervention cohort study of women with cervical intraepithelial neoplasia. *J Natl Cancer Inst* **89:**630-8.

de Gruijl TD, Bontkes HJ, Walboomers JM, *et al.* (1999). Immune responses against human papillomavirus (HPV) type 16 virus-like particles in a cohort study of women with cervical intraepithelial neoplasia. I. Differential T-helper and IgG responses in relation to HPV infection and disease outcome. *J Gen Virol* **80:**399-408.

De Marco F, Di Carlo A, Poggiali F, Muller A and Marcante ML (2001). Detection of HPV in genital condylomata: correlation between viral load and clinical outcome. *J Exp Clin Cancer Res* **20:**377-83.

De Sanjose S, Santamaria M, de Ruiz PA (1992). HPV types in women with normal cytology. In: Munoz N, Bosch FX, Shah KV, Meheus A. The epidemiology of human papillomavirus and cervical cancer. IARC, Lyon.

de Vet HC, Sturmans F and Knipschild PG (1994). The role of cigarette smoking in the etiology of cervical dysplasia. *Epidemiology* **5:**631-3.

de Villiers EM, Fauquet C, Broker TR, Bernard HU and zur Hausen H (2004). Classification of papillomaviruses. *Virology* **324:**17-27.

Deacon JM, Evans CD, Yule R, *et al.* (2000). Sexual behaviour and smoking as determinants of cervical HPV infection and of CIN3 among those infected: a case-control study nested within the Manchester cohort. *Br J Cancer* **83:**1565-72.

Depuydt CE, Benoy IH, Bailleul EJ, *et al.* (2006). Improved endocervical sampling and HPV viral load detection by Cervex-Brush Combi. *Cytopathology* **17:**374-81.

Diggle PJ, Liang KY, Zeger SL (2000). Analysis of Longitudinal Data. Oxford University Press, Oxford.

Doorbar J (2005). The papillomavirus life cycle. *J Clin Virol* **32 Suppl 1:**S7-15.

Doorbar J (2006). Molecular biology of human papillomavirus infection and cervical cancer. *Clin Sci (Lond)* **110:**525-41.

Draper NR, Smith H (1998). Applied regression analysis: third edition. John Wiley and Sons, New York.

Farrington CP (1996). Interval censored survival data: a generalized linear modelling approach. *Stat Med* **15:**283-92.

Ferguson M, Heath A, Johnes S *et al.* (2006). Results of the first WHO international collaborative study on the standardization of the detection of antibodies to human papillomaviruses. *Int J Cancer* **118:**1508-14.

Fleiss JL, Levin B, Paik MC (2003) Statistical methods for rates and proportions: third edition. John Wiley and Sons, New Jersey.

Fleury MJ, Touze A, de Sanjose S, *et al.* (2008). Detection of human papillomavirus type 31-neutralizing antibodies from naturally infected patients by an assay based on intracellular assembly of luciferase-expressing pseudovirions. *Clin Vaccine Immunol* **15:**172-5.

Flint SJ, Enquist LW, Racaniello VR, Shalka AM (2004). Principles of Virology - Molecular biology, pathology, and control of animal viruses: second edition. ASM Press, Washington DC.

Flores R, Papenfuss M, Klimecki WT and Giuliano AR (2006). Cross-sectional analysis of oncogenic HPV viral load and cervical intraepithelial neoplasia. *Int J Cancer* **118:**1187-93.

Fontaine J, Gravitt P, Duh LM, *et al.* (2005). High level of correlation of human papillomavirus-16 DNA viral load estimates generated by three real-time PCR assays applied on genital specimens. *Cancer Epidemiol Biomarkers Prev* **14:**2200-7.

Franco E, Villa L, Rohan T, *et al.* (1999a). Design and methods of the Ludwig-McGill longitudinal study of the natural history of human papillomavirus infection and cervical neoplasia in Brazil. Ludwig-McGill Study Group. *Rev Panam Salud Publica* **6:**223-33.

Franco EL, Villa LL, Sobrinho JP, *et al.* (1999b). Epidemiology of acquisition and clearance of cervical human papillomavirus infection in women from a high-risk area for cervical cancer. *J Infect Dis* **180:**1415-23.

Franco EL, Ferenczy A (2002). Cervix. In: Cancer precursors: epidemiology, detection, and prevention. Franco EL, Rohan TE, eds. Springer, New York.

FUTURE II study group (2007a). Quadrivalent vaccine against human papillomavirus to prevent high-grade cervical lesions. *N Engl J Med* **356:**1915-27.

FUTURE II study group (2007b). Effect of prophylactic human papillomavirus L1 virus-like-particle vaccine on risk of cervical intraepithelial neoplasia grade 2, grade 3, and adenocarcinoma in situ: a combined analysis of four randomised clinical trials. *Lancet* **369:**1861-8.

Garland SM, Hernandez-Avila M, Wheeler CM, *et al.* (2007). Quadrivalent vaccine against human papillomavirus to prevent anogenital diseases. *N Engl J Med* **356:**1928-43.

Giuliano AR, Sedjo RL, Roe DJ, *et al.* (2002a). Clearance of oncogenic human papillomavirus (HPV) infection: effect of smoking (United States). *Cancer Causes Control* **13:**839-46.

Giuliano AR, Harris R, Sedjo RL, *et al.* (2002b). Incidence, prevalence, and clearance of type-specific human papillomavirus infections: The Young Women's Health Study. *J Infect Dis* **186:**462-9.

Giuliano AR, Siegel EM, Roe DJ, *et al.* (2003). Dietary intake and risk of persistent human papillomavirus (HPV) infection: the Ludwig-McGill HPV Natural History Study. *J Infect Dis* **188:**1508-16.

Giuliano AR, Papenfuss M, De Galaz EM, *et al.* (2004). Risk factors for squamous intraepithelial lesions (SIL) of the cervix among women residing at the US-Mexico border. *Int J Cancer* **109:**112-8.

Goetghebeur E and Ryan L (2000). Semiparametric regression analysis of interval-censored data. *Biometrics* **56:**1139-44.

Goggins WB, Finkelstein DM, Schoenfeld DA and Zaslavsky AM (1998). A Markov chain Monte Carlo EM algorithm for analyzing interval-censored data under the Cox proportional hazards model. *Biometrics* **54:**1498-507.

Gram IT, Austin H and Stalsberg H (1992). Cigarette smoking and the incidence of cervical intraepithelial neoplasia, grade III, and cancer of the cervix uteri. *Am J Epidemiol* **135:**341-6.

Gram IT, Braaten T, Terry PD, *et al.* (2005). Breast cancer risk among women who start smoking as teenagers. *Cancer Epidemiol Biomarkers Prev* **14:**61-6.

Gravitt PE, Burk RD, Lorincz A, *et al.* (2003). A comparison between real-time polymerase chain reaction and hybrid capture 2 for human papillomavirus DNA quantitation. *Cancer Epidemiol Biomarkers Prev* **12:**477-84.

Gravitt PE, van Doorn LJ, Quint W, *et al.* (2007). Human papillomavirus (HPV) genotyping using paired exfoliated cervicovaginal cells and paraffin-embedded

tissues to highlight difficulties in attributing HPV types to specific lesions. *J Clin Microbiol* **45:**3245-50.

Green J, Berrington de Gonzalez A, Sweetland S, *et al.* (2003). Risk factors for adenocarcinoma and squamous cell carcinoma of the cervix in women aged 20-44 years: the UK National Case-Control Study of Cervical Cancer. *Br J Cancer* **89:**2078-86.

Hagensee ME, Koutsky LA, Lee SK, *et al.* (2000). Detection of cervical antibodies to human papillomavirus type 16 (HPV-16) capsid antigens in relation to detection of HPV-16 DNA and cervical lesions. *J Infect Dis* **181:**1234-9.

Hall S, Lorincz A, Shah F, *et al.* (1996). Human papillomavirus DNA detection in cervical specimens by hybrid capture: correlation with cytologic and histologic diagnoses of squamous intraepithelial lesions of the cervix. *Gynecol Oncol* **62:**353-9.

Harper DM, Franco EL, Wheeler C, *et al.* (2004). Efficacy of a bivalent L1 virus-like particle vaccine in prevention of infection with human papillomavirus types 16 and 18 in young women: a randomised controlled trial. *Lancet* **364:**1757-65.

Harris TG, Kulasingam SL, Kiviat NB, *et al.* (2004). Cigarette smoking, oncogenic human papillomavirus, Ki-67 antigen, and cervical intraepithelial neoplasia. *Am J Epidemiol* **159:**834-42.

Haverkos HW, Soon G, Steckley SL and Pickworth W (2003). Cigarette smoking and cervical cancer: Part I: a meta-analysis. *Biomed Pharmacother* **57:**67-77.

Heard I, Tassie JM, Schmitz V, *et al.* (2000). Increased risk of cervical disease among human immunodeficiency virus-infected women with severe immunosuppression and high human papillomavirus load(1). *Obstet Gynecol* **96:**403-9.

Hebner CM and Laimins LA (2006). Human papillomaviruses: basic mechanisms of pathogenesis and oncogenicity. *Rev Med Virol* **16:**83-97.

Herrero R, Schiffman MH, Bratti C, *et al.* (1997). Design and methods of a population-based natural history study of cervical neoplasia in a rural province of Costa Rica: the Guanacaste Project. *Rev Panam Salud Publica* **1:**362-75.

Hildesheim A, Schiffman MH, Gravitt PE, *et al.* (1994). Persistence of type-specific human papillomavirus infection among cytologically normal women. *J Infect Dis* **169:**235-40.

Hildesheim A, Schiffman M, Scott DR, *et al.* (1998). Human leukocyte antigen class I/II alleles and development of human papillomavirus-related cervical neoplasia: results from a case-control study conducted in the United States. *Cancer Epidemiol Biomarkers Prev* **7:**1035-41.

Hildesheim A, Herrero R, Castle PE, *et al.* (2001). HPV co-factors related to the development of cervical cancer: results from a population-based study in Costa Rica. *Br J Cancer* **84:**1219-26.

Ho GY, Palan PR, Basu J, *et al.* (1998a). Viral characteristics of human papillomavirus infection and antioxidant levels as risk factors for cervical dysplasia. *Int J Cancer* **78:**594-9.
Ho GY, Bierman R, Beardsley L, Chang CJ and Burk RD (1998b). Natural history of cervicovaginal papillomavirus infection in young women. *N Engl J Med* **338:**423-8.

Ho GY, Kadish AS, Burk RD, *et al.* (1998c). HPV 16 and cigarette smoking as risk factors for high-grade cervical intra-epithelial neoplasia. *Int J Cancer* **78:**281-5.

Ho GY, Studentsov Y, Hall CB, *et al.* (2002). Risk factors for subsequent cervicovaginal human papillomavirus (HPV) infection and the protective role of antibodies to HPV-16 virus-like particles. *J Infect Dis* **186:**737-42.

Ho GY, Studentsov YY, Bierman R and Burk RD (2004). Natural history of human papillomavirus type 16 virus-like particle antibodies in young women. *Cancer Epidemiol Biomarkers Prev* **13:**110-6.

Ho CM, Chien TY, Huang SH, Lee BH and Chang SF (2006). Integrated human papillomavirus types 52 and 58 are infrequently found in cervical cancer, and high viral loads predict risk of cervical cancer. *Gynecol Oncol* **102:**54-60.

Hoffman MS, Cavanagh D (1996). Cervical cancer. In: Cancer screening. Reintgen DS, Clark RA eds. Mosby, St Louis, 1996.

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

Hougaard (2000). Analysis of multivariate survival data. Springer-Verlag, New York.

Howley P, Lowy D (2001). Papillomaviruses and their replication. In: Knipe D, Howley P, Griffin D *et al.* eds. Fields virology: fourth edition. Lippincott Williams & Wilkins, New York.

Hubbard RA (2003). Human papillomavirus testing methods. *Arch Pathol Lab Med* **127:**940-5.

Iftner T and Villa LL (2003). Chapter 12: Human papillomavirus technologies. *J Natl Cancer Inst Monogr* **31:**80-8.

International Collaboration of Epidemiological Studies of Cervical Cancer. (2006a). Carcinoma of the cervix and tobacco smoking: collaborative reanalysis of individual data on 13,541 women with carcinoma of the cervix and 23,017 women without carcinoma of the cervix from 23 epidemiological studies. *Int J Cancer* **118:**1481-95.

International Collaboration of Epidemiological Studies of Cervical Cancer (2006b). Comparison of risk factors for invasive squamous cell carcinoma and adenocarcinoma of the cervix: collaborative reanalysis of individual data on 8,097 women with squamous cell carcinoma and 1,374 women with adenocarcinoma from 12 epidemiological studies. *Int J Cancer* **120:**885-91.

International Collaboration of Epidemiological Studies of Cervical Cancer (2006c). Cervical carcinoma and reproductive factors: collaborative reanalysis of individual data on 16,563 women with cervical carcinoma and 33,542 women without cervical carcinoma from 25 epidemiological studies. *Int J Cancer* **119:**1108-24.

International Collaboration of Epidemiological Studies of Cervical Cancer. (2007a). Cervical cancer and hormonal contraceptives: collaborative reanalysis of individual data for 16,573 women with cervical cancer and 35,509 women without cervical cancer from 24 epidemiological studies. *Lancet* **370:**1609-21.

Jacobs MV, Walboomers JM, Snijders PJ, *et al.* (2000). Distribution of 37 mucosotropic HPV types in women with cytologically normal cervical smears: the age-related patterns for high-risk and low-risk types. *Int J Cancer* **87:**221-7.

Jacobson DL, Peralta L, Farmer M, *et al.* (1999). Cervical ectopy and the transformation zone measured by computerized planimetry in adolescents. *Int J Gynaecol Obstet* **66:**7-17.

Jeon S, Allen-Hoffmann BL and Lambert PF (1995). Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells. *J Virol* **69:**2989-97.

Joly P, Commenges D and Letenneur L (1998). A penalized likelihood approach for arbitrarily censored and truncated data: application to age-specific incidence of dementia. *Biometrics* **54:**185-94.

Joly P, Letenneur L, Alioum A and Commenges D (1999). PHMPL: a computer program for hazard estimation using a penalized likelihood method with interval-censored and left-truncated data. *Comput Methods Programs Biomed* **60:**225-31.

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

Josefsson AM, Magnusson PK, Ylitalo N, *et al.* (2000). Viral load of human papilloma virus 16 as a determinant for development of cervical carcinoma in situ: a nested case-control study. *Lancet* **355:**2189-93.

Kahn JA, Rosenthal SL, Succop PA, Ho GY and Burk RD (2002a). The interval between menarche and age of first sexual intercourse as a risk factor for subsequent HPV infection in adolescent and young adult women. *J Pediatr* **141:**718-23.

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

Kalbfleisch JD, Prentice RL (2002). The statistical analysis of failure time data: second edition. John Wiley & Sons, New York.

Kawana K, Yoshikawa H, Taketani Y, Yoshiike K and Kanda T (1998). In vitro construction of pseudovirions of human papillomavirus type 16: incorporation of plasmid DNA into reassembled L1/L2 capsids. *J Virol* **72:**10298-300.

Khan MJ, Castle PE, Lorincz AT, *et al.* (2005). The elevated 10-year risk of cervical precancer and cancer in women with human papillomavirus (HPV) type 16 or 18 and the possible utility of type-specific HPV testing in clinical practice. *J Natl Cancer Inst* **97:**1072-9.

Kim MY, De Gruttola VG and Lagakos SW (1993). Analyzing doubly censored data with covariates, with application to AIDS. *Biometrics* **49:**13-22.

Kim MY and Xue X (2002). The analysis of multivariate interval-censored survival data. *Stat Med* **21:**3715-26.

Kiviat NB, Critchlow CW, Kurman RJ (1992). Reassessment of the morphological continuum of cervical intraepithelial lesions: does it reflect different stages in the progression to cervical carcinoma? In: Munoz N, Bosch FX, Shah KV, Meheus A eds. The epidemiology of human papillomavirus and cervical cancer. IARC, Lyon.

Kjaer SK, van den Brule AJ, Bock JE, *et al.* (1996). Human papillomavirus--the most significant risk determinant of cervical intraepithelial neoplasia. *Int J Cancer* **65:**601-6.

Kjaer SK, van den Brule AJ, Bock JE, *et al.* (1997). Determinants for genital human papillomavirus (HPV) infection in 1000 randomly chosen young Danish women with normal Pap smear: are there different risk profiles for oncogenic and nononcogenic HPV types? *Cancer Epidemiol Biomarkers Prev* **6:**799-805.

Kjaer SK, van den Brule AJ, Svare EI, *et al.* (1998). Different risk factor patterns for high-grade and low-grade intraepithelial lesions on the cervix among HPV-positive and HPV-negative young women. *Int J Cancer* **76:**613-9.

Kjaer SK, Chackerian B, van den Brule AJ, *et al.* (2001). High-risk human papillomavirus is sexually transmitted: evidence from a follow-up study of virgins starting sexual activity (intercourse). *Cancer Epidemiol Biomarkers Prev* **10:**101-6.

Kjaer SK, van den Brule AJ, Paull G, *et al.* (2002). Type specific persistence of high risk human papillomavirus (HPV) as indicator of high grade cervical squamous intraepithelial lesions in young women: population based prospective follow up study. *Bmj* **325:**572.

Klein JP, Moeschberger ML (1997). Survival analysis: techniques for censored and truncated data. Springer-Verlag, New York.

Koch A, Hansen SV, Nielsen NM, Palefsky J and Melbye M (1997). HPV detection in children prior to sexual debut. *Int J Cancer* **73:**621-4.

Kooperberg C, Stone CJ, Truong YK (1995). Hazard Regression. *Journal of the American Statistical Association* **90:**78-94.

Kooperberg C and Clarkson DB (1997). Hazard regression with interval-censored data. *Biometrics* **53:**1485-94.

Koshiol J, Schroeder J, Jamieson DJ, *et al.* (2006). Smoking and time to clearance of human papillomavirus infection in HIV-seropositive and HIV-seronegative women. *Am J Epidemiol* **164:**176-83.

Kotloff KL, Wasserman SS, Russ K, *et al.* (1998). Detection of genital human papillomavirus and associated cytological abnormalities among college women. *Sex Transm Dis* **25:**243-50.

Koutsky LA, Ault KA, Wheeler CM, *et al.* (2002). A controlled trial of a human papillomavirus type 16 vaccine. *N Engl J Med* **347:**1645-51.

Koutsky LA, Holmes KK, Critchlow CW, *et al.* (1992). A cohort study of the risk of cervical intraepithelial neoplasia grade 2 or 3 in relation to papillomavirus infection. *N Engl J Med* **327:**1272-8.

Kumar G, Howell R (2000). Cervical glandular neoplasia. *The Obstetrician and Gynaecologist* **2:**43-5.

La Vecchia C, Franceschi S, Decarli A, *et al.* (1986). Cigarette smoking and the risk of cervical neoplasia. *Am J Epidemiol* **123:**22-9.

Lai HC, Peng MY, Nieh S, *et al.* (2006). Differential viral loads of human papillomavirus 16 and 58 infections in the spectrum of cervical carcinogenesis. *Int J Gynecol Cancer* **16:**730-5.

Landis JR and Koch GG (1977). The measurement of observer agreement for categorical data. *Biometrics* **33:**159-74.

Law CG and Brookmeyer R (1992). Effects of mid-point imputation on the analysis of doubly censored data. *Stat Med* **11:**1569-78.

Lawless JF (2003). Statistical models and methods for lifetime data: second edition. Wiley, New Jersey.

Leffondre K, Abrahamowicz M, Siemiatycki J and Rachet B (2002). Modeling smoking history: a comparison of different approaches. *Am J Epidemiol* **156:**813-23.

Levi F, Lucchini F, Negri E, Franceschi S and la Vecchia C (2000). Cervical cancer mortality in young women in Europe: patterns and trends. *Eur J Cancer* **36:**2266-71.

Levi JE, Fernandes S, Tateno AF, *et al.* (2004). Presence of multiple human papillomavirus types in cervical samples from HIV-infected women. *Gynecol Oncol* **92:**225-31.

Ley C, Bauer HM, Reingold A, *et al.* (1991). Determinants of genital human papillomavirus infection in young women. *J Natl Cancer Inst* **83:**997-1003.

Liaw KL, Hildesheim A, Burk RD, *et al.* (2001). A prospective study of human papillomavirus (HPV) type 16 DNA detection by polymerase chain reaction and its association with acquisition and persistence of other HPV types. *J Infect Dis* **183:**8-15.

Lillo FB, Lodini S, Ferrari D, *et al.* (2005). Determination of human papillomavirus (HPV) load and type in high-grade cervical lesions surgically resected from HIV-infected women during follow-up of HPV infection. *Clin Infect Dis* **40:**451-7.

Lindsey JC and Ryan LM (1998). Tutorial in biostatistics methods for interval-censored data. *Stat Med* **17:**219-38.

Little RJA, Rubin DB (2002). Statistical analysis with missing data: second edition. John Wiley and Sons, New York.

Lorincz AT, Castle PE, Sherman ME, *et al.* (2002). Viral load of human papillomavirus and risk of CIN3 or cervical cancer. *Lancet* **360:**228-9.

Lydyard P, Whelan A, Fanger MW (2004). Instant notes Immunology: second edition. BIOS Scientific publishers, Abingdon.

Lyon JL, Gardner JW, West DW, Stanish WM and Hebertson RM (1983). Smoking and carcinoma in situ of the uterine cervix. *Am J Public Health* **73:**558-62.

Maciag PC, Schlecht NF, Souza PS, *et al.* (2002). Polymorphisms of the human leukocyte antigen DRB1 and DQB1 genes and the natural history of human papillomavirus infection. *J Infect Dis* **186:**164-72.

Mahdavi A and Monk BJ (2005). Vaccines against human papillomavirus and cervical cancer: promises and challenges. *Oncologist* **10:**528-38.

Mao C, Koutsky LA, Ault KA, *et al.* (2006). Efficacy of human papillomavirus-16 vaccine to prevent cervical intraepithelial neoplasia: a randomized controlled trial. *Obstet Gynecol* **107:**18-27.

Mayrand MH, Duarte-Franco E, Rodrigues I, *et al.* (2007). Human papillomavirus DNA versus Papanicolaou screening tests for cervical cancer. *N Engl J Med* **357:**1579-88.

McIntyre-Seltman K, Castle PE, Guido R, Schiffman M and Wheeler CM (2005). Smoking is a risk factor for cervical intraepithelial neoplasia grade 3 among oncogenic human papillomavirus DNA-positive women with equivocal or mildly abnormal cytology. *Cancer Epidemiol Biomarkers Prev* **14:**1165-70.

Miller AB (1999). Cervix cancer. In: Cancer screening: theory and practice. Kramer BS, Gohagan JK, Prorok PC, eds. Marcel-Dekker, New York.
Minkoff H, Feldman JG, Strickler HD, *et al.* (2004). Relationship between smoking and human papillomavirus infections in HIV-infected and -uninfected women. *J Infect Dis* **189:**1821-8.

Moberg M, Gustavsson I and Gyllensten U (2003). Real-time PCR-based system for simultaneous quantification of human papillomavirus types associated with high risk of cervical cancer. *J Clin Microbiol* **41:**3221-8.

Moberg M, Gustavsson I and Gyllensten U (2004). Type-specific associations of human papillomavirus load with risk of developing cervical carcinoma in situ. *Int J Cancer* **112:**854-9.

Moberg M, Gustavsson I, Wilander E and Gyllensten U (2005). High viral loads of human papillomavirus predict risk of invasive cervical carcinoma. *Br J Cancer* **92:**891-4.

Molano M, Van den Brule A, Plummer M, *et al.* (2003). Determinants of clearance of human papillomavirus infections in Colombian women with normal cytology: a population-based, 5-year follow-up study. *Am J Epidemiol* **158:**486-94.

Monnier-Benoit S, Dalstein V, Riethmuller D, *et al.* (2006). Dynamics of HPV16 DNA load reflect the natural history of cervical HPV-associated lesions. *J Clin Virol* **35:**270-7.

Morrison EA, Ho GY, Beardsley L, Bierman R and Burk RD (1998). Development of a multi-item scale to quantitatively assess sexual behaviors and the transmission of high- and low-risk human papillomaviruses. *Sex Transm Dis* **25:**509-15.

Moscicki AB, Shiboski S, Broering J, *et al.* (1998). The natural history of human papillomavirus infection as measured by repeated DNA testing in adolescent and young women. *J Pediatr* **132:**277-84.

Moscicki AB, Burt VG, Kanowitz S, Darragh T and Shiboski S (1999). The significance of squamous metaplasia in the development of low grade squamous intraepithelial lesions in young women. *Cancer* **85:**1139-44.

Moscicki AB, Hills N, Shiboski S, *et al.* (2001). Risks for incident human papillomavirus infection and low-grade squamous intraepithelial lesion development in young females. *Jama* **285:**2995-3002.

Moscicki AB, Shiboski S, Hills NK, *et al.* (2004). Regression of low-grade squamous intra-epithelial lesions in young women. *Lancet* **364:**1678-83.

Munoz N, Bosch FX (1992). HPV and cervical neoplasia: review of case-control and cohort studies. In: Munoz N, Bosch FX, Shah KV, Meheus A. The epidemiology of human papillomavirus and cervical cancer. IARC, Lyon.

Munoz N, Bosch FX, de Sanjose S, *et al.* (1993). Risk factors for cervical intraepithelial neoplasia grade III/carcinoma in situ in Spain and Colombia. *Cancer Epidemiol Biomarkers Prev* **2:**423-31.
Munoz N (2000). Human papillomavirus and cancer: the epidemiological evidence. *J Clin Virol* **19:**1-5.

Munoz N, Mendez F, Posso H, *et al.* (2004). Incidence, duration, and determinants of cervical human papillomavirus infection in a cohort of Colombian women with normal cytological results. *J Infect Dis* **190:**2077-87.

Nadais Rda F, Campaner AB, Piato S, *et al.* (2006). Langerhans' cells and smoking in intraepithelial neoplasia of the cervix. *Gynecol Oncol* **102:**356-60.

Nakagawa M, Stites DP, Patel S, *et al.* (2000). Persistence of human papillomavirus type 16 infection is associated with lack of cytotoxic T lymphocyte response to the E6 antigens. *J Infect Dis* **182:**595-8.

Nakagawa M, Viscidi R, Deshmukh I, *et al.* (2002). Time course of humoral and cell-mediated immune responses to human papillomavirus type 16 in infected women. *Clin Diagn Lab Immunol* **9:**877-82.

Nakazato H, Koike A, Saji S, Ogawa N and Sakamoto J (1994). Efficacy of immunochemotherapy as adjuvant treatment after curative resection of gastric cancer. Study Group of Immunochemotherapy with PSK for Gastric Cancer. *Lancet* **343:**1122-6.

Nanda K, McCrory DC, Myers ER, *et al.* (2000). Accuracy of the Papanicolaou test in screening for and follow-up of cervical cytologic abnormalities: a systematic review. *Ann Intern Med* **132:**810-9.

Naucler P, Ryd W, Tornberg S, *et al.* (2007). Human papillomavirus and Papanicolaou tests to screen for cervical cancer. *N Engl J Med* **357:**1589-97.

NHSCSP website. Website of the National Health Service cervical screening programme. http://www.cancerscreening.nhs.uk/cervical/.

Nindl I, Greinke C, Zahm DM, *et al.* (1997). Human papillomavirus distribution in cervical tissues of different morphology as determined by hybrid capture assay and PCR. *Int J Gynecol Pathol* **16:**197-204.

Nobbenhuis MA, Walboomers JM, Helmerhorst TJ, *et al.* (1999). Relation of human papillomavirus status to cervical lesions and consequences for cervical-cancer screening: a prospective study. *Lancet* **354:**20-5.

Odell PM, Anderson KM and D'Agostino RB (1992). Maximum likelihood estimation for interval-censored data using a Weibull-based accelerated failure time model. *Biometrics* **48:**951-9.

Olsen AO, Dillner J, Skrondal A and Magnus P (1998). Combined effect of smoking and human papillomavirus type 16 infection in cervical carcinogenesis. *Epidemiology* **9:**346-9.

Olsen AO, Gjoen K, Sauer T, *et al.* (1995). Human papillomavirus and cervical intraepithelial neoplasia grade II-III: a population-based case-control study. *Int J Cancer* **61:**312-5.
Onda T, Carter JJ, Koutsky LA, *et al.* (2003). Characterization of IgA response among women with incident HPV 16 infection. *Virology* **312:**213-21.

Paavonen J, Jenkins D, Bosch FX, *et al.* (2007). Efficacy of a prophylactic adjuvanted bivalent L1 virus-like-particle vaccine against infection with human papillomavirus types 16 and 18 in young women: an interim analysis of a phase III double-blind, randomised controlled trial. *Lancet* **369:**2161-70.

Pan W (2000). A multiple imputation approach to Cox regression with interval-censored data. *Biometrics* **56:**199-203.

Parazzini F, La Vecchia C, Negri E, *et al.* (1992). Risk factors for cervical intraepithelial neoplasia. *Cancer* **69:**2276-82.

Parkin DM, Bray F, Ferlay J and Pisani P (2005). Global cancer statistics, 2002. *CA Cancer J Clin* **55:**74-108.

Parkin DM, Bray F (2006). Chapter 2: The burden of HPV-related cancers. *Vaccine* **24S3:**S3/11-S3/25.

Pastrana DV, Buck CB, Pang YY, *et al.* (2004). Reactivity of human sera in a sensitive, high-throughput pseudovirus-based papillomavirus neutralization assay for HPV16 and HPV18. *Virology* **321:**205-16.

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

Peto R (1973). Experimental survival curves for interval-censored data. *Applied Statistics* **22:**86-91.

Peto J, Gilham C, Fletcher O and Matthews FE (2004). The cervical cancer epidemic that screening has prevented in the UK. *Lancet* **364:**249-56.

Petroni GR and Wolfe RA (1994). A two-sample test for stochastic ordering with interval-censored data. *Biometrics* **50:**77-87.

Pinheiro JC, Bates DM (2002). Mixed-effects models in S and S-plus. Springer-Verlag, New York.

Plummer M, Herrero R, Franceschi S, *et al.* (2003). Smoking and cervical cancer: pooled analysis of the IARC multi-centric case-control study. *Cancer Causes Control* **14:**805-14.

Quinn M, Babb P, Jones J and Allen E (1999). Effect of screening on incidence of and mortality from cancer of cervix in England: evaluation based on routinely collected statistics. *BMJ* **318:**904-8.

Quinn M, Babb P, Brock A, Kirby L, Jones J (2001). Cancer trends in England and Wales 1950-1999. The Stationery Office, London.

Rapley R (2004). Molecular biology and basic techniques. In: Keith Wilson, John Walker (eds). Principles and techniques of practical biochemistry: fifth edition. Cambridge University Press, Cambridge.

Remmink AJ, Walboomers JM, Helmerhorst TJ, *et al.* (1995). The presence of persistent high-risk HPV genotypes in dysplastic cervical lesions is associated with progressive disease: natural history up to 36 months. *Int J Cancer* **61:**306-11.

Richardson H, Franco E, Pintos J, *et al.* (2000). Determinants of low-risk and high-risk cervical human papillomavirus infections in Montreal University students. *Sex Transm Dis* **27:**79-86.

Richardson H, Abrahamowicz M, Tellier PP, *et al.* (2005). Modifiable risk factors associated with clearance of type-specific cervical human papillomavirus infections in a cohort of university students. *Cancer Epidemiol Biomarkers Prev* **14:**1149-56.

Roden RB, Greenstone HL, Kirnbauer R, *et al.* (1996). In vitro generation and type-specific neutralization of a human papillomavirus type 16 virion pseudotype. *J Virol* **70:**5875-83.

Ronco G, Giorgi-Rossi P, Carozzi F, *et al.* (2008). Results at recruitment from a randomized controlled trial comparing human papillomavirus testing alone with conventional cytology as the primary cervical cancer screening test. *J Natl Cancer Inst* **100:**492-501.

Rossi JL, Gissmann L, Jansen K and Muller M (2000). Assembly of human papillomavirus type 16 pseudovirions in Saccharomyces cerevisiae. *Hum Gene Ther* **11:**1165-76.

Rousseau MC, Franco EL, Villa LL, *et al.* (2000). A cumulative case-control study of risk factor profiles for oncogenic and nononcogenic cervical human papillomavirus infections. *Cancer Epidemiol Biomarkers Prev* **9:**469-76.

Rousseau MC, Pereira JS, Prado JC, *et al.* (2001). Cervical coinfection with human papillomavirus (HPV) types as a predictor of acquisition and persistence of HPV infection. *J Infect Dis* **184:**1508-17.

Rousseau MC, Abrahamowicz M, Villa LL, *et al.* (2003a). Predictors of cervical coinfection with multiple human papillomavirus types. *Cancer Epidemiol Biomarkers Prev* **12:**1029-37.

Rousseau MC, Villa LL, Costa MC, *et al.* (2003b). Occurrence of cervical infection with multiple human papillomavirus types is associated with age and cytologic abnormalities. *Sex Transm Dis* **30:**581-7.

Rousson V, Gasser T and Seifert B (2002). Assessing intrarater, interrater and test-retest reliability of continuous measurements. *Stat Med* **21:**3431-46.
Royston P, Altman DG and Sauerbrei W (2006). Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med* **25:**127-41.

Rozendaal L, Walboomers JM, van der Linden JC, *et al.* (1996). PCR-based high-risk HPV test in cervical cancer screening gives objective risk assessment of women with cytomorphologically normal cervical smears. *Int J Cancer* **68:**766-9.

Rozendaal L, Westerga J, van der Linden JC, *et al.* (2000). PCR based high risk HPV testing is superior to neural network based screening for predicting incident CIN III in women with normal cytology and borderline changes. *J Clin Pathol* **53:**606-11.

Sakamoto J, Teramukai S, Nakazato H and Ohashi Y (1997). A re-analysis of a randomized clinical trial for gastric cancer using interval censoring. *Jpn J Clin Oncol* **27:**445-6.

Samuelsen SO and Kongerud J (1994). Interval censoring in longitudinal data of respiratory symptoms in aluminium potroom workers: a comparison of methods. *Stat Med* **13:**1771-80.

Sasagawa T, Dong Y, Saijoh K, *et al.* (1997). Human papillomavirus infection and risk determinants for squamous intraepithelial lesion and cervical cancer in Japan. *Jpn J Cancer Res* **88:**376-84.

Sawaya GF and Smith-McCune K (2007). HPV vaccination--more answers, more questions. *N Engl J Med* **356:**1991-3.

Schiff M, Becker TM, Masuk M, *et al.* (2000). Risk factors for cervical intraepithelial neoplasia in southwestern American Indian women. *Am J Epidemiol* **152:**716-26.

Schiffman MH, Bauer HM, Hoover RN, *et al.* (1993). Epidemiologic evidence showing that human papillomavirus infection causes most cervical intraepithelial neoplasia. *J Natl Cancer Inst* **85:**958-64.

Schiffman M and Kjaer SK (2003). Chapter 2: Natural history of anogenital human papillomavirus infection and neoplasia. *J Natl Cancer Inst Monogr*14-9.

Schlecht NF, Kulaga S, Robitaille J, *et al.* (2001). Persistent human papillomavirus infection as a predictor of cervical intraepithelial neoplasia. *Jama* **286:**3106-14.

Schlecht NF, Trevisan A, Duarte-Franco E, *et al.* (2003a). Viral load as a predictor of the risk of cervical intraepithelial neoplasia. *Int J Cancer* **103:**519-24.

Schlecht NF, Platt RW, Negassa A, *et al.* (2003b). Modeling the time dependence of the association between human papillomavirus infection and cervical cancer precursor lesions. *Am J Epidemiol* **158:**878-86.

Schlecht NF, Platt RW, Duarte-Franco E, *et al.* (2003c). Human papillomavirus infection and time to progression and regression of cervical intraepithelial neoplasia. *J Natl Cancer Inst* **95:**1336-43.

Schneider A, Oltersdorf T, Schneider V and Gissmann L (1987). Distribution pattern of human papilloma virus 16 genome in cervical neoplasia by molecular in situ hybridization of tissue sections. *Int J Cancer* **39:**717-21.

Scott M, Stites DP and Moscicki AB (1999). Th1 cytokine patterns in cervical human papillomavirus infection. *Clin Diagn Lab Immunol* **6:**751-5.

Seaman SR and Bird SM (2001). Proportional hazards model for interval-censored failure times and time-dependent covariates: application to hazard of HIV infection of injecting drug users in prison. *Stat Med* **20:**1855-70.

Sellors JW, Karwalajtys TL, Kaczorowski J, *et al.* (2003). Incidence, clearance and predictors of human papillomavirus infection in women. *Cmaj* **168:**421-5.

Severson J, Evans TY, Lee P, *et al.* (2001). Human papillomavirus infections: epidemiology, pathogenesis, and therapy. *J Cutan Med Surg* **5:**43-60.

Sherman ME, Wang SS, Wheeler CM, *et al.* (2003a). Determinants of human papillomavirus load among women with histological cervical intraepithelial neoplasia 3: dominant impact of surrounding low-grade lesions. *Cancer Epidemiol Biomarkers Prev* **12:**1038-44.

Sherman ME, Wang SS, Tarone R, Rich L and Schiffman M (2003b). Histopathologic extent of cervical intraepithelial neoplasia 3 lesions in the atypical squamous cells of undetermined significance low-grade squamous intraepithelial lesion triage study: implications for subject safety and lead-time bias. *Cancer Epidemiol Biomarkers Prev* **12:**372-9.

Sherman ME, Lorincz AT, Scott DR, *et al.* (2003c). Baseline cytology, human papillomavirus testing, and risk for cervical neoplasia: a 10-year cohort analysis. *J Natl Cancer Inst* **95:**46-52.

Shew ML, Fortenberry JD, Miles P and Amortegui AJ (1994). Interval between menarche and first sexual intercourse, related to risk of human papillomavirus infection. *J Pediatr* **125:**661-6.

Singer A (1989) Cancer of the cervix: diagnosis and treatment. WB Saunders, London.

Singer JD, Willett JB (2003). Applied longitudinal data analysis: modelling change and event occurrence. Oxford University Press, New York.

Silverberg MJ, Schneider MF, Silver B, *et al.* (2006). Serological detection of human papillomavirus type 16 infection in human immunodeficiency virus (HIV)-positive and high-risk HIV-negative women. *Clin Vaccine Immunol* **13:**511-9.

Smith PJ, Thompson TJ and Jereb JA (1997). A model for interval-censored tuberculosis outbreak data. *Stat Med* **16:**485-96.

Snijders PJ, Hogewoning CJ, Hesselink AT, *et al.* (2006). Determination of viral load thresholds in cervical scrapings to rule out CIN 3 in HPV 16, 18, 31 and 33-positive women with normal cytology. *Int J Cancer* **119:**1102-7.

Song SH, Lee JK, Oh MJ, *et al.* (2006). Persistent HPV infection after conization in patients with negative margins. *Gynecol Oncol* **101:**418-22.

Souhami R, Tobias J (2005). Cancer and its management: fifth edition. Blackwell, Oxford.

Spartz H, Lehr E, Zhang B, Roman A and Brown DR (2005). Progression from productive infection to integration and oncogenic transformation in human papillomavirus type 59-immortalized foreskin keratinocytes. *Virology* **336:**11-25.

Spink KM and Laimins LA (2005). Induction of the human papillomavirus type 31 late promoter requires differentiation but not DNA amplification. *J Virol* **79:**4918-26.

Stanley M (2006). Immune responses to human papillomavirus. *Vaccine* **24 Suppl 1:**S16-22.

Stauffer Y, Raj K, Masternak K and Beard P (1998). Infectious human papillomavirus type 18 pseudovirions. *J Mol Biol* **283:**529-36.

Steele J, Collins S, Wen K, *et al.* (2008). Measurement of the humoral immune response following an incident human papillomavirus type 16 or 18 infection in young women by a pseudovirion-based neutralizing antibody assay. *Clin Vaccine Immunol* **15:**1387-90.

Sun CA, Lai HC, Chang CC, *et al.* (2001). The significance of human papillomavirus viral load in prediction of histologic severity and size of squamous intraepithelial lesions of uterine cervix. *Gynecol Oncol* **83:**95-9.

Sun J (2006). The statistical analysis of interval-censored failure time data. Springer, New York.

Swan DC, Tucker RA, Tortolero-Luna G, *et al.* (1999). Human papillomavirus (HPV) DNA copy number is dependent on grade of cervical disease and HPV type. *J Clin Microbiol* **37:**1030-4.

Syrjanen S, Shabalova I, Petrovichev N, *et al.* (2005a). Factors predicting persistence of high-risk human papillomavirus (HPV) infections in women prospectively followed-up in three New Independent States (NIS) of the former Soviet Union. *Eur J Gynaecol Oncol* **26:**491-8.

Syrjanen S, Shabalova IP, Petrovichev N, *et al.* (2005b). Clearance of high-risk human papillomavirus (HPV) DNA and PAP smear abnormalities in a cohort of women subjected to HPV screening in the New Independent States of the former Soviet Union (the NIS cohort study). *Eur J Obstet Gynecol Reprod Biol* **119:**219-27.

Tanaka H, Sato H, Sato N, *et al.* (2004). Adding HPV16 testing to abnormal cervical smear detection is useful for predicting CIN3: a prospective study. *Acta Obstet Gynecol Scand* **83:**497-500.

Therneau TM, Grambsch PM (2002). Modeling survival data: extending the Cox model. Springer, New York.

Thomas KK, Hughes JP, Kuypers JM, *et al.* (2000). Concurrent and sequential acquisition of different genital human papillomavirus types. *J Infect Dis* **182:**1097-102.

Touze A and Coursaget P (1998). In vitro gene transfer using human papillomavirus-like particles. *Nucleic Acids Res* **26:**1317-23.

Turnbull BW (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. *J R Statistic Soc B* **38:**290-5.

Unckell F, Streeck RE and Sapp M (1997). Generation and neutralization of pseudovirions of human papillomavirus type 33. *J Virol* **71:**2934-9.

Ung A, Kramer TR, Schiffman M, *et al.* (1999). Soluble interleukin 2 receptor levels and cervical neoplasia: results from a population-based case-control study in Costa Rica. *Cancer Epidemiol Biomarkers Prev* **8:**249-53.

Vaccarella S, Herrero R, Dai M, *et al.* (2006a). Reproductive factors, oral contraceptive use, and human papillomavirus infection: pooled analysis of the IARC HPV prevalence surveys. *Cancer Epidemiol Biomarkers Prev* **15:**2148-53.

Vaccarella S, Franceschi S, Herrero R, *et al.* (2006b). Sexual behavior, condom use, and human papillomavirus: pooled analysis of the IARC human papillomavirus prevalence surveys. *Cancer Epidemiol Biomarkers Prev* **15:**326-33.

van Duin M, Snijders PJ, Schrijnemakers HF, *et al.* (2002). Human papillomavirus 16 load in normal and abnormal cervical scrapes: an indicator of CIN II/III and viral clearance. *Int J Cancer* **98:**590-5.

Van Tine BA, Kappes JC, Banerjee NS, *et al.* (2004). Clonal selection for transcriptionally active viral oncogenes during progression to cancer. *J Virol* **78:**11172-86.

Vernon SD, Reeves WC, Clancy KA, *et al.* (1994). A longitudinal study of human papillomavirus DNA detection in human immunodeficiency virus type 1-seropositive and -seronegative women. *J Infect Dis* **169:**1108-12.

Villa LL, Sichero L, Rahal P, *et al.* (2000). Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia. *J Gen Virol* **81:**2959-68.

Villa LL, Costa RL, Petta CA, *et al.* (2005). Prophylactic quadrivalent human papillomavirus (types 6, 11, 16, and 18) L1 virus-like particle vaccine in young women: a randomised double-blind placebo-controlled multicentre phase II efficacy trial. *Lancet Oncol* **6:**271-8.

Viscidi RP, Schiffman M, Hildesheim A, *et al.* (2004). Seroreactivity to human papillomavirus (HPV) types 16, 18, or 31 and risk of subsequent HPV infection: results from a population-based study in Costa Rica. *Cancer Epidemiol Biomarkers Prev* **13:**324-7.

Vizcaino AP, Moreno V, Bosch FX, *et al.* (1998). International trends in the incidence of cervical cancer: I. Adenocarcinoma and adenosquamous cell carcinomas. *Int J Cancer* **75:**536-45.

Vizcaino AP, Moreno V, Bosch FX, *et al.* (2000). International trends in incidence of cervical cancer: II. Squamous-cell carcinoma. *Int J Cancer* **86:**429-35.

Walboomers JMM, Jacobs MV, ManosMM, *et al.* (1999). Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathology* **189:**12-9.

Wallin KL, Wiklund F, Angstrom T, *et al.* (1999). Type-specific persistence of human papillomavirus DNA before the development of invasive cervical cancer. *N Engl J Med* **341:**1633-8.

Wang SS, Wheeler CM, Hildesheim A, *et al.* (2001). Human leukocyte antigen class I and II alleles and risk of cervical neoplasia: results from a population-based study in Costa Rica. *J Infect Dis* **184:**1310-4.

Wang SS, Schiffman M, Shields TS, *et al.* (2003). Seroprevalence of human papillomavirus-16, -18, -31, and -45 in a population-based cohort of 10000 women in Costa Rica. *Br J Cancer* **89:**1248-54.

Wang SS, Schiffman M, Herrero R, *et al.* (2004a). Determinants of human papillomavirus 16 serological conversion and persistence in a population-based cohort of 10 000 women in Costa Rica. *Br J Cancer* **91:**1269-74.

Wang SS, Trunk M, Schiffman M, *et al.* (2004b). Validation of p16INK4a as a marker of oncogenic human papillomavirus infection in cervical biopsies from a population-based cohort in Costa Rica. *Cancer Epidemiol Biomarkers Prev* **13:**1355-60.

Weissenborn SJ, Funke AM, Hellmich M, *et al.* (2003). Oncogenic human papillomavirus DNA loads in human immunodeficiency virus-positive women with high-grade cervical lesions are strongly elevated. *J Clin Microbiol* **41:**2763-7.

Wickenden C, Hanna N, Taylor-Robinson D, *et al.* (1988). Sexual transmission of human papillomaviruses in heterosexual and male homosexual couples, studied by DNA hybridisation. *Genitourin Med* **64:**34-8.

Wiencke JK, Thurston SW, Kelsey KT, *et al.* (1999). Early age at smoking initiation and tobacco carcinogen DNA damage in the lung. *J Natl Cancer Inst* **91:**614-9.

Wiencke JK and Kelsey KT (2002). Teen smoking, field cancerization, and a "critical period" hypothesis for lung cancer susceptibility. *Environ Health Perspect* **110:**555-8.

Wiley DJ, Wiesmeier E, Masongsong E, *et al.* (2006). Smokers at higher risk for undetected antibody for oncogenic human papillomavirus type 16 infection. *Cancer Epidemiol Biomarkers Prev* **15:**915-20.

Wilson R, Fehrmann F and Laimins LA (2005). Role of the E1--E4 protein in the differentiation-dependent life cycle of human papillomavirus type 31. *J Virol* **79:**6732-40.

Winer RL, Lee SK, Hughes JP, *et al.* (2003). Genital human papillomavirus infection: incidence and risk factors in a cohort of female university students. *Am J Epidemiol* **157:**218-26.

Winer RL, Kiviat NB, Hughes JP, *et al.* (2005). Development and duration of human papillomavirus lesions, after initial infection. *J Infect Dis* **191:**731-8.

Winer RL, Hughes JP, Feng Q, *et al.* (2006). Condom use and the risk of genital human papillomavirus infection in young women. *N Engl J Med* **354:**2645-54.

Winer RL, Feng Q, Hughes JP, *et al.* (2008). Risk of female human papillomavirus acquisition associated with first male sex partner. *J Infect Dis* **197:**279-82.

Woodman CB, Collins S, Winter H, *et al.* (2001). Natural history of cervical human papillomavirus infection in young women: a longitudinal cohort study. *Lancet* **357:**1831-6.

Woodman CB and Collins S (2002). A critique of cohort studies examining the role of human papillomavirus infection in cervical neoplasia. *Bjog* **109:**1311-8.

Woodman CB, Collins S, Rollason TP, *et al.* (2003). Human papillomavirus type 18 and rapidly progressing cervical intraepithelial neoplasia. *Lancet* **361:**40-3.

Woodman CB, Collins SI, Vavrusova N, *et al.* (2005). Role of sexual behavior in the acquisition of asymptomatic Epstein-Barr virus infection: a longitudinal study. *Pediatr Infect Dis J* **24:**498-502.

Woodman CB, Collins SI and Young LS (2007). The natural history of cervical HPV infection: unresolved issues. *Nat Rev Cancer* **7:**11-22.

Wright TC Jr. (2003). Chapter 10: Cervical cancer screening using visualization techniques. *J Natl Cancer Inst Monogr*66-71.

Xi LF, Demers GW, Koutsky LA, *et al.* (1995). Analysis of human papillomavirus type 16 variants indicates establishment of persistent infection. *J Infect Dis* **172:**747-55.

Xi LF, Koutsky LA, Galloway DA, *et al.* (1997). Genomic variation of human papillomavirus type 16 and risk for high grade cervical intraepithelial neoplasia. *J Natl Cancer Inst* **89:**796-802.

Xi LF, Carter JJ, Galloway DA, *et al.* (2002). Acquisition and natural history of human papillomavirus type 16 variant infection among a cohort of female university students. *Cancer Epidemiol Biomarkers Prev* **11:**343-51.

Ylitalo N, Sorensen P, Josefsson A, *et al.* (1999). Smoking and oral contraceptives as risk factors for cervical carcinoma in situ. *Int J Cancer* **81:**357-65.

Ylitalo N, Sorensen P, Josefsson AM, *et al.* (2000a). Consistent high viral load of human papillomavirus 16 and risk of cervical carcinoma in situ: a nested case-control study. *Lancet* **355:**2194-8.

Ylitalo N, Josefsson A, Melbye M, *et al.* (2000b). A prospective study showing long-term infection with human papillomavirus 16 before the development of cervical carcinoma in situ. *Cancer Res* **60:**6027-32.

Young LS, Bevan IS, Johnson MA, *et al.* (1989). The polymerase chain reaction: a new epidemiological tool for investigating cervical human papillomavirus infection. *Bmj* **298:**14-8.

Zerbini M, Venturoli S, Cricca M, *et al.* (2001). Distribution and viral load of type specific HPVs in different cervical lesions as detected by PCR-ELISA. *J Clin Pathol* **54:**377-80.

zur Hausen H (2002). Papillomaviruses and cancer: from basic studies to clinical application. *Nat Rev Cancer* **2:**342-50.