Human papillomavirus and biomarker analysis to predict high-grade cervical disease in women with persistent low-grade squamous cytological abnormalities

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Medicine at Cardiff University

By

Jonathan Andrew Lippiatt MRCOG MBBCh

2013

HPV Research Group
Institute of Cancer & Genetics
School of Medicine
Cardiff University

Summary

Persistent HPV infection can cause cervical intraepithelial neoplasia (CIN) and ultimately cervical cancer. Cervical cytology is currently used to screen for CIN with HPV testing recently emerging as an adjunct to cytology. Most women with HPV infection, however, will not go on to develop cancer. Therefore an additional biomarker is required to identify those women most at risk.

The BD SurePath Plus™ test (SPP) is a novel immunocytochemistry assay based upon the detection of minichromosome maintenance protein 2 (MCM2) and MCM7. Studies have shown both MCM2 and MCM7 have the potential for use as a biomarker for CIN2+ and cervical cancer. A two-centre, prospective, observational study was devised to test SPP as triage test in women with persistent low-grade cytology. Two commercial HPV tests were also examined in their role in triage of these women.

A further study was conducted to examine viral integration and viral DNA methylation as potential biomarkers of high-grade disease. An assay that determines the status of the HPV E2 gene and the Detection of Integrated Papillomavirus Sequences were used to assess integration. Bisulfite conversion followed by pyrosequencing was used to assess DNA methylation within two regions of the HPV genome (E2 and L1L2).

Following the clinical study BD SurePath Plus™ was found to be inferior to HPV testing in the discrimination of high-grade cervical disease. The age of women was found to significantly affect the results of all three tests.

Viral integration and viral DNA methylation were both associated with high-grade disease. New sites of viral integration were found in the study and the predilection of common fragile sites and repeat sequences as sites of integration was also reinforced. It was also discovered that integration could affect the accuracy of some HPV and biomarker tests.

To Katy, William, Edward, and Rebekah

Acknowledgements

First and foremost, I would like to thank my loving wife Katy. Her patience and understanding at all times were incredible and her ability to maintain family life throughout was extraordinary. I certainly could not have done this without her unwavering love, support and encouragement.

I owe a great deal of thanks to my supervisors, Dr Amanda Tristram, Dr Sam Hibbitts, and Dr Ned Powell. I am thankful for their help, guidance, friendship, and, in particular, their hard work at reading and editing chapters in the final weeks of writing. I would also like to thank Professor Alison Fiander, Mr Kenneth Lim, Mr Robert Howells and the rest of the staff at the South East Wales Gynaecological Oncology Centre for inspiring me to start and allowing me to finish this MD. I appreciated also the statistical advice from Professor Robert Newcombe, the expertise of the SuPerLy project management team, and the commitment of the staff in colposcopy services.

I am grateful to Dr Sadie Jones, Dr Dean Bryant, Mrs Angharad Edwards and all other past and present members of the HPV Research Group for sharing in the experience, the highs and the lows, and for their combined help and wisdom in the lab.

I extend my gratitude to my family and close friends who have also demonstrated their support. In particular, I would like to thank my parents for their love and guidance for the last 32 years; I would not have come this far without them.

I would like also to acknowledge my funders; the Emma Jane Demery Bequest Fund and the Tom Owen Memorial Fund for supporting the work in this MD. Finally, I am extremely grateful to all the women that took part in the study.

Throughout this experience, as in indeed throughout life, 'I can do everything through Christ, who gives me strength' (Philippians 4:13, New Living Translation).

Abbreviations

Abbreviation	Definition
AIN	Anal intraepithelial neoplasia
ANOVA	Analysis of variance
APOT	Amplification of papillomavirus oncogene transcripts
AUC	Area under the curve
bp	Base-pair
BS	Bisulfite (i.e. Sodium bisulfite)
CDK	Cyclin-dependent kinase
CI	Confidence interval
CIN	Cervical intraepithelial neoplasia (also grade 1,2 and 3)
CFS	Common fragile site
CpG	Cytosine-guanine dinucleotide
CSW	Cervical Screening Wales
DIPS	Detection of integrated papillomavirus sequences
DNMT	DNA methyltransferase
dNTP	Deoxyribonucleotide triphosphates
E1-E8	The HPV early genes
E2-(1-8)	The 8 E2 CpGs tested
E2BS	E2 binding sites
HC2	Hybrid capture 2 test
HPV	Human papillomavirus
hrHPV	High-risk HPV
HSIL	High grade squamous intraepithelial lesion
ICC	Intra-class correlation coefficient
L1 and L2	The HPV late genes
L1L2	The overlap ORF between L1 and L2
L1L2-(1-4)	The 4 CpGs of the L1L2 region tested
LBC	Liquid based cytology
LCR	Long control region
lrHPV	Low-risk HPV
LSIL	Low-grade squamous intraepithelial lesion
M	Mean
MCM	Minichromosome maintenance proteins
Md	Median
NHSCSP	NHS cervical screening programme
NPV	Negative predictive value
nt	Nucleotide
OR	Odds ratio
ORF	Open reading frame
P97	The HPV16 early promoter
Pap	Papanicolaou (stain)
PC	PapilloCheck®
PPV	Positive predictive value
RLU	Relative light unit
SCC	Squamous cell carcinoma
SD	Standard deviation
SPP	BD SurePath Plus™
SuPerLy	BD SurePath Plus™ in persistent low-grade cytology study
SuPerLy–HIM	SuPerLy–HPV integration and methylation study

Table of Contents

Chapter 1 – HUMAN PAPILLOMAVIRUSES 1

1.1 Introduction 1

1.2 HPV Taxonomy 1

1.3 The HPV Genome 2

1.3.1 Early region 3

1.3.2 Late region 5

1.3.3 The long control region 6

1.4 The HPV Life Cycle 6

1.4.1 Infection and uncoating 6

1.4.2 Genome maintenance 7

1.4.3 Proliferative phase 8

1.4.4 Genome amplification 9

1.4.5 Virus synthesis 10

1.5 Human Immune Response 10

1.5.1 Innate immune response 10

1.5.2 Adaptive immune response 11

1.5.3 Host immune response to HPV infection 11

1.5.4 Immune evasion mechanisms 11

Chapter 2 – HPV ASSOCIATED DISEASE 13

2.1 Introduction 13

2.2 Anatomy of the cervix 13

2.2.1 The Transformation Zone 13

2.3 Cervical Cancer 14

2.3.1 Incidence and Prevalence 14

2.3.2 Aetiology 16

2.3.3 Presentation 16

2.3.4 Investigation 17

2.3.5 Management 18
2.3.6 Prognosis 18
2.4 Cervical Intraepithelial Neoplasia 18
2.4.1 Incidence and Prevalence 20
2.4.2 Progression to cancer 20
2.5 Non-Cervical HPV Associated Disease 23
Chapter 3 – CERVICAL CANCER PREVENTION 25
3.1 Introduction 25
3.2 Definition and Principles 25
3.2.1 Screening tests 25
3.3 Secondary prevention 26
3.3.1 Cytological classification 27
3.3.2 Changes in policy 28
3.3.3 HPV testing within the screening programme 29
3.4 Primary Prevention 31
Chapter 4 – BIOMARKERS OF HPV ASSOCIATED DISEASE 33
4.1 Introduction 33
4.2 Definition 33
4.2.1 Screening biomarkers 33
4.2.2 Rationale for biomarkers in cervical screening 34
4.3 HPV DNA-based tests 34
4.3.1 HPV genotyping 35
4.4 Biomarkers of transforming HPV infections 35
4.5 Biomarkers of aberrant S-phase induction 36
4.5.1 BD ProEx™ C 37
4.5.2 BD SurePath Plus™ 37
4.6 Viral integration 38
4.6.1 Integration assays 39
4.7 DNA methylation 41

4.7.1 Methylation assays....42
Chapter 5 – AIMS AND HYPOTHESES....45
5.1 Aims....45
5.2 Hypotheses....45
Chapter 6 – STUDY DESIGN....47
6.1 Introduction....47
6.2 SuPerLy Study....47
6.2.1 Aims and objectives....47
6.2.2 Study design....47
6.2.3 Study population....48
6.2.4 Inclusion and exclusion criteria....48
6.2.5 Recruitment and consent....48
6.2.6 Study plan....49
6.2.7 Sample management....50
6.2.8 Endpoints....51
6.2.9 Blinding....51
6.2.10 Patient confidentiality....52
6.2.11 Ethical considerations....52
6.2.12 Regulatory approvals....52
6.2.13 Sample size calculations....52
6.2.14 Funding....53
6.3 SuPerLy – HIM study....53
6.3.1 Aims and objectives....54
6.3.2 Study design and study population....54
6.3.3 Inclusion and exclusion criteria....54
6.3.4 Ethical considerations....54
6.3.5 Funding....54
Chapter 7 – METHODS AND MATERIALS....55
7.1 Introduction....55

7.2 Sample Reception and Login 55
7.3 Sample Processing 55
7.3.1 Sample Processing Procedure 55
7.4 General 56
7.4.1 Sample Handling 56
7.4.2 DNA Handling 56
7.4.3 Quantification and purity of DNA 56
7.4.4 PCR 56
7.4.5 Agarose gel electrophoresis 57
7.4.6 Positive controls 57
7.5 DNA Extraction 58
7.5.1 QIAamp MinElute DNA Extraction Procedure 58
7.5.2 Proteinase K DNA Extraction Procedure 59
7.5.3 ß-globin PCR 60
7.6 HPV Typing 60
7.6.1 PapilloCheck Microarray Assay 60
7.6.2 Hybrid Capture 2 63
7.7 Integration assays 63
7.7.1 E2 PCR 63
7.7.2 Detection of Integrated Papillomaviruses Sequences 64
7.8 Methylation assay 68
7.8.1 Bisulfite Treatment 68
7.8.2 Pyrosequencing PCR 69
7.8.3 Pyrosequencing procedure 70
7.9 PCR conditions and primers 73
7.9.1 PCR conditions 73
7.9.2 Primers 74
7.10 Statistical analysis 75
7.10.1 Statistical techniques 76

7.11 Method development 77
7.11.1 DNA extraction and the quality of the DNA 77
7.11.2 E6 and E2 PCRs 79
Chapter 8 – RESULTS AND DISCUSSIONS – SUPERLY STUDY 83
8.1 Introduction 83
8.2 Study population 83
8.3 Referral cytology 84
8.4 Patient age 85
8.5 DNA extraction 85
8.6 Histological outcomes 85
8.6.1 Referral cytology correlation with histology 87
8.6.2 BD SurePath Plus™ correlation with histology 88
8.6.3 HPV testing correlation with histology 89
8.7 Use of BD SurePath Plus™ and HPV testing as a triage screening test 91
8.7.1 Comparison with other studies 93
8.8 Further analysis of HPV testing results 95
8.8.1 Comparison of HC2 and PapilloCheck® 98
8.8.2 HPV type-specific PCRs 101
8.9 Discussion 102
8.9.1 BD SurePath Plus™ 102
8.10 HPV typing 104
8.10.1 Strengths and weaknesses 106
8.11 Conclusions 106
Chapter 9 RESULTS AND DISCUSSION – INTEGRATION 107
9.1. Introduction 107
9.2. Viral Integration – E2 107
9.2.1. Study population 107
9.2.2. HPV 16 E2 status and histology 109
9.2.3. HPV16 E2 status and age of patient 109

9.2.4. HPV16 E2 status and referral cytology 111
9.2.5. HPV16 E2 status as a screening biomarker 112
9.3. Viral integration – DIPS 113
9.3.1. Study population 113
9.3.2. Overall DIPS data 116
9.3.3. Viral integration detected by DIPS and histology 117
9.3.4. Viral integration detected by DIPS and age 117
9.3.5. Viral integration detected by DIPS and referral cytology 118
9.3.6. Viral integration as a biomarker 119
9.4. Site of integration 120
9.4.1. Integration and correlation with tiling PCRs 121
9.4.2. Site of integration into human genome 121
9.5. Discussion 124
9.5.1. HPV16 E2 status 124
9.5.2. DIPS 126
9.5.3. Use as a biomarker 127
9.5.4. Site of integration 127
9.5.5. Strengths and weaknesses 129
9.6. Conclusion 130
Chapter 10 RESULTS AND DISCUSSION – METHYLATION 131
10.1. Introduction 131
10.2. Study population 131
10.3. Initial analyses of data 132
10.3.1. Viral DNA methylation and histology 139
10.3.2. Viral DNA methylation and referral cytology 141
10.4. Relationship between viral integration and methylation 143
10.4.1. Site of integration and methylation 145
10.4.2. Viral DNA methylation and E2 status 145
10.5. Discussion 147

10.5.1. Reproducibility....147
10.5.2. Methylation and disease grade....149
10.5.3. Methylation and cytology....150
10.5.4. Methylation and integration....150
10.5.5. Strengths and weaknesses....151
10.6. Conclusions....152
Chapter 11 – GENERAL DISCUSSION AND CONCLUSIONS....153
11.1 Introduction....153
11.2 General discussion....153
11.3 Hypothesis 1: BD SurePath Plus™ can predict the presence of high-grade cervical disease in women with persistent low-grade cytological abnormalities....153
11.4 Hypothesis 2: HPV testing can predict the presence of high-grade cervical disease in women with persistent low-grade cytological abnormalities....154
11.5 Hypothesis 3: BD SurePath Plus™ will predict with higher positive predictive value, but lower negative predictive value than HPV testing....154
11.6 Hypothesis 4: E2 disruption is a marker of a transforming HPV infection and, therefore, increases the risk of having high-grade cervical disease....155
11.7 Hypothesis 5: Viral integration is associated with a transforming HPV infection and, therefore, an increased risk of having high-grade cervical disease....155
11.8 Hypothesis 6: Hypermethylation within the viral genome correlates with high-grade cervical disease....155
11.9 HPV testing....156
11.10 Development of biomarkers and reproducibility of assays....157
11.11 Future work....157
11.12 Conclusions....158
References....159
Appendix I – REVIEW ARTICLE ACCEPTED FOR PUBLICATION....183
Appendix II – PATIENT INFORMATION, CONSENT FORM, CLINIC STUDY FLOW CHART....194
Appendix III – HYBRID CAPTURE 2 PROCEDURE....199

Chapter 1 – HUMAN PAPILLOMAVIRUSES

1.1 Introduction

Human papillomaviruses are non-enveloped double-stranded DNA viruses that infect the epithelial basal layer. The majority of HPV infections occur without symptoms and are cleared by the host within 8-12 months. However, in a small number of cases infection may persist resulting in intraepithelial neoplasia and, over time, progression to invasive carcinoma. This chapter covers the classification and molecular biology of HPV and the host's response to infection.

1.2 HPV Taxonomy

To date, 189 papillomavirus (PV) types have been isolated. The majority have been isolated from humans (120 types) but many have been found in other mammals, birds and reptiles (64, 3 and 2 types, respectively) (Bernard et al., 2010). The classification of the different types is based on the nucleotide sequence of part of the HPV genome (L1) and is divided into genera, species, and types. The amount of nucleotide sequence identity shared within this taxonomy is: genera – less than 60%, species – between 60% and 70%, types – between 71% and 89% (de Villiers et al., 2004). The two main genera are the Alpha and Beta papillomaviruses, accounting for approximately 90% of the known HPV types (Doorbar, 2006). The larger Alpha genus contains the genital/mucosal HPV types. These types can be further classified into low and high risk depending on the frequency with which they are found in cancer (Figure 1.1).

Beta papillomaviruses are associated with cutaneous lesions in humans. These infections remain unnoticed amongst the general population; however, in immunocompromised individuals and patients with the inherited disease epidermodysplasia verruciformis, they can spread and result in the development of non-melanoma skin cancer (Harwood and Proby, 2002, Pfister, 2003). The other HPV genera include Gamma, Mu and Nu. They are associated with mainly benign cutaneous lesions (de Villiers et al., 2004).

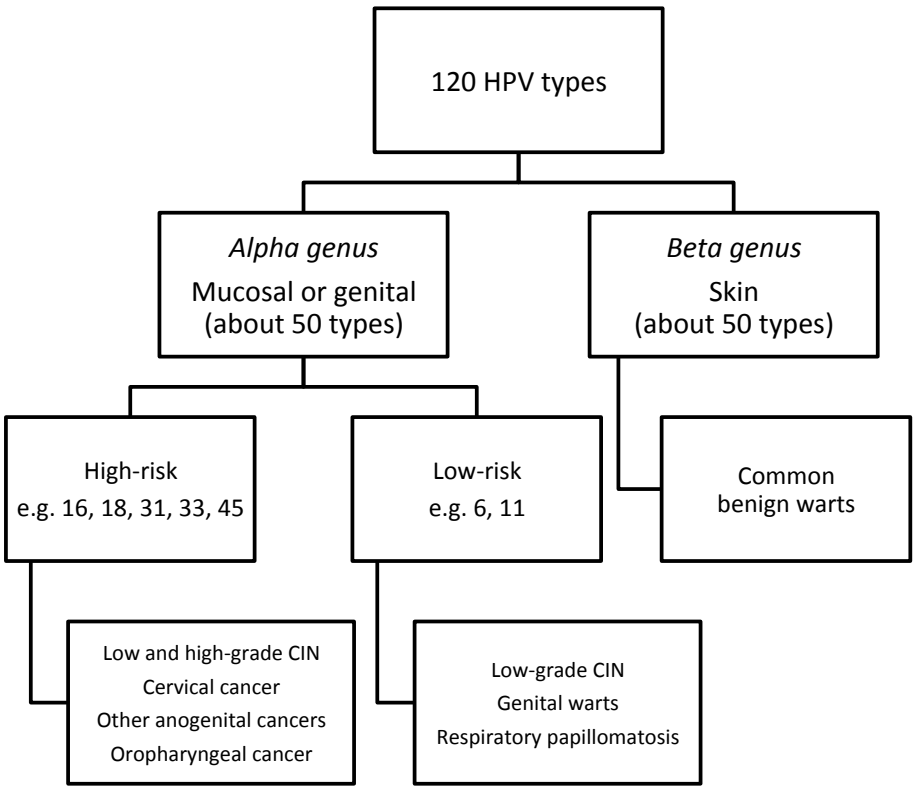

Figure 1.1: Features of the two largest papillomavirus genera.

The much smaller Gamma, Mu, and Nu genera are not shown; they are associated with benign and malignant skin lesions.

1.3 The HPV Genome

All papillomaviruses have a similar genomic organisation with approximately 8000 base-pairs of double-stranded circular DNA within an icosahedral capsid. The length of the HPV16 genome is 7904 base-pairs (GenBank® accession number NC_001526). The HPV genome includes eight open-reading frames (ORFs), which are divided into two different regions, according to when they are expressed during the virus' life cycle. The early (E) region is mainly involved in regulating viral DNA replication and is made up of E1, E2, E4, E5, E6, and E7. The late (L) region is responsible for the virus structure and consists of L1 and L2. The genes within each region can overlap. There is also a third region of the genome known as the Long Control Region (LCR) or regulatory region.

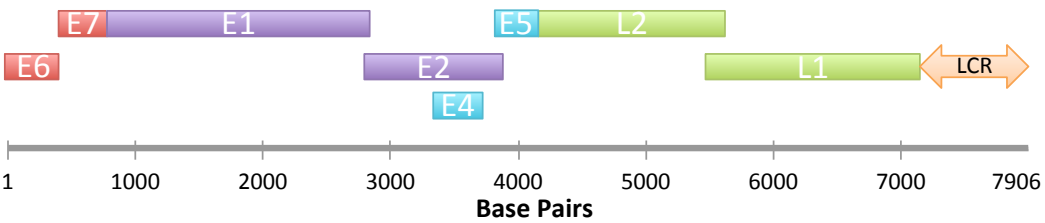

Figure 1.2: The HPV16 Genome. (Genbank® accession no. NC_001526)

1.3.1 Early region

1.3.1.1 E1

The E1 ORF encodes a protein that binds to the viral origin of replication (ori) within the LCR. It has adenosine triphosphatase (ATPase) and DNA helicase activity that prepares the viral genome for replication by forcing the DNA strands apart, and is an essential part of initiating viral DNA replication (Wilson et al., 2002). E1 has a low affinity to bind to the ori and in order to achieve its crucial part in the replication process it needs to be bound to E2, thereby increasing its affinity and ability to unwind the DNA (Liu et al., 1995, Desaintes and Demeret, 1996).

1.3.1.2 E2

The E2 ORF can give rise to multiple gene products including the full length E2 protein and truncated E2 polypeptides (Lambert et al., 1990). The E2 protein is well characterised as a viral transcription factor and can function as both a transcriptional activator and repressor (Bouvard et al., 1994, Stenlund, 2003). Truncated E2 has been shown to repress full length E2 activity via competitive exclusion at E2 binding sites (E2BS) or by the formation of inactive heterodimers (Lambert et al., 1990). The activity of E2 also appears to be dependent on the amount of E2 protein within the cell (Bouvard et al., 1994)(see 1.4.2).

1.3.1.3 E4

The E4 ORF is located within the E2 ORF; however, translation occurs in separate reading frames. E4 is expressed via a splice from E1 ORF (first 5 amino acids) to E4 ORF (last 85 amino acids) creating an E1^E4 fusion protein (Doorbar et al., 1986). The role of E1^E4 is not yet fully understood; however, it appears to be involved in viral DNA replication (Nakahara et al., 2005), disruption of the cytokeratin network (Doorbar et al., 1991), induction of G2 cell cycle arrest in keratinocytes (Davy et al., 2006, Davy et al., 2005) and release from keratinocytes (Doorbar et al., 2012, Doorbar et al., 1991).

1.3.1.4 E5

The E5 protein is involved in genome amplification through an ability to modulate the cell's signalling pathways. Over expression of E5 inhibits the degradation of and increases the

phosphorylation of Epidermal Growth Factor (EGF) receptors (Straight et al., 1993, Fehrmann et al., 2003) resulting in a replication competent environment being maintained in the higher differentiated layers of the epithelium (Doorbar, 2006). E5 is also implicated in the promotion of cellular survival, via the prevention of DNA damaged induced apoptosis (Krawczyk et al., 2008, Zhang et al., 2002) and inhibition of the host immune response, via the down-regulation of major histocompatibility complexes (Ashrafi et al., 2006).

1.3.1.5 E6

Within high-risk HPV types the E6 gene encodes a transforming protein, which is capable of immortalising epithelial cells, and initiating an oncogenic process. One of the main ways it achieves this is by mediating the degradation of p53 (Scheffner et al., 1990). The p53 protein is a major tumour suppressor protein that regulates cell growth in response to cellular stress. When the p53 pathway is activated there is transcriptional upregulation of cell cycle arrest and pro-apoptotic genes that ultimately result in the self-destruction of the cell. E6 has also been shown to degrade Bak, a pro-apoptotic protein that belongs to the Bcl-2 protein family (Thomas and Banks, 1999). Bak is located in the mitochondrial membrane, where in response to cellular stress, it forms pores within the membrane that releases cytochrome c, which is a key modulator of the apoptosis pathway (Howie et al., 2009).

E6 further promotes the immortalisation of human cells via its action on telomerase (Klingelhutz et al., 1996). Telomeres are found on the chromosome termini and typically get progressively shorter with successive cell division. When telomere length reaches a critical point a protective cellular senescence pathway is triggered. Telomerase is a ribonucleoprotein that lengthens telomeres. E6 indirectly activates the expression of the catalytic component of telomerase, human Telomerase Reverse Transcriptase (hTERT) (Oh et al., 2001). A direct post-transcriptional interaction between E6 and hTERT has been described, demonstrating an alternative mechanism by which E6 targets this enzymatic activity (Liu et al., 2009). Heightened hTERT activity permits cancer cells to undergo repeated rounds of replication and resist cellular senescence (Yugawa and Kiyono, 2009).

1.3.1.6 E7

The E7 ORF encodes a protein that interacts with and degrades the Retinoblastoma tumour suppressor protein (pRb) (Boyer et al., 1996b). The affinity of this interaction depends on the HPV type, with hrHPV types having a far greater affinity for pRb (Boyer et al., 1996b). Proteosomal degradation of pRb results in expression of DNA synthesis genes leading to

unscheduled cell proliferation and immortalisation (Flores et al., 2000). Further pRb independent activity of E7 has been described. E7 has been shown to increase cyclins A and E, thus enhancing cell cycle progression (Longworth and Laimins, 2004). E7 also binds to and inactivates p21 and p27, both cyclin-dependent kinase (CDK) inhibitors, thus preventing cell cycle arrest at the G1/S checkpoint (Cho et al., 2002).

1.3.2 Late region

Both L1 and L2 proteins play a crucial role in mediating virus infectivity and both have been implicated as targets for vaccine development (see section 3.4). They are late proteins and, thus, only expressed when the virus is preparing to be released from the cell (Doorbar, 2005).

1.3.2.1 L1

This encodes the major viral structural protein. 360 copies of this protein are organised into 72 capsomeres to make up the 55–60 nanometre icosahedral capsid (Modis et al., 2002); Figure 1.3). When L1 proteins are expressed as eukaryotic recombinant proteins they are capable of self-assembly into virus-like particles (VLPs)(Hagensee et al., 1994). When VLPs are used for immunisation they produce an immune response in the infected host (Breitburd et al., 1995, Suzich et al., 1995, Kahn and Burk, 2007).

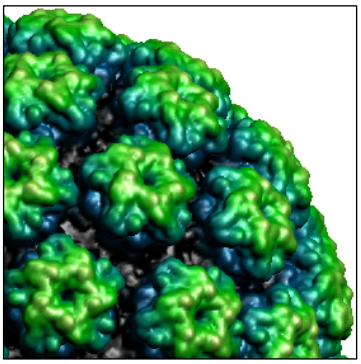

Figure 1.3: Atomic model showing a section of an HPV virion.

A small proportion of the 72 capsomeres formed by HPV L1 proteins is shown. Adapted from (Modis et al., 2002).

1.3.2.2 L2

This encodes the minor viral structural protein. The exact number of L2 particles that are present in the middle of the pentavalent capsomeres at the virion vertices is unknown, but it is estimated that there are 12 (Conway et al., 2009). The L2 protein plays a crucial part in virion assembly and is required for efficient encapsidation of viral DNA (Holmgren et al., 2005).

1.3.3 The long control region

The LCR is a non-coding region of approximately 1000 base-pairs, which separates the early and late gene clusters. It contains transcriptional, post-transcriptional and replicative cis-regulatory elements (Thierry, 2009). The ori and early promoter (p97 for HPV 16) are located at the 3' end of the LCR and it is the activation of the early promoter that triggers the transcription of the early genes that are required for viral replication. Within the LCR there are binding sites for CCAAT displacement protein (CDP), Ying Yang 1 (YY1), Activator protein 1 (AP1), Specificity factor 1 (Sp1), Transcription factor IID (TFIID) as well as binding sites for E1 and E2 (Spink and Laimins, 2005, Ai et al., 2000, O'Connor, 2000). Both YY1 and CDP have been shown to block p97 activity and their levels remain high until differentiation commences, at which point their expression is downregulated and their repressive effects reversed (O'Connor et al., 2000, O'Connor et al., 1996). AP1 also varies with differentiation and is thought to play a role in E6/E7 transcription (Sen et al., 2004). Binding of Sp1, TFIID, E1 and E2 all appear to interact in p97 activation and initiation of viral replication (see 1.4.2)

1.4 The HPV Life Cycle

The life cycle of HPV is dependent on the differentiation of the epithelial cell that it invades. Doorbar (2005) has divided the organisation of the HPV life cycle into several distinct phases: infection and uncoating, genome maintenance, proliferative phase, and virus synthesis.

1.4.1 Infection and uncoating

The virus first infects epithelial stem cells found in the basal layer. For most HPV types this would require a breach in the upper layers of the epithelium via small wounds or micro-abrasions. Sulphated sugars, particularly heparin sulphate, on the cell surface interact with L1 protein allowing initial attachment (Giroglou et al., 2001). Another potential receptor for HPV is laminin, which is found in the extracellular matrix and has demonstrated a high affinity for HPV (Selinka et al., 2007). The cleaving enzyme furin causes structural changes in the virion capsid allowing transfer to a secondary receptor on the basal keratinocyte, a necessary step for internalisation of the virus and consequent nuclear entry (Schiller et al., 2010, Kines et al., 2009, Doorbar, 2005). The endosomal pathway appears to complete the internalisation process releasing the virion into the cell (Selinka et al., 2002). The virions are finally uncoated by a membrane destabilising L2 peptide (Kamper et al., 2006) allowing the viral DNA to be transported into the nucleus. The L1 protein is retained in the endosome and eventually undergoes lysosomal degradation (Schelhaas et al., 2012, Doorbar et al., 2012).

1.4.2 Genome maintenance

Once infection has occurred and the viral DNA is in the cell nucleus genomic replication can take place. It is believed that viral DNA is held in low copy number episomes within the basal cells of the epithelium (50-100 copies per cell) (Flores et al., 1999). The pattern of gene expression involved in this process is not fully known but it seems most likely that E1 and E2 are required for replicating and maintaining the viral DNA as an episome (Doorbar, 2005). E2 recruits E1 to the ori, which as a result of E1's helicase activity separates the viral DNA, allowing further recruitment of E1 and eventual displacement of E2 (Sarafi and McBride, 1995, Moscufo et al., 1999, Sedman and Stenlund, 1998).

The stability of the genome is maintained by the minichromosome maintenance (MCM) complex. The complex is made up of six MCM proteins (MCM2–7) that form a hetero-hexamer helicase that is essential for licensing and then restricting DNA replication to only once per cell (Laskey and Madine, 2003, Kearsey and Labib, 1998). MCM proteins interact with both checkpoint and recombination proteins to promote S-phase stability (Bailis et al., 2008). They are abundant in the nucleus during the cell cycle and then appear to be degraded following exit from the cell cycle (Musahl et al., 1998). This ordered replication is seen in another virus that causes cancer in humans; the Epstein-Barr virus (EBV). Latently infected cells with EBV DNA are associated with MCM protein, thus, controlling replication and minimising viral protein expression in the host cell (Chaudhuri et al., 2001). The MCM proteins, therefore, may assist in establishment of latency in an HPV infection. It is unclear, however, given that HPV has its own helicase to what extent HPV utilises the host's helicase. Evidence has suggested that, unlike MCM proteins, E1 protein is able to license HPV 16 DNA replication continuously and in a random fashion (Hoffmann et al., 2006). However, the same study found that in some cell lines HPV would replicate only once per cell. Furthermore, HPV 16 DNA can be replicated and maintained without the presence of E1 (Kim and Lambert, 2002) and even in the absence of any viral gene expression (Kim et al., 2005). Conversely, overexpression of E1 appears to be a trigger for excess replication suggesting that HPV uses both mechanisms for viral DNA replication (Hoffmann et al., 2006).

There are four E2 binding sites (E2BS1–4) within the LCR that regulate the early promoter (p97) (Romanczuk et al., 1990). At low concentrations E2 binds to the E2BS that it has the highest affinity for: E2BS4 (located 500bp upstream of the promoter) and to a lesser extent

E2BS3, resulting in increased promoter activity and early protein transcription (Demeret et al., 1998, Hines et al., 1998, Steger and Corbach, 1997). However, as a consequence of increased transcription the levels of E2 accumulate and E2 binds to the other lower affinity binding sites. E2BS1 and 2 are positioned in close proximity to two other transcription factor binding sites, the TATA box and Sp1. As a result of E2 binding at these two sites, the TATA binding protein (TBP) that mediates recruitment of the TFIID complex (transcription initiation complex) is inhibited and transcription is downregulated (Dostatni et al., 1991). Furthermore, E2 binding at E2BS2 has been shown to result in the displacement of Sp1 and TFIID from the adjacent Sp1 binding site (Tan et al., 1994). Consequently, Sp1 is prevented from activating p97 and transcription of E6 and E7 is reduced (Gloss and Bernard, 1990).

E2 also has a significant role in the partitioning of episomal genomes. During mitosis the viral genomes are secured to the cellular chromatin ensuring that there is appropriate viral genome segregation into daughter cells (McBride, 2008, Dao et al., 2006).

The role of E6 and E7 proteins in the infected basal cells has not been fully defined, especially in infections caused by low-risk HPV types. In these infections it is believed that the inherent wound healing response is responsible for the early propagation of the infected cells (Valencia et al., 2008), with signalling for viral gene expression and protein function from the epidermal growth factor (EGF) pathway (Doorbar et al., 2012, Rosenberger et al., 2010). In high-risk HPV infections the E6 and E7 proteins act on the basal and parabasal cell layers by augmenting cell proliferation (Doorbar, 2006).

1.4.3 Proliferative phase

In normal epithelium only basal cells are actively able to divide. Once a daughter cell exits the basal cell layer it migrates to the suprabasal layers where it undergoes terminal differentiation. When the HPV DNA is in the daughter cell the production of the E7 oncoprotein inactivates the tumour suppressor protein pRb, which drives the quiescent cell back into S phase allowing further replication (see section 1.3.1.6)(Cheng et al., 1995, Boyer et al., 1996a). It does this through down regulation of the E2F-dependent DNA replication genes. Consequently there is increased expression of cyclin E and cyclin A resulting in aberrant CDK2 activity (Duensing and Munger, 2004). It is this activity that induces cellular division and the transcription of the host genes required for continuous G1/S phase DNA replication (Munger and Howley, 2002, Doorbar, 2006).

The E2F pathway is normally controlled by a set of CDK inhibitors. One of them, $p16^{INK4a}$, represses cell cycle progression by blocking CDKs from phosphorylating pRb (Khleif et al., 1996, Serrano et al., 1993). In cells with transforming HPV infections E7 counteracts the effect of $p16^{INK4a}$ resulting in significant overexpression and accumulation of $p16^{INK4a}$(Sano et al., 1998). Excess $p16^{INK4a}$ has been found in the vast majority of cervical precancers and cancers, whereas it has rarely been discovered in normal tissue (Cuschieri and Wentzensen, 2008, Klaes et al., 2001)

Unscheduled DNA replication would normally be prevented by another tumour suppressor protein, p53, causing apoptosis; however, this is neutralised by the actions of the E6 oncoprotein (Scheffner et al., 1990). The E6 protein of both high- and low-risk types disable elements of p53 function, but only the high-risk types cause its ubiquitination and proteasome-dependent degradation (Fu et al., 2010, Zanier et al., 2012, Pim and Banks, 2010, Doorbar et al., 2012). As previously mentioned DNA replication is also permitted to occur freely by the ability of E6 to increase telomerase activity, and maintain telomere length.

Another CDK inhibitor that regulates the cell cycle is p21. The p21 protein inhibits the activity of cyclin/CDK2 complexes, binds to proliferating cell nuclear antigen (PCNA) and causes growth arrest promoting normal keratinocyte differentiation (Gartel and Radhakrishnan, 2005). The transcription of this protein is tightly controlled by p53 and if p53 is downregulated it results in reduced p21 activity (Gartel and Radhakrishnan, 2005).

It is thought that the other HPV early proteins continue to be expressed in the proliferation phase in order to maintain the viral episomes at a low copy number (Middleton et al., 2003).

1.4.4 Genome amplification

Further genomic growth occurs in the mid to upper layers of the epithelium following an increase in the activity of the late (differentiation dependent) promoter. This promoter (P670 in HPV 16), located within the E7 ORF, increases expression of E1, E2, E4, and E5 in order to facilitate genome amplification. Consequently, the viral cell copy number increases to approximately 1000 copies per cell (Flores and Lambert, 1997). The infected differentiating epithelial cell is subject to differentiation signals and can express markers of both differentiation (keratins 1 and 10) and cell cycle entry (MCM, Ki-67, (PCNA), Cyclin E and Cyclin

A) (Doorbar, 2006). The roles of E4 and E5 in genome amplification have been described already (see 1.3.1.3 and 1.3.1.4).

1.4.5 Virus synthesis

The structural capsid proteins L1 and L2 are expressed following genomic amplification in the upper layers of the infected epithelium. Viral DNA is packaged into the capsid and the virions are only released once they reach the epithelial surface in the terminally differentiated cell. There is no separate promoter for L1 and L2 expression; instead it requires initial activation by the late promoter followed by an alteration in splice site (Doorbar, 2006) that is facilitated by elevated levels of E2 expression (Ozbun and Meyers, 1998, Johansson et al., 2012). The alternative E1^E4 protein promotes L1 production and genome packaging (Johansson et al., 2012, Milligan et al., 2007, Doorbar, 2005). In the lower layers of the epithelium there is a splicing silencer preventing premature expression of the L1 gene (Zhao et al., 2004). In BPV E2 has been shown to improve the efficiency of genome encapsidation during natural infection (Zhao et al., 2000). The binding of L2 to viral DNA through its association with promyelocytic leukaemia bodies is thought to require E2 (Doorbar, 2006). The final egress of the virus has been mainly attributed to E4 (Doorbar, 2006, Doorbar et al., 2012). High levels of E4 in the upper epithelial layers have been shown to assemble into amyloid fibres that can disturb the keratin structure (McIntosh et al., 2008), thus affecting the normal assembly of the cornified envelope (Wang et al., 2004, Brown et al., 2006). Once released, the virions must survive outside the cell and begin the infective cycle again.

1.5 Human Immune Response

The immune response to any pathogen usually involves two systems: the innate and the adaptive immune responses. Both of these are important in the control of HPV infection.

1.5.1 Innate immune response

Innate immunity identifies the pathogen and is the human body's first line of defence. It is normally activated by cell damage or cell death resulting in inflammation (the local vascular response to injury). Firstly, local parenchymal cells are employed, and then local phagocytes are activated. The phagocytes release inflammatory cytokines and other soluble defence proteins. Pathogens contain some common molecular targets, known as pathogen-associated

molecular patterns, which are identified by receptors within the innate immune system resulting in its activation and subsequent induction of the adaptive immune response.

1.5.2 Adaptive immune response

As part of the innate immune response dendritic cells are activated. The dendritic cells use enzymes to split the pathogen into smaller pieces, known as antigens. These cells then present the antigen on the surface of the cell by binding them to a receptor called the major histocompatibility complex. Dendritic cells are the only antigen-presenting cells that can activate naïve T lymphocytes (T cells) which begins the cell-mediated adaptive immune response (Stanley, 2006). There are also B lymphocytes (B cells) that wait in the blood and lymph for circulating antigens. They also can present antigen to T cells, which stimulate the B cell to differentiate into antibody-producing plasma cells. The ability to produce a specific antibody is then retained so that a second exposure to the same antigen will result in rapid release of the appropriate antibody. This is known as the humoral immune system.

1.5.3 Host immune response to HPV infection

Most HPV infections are cleared without any overt clinical disease. The evidence for how they are cleared comes mainly from animal models with PV-associated disease (Nicholls et al., 2001, Wilgenburg et al., 2005, Nicholls et al., 1999, Monnier-Benoit et al., 2006). The immune response appears to be modulated by antigen-specific CD4+ T cell dependent mechanisms and commonly results in seroconversion and antibody production to the viral capsid protein L1 (Stanley, 2010).

1.5.4 Immune evasion mechanisms

The immune response is inherently restricted due to the fact that HPV infects keratinocytes that are cells destined for death and desquamation. There is no warning signal to the host such as viraemia, inflammation or virus induced cytolysis (Stanley, 2012). Furthermore, the innate immune system is significantly downregulated by HPV. Pro-inflammatory cytokines are not released and the signals required for antigen-presenting cells are either not present or absent (Kanodia et al., 2007). Moreover, complete HPV virions are only found in the upper layers of the epithelium separated from the circulating immune cells.

Chapter 2 – HPV ASSOCIATED DISEASE

2.1 Introduction

This chapter explores the most common HPV associated diseases, cervical cancer and cervical intraepithelial neoplasia (CIN). It describes how HPV infects the cervix, the progression of CIN to cancer and the impact these diseases have on the individual and the wider society.

2.2 Anatomy of the cervix

The word cervix comes from the Latin *cervix uteri*, meaning "neck of the womb". It is separated from the upper two-thirds of the uterus or *corpus uteri* by a fibromuscular junction: the *internal os*. The cylindrical cervix protrudes through the anterior vaginal wall at the vaginal vault resulting in the visible part of the cervix; the *portio vaginalis*. This part, on average, is 3 cm long and 2 cm wide (Singer and Jordan, 2006). The opening of the cervix, referred to as the *external os*, is small and circular in the nulliparous cervix, whereas in the parous cervix it appears wider and slit-like; furthermore the cervix itself is bulkier. The portion of cervix that is exterior to the external os is the *ectocervix*. The endocervical canal connects the uterine cavity and the vagina.

The type of epithelial lining that covers the cervix varies primarily according to location, but also to a number of other factors. The majority of the ectocervix is lined by stratified squamous epithelium and is in continuity with the vaginal epithelium; i.e. multiple layers to protect against the relatively hostile environment of the vagina. The epithelium covering the endocervical canal and sometimes part of the ectocervix is the columnar epithelium. This is a single layer of rectangular cells that includes some secretory cells and is in continuity with the lining of the uterine cavity. One feature of the columnar epithelium within the endocervical canal is that it forms folds and invaginations that are referred to as glands; hence it may also be called the glandular epithelium. These folds can present a problem when trying to clinically assess the endocervix either by cytological screening or at colposcopy.

2.2.1 The Transformation Zone

The junction where the squamous epithelium and the columnar epithelium meet is called the squamocolumnar junction (SCJ). The location of the SCJ can vary throughout life and

consequently two SCJs are referred to: the *original* SCJ that existed in the fetus and the *new* SCJ where the process of squamous metaplasia has replaced columnar epithelium with squamous epithelium. Squamous metaplasia occurs during late fetal life, adolescence and in the first pregnancy usually driving the new SCJ further into the vagina. The area between the original and new SCJs, where metaplasia occurs is called the transformation zone (TZ).

Epithelial carcinogenesis frequently occurs in areas of metaplasia where it is believed there are alterations in stem cell fate decision and epithelial-stromal tissue remodelling (Singer and Jordan, 2006, Birchmeier et al., 1995). Carcinogenic metaplasia occurs in other parts of the body as an adaptive response to harmful stimuli; e.g., gastro-oesophageal reflux disease can cause columnar cell metaplasia (Barrett's oesophagus) in the oesophagus that can give rise to oesophageal cancer, infection with *Helicobacter pylori* can cause gastric intestinal metaplasia and may result in gastric cancer. In the case of cervical carcinogenesis HPV primarily affects the squamous metaplasia within the TZ. Despite an increasing understanding of the molecular biology of HPV infection, the precise mechanisms of carcinogenesis remain unclear (Stanley, 2010)(see 2.4.2).

2.3 Cervical Cancer

Invasive cervical cancer is a disease where the cells within the epithelium of the *cervix uteri* have become abnormal, have grown uncontrollably and have spread from the epithelium into the underlying connective tissue. It is one of the most common malignant neoplastic diseases affecting women. The vast majority of cervical cancers arise from one of two lineages depending on whether they originate in squamous or in glandular epithelium (Koushik and Franco, 2006). Squamous cell carcinomas (SCC) account for 80%, whereas glandular malignancies, or adenocarcinomas (ADC), account for 10-15%(Schiffman and Brinton, 1995).

2.3.1 Incidence and Prevalence

Cervical cancer is the third most frequent cancer in women worldwide and seventh overall, with an estimated 530 000 new cases in 2008 (Ferlay J, 2010). This is due to the high incidence in the developing countries where access to healthcare is restricted and there are no screening programmes. In excess of 85% of the worldwide burden occurs in developing countries, where it is responsible for 13% of all female cancers (Ferlay J, 2010). The highest incidences can be found in Eastern and Western Africa where the age-standardised rates (ASR) are greater than

30 per 100 000. In countries in Eastern Africa, South-Central Asia and Melanesia cervical cancer remains the most common female cancer (Figure 2.1).

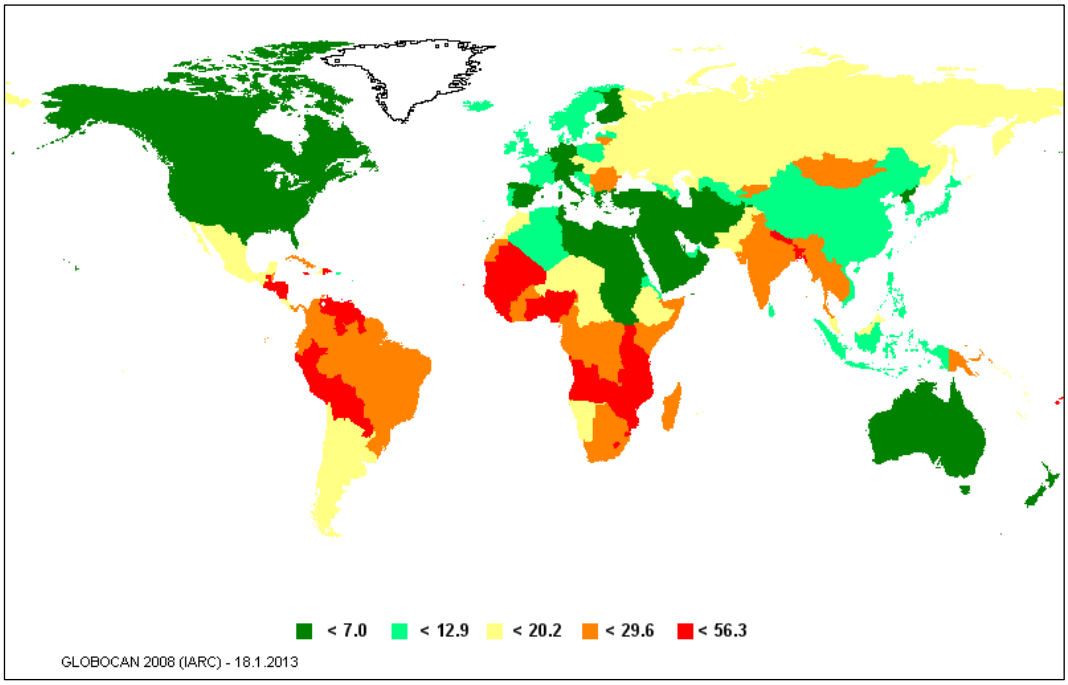

Figure 2.1: Map showing the worldwide incidence of cervical cancer

The figures given are estimated age-standardised incidence rate per 100 000. Incidence map constructed on the GLOBOCAN website (Ferlay J, 2010).

In the UK, cervical cancer is the twelfth most common cancer in women and the third most common gynaecological cancer after uterus and ovary. In 2010 there were 2 900 new cases of cervical cancer in the UK with 60% of them in women under the age of 50 (Cancer Research UK, 2013).

2.3.1.1 Prevalence of HPV infection

HPV infection is extremely common; most women in the world being infected at some point in their lives, with a lifetime risk of between 50–80% (Koutsky, 1997). The prevalence can vary with differences in geography and age; peak infection rate is seen in women less than 25 years of age with a decline that plateaus around 30–35 years and, in some countries, a much smaller second peak in women aged 50+ years (Franceschi et al., 2006). According to a meta-analysis that included one million women worldwide with normal cytological findings, the worldwide point prevalence for a woman carrying an HPV infection can be estimated at 11–12% (Bruni et al., 2010). The five most common types found worldwide were HPV 16 (3.2%), HPV 18 (1.4%), HPV 52 (0.9%), HPV 31 (0.8%) and HPV 58 (0.7%) (Bruni et al., 2010).

2.3.1.2 HPV prevalence in cervical cancer and CIN

HPV prevalence corresponds with the severity of cervical disease, such that the higher the grade of lesion the greater the number that are HPV positive (Guan et al., 2012). The most recent meta-analysis reports that of the HPV positive cervical cancers, the three most prevalent HPV types were HPV 16 (63%), HPV 18 (16%) and HPV 45 (5%) (Guan et al., 2012).

2.3.2 Aetiology

Since the 1970s, human papillomavirus (HPV) has been proposed as a causative factor in a variety of benign and malignant diseases. Harald zur Hausen, who won the Nobel Prize for Medicine in 2008, was the first to make the causal link between HPV and cervical cancer (zur Hausen et al., 1974, zur Hausen, 1976, zur Hausen, 1977). This led to the first epidemiological study to investigate the link (Munoz et al., 1992), followed by a landmark worldwide HPV prevalence study (Bosch et al., 1995) that reported the worldwide prevalence of HPV in cervical carcinomas is 99.7% and, therefore concluded it is a necessary cause of invasive cervical cancer (Walboomers et al., 1999). HPV is the single most important risk factor; however, it is not considered a sufficient cause because the host normally clears the infection. Other factors such as smoking, age at first intercourse, oral contraceptive use, other sexually transmitted infections (STIs; e.g., *Chlamydia trachomatis* and herpes simplex virus), parity, immunosuppressive conditions including HIV infection and polymorphisms in the human leucocyte antigen system are also considered to be involved in the development of cervical cancer (Baseman and Koutsky, 2005, Richardson et al., 2003, Sellors et al., 2003, Winer et al., 2003, Schiffman and Castle, 2003, Moscicki et al., 2001, Koutsky, 1997).

2.3.2.1 Acquisition of cervical HPV infection

There is strong evidence to show that cervical HPV infection is transmitted as a consequence of penetrative vaginal sexual intercourse (Koch et al., 1997, Ley et al., 1991, Andersson-Ellstrom et al., 1996). The majority of women acquire cervical HPV infection shortly after their sexual debut.

2.3.3 Presentation

The most common symptom of cervical cancer is unexpected vaginal bleeding; i.e., between periods, after or during sexual intercourse or at any time if the woman is postmenopausal. Another common symptom is abnormal, often offensive, vaginal discharge. However, both of

these symptoms are not very specific and in young women infections such as C*hlamydia trachomatis* should be ruled out first (Herod, 2010). Women with cervical cancer may also complain of dyspareunia (discomfort during sexual intercourse).

Early stage cervical cancer is often asymptomatic and the woman may only be alerted to a problem following an abnormal smear result or an abnormal appearance (or feel) of the cervix during a vaginal examination. It is not possible to diagnose an invasive process on a smear; however, if highly atypical cells are seen it can be highly suggestive of cancer. Similarly, if a smear result reports a preinvasive lesion cervical cancer is still a possibility. Moreover, a negative smear cannot be relied upon to exclude cancer, as a necrotic tumour may not desquamate abnormal cells.

Current UK standards dictate that: (i) a clinically suspicious cervix should prompt urgent referral to a gynaecologist, (ii) smears suggestive of malignancy or glandular neoplasia should be referred for colposcopy within two weeks, (iii) smears suggestive of high-grade dysplasia should be referred for colposcopy within four weeks (Luesley and Leeson, 2010).

2.3.4 Investigation

The diagnosis of cervical cancer is based on histological examination of a biopsy. This would usually be done using a colposcope, which magnifies the cervix. The cervix is swabbed with acetic acid to highlight any abnormal areas, and biopsies are taken for histological assessment. For a colposcopic assessment to be satisfactory, the entire TZ must be visualised. For larger lesions a punch biopsy is adequate; however, if a microinvasive cancer is suspected then a cone biopsy or large loop excision of the transformation zone should be undertaken in order to obtain adequate tissue to correctly stage the cancer (Herod, 2010). All histology showing invasive cervical disease should be reviewed in a cancer centre by a specialist histopathologist to ensure accurate staging.

Once the diagnosis is confirmed there are further investigations required depending on the initial histology. A stage 1a cervical cancer can be managed in a cancer unit and is unlikely to require any further investigations. If the cancer is of a higher stage or there is any doubt over the diagnosis of a microinvasive disease then the patient should be managed in a cancer centre. The next step is to accurately stage the disease, which in the case of cervical cancer is

based on clinical examination. This involves a detailed examination of the cervix, hysteroscopy, cystoscopy and proctoscopy. A chest x-ray is essential and intravenous urography is advised; however, most centres now use magnetic resonance imaging to assist with the clinical staging, in particular to assess lymph node metastases and hydroureter (Hricak et al., 2005, Bipat et al., 2003).

2.3.5 Management

Very early stage 1a (microinvasive) cancers can be treated surgically; either by LLETZ, cone biopsy, or hysterectomy. Whereas, locally advanced stage 2b or greater cancers (indicating that cancer has spread to the parametrium) are not appropriate for surgery and instead require a combination of chemotherapy and radiotherapy. For the cancers staged in between; i.e. stage 1b–2a, no advantage has yet been shown indicating whether radical surgery with adjuvant radiotherapy or radiotherapy alone or chemotherapy and radiotherapy together are best (Herod, 2010).

2.3.6 Prognosis

Overall, the mortality-to-incidence ratio is 52%, and in 2008 cervical cancer caused 275 000 deaths worldwide, of which 88% were in developing countries (Ferlay J, 2010). In 2000, there were 2.7 million years of life lost between the ages of 25 and 64 attributable to cervical cancer, 2.4 million of which occurred in developing countries and 0.3 million in developed countries (Yang et al., 2004). In the UK, around 940 women died from cervical cancer in 2010, which is a 70% improvement on the early 1970s (Cancer Research UK, 2013). Two thirds of women with cervical cancer survive their disease for five years or more (Cancer Research UK, 2013).

2.4 Cervical Intraepithelial Neoplasia

Cervical intraepithelial neoplasia (CIN) is used to describe proliferative intraepithelial squamous lesions that contain cytonuclear atypia and abnormal maturation. The term was first introduced to describe a continuum of cervical dysplasia to cancer (Richart, 1973). CIN can be divided into grades 1, 2 and 3. These grades correspond to the amount of dysplasia in the epithelium: mild dysplasia = CIN1, moderate dysplasia = CIN2, severe dysplasia and *carcinoma in situ* = CIN3 (Figure 2.2).

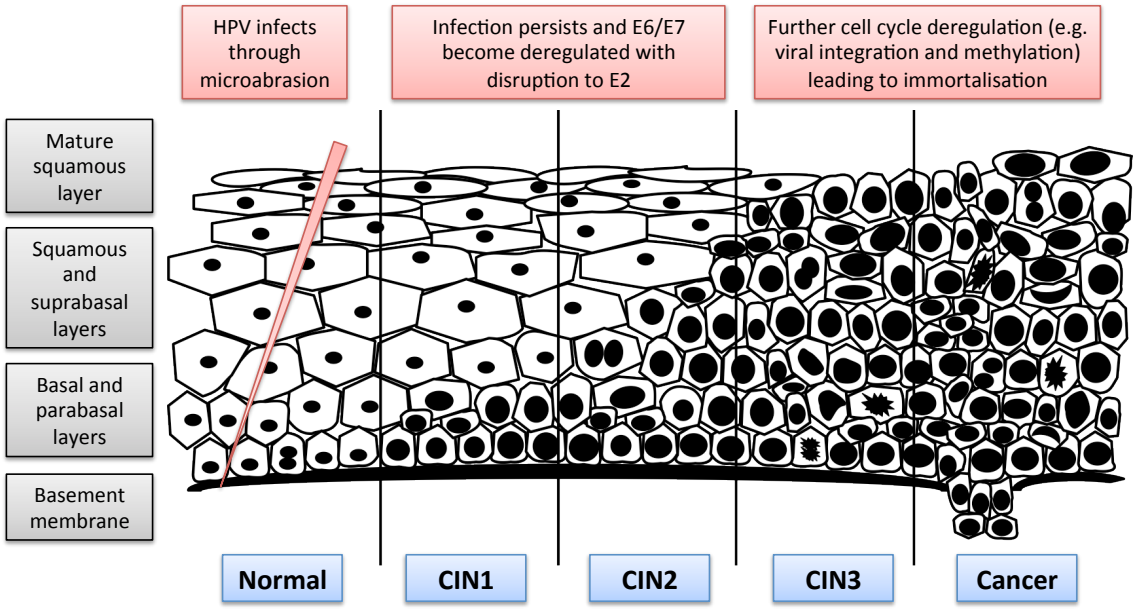

Figure 2.2: The grading of cervical intraepithelial neoplasia.

CIN is graded according to the proportion of the epithelium that is occupied by basaloid, undifferentiated cells (Kiviat et al., 1992). CIN1 – Nuclear atypia (including coarse chromatin pattern, abnormal chromatin distribution, hyperchromaticity and increased nuclear-to-cytoplasmic ratio) most marked in the basal third of the epithelium. There is good amount of differentiation in the upper two-thirds of the epithelium and mitotic figures only feature in the lower third. CIN2 – Nuclear atypia extends to the middle third. There is still maturing and differentiation of cells occurring in the middle third of the epithelium and mitotic figures, which may be abnormal, occur anywhere in the lower two-thirds of the epithelium. CIN3 – Nuclear atypia is marked throughout the full thickness of the epithelium. Differentiated cells may be found in the upper third or may be completely absent. Multiple mitotic figures feature throughout the epithelium and are commonly abnormal. Cancer – The basement membrane is breached and a microinvasive cervical carcinoma is formed. The pink boxes in the diagram summarise the HPV life cycle.

Since this classification was introduced much more is understood about the development of cervical cancer. As previously mentioned, the majority of HPV infections are transient and it is believed that most CIN1 lesions would fall into this group. When a combination of factors interact the HPV infection will persist and the dysplasia will become more advanced; i.e., CIN2–3 (see 2.4.2). Distinguishing CIN2 from CIN1 or CIN3 has been consistently problematic and not as reproducible as the other CIN diagnoses (Carreon et al., 2007, Ismail et al., 1989). For this reason it has been suggested that there should be a two-tier system in line with the Bethesda classification system (Smith and Desai, 2007). CIN1 becomes low-grade CIN, and CIN2–3 becomes high-grade CIN (or CIN2+). However, CIN3 has a higher rate of progression (see 2.4.2) and is seen as a better marker of the effectiveness of any new cervical cancer prevention strategies (International Agency for Research on Cancer, 2005). Furthermore, it is postulated that a number of women will be overtreated as a consequence of just having a simplified two-tier system (Schneider, 2003).

2.4.1 Incidence and Prevalence

CIN is asymptomatic and is, therefore, only picked up as a result of screening tests. Incidence and prevalence rates will therefore be significantly influenced by the quality of the screening test used and rates of attendance. In Wales, the incidence of CIN1 in women that had been referred to colposcopy for an inadequate or abnormal smear and had an adequate biopsy (between April 2010–March 2011) was 18.2% (1145/6497) which was 0.5% (1145/227 597) of all women screened (Cervical Screening Wales, 2007). For CIN2 the incidences were 13.7% (893/6497) and 0.4% (893/227 597) respectively, and for CIN3 the incidences were 27.8% (1804/6497) and 0.8% (1804/227 597) respectively (Cervical Screening Wales, 2007). In a large cohort study the prevalence of CIN3 at enrolment was 0.5% (70/13 084) (Peto et al., 2004a).

2.4.2 Progression to cancer

The estimated risk of progression for CIN3 comes from the notorious unethical study where women with biopsy-proven CIN3 were not informed of the potential risks and were left inadequately treated over many years (McIndoe et al., 1984, McCredie et al., 2010, McCredie et al., 2008). They reported that the risk of CIN3 progressing to cervical cancer was 30–50% over 30 years and that untreated women had a 50–100-times risk of invasion compared to adequately treated women. Meta-analysis of 15 longitudinal studies with follow-up periods of between two to five years, estimated the progression of CIN2 to CIN3 at 20%, and to invasive cancer at 5%, whereas 40% of CIN2 regress and 40% persist (Oster, 1993). For CIN3 the same study concluded, the risk of progression to cancer was 14%, regression was 32%, and the risk of persisting was 55%. The risk of progression in CIN1 is small, with just 10% progressing to CIN3 and only 1% to invasive cancer, whereas regression is common, with up to 60% regression rates (Oster, 1993, Melnikow et al., 1998).

In the majority of HPV cervical infections there is ordered viral gene expression resulting in viral synthesis and release from the upper epithelial layers (productive infection or CIN1). When the viral gene expression becomes disordered as a result of a persistent infection HPV-associated neoplasia (CIN2–3) can develop. It is generally believed that the levels of E6 and E7 expression increase in relation to the severity of CIN (Doorbar et al., 2012). High levels of E6 and E7 can cause the accumulation of genetic changes and genetic instability within the epithelial cells, contributing to cancer progression.

It is still not clear precisely how the viral gene expression becomes deregulated. A number of mechanisms have been postulated and it remains a pertinent area for research. One widely held perception is that it requires a persistent infection (Doorbar, 2006, Stanley, 2010). The maximal incidence of CIN (around 25–34 years) occurs about twenty years before the maximal incidence of cervical cancer (Bosch and de Sanjose, 2003). Persistence, is however, both difficult to define and difficult to prove. A persistent infection may not be always productive and it is now known that HPV can remain latent in the basal layer, with its genomic expression suppressed and no apparent disease present (Doorbar et al., 2012, Hopman et al., 2000). Two consecutive positive HPV tests do not prove a persistent infection and, furthermore, a negative repeat test may not definitely mean the viral DNA has been cleared. Although there is a general consensus that a number of years are required to develop high-grade disease, studies have shown CIN2+ present in young women very soon after an HPV infection has been identified (Woodman et al., 2001, Paavonen et al., 2007, Paavonen et al., 2009, Szarewski et al., 2012). Overall, a relatively small number of women that have HPV infection go on to get cervical cancer, indicating that other cofactors are involved along with hrHPV to generate cervical carcinogenesis.

Viral integration of the viral episome into the human cell chromosome appears to be an important sequela of HPV gene deregulation and is thought to play a significant role in carcinogenesis (Melsheimer et al., 2004). HPV viral DNA methylation appears to play an equally important role in carcinogenesis (Clarke et al., 2012). The roles and timing of viral integration and viral methylation within carcinogenesis are widely debated and are explored in more detail in sections 4.6 and 4.7 (Figure 2.2.3).

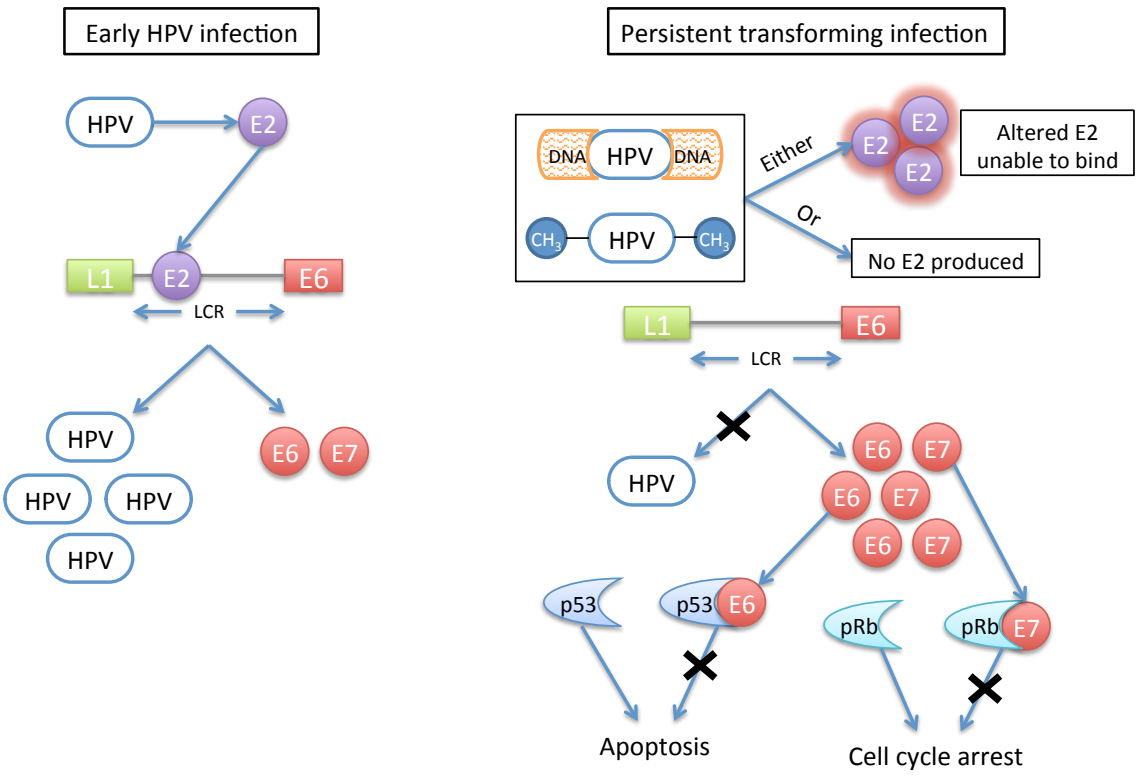

Figure 2.2.3: Diagram showing potential mechanism of HPV oncogene deregulation.

In early HPV infection the E2 protein regulates viral transcription. In excess E2 binds to all four E2 binding sites and E6 and E7 expression is downregulated. In a persistent transforming infection it is believed that the loss of functioning E2, possibly by viral integration or DNA methylation, results in uncontrolled E6 and E7 expression leading to the silencing of tumour suppressor genes.

As previously discussed, the host immunity is crucial in clearing the HPV infection (see 1.5.3). It follows, therefore, that processes that affect host immunity have the potential to play an influential role in carcinogenesis. There are several examples of this in the literature. Immunosuppressed women with HIV have an increased prevalence of HPV infections and higher risk of progression to high-grade CIN and cervical cancer (Moscicki et al., 2004, Denny et al., 2012). Iatrogenic immunosuppression, such as following renal transplant, also significantly increases the risk of developing HPV infection and CIN (Rudlinger et al., 1986).

Cigarette smoking has long been associated with the risk of cervical neoplasia; however, it has been difficult to prove because smoking strongly correlates with sexual behaviour (Szarewski and Cuzick, 1998, Appleby et al., 2006). Studies have shown that smoking doubles the risk of developing high-grade CIN (Collins et al., 2010) and cancer (Appleby et al., 2006). Whether smoking increases the risk of acquiring a cervical HPV infection, impairs the antibody response to the infection, or both is still unclear (Collins et al., 2010). Recent evidence suggests that smoking may induce epigenetic changes that result in altered HPV gene expression (Calleja-Macias et al., 2009, Lea et al., 2004).

There is increasing evidence that oestrogen plays a significant role in cervical carcinogenesis. A systematic review of the literature concluded that use of hormonal contraceptives is associated with an increased risk of cervical cancer (Smith et al., 2003). The mechanism of this is still not clear, although it appears the longer the duration of use the higher the risk and once use of hormonal contraception is stopped the risks may decrease (Smith et al., 2003). Experimental evidence from mouse models have demonstrated that oestrogens upregulate HPV E6 and E7 oncogene expression, stimulate cell proliferation, block apoptosis, and lead to DNA damage (Gariglio et al., 2009, Brake and Lambert, 2005).

2.5 Non-Cervical HPV Associated Disease

There is strong evidence in the literature that HPV has a causal aetiology for other non-cervical cancers. The International Agency for Research on Cancer (IARC) considers HPV as a carcinogen in the following cancer sites: vulva, vagina, anus, penis and the oropharynx (Bouvard et al., 2009). Whilst currently, the greatest amount of research into HPV is associated with the cervix there does appear to be an upsurge in research at different sites. Certainly, it is hoped that increased knowledge of HPV carcinogenesis in the cervix will be equally relevant at non-cervical sites. For a more detailed look at non-cervical HPV associated disease see Appendix I.

Chapter 3 – CERVICAL CANCER PREVENTION

3.1 Introduction

Many would argue that cervical screening is one of the most successful examples of a screening programme in modern medicine. Furthermore, with the development of the HPV vaccines the notion of potentially eradicating HPV associated disease does not seem altogether irrational. This chapter describes the different types of prevention, what prevention strategies are in current practice, and some of the proposed refinements. It will also introduce the rationale for the two studies that are contained within this thesis.

3.2 Definition and Principles

The prevention of disease is divided into three levels. The purpose of primary prevention is to stop a disease from occurring, reducing both the incidence and prevalence of a disease. Secondary prevention aims to detect and treat disease in its earliest stages before it causes significant morbidity. The goal of tertiary prevention is to reduce the morbidity of existent disease and provide rehabilitation to restore function and lessen the associated sequelae.

3.2.1 Screening tests

One of the fundamental concepts of secondary prevention is the ability to identify a disease in its early stages and then instigate an effective treatment. A screening test is used to test a population at risk that have no symptoms but are in a preclinical stage of the disease. Wilson and Jungner (1968) set out the fundamental requirements for developing a screening programme (Figure 3.1).

The usefulness of the screening test is evaluated by its *sensitivity* and *specificity*. Sensitivity is the proportion of people with the disease that are correctly identified by the test. A higher sensitivity means there are less false negatives. Specificity is the proportion of people that do not have the disease and test negative. A highly specific test produces a small percentage of false positive results. Ideally a test would be both highly sensitive and specific, however, this is frequently not possible and typically there is a trade-off. Using sequential tests; e.g., a test with high sensitivity is followed by a test with high specificity, can sometimes help overcome this. The predictive value of a test is the probability of having the disease, given the results of the test. A positive predictive value (PPV) is the probability of a person with a positive test result actually having the disease. A negative predictive value (NPV) is the probability that a

person, who tests negative, genuinely does not have the disease. It is not only the sensitivity and specificity of the test, but also the prevalence of the disease in the tested population that determines the predictive value. When the prevalence of a disease is low, the PPV will also be low even if the test has high sensitivity and specificity. In order to increase the PPV the test can be targeted to populations that have a higher risk of having the disease. When a patient or clinician considers a positive screening test result it is important to remember that a screening test is not diagnostic and a further test or procedure is required in order to confirm the presence of the disease.

- the condition should be an important health problem
- the natural history of the condition should be understood
- there should be a recognisable latent or early symptomatic stage
- there should be a test that is easy to perform and interpret, acceptable, accurate, reliable, sensitive and specific
- there should be an accepted treatment recognised for the disease
- treatment should be more effective if started early
- there should be a policy on who should be treated
- diagnosis and treatment should be cost-effective
- case-finding should be a continuous process

Figure 3.1: Wilson and Jungner's criteria for a screening programme. Adapted from (Wilson and Jungner, 1968).

3.3 Secondary prevention

The discovery of HPV DNA in almost all cervical cancers has provided great opportunity to develop strategies for prevention and treatment of cervical cancer (Walboomers et al., 1999). Prevention has involved cervical screening for pre-cancerous lesions using cytological testing. Papanicolaou (Papanicolaou, 1946) first developed the technique of sampling the exfoliated cells from the transformation zone of the cervix. Most commonly a wooden spatula would be used as the sampling device and then the cytological material would be “smeared” evenly on to a glass slide, and a fixative would then be applied. The cervical smear would then be sent to a laboratory where it would be examined by a cytopathologist under a microscope.

In developed countries, screening programmes based on this technique have proved very successful in reducing the incidence and mortality of cervical cancer (Levi et al., 2000, Robles et al., 1996). One of the most important aspects of the screening programme in England and

Wales was the introduction of the systematic computerised call/recall system, which correlates with a large increase in coverage and a dramatic reduction in death rates from cervical cancer (Quinn et al., 1999, Sasieni and Adams, 1999). It has been estimated that the NHS Cervical Screening Programme (NHSCSP), regarded as one of the most successful cancer prevention programmes in the world, has prevented 80% of deaths from cervical cancer in the UK (Peto et al., 2004b).

There are a number of drawbacks with Papanicolaou-based cytological screening. One study found that 47% of the women in the UK that had a cervical cancer (stage 1B1 or worse) under the age of 70 had a full screening history (Sasieni et al., 1996). The sensitivity of cervical screening to detect cervical cancers and its precursors has been reported to be between 22–99% (Martin-Hirsch et al., 2002). The specificity, on the other hand, is more consistent ranging from 85-100% (Martin-Hirsch et al., 2002). More recent data on the use of cytology have been reported in studies assessing the use of HPV as a screening test. Pooled data including 60 000 women found the sensitivity in detecting CIN2+ was 53%, whereas the specificity was 96% (Cuzick et al., 2006).

Cytological assessment is labour intensive, subject to inter-observer variations and requires continuous quality control and assurance. One improvement in the cytological assessment has been the development of liquid-based cytology (LBC). The sample is taken from the cervix using a brush, which is placed into a liquid preservative. The slide is prepared with the liquid in the laboratory and examined by the cytopathologist. The LBC method has now superseded conventional cytology in the UK and it was the method used in the study described in this thesis. LBC has the same sensitivity as conventional cytology for detection of CIN2+, however, more CIN1 lesions are detected and consequently it has a lower PPV (Ronco et al., 2007). The main advantage of LBC is that there is a large reduction in inadequate smears (Ronco et al., 2007, NICE, 2003).

3.3.1 Cytological classification

There are several different classification systems that have been developed over the years. In general, most systems around the world are either based on the British Society for Clinical Cytology or the North American Bethesda terminologies (Table 3.1). For the study presented in this thesis the Cervical Screening Wales (CSW) terminologies were used.

CSW[a] (BSCC 1986)	Bethesda System 2001[b]	NHSCSP 2013[c] (BSCC 2008)
Borderline changes	Atypical squamous cells of undetermined significance (ASC-US)	Borderline change in squamous cells
Borderline ?high-grade	Atypical squamous cells suggesting high-grade squamous intraepithelial lesion (ASC-H)	
Borderline endocervical		Borderline change in endocervical cells
Borderline endometrial	Atypical endocervical, endometrial or glandular cells	
Borderline other glandular		
Mild dyskaryosis	Low-grade squamous intraepithelial lesions (LSIL)	Low-grade dyskaryosis
Moderate dyskaryosis	High-grade squamous intraepithelial lesions (HSIL)	High-grade dyskaryosis (moderate)
Severe dyskaryosis		High-grade dyskaryosis (severe)
Severe dyskaryosis ?invasive	Squamous cell carcinoma	High-grade dyskaryosis ?invasive squamous cell carcinoma
?Glandular neoplasia	Endocervical carcinoma *in situ*	?Glandular neoplasia of endocervical type
	Adenocarcinoma	?Glandular neoplasia (non-cervical)

Table 3.1: The different cytological classification terminologies.

a. Cervical Screening Wales terminology, which was based on the original BSCC 1986 classification system (Cervical Screening Wales, 2012b).

b. The Bethesda system, which was revised in 2001 (Solomon et al., 2002).

c. The recently published NHSCSP guidelines (NHSCSP, 2013), which were based on the revised BSCC 2008 system.

3.3.2 Changes in policy

Historically, cervical screening was offered to all women aged between 20 and 64 years every three to five years. Since 2005, women living in England have been invited for screening from the ages of 25 to 50 every three years and from 50 to 64 every five years. However, the screening policies in Wales and Scotland continued to recommend three yearly screening should start at 20 years of age. The devolved health departments reviewed their policies for cervical screening following a recommendation from the UK Nation Screening Committee and both Scotland and Wales have made plans to alter their policy in line with England and Northern Ireland (UK National Screening Committee, 2012, Welsh Government, 2013, Scottish Government, 2012). The decision was informed largely by four studies that showed the risk of

screening and subsequent intervention to women aged 20–25 was greater than the benefit (Sasieni et al., 2009a, Sasieni et al., 2009b, Arbyn et al., 2008, Sasieni et al., 2003).

Another fundamental change in policy has been the introduction of hrHPV testing as part of the screening programme. This is discussed in section 3.3.3.

3.3.3 HPV testing within the screening programme

There are three apparent settings for HPV testing within a screening programme:

(i) triage of mild cytological abnormalities to select for colposcopy

(ii) 'test of cure' following treatment of CIN

(iii) primary screening test.

3.3.3.1 HPV Triage

Approximately 240 000 smears in the UK per year are reported as having low-grade cytological abnormalities (Cervical Screening Wales, 2012a, ISD Scotland, 2012, Health and Social Care Information Centre, 2012). These abnormalities include: borderline changes and mild dyskaryosis. Approximately one quarter of these abnormalities are associated with CIN2+ and, historically, all these women are referred to colposcopy to exclude high grade disease (NHS Cervical Screening Programmes, 2010, Cervical Screening Wales, 2012a). It has been proposed that hrHPV testing could be used to triage these women in order to improve efficiency and alleviate stress for the women that do not have CIN. Two of the largest randomised controlled trials investigating HPV triage have shown the sensitivity of hrHPV testing to vary considerably from 69.9-95.7% for detecting CIN2+ in borderline changes and 75.2-96.1% in mild dyskaryosis (Cotton et al., 2010, ALTS Group, 2000, ALTS Group, 2003). Meta-analysis of all studies concluded that HPV triage in borderline changes improves screening accuracy for an outcome of CIN2+ compared to cytology alone (Arbyn et al., 2010, Arbyn et al., 2009a, Arbyn et al., 2012). The benefit of triage in mild dyskaryosis is not as apparent due to the high HPV positivity rate. In Europe and the USA, it is not considered cost-effective to test all of these women and current practice is to only perform HPV testing on women over 30 years old (Jordan et al., 2008, Solomon et al., 2009). The Sentinel Sites study, which included 25-64 year-old women in six centres in England, showed that a third of women with borderline changes or mild dyskaryosis, when triaged with a hrHPV test, were returned to routine recall (Kelly et al.,

2011). Overall, there is increased referral to colposcopy, however, CIN2+ should be identified sooner in affected women and the number of repeat smears is reduced along with the risk of loss to follow-up. Consequently, HPV triage has been recommended by the Department of Health (Department of Health, 2011). At the time of writing (January 2013) HPV triage was still in the process of rolling out across the UK (Smith, 2012).

3.3.3.2 HPV test of cure

Women that are treated for CIN must be followed-up in order to check for residual or recurrent disease. After women have been treated for CIN1, it was previously recommended that they should have cytological follow-up at 6, 12 and 24 months post treatment (Luesley and Leeson, 2010). If these results were negative then the women would be returned to routine screening intervals. Following the treatment of CIN2 or CIN3 cytological assessment was recommended at 6 and 12 months and thereafter annually for the subsequent 9 years before returning to routine screening (Luesley and Leeson, 2010). However, the usefulness of hrHPV testing in these women has been demonstrated. High-risk HPV testing has a higher sensitivity and slightly lower specificity compared to follow-up cytology for identifying high-grade CIN (Arbyn et al., 2012). The Sentinel Site study reported that women treated for CIN were at very low risk of residual or recurrent disease if they were hrHPV negative (Kelly et al., 2011). It is now policy that if a woman has a normal or low-grade cytology result 6 months following treatment for CIN and a subsequent hrHPV test is negative then she should be returned to 3-yearly recall (NHSCSP, 2013).

3.3.3.3 HPV in primary screening

There is a large amount of evidence from randomised clinical trials supporting the use of hrHPV testing in a primary screening setting (Ronco et al., 2006, Bulkmans et al., 2007, Naucler et al., 2007, Naucler et al., 2009, Kitchener et al., 2009, Kitchener et al., 2011). They have shown improved sensitivity in identifying CIN2+ at the expense of a loss in specificity, in particular among younger women. The true success of primary screening is measured ultimately by cervical cancer incidence rates. As the longer-term benefits become clearer it is likely that use of HPV testing in primary screening will become national health policy.

It is widely accepted that triage of HPV positive tests needs further improvement. There still remains a low chance that a woman has CIN2+ even if she is HPV positive; the PPV in most

studies is <30% (Arbyn et al., 2010). This is particularly true for younger women. Transient infections are more common in this population and it is thought that CIN2+ takes several years to develop following a persistent HPV infection. Improving the specificity of HPV-based screening has the potential to make a significant difference. A number of biomarkers have been investigated for this purpose and they are discussed in the next chapter.

3.4 Primary Prevention

One of the most promising outcomes of the discovery that HPV is a necessity for the development of cervical cancer is the HPV vaccine. Primary prevention of a disease is undoubtedly the best form of prevention and tends to be far more cost-effective than secondary forms of prevention.

Two prophylactic vaccines targeting HPV have been developed and extensively tested: Gardasil® (Merck and Company), a quadrivalent vaccine against HPV 16, HPV 18, HPV 6, and HPV 11 (Garland et al., 2007, Future II Study Group, 2007) Cervarix® (GlaxoSmithKline PLC), a bivalent vaccine against HPV 16 and HPV 18 (Paavonen et al., 2009). Full results of the trial with four-year follow-up data have been reported (Dillner et al., 2010, Munoz et al., 2010, Lehtinen et al., 2012). Immunity, measured by antibody titres, was as high after four years compared to immediately post-vaccination. Their efficacy for preventing HPV 16 or 18 associated CIN2+ was over 98% in both vaccines. There was some cross-protection against other HPV types in both cases and the quadrivalent vaccine was also near 100% effective at preventing genital warts. The full impact of the vaccines on non-cervical cancers associated with HPV is not yet known but there have been equally promising reduction in the various precancerous stages (Dillner et al., 2010, Kjaer et al., 2009, Joura et al., 2007). It will take a few decades to show how effective any cervical cancer primary prevention strategy is.

In the UK, the Cervarix™ was initially chosen, and a nationwide vaccination programme for 12–13 year old girls commenced in 2008. From September 2012, the UK health policy changed over to Gardasil®. There is a considerable benefit in prevention of genital warts and the rare but debilitating recurrent respiratory papillomatosis (see Appendix I). By the beginning of 2012, over 40 countries had introduced national HPV vaccination programmes (Markowitz et al., 2012). Their use in developing countries is still limited by the need for a cold chain, compliance with a three-dose course, and, probably the biggest of all, cost. Researchers and drug companies are, however, developing a number of new vaccines that as well as covering more HPV types should be more stable and cheaper to make (Peres, 2011).

Chapter 4 – BIOMARKERS OF HPV ASSOCIATED DISEASE

4.1 Introduction

This chapter defines and describes the different types of biomarkers. It goes on to explain the process for developing and testing a biomarker in a clinical context. The use of biomarkers in cervical screening is explored and in particular, the biomarkers that have been studied in this thesis will be discussed.

4.2 Definition

The term *biomarker* has been defined as "a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention."(Biomarkers Definitions Working Group, 2001).

A number of different types of biomarkers have been described (Puntmann, 2009):

(i) antecedent biomarkers – identify people at increased risk of developing a disease

(ii) screening biomarkers – identify a disease at an early preclinical stage

(iii) diagnostic biomarkers – identify a disease state or stage

(iv) prognostic biomarker – predict future disease course or response to treatment

The focus of this thesis is on the use of screening biomarkers in the secondary prevention of cervical cancer.

4.2.1 Screening biomarkers

The process of developing and testing a biomarker can be long and expensive. Only relatively few biomarkers will complete the process and be used regularly in clinical practice. A useful screening biomarker framework has been developed to ensure logical vigorous testing of any potential biomarker (Figure 4.1)(Arbyn et al., 2009b).

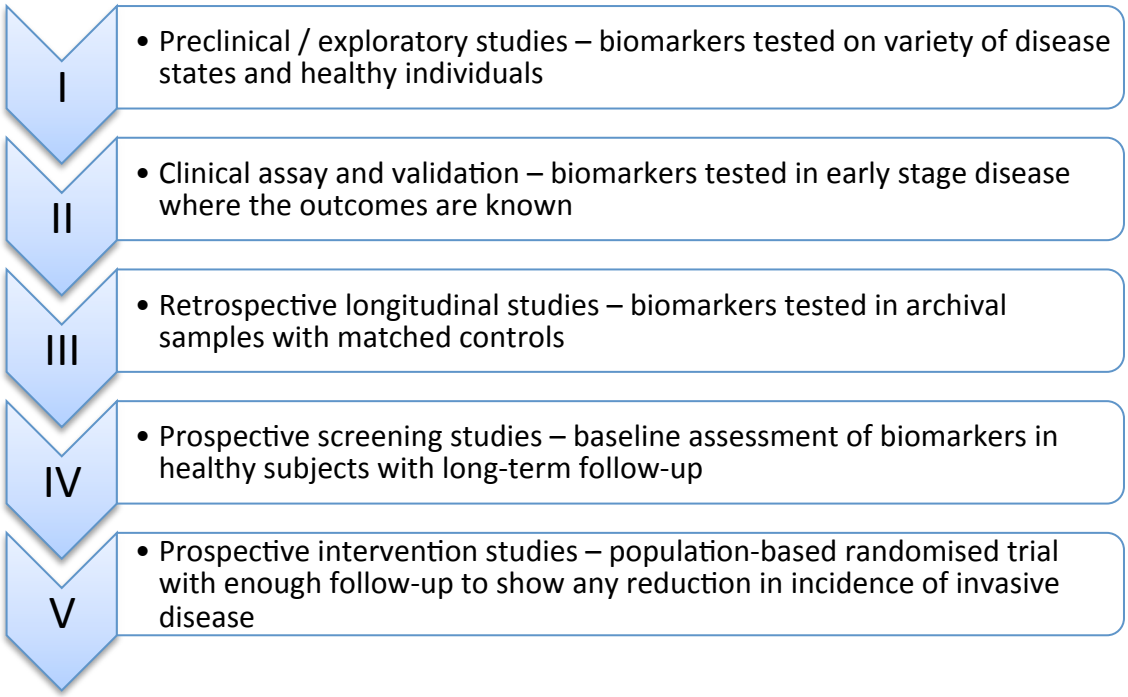

Figure 4.1: A framework for the development of screening biomarkers. Adapted from (Arbyn et al., 2009b).

4.2.2 Rationale for biomarkers in cervical screening

There are a number of limitations to cytological screening that have already been discussed (see section 3.3). Testing for hrHPV, which is biomarker of CIN, has been established to varying extents in screening programmes in a large proportion of the developed world (see section 3.3.3). Moreover, HPV vaccination programmes have also been running for several years now. However, HPV infection remains a common problem resulting in precancerous disease, especially in the unvaccinated generation. Not all HPV infections lead to CIN (as shown by the low specificity) and, therefore, if a positive HPV test results in a referral to colposcopy a lot of women will be seen unnecessarily. This is especially true for younger women. Colposcopy itself generates anxiety in women (Jones et al., 1996) and it uses valuable resources. Hence, alternative or additional screening methods using novel biomarkers are required.

4.3 HPV DNA-based tests

There are over 125 different commercially available HPV tests (Poljak et al., 2012). The majority of the clinically validated and nationally approved HPV tests are hrHPV DNA tests. They are qualitative or semi-quantitative assays that test for a number of HPV types that are considered to be oncogenic. They do not specify which individual type has been found. The

original, FDA approved and "gold standard" hrHPV test is the Hybrid Capture® 2 (HC2) HPV DNA Test (QIAGEN)(see section 7.6.2 and Appendix III for HC2 methodology). This test has been used in a large number of randomised, controlled and cohort studies and has demonstrated the value of HPV testing in cervical screening (see section 3.3.3)(Poljak et al., 2012). It is for this reason, that new HPV tests should be compared to HC2 and only used in a screening setting if they are equivalent or better than HC2 (Meijer et al., 2009). One of the potential ways of improving the specificity of hrHPV testing with HC2 is to use higher viral load cut-offs (Rebolj et al., 2011, Origoni et al., 2012). However, for accurate measurement this would require a standardised amount of sample input, which poses a significant technical challenge. Another option is to perform more detailed genotyping tests.

4.3.1 HPV genotyping

HPV16 and HPV18 are the most carcinogenic HPV types and account for 71% of the global burden of invasive cervical cancer (de Sanjose et al., 2010). In a study with 10 years of follow-up data the risk of CIN3 and cancer was significantly greater in women who tested positive for HPV16 or HPV18 than in those women who were positive for other high-risk types (Khan et al., 2005). In more recent studies the risk to women with HPV31 and HPV33 appears to be also high (Kjaer et al., 2010, Berkhof et al., 2006). HPV genotyping has been studied in a triage setting (Castle et al., 2011). This study showed that in the triage of women that are hrHPV positive, genotyping for HPV16/18 was as sensitive as cytology for the detection of CIN3+. However, type distribution varies with age (de Sanjose et al., 2010, Wright et al., 2011) and this would have to be considered in any type-specific triage screening strategy. It is also unclear how multiple concomitant HPV infections can affect the outcome.

4.4 Biomarkers of transforming HPV infections

The CDK inhibitor $p16^{INK4A}$ is over expressed in cervical neoplasia and, therefore, has been well studied (see section 1.4.3 and (Tsoumpou et al., 2009, Cuzick et al., 2012). The results from clinical trials, in which $p16^{INK4A}$ has been applied as a triage test, show a higher specificity but an equivalent or lower sensitivity when compared to HPV testing (Denton et al., 2010, Szarewski et al., 2008). However, overexpression of $p16^{INK4A}$ can sometimes occur in normal squamous metaplasia, and consequently morphological interpretation is required at the same time (Agoff et al., 2003). Therefore, an alternative dual-stained immunocytochemical test was developed to include Ki67, a known cell cycle progression marker. In a triage setting this dual-stain cytology gave equivalent sensitivity for the detection of CIN2+ in women with low-grade

conventional cytology when compared to HPV testing or $p16^{INK4A}$ alone (Schmidt et al., 2011). However, the dual-stained test showed the highest specificity and, furthermore, interpretation of the staining was morphologically independent. In primary screening trials, $p16^{INK4A}$/Ki67 dual-stained cytology significantly outperformed conventional cytology in women aged <30 years, whilst in women >30 years HPV testing was more sensitive, but significantly less specific (Denton et al., 2010).

4.5 Biomarkers of aberrant S-phase induction

Markers of cell proliferation have the potential to be useful in separating high-risk progressing HPV infections from low-risk regressing HPV infections. The role of MCMs in control of cell replication has been discussed (see 1.4.2). Dysplastic and malignant cells from a number of different tissue sites; including cervix, oesophagus, larynx, lung, and skin showed diffuse staining with antibodies to MCM proteins (Freeman et al., 1999). Immunocytochemistry using antibodies to many of the different MCM proteins to specifically identify high-grade or progressive low-grade CIN is an evolving area of research.

The use of MCM5 antibodies alongside Cdc6 (protein that permits MCM binding) showed increased sensitivity for CIN2+, when compared to Ki67 and proliferation cell nuclear antigen (PCNA)(Williams et al., 1998). Antibodies to MCM2 and MCM7 have also been shown to stain the nuclei of proliferating cells (Freeman et al., 1999, Brake et al., 2003, Henderson et al., 2011). Both the expression of MCM6 (Chen et al., 2003) and MCM6 antibody staining (Henderson et al., 2011) appear to correlate with high-grade CIN and cervical cancer. MCM7 staining in a series of archival cervical tissue demonstrated a good correlation with CIN progression (Lobato et al., 2012). It has been postulated that MCM7 may also contribute to cervical carcinogenesis. MCM7 binds directly to the E6 oncoprotein, with a greater affinity for HPV16 and HPV18 types compared to low-risk type HPV6 and HPV11, suggesting a possible regulatory role (Kukimoto et al., 1998). Furthermore, using a mouse model, a study has implicated deregulated MCM7 expression as a contributing factor in oncogene driven tumourigenesis (Honeycutt et al., 2006).

In a prospective trial involving analysis of the cervical smear samples of 455 Indian women where MCM2 and MCM5 antibody staining was analysed alongside Papanicolou counterstain an additional 10 previously missed cases of biopsy-proven cervical cancer or high-grade CIN were identified (Mukherjee et al., 2007).

4.5.1 BD ProEx™ C

BD ProEx™ C (Becton-Dickinson, New Jersey, USA) was the first commercial assay developed that uses antibodies to MCMs. It combines antibodies to MCM2 together with antibodies to DNA topoisomerase 2-alpha (TOP2A). TOP2A is an enzyme that controls and alters the topologic states of DNA during transcription and acts as a biomarker of aberrant S-phase induction (Champoux, 2001, Lang et al., 1998). BD ProEx™ C has shown increased sensitivity and specificity for detecting CIN2+ when compared to cytology (Tambouret et al., 2008, Kelly et al., 2006). This test has also been compared to hrHPV testing for borderline smears where it showed a much increased sensitivity and specificity for detecting CIN2+ (Siddiqui et al., 2008). In a primary screening setting where BD ProEx™ C was used as a triage for hrHPV positive women there was significant increases in sensitivity and PPV, resulting in 55% fewer referrals to colposcopy (Depuydt et al., 2011). One of the drawbacks of this technology is that normal proliferating cells also express MCM2 and TOP2A to some extent, which may lead to false positive staining of cytological slides (Oberg et al., 2010).

4.5.2 BD SurePath Plus™

The BD SurePath Plus™ test is a novel immunocytochemistry assay based upon the detection of MCM2 and MCM7. It has been developed, like the BD ProEx™ C, to enable straightforward assessment of LBC material. Once the stain that contains the MCM2 and MCM7 antibodies has been applied, interpretation is claimed to be much more efficient than standard cytological assessment. Furthermore, with the stain applied it is still possible to comment on the morphology of the cells (Figure 4.2).

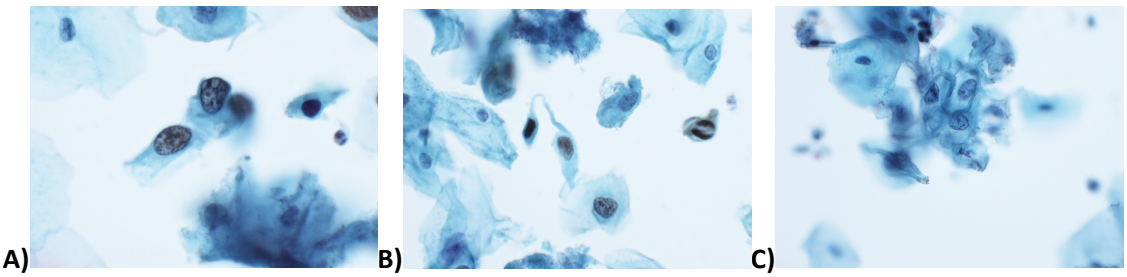

Figure 4.2: Example high power images from cytology slides stained with BD SurePath Plus™.

A, positively stained squamous cells with increased nuclear-to-cytoplasmic ratio; B, positively stained abnormally shaped squamous cell and a positively stained mitotic figure; C, no SPP staining in normal squamous cells.

To date, two studies have used this new test and presented data in abstracts at national conferences (Whitehead et al., 2011, Whitehead et al., 2012). In the first prospective study of 996 cytology specimens BD SurePath Plus™ showed improved identification of CIN2+ compared to LBC; the number of CIN2+ cases within the HSIL group increased 154% (77 cases), while the number of CIN2+ in the LSIL and ASCUS groups decreased by 62% (52 cases) and 62% (13 cases) respectively (Whitehead et al., 2011). In the preliminary data of the second prospective multicentre study (*n* = 3613) comparing BD SurePath Plus™ to LBC in a primary screening setting the sensitivity was increased (relative increase to LBC of 13.9%–21.7%), but there was a slight loss of specificity (relative loss of 3.7%–1.1%)(Whitehead et al., 2012).

4.6 Viral integration

Most HPV infections will remain in episomal form throughout their life cycle and eventually the host will clear the virus, or the virus may switch to a latent phase. However, in about 10% of cases the HPV infection develops into a transforming type. A transforming HPV infection is characterised by deregulation of the viral oncogenes E6 and E7 resulting in genomic instability, mutations and potentially immortality (see sections 1.3–1.5 and 2.4.2). Integration is likely to be one of the many mechanisms involved in the deregulation of the viral oncogenes. Integration of HPV DNA into the host cell genome is found in almost 90% of cervical cancers (Hafner et al., 2008, Arias-Pulido et al., 2006, Melsheimer et al., 2004, Corden et al., 1999, Pirami et al., 1997, Cullen et al., 1991). Integration also appears to be present in varying degrees in high-grade CIN (Matovina et al., 2009, Hafner et al., 2008, Hudelist et al., 2004, Hopman et al., 2004, Melsheimer et al., 2004, Klaes et al., 1999), and in a few cases integration has been described in CIN1 (Li et al., 2012, Huang et al., 2008).

Integration is not a natural part of the HPV life cycle; it is characterised by the deletion of genes that are essential for synthesis of an infectious virus. Hence, the virus may often persist in a mixed form with both integrated DNA and episomal DNA present. *In vitro* studies have shown that host cells containing integrated viral DNA have a selective growth advantage (Jeon et al., 1995). Loss of functioning E2 protein appears to correlate with deregulation of E6 and E7 (Thierry, 2009, Jeon et al., 1995, Romanczuk and Howley, 1992), and the E2 region of the viral genome appears to be a common site of disruption and integration (Collins et al., 2009, Arias-Pulido et al., 2006, Luft et al., 2001, Choo et al., 1987). Moreover, telomerase activation by E6 (Veldman et al., 2001) and the loss of inhibitory action of E2 (Lee et al., 2002), combine with the effects of E7 to immortalise epithelial cells (Klingelhutz et al., 1996, Kiyono et al., 1998). A growth advantage may occur indirectly as a consequence of HPV integration into a region of

the human genome responsible for cell cycle regulation. Increased expression of the human transcription factor *MYC* has been shown where HPV integration occurred in the gene's coding region of cervical cancer cell lines (Herrick et al., 2005, Peter et al., 2006). Therefore, despite the inevitably terminal loss of DNA from the virus, if the disruption occurs in the right place within the virus and the host the result is deregulated proliferation, cellular immortalisation, insertional host mutagenesis and ultimately a malignant phenotype.

Considering that integration is a common event in cervical cancer and to a lesser extent high-grade CIN it has been proposed as a potential molecular biomarker (Pett and Coleman, 2007). Identifying integration events, if caught early enough may be an efficient way of identifying which are the progressive infections with high carcinogenic risk and which are not.

4.6.1 Integration assays

There are several different ways of identifying HPV integration, but they broadly fit into two groups: (i) those that detect HPV DNA within the human genome – e.g. detection of integrated papillomavirus sequences (DIPS); and (ii) those that detect transcriptionally active viral-host integrants – e.g. amplification of papillomavirus oncogene transcripts (APOT). A third slightly less sophisticated method investigates the status of E2 as a surrogate marker of integration.

4.6.1.1 E2 PCRs

A simple method of detecting E2 disruption is to use overlapping PCRs that cover the whole E2 ORF (Figure 4.3)(Collins et al., 2009). A disruption in E2 is shown by one or more of the PCRs failing to produce the correctly sized amplicon. Similar E2 PCR methods have been used on clinical material showing potential for its use as a biomarker of transforming HPV infections (Li et al., 2012, Li et al., 2008, Alazawi et al., 2004, Tonon et al., 2001).

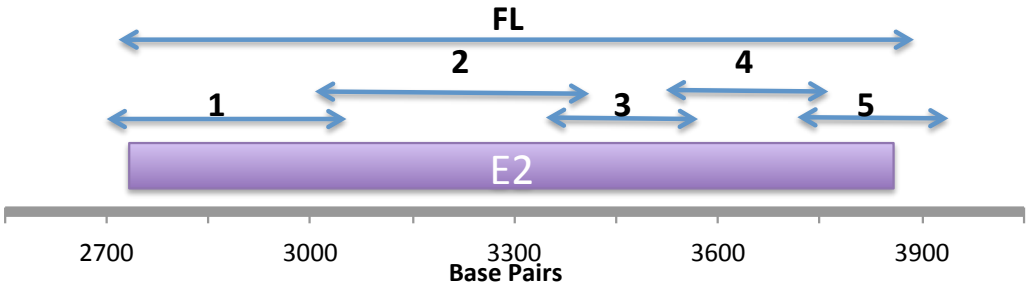

Figure 4.3: Schematic representation of the E2 tiling PCR method.

There are five primer sets covering different section of the E2 gene and a sixth primer set that covers the whole gene. The base pairs refer to the position within the whole HPV 16 genome. For exact primer locations see section 7.7.1.

4.6.1.2 DIPS

DIPS is a DNA based method for detecting integration (Luft et al., 2001). The process involves digestion of the DNA using restriction endonucleases followed by ligating a known adapter to the sticky end (Figure 4.4). Any integrated HPV DNA is then identified using a series of linear and nested PCRs with HPV primers and an adapter specific primer. The PCR products are separated by gel electrophoresis, extracted, purified and sequenced. (For more details see section 7.7.2.)

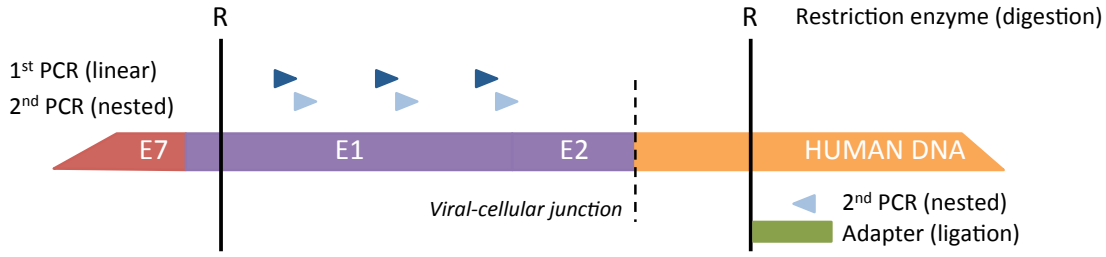

Figure 4.4: Diagrammatic representation of the DIPS method.

The different steps involved in DIPS are shown here. Digestion using restriction enzymes (R indicates example cut site), ligation with an adapter (green box), and linear and nested PCRs (blue triangles) using HPV and adapter specific primers.

4.6.1.3 RS-PCR

Restriction site PCR (RS-PCR) is another DNA based method that uses a series of HPV specific primers and restriction enzyme site-specific primers. These target viral integrants that have occurred proximal to a restriction site that are commonly found throughout the human genome (Thorland et al., 2000). This technique, however, is labour intensive and requires large concentrations of DNA (Raybould et al., 2011).

4.6.1.4 APOT

This assay allows amplification of mRNA transcripts that can contain either viral and cellular sequences derived from integrants or exclusively viral sequences derived from episomes (Klaes et al., 1999). This assay relies on high quality RNA, is labour intensive and expensive (Raybould et al., 2011). Consequently, it has been used mainly on biopsy material rather than cervical smears (Hafner et al., 2008, Ziegert et al., 2003, Klaes et al., 1999).

4.7 DNA methylation

DNA methylation is one of several epigenetic mechanisms that can regulate gene expression (Feinberg and Tycko, 2004). It can affect the structural integrity of the genome and may offer a defence mechanism against foreign agents (Lorincz, 2011, Robertson, 2005).

Methylation of DNA occurs when a methyl group is covalently added to the 5' position of cytosine. The most common place for this to occur within DNA is where cytosine (C) and guanosine (G) are connected by a phosphodiester bond (p) to form a CpG dyad. Methylation of CpGs is facilitated by a number of DNA Methyltransferases (DNMTs), of which DNMT1 is the principal one (Portela and Esteller, 2010). Within the human genome there are segments of DNA that contain a concentrated number of CpGs; referred to as "CpG islands" (Portela and Esteller, 2010). There are upwards of 45 000 CpG islands contained within the human genome, mostly found within or proximal to 5' untranslated gene promoter regions or in the first exons of genes (Antequera and Bird, 1993). Most of the CpGs in the human genome are hypermethylated, whereas those within CpG islands tend to be less methylated (Takai and Jones, 2002). In tumour cells the normal methylation patterns are frequently disrupted with global hypomethylation accompanied by region-specific hypermethylation (Robertson and Jones, 2000). When hypermethylation, or sometimes hypomethylation, occurs within the promoter of a cell cycle control gene, a DNA repair gene or tumor suppressor gene it can silence the gene and provide the cell with a growth advantage, thus, playing an important role in carcinogenesis (Esteller et al., 2001).

Many studies have investigated DNA methylation as a marker of cervical neoplasia. Certain genes, such as CADM1, which encodes a cell adhesion molecule that has been reported in a wide variety of tumour types, have shown significant potential as a biomarker of cervical

cancer development (Steenbergen et al., 2004) and to triage hrHPV positive women for CIN3+ (Hesselink et al., 2011).

DNA methylation has also been studied in the HPV genome, mainly in HPV16 and to a lesser extent HPV18. HPV16 has 113 CpG sites but no identifiable CpG islands (Cuzick et al., 2012). A number of small, mainly convenience and case-control studies have found differing degrees of methylation in parts of the early and late ORFs, and the LCR of hrHPV in cervical cancer and high-grade CIN (Kalantari et al., 2010, Fernandez et al., 2009, Brandsma et al., 2009). Elevated methylation in both the L1 and L2 ORFs appears to be the most promising candidates for use as a biomarker (Mirabello et al., 2013, Lorincz et al., 2013).

4.7.1 Methylation assays

The vast majority of DNA methylation study techniques rely on sodium bisulfite conversion. Sodium bisulfite treatment deaminates the cytosine residues found in DNA and converts them to uracils, whereas 5-methyl cytosines do not get converted (Figure 4.5). During subsequent rounds of PCR, cytosine nucleotides will be replaced by thymine nucleotides (Frommer et al., 1992). Any cytosine residues that remain represent methylcytosines that were present in the original sequence. The methylation status can then be analysed by a variety of methods (Fraga and Esteller, 2002).

4.7.1.1 Pyrosequencing

This technique is a high-throughput quantitative technique involving photon-base detection of released inorganic phosphate during nucleotide incorporation (Ronaghi, 2001). The amount of light produced is proportional to the number of deoxyribonucleotide triphosphates (dNTPs) incorporated. The methylation level at a specific site is calculated by the percentage incorporation of C versus T in bisulfite-treated DNA (Dupont et al., 2004).

4.7.1.2 Other DNA methylation assays

Other techniques include methylation-specific PCR (MSPCR) and bisulfite sequencing. MSPCR analyses only a limited range of CpGs, while bisulfite sequencing requires cloning of DNA and can be costly and labour intensive. Furthermore, the results are only qualitative or semi-quantitative at best. Newer highly sensitive and high-throughput technologies are being developed and could be ideal for screening in the future (Clarke et al., 2012).

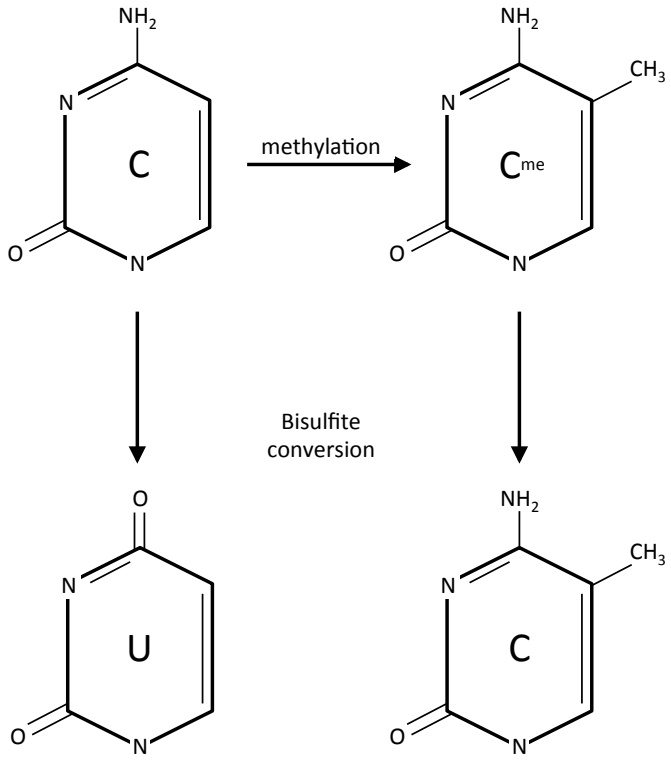

Figure 4.5: Diagram showing the effect of bisulfite treatment on unmethylated and methylated cytosines

Chapter 5 – AIMS AND HYPOTHESES

5.1 Aims

This thesis aims to evaluate a number of viral biomarkers at predicting high-grade cervical disease in women with persistent low-grade cytological abnormalities. Two studies were proposed that suitably fit with this aim and were designed to test the following hypotheses.

5.2 Hypotheses

1. BD SurePath Plus™ can predict the presence of high-grade cervical disease in women with persistent low-grade cytological abnormalities.
2. HPV testing can predict the presence of high-grade cervical disease in women with persistent low-grade cytological abnormalities.
3. BD SurePath Plus™ will predict with higher positive predictive value, but lower negative predictive value than HPV testing.
4. HPV16 E2 disruption is a marker of a transforming HPV infection and, therefore, an increased risk of having high-grade cervical disease.
5. Viral integration is associated with a transforming HPV infection and, therefore, an increased risk of having high-grade cervical disease.
6. Hypermethylation within the viral genome correlates with high-grade cervical disease.

Chapter 6 – STUDY DESIGN

6.1 Introduction

This chapter details the important features of the study designs for both the BD SurePath Plus™ in persistent low-grade cytology (SuPerLy) study and the SuPerLy–HPV integration and methylation (SuPerLy–HIM) study.

6.2 SuPerLy Study

6.2.1 Aims and objectives

The aim of the study was to establish whether BD SurePath Plus™ could predict high-grade cervical abnormality in women with persistent low-grade squamous cytological abnormalities.

6.2.1.1 Primary outcome

To establish the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the BD SurePath Plus™ test for predicting high-grade cervical disease in women with persistent low-grade squamous cytological abnormalities.

6.2.1.2 Secondary outcomes

To compare the sensitivity, specificity, PPV, and NPV for biopsy-proven high-grade cervical disease in this study population using:

a) BD SurePath Plus™ test alone

b) HPV testing alone

c) A combination of BD SurePath Plus™ test and HPV testing

6.2.2 Study design

The key objective of this study was to examine the benefits of BD SurePath Plus™ and HPV testing within a clinical cohort. The prospective, observational design of the study was intended to reflect, as far as possible, the setting in which either test might be used. Analysis of samples, therefore, would be on the same samples that were currently being used as part of the cervical screening programme in Wales. This also meant that no additional samples would be required from the women taking part. It was decided that because the test had not been used in this setting before the study should be observational rather than interventional.

6.2.3 Study population

The study population consisted of those women with borderline or mild abnormalities on a cervical smear, as defined by the inclusion criteria. The women were identified when a low-grade smear triggered a referral to colposcopy clinics in either the Cardiff and Vale University Health Board (Cardiff) or the Aneurin Bevan Health Board (Newport).

6.2.4 Inclusion and exclusion criteria

6.2.4.1 Inclusion criteria

Women were eligible for inclusion if they had a low-grade cervical smear that met with the criteria for referral for colposcopy in Wales (Figure 6.1).

- persistent low-grade squamous dyskaryosis (borderline x 3, mild x 2)
- borderline changes, high-grade dyskaryosis not excluded
- borderline changes in glandular cells x 2
- intermittent borderline changes x 3 in 10 years
- borderline changes during follow-up after treatment for CIN 2+, after the woman has been discharged from colposcopy.

Figure 6.1: Referral criteria for colposcopy in women with low-grade cytology in Wales.

6.2.4.2 Exclusion criteria

Women who were unable to give informed written consent.

6.2.5 Recruitment and consent

The Cervical Screening Administration Departments (CSAD) identified women eligible to take part in the study on their computer database system. The colposcopy clinic sent out patient information sheets in advance with the women's colposcopy appointment details. This enabled women to have plenty of time to read and absorb the information and prepare any questions regarding the study in advance of the clinic appointment. These women were then invited to take part in the study when they attended for colposcopy at either of the colposcopy clinics in Cardiff or Newport. After checking the inclusion and exclusion criteria written informed consent was obtained. The patient information sheets, consent forms and a study flow chart for the colposcopy clinic are included in Appendix II.

6.2.6 Study plan

1. All cervical liquid based samples (SurePath, Source BioScience, Nottingham, UK) were processed in the normal manner and a slide was prepared for cytological assessment. Once the slide had been prepared, 2ml of BD SurePath™ Preservative Fluid was added to the cell deposits and stored at room temperature for up to 4 weeks. For each specimen there was a residual cell pellet and residual vial.

2. Cervical cytology was read and reported using the British Society of Cervical Cytology's classification. Once the coding for the result and the intention to refer to colposcopy was established in the laboratory, the specimen was identified as a potential trial sample. The residual cell pellet and the residual vial were transferred to refrigerated storage. The only identifying details were a pseudonymous laboratory number.

3. The results of the smear tests were transferred to the CSAD from the laboratory in the usual manner. Women were informed of the test result and colposcopy referral as per Cervical Screening Wales (CSW) standard operating procedures.

4. Colposcopy was undertaken according to normal practice as set out in the current version of the CSW Colposcopy Quality Manual (Cervical Screening Wales, 2012b). It advises that all patients who have had two or more borderline or mildly dyskaryotic smears and have a recognisable atypical transformation zone should have biopsy material submitted for histological interpretation.

5. Once a copy of the written consent had been received in the cytology laboratory, the residual cell pellet was released to Source BioScience (Nottingham, UK) for processing. The sample contained a pseudonymous number. A BD SurePath Plus™ Pap slide was prepared using the original cell pellet and the BD PrepStain Plus™ Processor. The BD SurePath Plus™ slide was returned to originating Laboratory for interpretation by a suitably trained cytologist.

6. The residual vials from samples with consent were forwarded to the HPV research laboratory at Cardiff University for HPV testing.

7. HPV testing was performed by the HPV Research Group at Cardiff University.

8. Once the results of the histology and BD SurePath Plus™ were available, data analysis was performed for the primary outcome. Subsequent analysis will be performed as follow-up data within the two years of recruitment becomes available.

6.2.7 Sample management

The study flow chart in Figure 6.2 details how samples were managed and which department had responsibility for specific data.

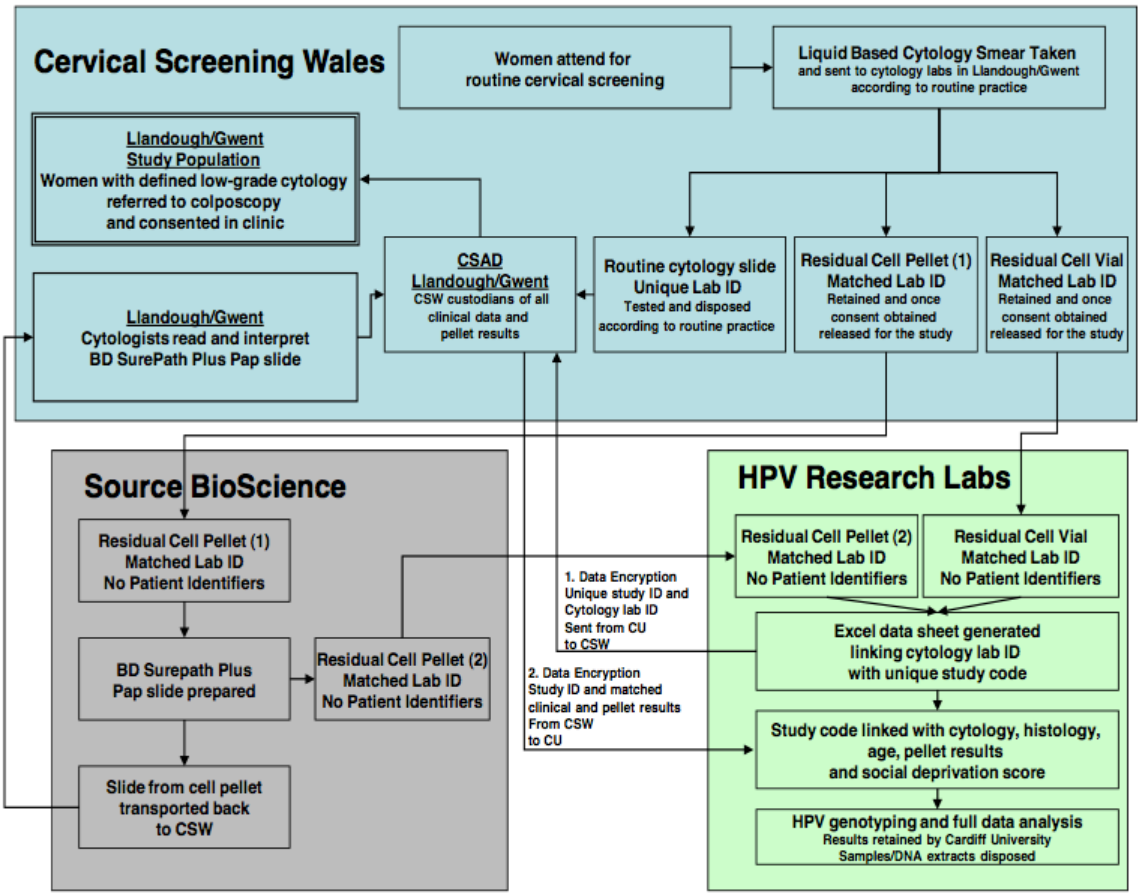

Figure 6.2: Sample and data flow chart for the SuPerLy study

6.2.7.1 Cytological analysis

The cytology slides were analysed in the normal manner, according to CSW policy. They underwent primary screening, checking and referral for consultant reporting as required. When a specimen was identified as suitable for the study, the residual cell pellet and vial were refrigerated and stored by the cytology laboratory, until a copy of the written consent was received. The residual cell pellet was then released to Source BioScience for processing. If the woman did not consent to take part in the study, the sample was disposed of according to usual practice.

6.2.7.2 BD SurePath Plus™

The preparation of the BD SurePath Plus™ slides was performed by the Source BioScience laboratory, which is Good Clinical Practice (GCP) and Continuing Professional Development (CPD) accredited. A BD SurePath Plus™ Pap slide was prepared using the original cell pellet and the BD PrepStain Plus™ Processor. This instrument performed both the immunocytochemistry and Pap staining procedures. The BD SurePath Plus™ slide was returned to the sending laboratory for interpretation by a suitably trained and experienced cytologist. A specimen was called positive when a moderate-to-intense brown nuclear staining was observed in atypical epithelial cells. The residual cell pellet was refrigerated and forwarded to the HPV research laboratory at Cardiff University for HPV testing. Refrigeration was not required for sample transport at any time point.

6.2.7.3 HPV testing

HPV testing was performed using two different platforms: Hybrid Capture 2 (HC2); and Greiner PapilloCheck® microarray assay (PapilloCheck). The methods for these tests can be found in the next chapter.

6.2.8 Endpoints

Primary endpoint: histologically proven high-grade disease (CIN2+) within 26 weeks of first colposcopy visit.
Secondary endpoint: histologically proven high-grade disease during subsequent visits within two years from recruitment.

Histological analysis was carried out according to usual practice at each centre. Histological analysis was subject to routine quality assurance as part of CSW standards and was not, therefore, repeated. This ensured that participation in the trial did not change patient management and that it reflected true clinical practice. High-grade disease was defined as the presence of CIN 2 or worse.

6.2.9 Blinding

All samples that left the cytology laboratory had only the cytology reference number on the sample. Samples from the Royal Gwent cytology laboratory were prefixed with a location code to prevent the possibility of two samples having the same number. Different personnel

analysed the samples for histology, BD SurePath Plus™ and HPV testing and were not aware of any of the results of the other tests. This ensured that no bias was introduced.

6.2.10 Patient confidentiality

The study was conducted in such a way as to preserve patient confidentiality and the study team did not disclose or reproduce any information by which patients could be identified. Each sample was given a unique study number and entered into a database. All electronic files were password protected and hard copies maintained in locked, secure areas. With the exception of consent forms, only anonymised data was stored. The HPV laboratory did not have access to patient identifiable information and CSW did not have access to individual HPV results. Data files exchanged between CSW and HPV laboratory were encrypted.

6.2.11 Ethical considerations

Women participating in this study were not required to provide extra samples and the study was designed so that it would not influence their clinical management. They were only required to permit extra testing on the sample they had already provided for cervical screening. They were also asked to allow the study team to be provided with anonymised results related to their cervical smears and colposcopy for the following two years. The women were sent information in advance and were consented during their visit to the colposcopy clinic, having had an opportunity to discuss the study with an appropriately trained health professional. There was no reason to inform women of their SPP or HPV result because they were being examined as potential screening tests.

6.2.12 Regulatory approvals

The study was reviewed by the South East Wales Regional Ethics Committee and given a favourable ethical opinion on 30th November 2011 (REC Reference number: 10/WSE03/36). Regulatory approval was also granted from the Cardiff and Vale University Health Board and the Aneurin Bevan Health Board to conduct the study. Contracts were set in place to cover all transfer of clinical material and data between the relevant organisations. The study was sponsored by Cardiff University.

6.2.13 Sample size calculations

From prevalence data it was expected that approximately 20% of women with persistent low-grade cytological abnormalities would have high-grade disease (Cervical Screening Wales,

2007). In order to predict how BD SurePath Plus™ might function with clinical material, data relating to its forerunner, BD ProEx™ C was studied. BD ProEx™ C worked using the same principles, the only difference being one of the markers used in the staining process (see section 4.5). In a previous study (Kelly et al., 2006), with a similar prevalence of high-grade disease (21%), the effect of using BD ProEx™ C test was to split the cohort into two groups with levels of risk sufficiently different to warrant different management:

1. High-risk group with positive BD SurePath Plus™ result, comprising 40% of women, with posterior risk = PPV = 44%.
2. Low-risk group with negative BD SurePath Plus™ result, comprising 60% of women, with posterior risk = 1-NPV = 5.2%.

Based on a total of 600 women, the 95% CIs for these proportions were anticipated to be from 38.0 to 50.5% and from 3.2 to 7.8% respectively. These interval widths were regarded sufficiently narrow that such findings, if replicated in this study, would constitute adequate evidence for use of the BD SurePath Plus™ test to guide management. Doing so would reduce referrals to colposcopy by approximately 60%.

One of the secondary outcomes was to compare the PPV and NPV of BD SurePath Plus™ with HPV testing. Assuming that approximately 64% of the samples would test positive for HPV (You et al., 2007) and that HPV testing would have a NPV of approximately 99% (Prinsen et al., 2007, You et al., 2007) the PPV for HPV testing was anticipated to be approximately 30.7%. With a total of 600 samples, the 95% confidence intervals would be 26.3–35.5%. This would not overlap with the 95% confidence interval in PPV expected for BD SurePath Plus™.

6.2.14 Funding

The main source of funding for this study was a large project grant. The grant was awarded from the Emma Jane Demery Fund (an endowment fund administered by Cardiff University) following a competitive peer review process.

6.3 SuPerLy – HIM study

An extension of the SuPerLy study was designed in order to investigate the use of novel biomarkers within the same study population. HPV integration and HPV DNA methylation (HIM) were chosen for further analyses as potential biomarkers.

6.3.1 Aims and objectives

The aim of the study was to investigate viral integration and HPV DNA methylation as potential biomarkers of high-grade disease in women with persistent low-grade cytological abnormalities.

6.3.2 Study design and study population

This study was designed to extend the worth of the SuPerLy study; a large prospective observational study investigating the use of BD SurePath Plus™ and HPV testing in women with persistent low-grade cytological abnormalities. The study population included women with persistent low-grade cytological abnormalities that had been recruited to the SuPerLy study.

6.3.3 Inclusion and exclusion criteria

6.3.3.1 Inclusion criteria

1. The woman had given informed consent and was enrolled in the SuPerLy study.
2. DNA had been successfully extracted – confirmed by B-globin PCR or presence of human gene ADAT1 (an in-built control within the PapilloCheck® assay).
3. The DNA extract was positive for HPV 16 by two independent laboratory tests.

6.3.3.2 Exclusion criteria

Samples were excluded if an inadequate amount of DNA was available.

6.3.4 Ethical considerations

The ethical approval for the SuPerLy study covered the use of DNA in HPV molecular testing and, therefore, no ethical amendment was required for the SuPerLy – HIM study. Obtaining any further consent was also, consequently, considered unnecessary.

6.3.5 Funding

Funding for consumables was awarded following competitive peer review from the Tom Owen Memorial Fund (administered by Cardiff University).

Chapter 7 – METHODS AND MATERIALS

7.1 Introduction

This chapter describes the methods and materials that were used. Details are given for sample processing, DNA extraction techniques, HPV typing assays, HPV integration assays and HPV methylation assays. The statistical methods employed are also provided. At the end of this chapter there is a section on method development. All the laboratory work herein was completed by myself with the exception of: a large proportion of the sample processing; the majority of DNA extraction using the QIAamp MinElute Media Kit; the majority of the B-globin PCR for the QIAamp MinElute extracted DNA; and all of the Hybrid Capture 2 hrHPV testing (this work was performed by Mrs Angharad Edwards, who is one of the research assistants in the HPV Research Group).

7.2 Sample Reception and Login

The samples reached the HPV laboratory via the process outlined in section 6.2.7. They were initially stored in the cold room (4 °C) until ready for processing. The laboratory number and location were recorded on an electronic spreadsheet and a study identification number was allocated to each sample.

7.3 Sample Processing

7.3.1 Sample Processing Procedure

Fifteen millilitre falcon tubes were labelled with the study ID numbers. A 3 ml plastic bulb pipette was used to transfer the entire residual sample from the SurePath pot into the corresponding labelled falcon tube. If samples had evaporated in the tubes, then 5 ml Tris (10 mM pH 7.4) was added prior to transferring to the falcon tubes. The falcon tubes were centrifuged at 5000 rpm for 10 minutes at 4 °C. Three 1.5 ml microfuge tubes were labelled for each sample. Two with just the study number, the other with the study number plus ‘P’ (for pellet). The supernatant was aspirated from the falcon tubes and discarded into a previously prepared bleach pot using a 3 ml plastic bulb pipette. Using a 3 ml plastic bulb pipette, the cell pellet was re-suspended in 2 ml Tris (10 mM pH 7.4). A 1 ml Gilson pipette was used to transfer 1 ml of cell suspension to two of the previously labelled microfuge tubes (one labelled P). These were centrifuged for 5 minutes at 13 200 rpm at room temperature. The supernatant was aspirated again using a 1 ml Gilson and discarded into a previously prepared bleach pot.

The tubes labelled ‘P’ were closed and placed to one side. In the other 1.5 ml microfuge tube the pellet was re-suspended by adding 0.5 ml Tris (10 mM pH 7.4) and repeat pipetting. Finally, 250 µl of cell suspension was transferred to the previously labelled 1.5 ml tube. All the samples were stored in a -80 °C freezer.

7.4 General

7.4.1 Sample Handling

All samples were handled using gloves and laboratory coats at all times. Any materials in contact with clinical samples were disposed of in orange “potentially infective” clinical waste bags or pipette tip boxes.

7.4.2 DNA Handling

DNA was stored in a -80 °C freezer. All PCR and DNA work was carried out in a HEPA filtered PCR cabinet that was sterilised between experiments using 15 minutes of ultraviolet (UV) light. All water used in PCR reactions was DNA grade water that had also received UV light treatment. UV sterilised tubes and plates were used and DNase free filter pipette tips.

7.4.3 Quantification and purity of DNA

Quantification of DNA was performed using a Thermo Scientific NanoDrop 1000 Spectrophotometer (Fisher Scientific UK Ltd, Loughborough, UK) using the manufacturer’s guidelines. 1.5 µl DNA was tested against a blank of DNA free elution buffer/water.

7.4.4 PCR

PCRs were performed on Techne TC-412 and Techne TC-512 thermocyclers (Bibby Scientific Ltd, Staffordshire, UK). The GeneAmp 9700 thermocycler (Perkin Elmer, Beaconsfield, UK) was also used. The thermocyclers and all post-PCR analyses were performed in a separate laboratory. Separate lab coats were worn and hand washing was performed when switching between the laboratories. Deoxyribonucleotide tirphosphates (dNTPs) were made to 2 mM working concentration by adding 10 µl each of dATP, dCTP, dGTP, and dTTP (all in 100 mM stocks) to 460 µl of sterile water, and were stored at -20 °C.

7.4.5 Agarose gel electrophoresis

Agarose gels were prepared with 2% w/v agarose in 1xTBE with 2 µl/100ml of ethidium bromide (10 mg/ml). PCR product was added to an appropriate amount of Orange G loading buffer depending on whether any further reactions were intended for the PCR product (typically the ratio of PCR product to Orange G was 2:1). DNA ladders were also loaded onto the gel to indicate fragment sizes. In most cases 100 bp ladders (Life Technologies, Paisley, UK) were used, however, for DIPS a wide range DNA marker (200 bp–10 Kbp) (Bioland Scientific LLC, California, USA) was preferred. The gels were run at 100–150 V depending on their size, and the degree of separation required dictated the length of time given. The DNA was visualised under UV light on a transilluminator and a digital image was taken.

7.4.6 Positive controls

CaSki was used as the positive control in all assays apart from the DIPS integration assay. SiHa was used for DIPS because it has a limited number of integration events.

7.4.6.1 CaSki

The CaSki cell line was originally derived from a metastatic squamous cell carcinoma of the cervix found within the small bowel mesentery of a 40-year-old Caucasian. The cells contain an integrated HPV16 genome (approximately 600 copies per cell). There are 30 reported point mutations and a single 1 nt deletion in the vast majority of the HPV16 genomes (Meissner,1999). It was sourced from the European Collection of Cell Cultures (ECACC)(catalogue number 87020501).

7.4.6.2 SiHa

The SiHa cell line was originally derived from a squamous cell carcinoma of the cervix found in a 55 year old Asian. SiHa is reported to contain integrated HPV16, with between 1 and 2 copies per cell. Integration is associated with disruption of the HPV16 genome and a deletion at nt 3133–3385, there are also deletions at nt 3460–3512 and nt 7757–7794 (Agoff et al., 2003). It was sourced from the American Type Culture Collection (ATCC – number HTB-35).

7.4.6.3 HPV16 Plasmid Vector

HPV16, complete intact genome, in vector was obtained from the HPV research group at Manchester University. Genomic DNA of HPV16 was originally obtained by the HPV research group from the World Health Organisation, inserted into vector pBR322 (Sutcliffe, 1979) at BamHI restriction site (357bp), transformed into HB101 E. coli HB101 and cultured.

7.5 DNA Extraction

The majority of molecular assays cannot be performed on a cell that is structurally intact: the DNA content must first be separated from the remainder of the cell. For the SuPerLy study a commercially available kit was used for this procedure. The reasons for this were to be as consistent and reliable as possible and to use approved systems that could also be used in clinical laboratory settings. The QIAamp MinElute Media Kit (Qiagen, Hilden, Germany) was chosen because it is suitable for use with liquid media containing nucleic acids.

The QIAamp MinElute Media Kit is designed to ensure that there is no sample-to-sample cross contamination and allow safe handling of potentially infectious samples. There are four steps to the procedure including: lyse, bind, wash, and elute. The samples are lysed at high temperature denaturing conditions using proteinase K and two lysis buffers. The buffers increase lysis efficiency and ensure inactivation of RNases. Binding of the nucleic acids to the QIAamp MinElute column membrane is facilitated by adding ethanol to the lysates followed by high-speed centrifugation. Two wash steps are required to remove any contaminants from the membrane. Finally, the pure nucleic acids are eluted into a buffer.

One of the most simple and inexpensive ways of DNA extraction involves proteinase K. Contaminating proteins, including a number of nucleases, are degraded by the addition of proteinase K. This method was applied to some of the residual clinical material.

7.5.1 QIAamp MinElute DNA Extraction Procedure

The samples were removed from the -80 °C freezer and defrosted prior to extraction. The samples were extracted in batches of 24 including one positive control (CaSki) and one negative control (water). If the pellet sample was used then 500 µl Tris (10 mM pH 7.4) was added to resuspend the pellet. A 250-µl aliquot of each sample was pipetted using a 1 ml Gilson pipette into a 1.5 ml microfuge tube. Into each sample 80 µl of ATL buffer and 20 µl of Qiagen proteinase K (Qiagen, Hilden, Germany) was added and the samples were vortexed for 10 seconds. They were then incubated at 56 °C for 30 minutes on a heated block. During this incubation period the samples were vortexed every 10 minutes.

Carrier RNA was prepared by mixing 310 µl AVE Buffer to 310 µg lyophilized carrier RNA to obtain a solution of 1 µg/µl. This was divided into 150 µl aliquots and stored at -20 °C. For each

sample 3 µl of the carrier RNA/AVE Buffer solution was first mixed with 300 µl AL Buffer and then 250 µl of this solution was added to the sample. The samples were vortexed for 10 seconds and incubated at 70 °C for 15 minutes on a heated block.

300 µl of ethanol (96–100%) was added to the samples and then they were vortexed for 15 seconds. The ethanol lysate was incubated for 5 minutes at room temperature (15–22 °C) (RT). Half of the lysate was pipetted into the QIAamp MinElute column and centrifuged at 8 000 rpm for 3 minutes (RT). The eluate that had collected in the collection tube was discarded and the process was repeated with the remaining lysate.

Into each column 750 µl of buffer AW2 was added and the sample was then centrifuged as above. The eluate was again discarded from the collection tube. A further 750 µl of ethanol (96–100%) was added and same process of centrifugation and elution repeated. The membrane in the QIAamp MinElute column was then dried completely by centrifuging at 14 000 rpm for 3 minutes (RT). The columns were then placed into clean appropriately labelled 1.5 ml microfuge tubes and placed onto the heated block and incubated at 56 °C for 3 minutes with the lid of the column open to evaporate residual ethanol. Into the centre of each membrane 120 µl AVE Buffer was added and then left to incubate for 1 minute at room temperature with the lid closed. Finally, the samples were centrifuged at 14 000 rpm for 1 minute. The eluate, now containing pure nucleic acids, was stored at -80 °C until required for further analysis.

7.5.2 Proteinase K DNA Extraction Procedure

In this study the cell pellet was used in the proteinase K extraction. The pellet was resuspended with 0.5 ml of Tris (10 mM pH 7.4) and centrifuged at 13 000 rpm for 3 minutes. The supernatant was discarded and the pellet was again resuspended in 100 µl of Tris (10 mM pH 7.4). One hundred microlitres of positive (CaSki) and negative controls were prepared. To all the samples 10 µl of Proteinase K (10 mg/ml, stored at 4 °C) was added and vortexed. The samples were then placed in a shaking incubator at 56 °C for 3 hours. After that the samples were placed in a preheated block at 80 °C for 10 minutes. The tubes were transferred into a previously chilled rack (-20 °C) and placed in the fridge (4 °C) for 10 minutes. Following that the samples were spun in the refrigerated microfuge (4 °C) at 13 000 rpm for 10 minutes. Finally, the supernatant containing the DNA, was pipetted into labelled 1.5 ml tubes and stored in the freezer at -80 °C.

7.5.3 ß-globin PCR

ß-globin, the human housekeeping gene was PCR amplified to ensure that the there was amplifiable human DNA and serve as an indicator of successful DNA extraction. The PCR was set up according to Table 7.1. The conditions for the thermocycler are shown in Table 7.12.

Reagents	Volume per reaction (µl)
dNTP 2 mM	2.5
10x PCR buffer	2.5
$MgCl_2$ 50 mM	0.875
PCO3 Primer 5 µM	2.5
PCO5 Primer 5 µM	2.5
Water	9.025
Taq 5 U/µl	0.1
DNA	5
Total volume	25

Table 7.1: Reagents for ß-globin PCR

7.6 HPV Typing

HPV genotyping was performed using Greiner PapilloCheck® Microarray (PapilloCheck, Greiner Bio-One GmbH, Germany). High-risk HPV typing was performed using the commercial assay: Hybrid Capture 2 (HC2; Qiagen, Hilden, Germany).

7.6.1 PapilloCheck Microarray Assay

PapilloCheck is a broad-spectrum PCR-based method that can detect and differentiate 24 HPV types (17 hrHPV types (HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 70, 73, 82), and 7 lrHPV types (HPV6, 11, 40, 42, 43, 44/55, 70)). It uses a consensus primer set that targets a 350bp fragment of the E1 region of HPV and a fragment from the human ADAT1 gene. The human gene fragment is also amplified in order to provide an internal PCR control. The amplified products are hybridised to specific DNA probes on the DNA chip. During hybridisation the bound DNA is fluorescence-labelled with Cy5. The amplification quality is determined by PCR product binding to 5 control spots and their subsequent signal strength on the array. After hybridization and subsequent washing, the PapilloCheck® DNA chip is scanned with the CheckScanner™ at excitation wavelengths of 532 and 635 nm.

7.6.1.1 PapilloCheck procedure

The PCR was set up using the PapilloCheck® MasterMix. This includes all the required buffers, $MgCl_2$, dNTPs, DNase-free water and fluorophore-labelled primers. In addition, Uracil-N-DNA Glycosylase (UNG) (Fermentas GmBH, Germany) is used to eliminate carry over contaminations from previous PCR reactions (Longo et al., 1990). It was diluted 1:200 in water and 1 µl was added to each reaction. The heat-stable polymerase HotStarTaq (Qiagen, Hilden, Germany) was used, as recommended by the manufacturers. The PCR was performed in a total volume of 25 µl, with each component added as shown in Table 7.2.

Reagents	**Volume per reaction (µl)**
PapilloCheck MasterMix	19.8
HotstarTaq Polymerase (5 U/µl)	0.2
Uracil-N-Glycosylase	1
DNA	5
Total volume	25

Table 7.2: Reagents for PapilloCheck PCR

For the PCR reaction a GeneAmp 9700 (Perkin Elmer, Beaconsfield, UK) thermocycler was used with the conditions programmed as in Table 7.12. Following the PCR reaction the washing solution were prepared using 140 ml of double-distilled water with 14 ml PapilloCheck® Buffer A and 1.75 ml PapilloCheck® Buffer B. This was divided equally into three 50 ml falcon tubes and labelled 1, 2, and 3. Wash buffer 2 was heated in a water bath at 50°C for 20 minutes. In a new 96-well PCR plate 30 µl of PapilloCheck® Hybridisation Buffer was mixed with 5 µl of PCR product and spun down. Onto each compartment of the chip 25 µl of the hybridisation mix was added using a multipipette. It was very important to avoid any air bubbles that could affect the reading of the chip (see Figure 7.1).

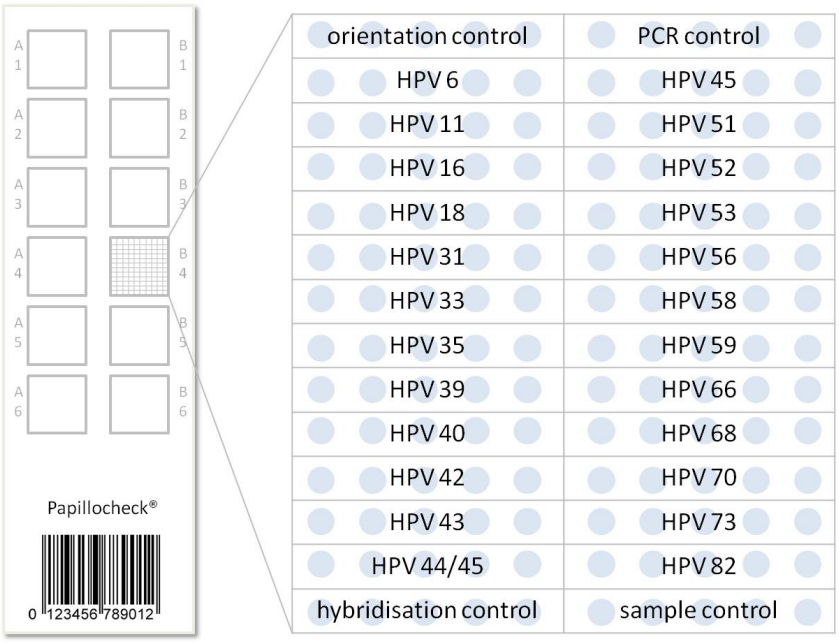

Figure 7.1: Diagram showing the PapilloCheck® chip layout.

The chip was incubated at room temperature in a humid atmosphere within a hybridisation chamber for 15 minutes (Figure 7.2).

Figure 7.2: PapilloCheck® chips loaded in the hybridisation chamber

The details of all the samples on the chip were entered into the CheckReport™ software. Following this the chip was loaded into the CheckScanner™ and the laser was initiated to read the chip. The chip was scanned at two wavelengths 532nm (green) and 635nm (red). A report was generated using the software and exported into an Excel spreadsheet.

7.6.2 Hybrid Capture 2

The Hybrid Capture 2 (HC2) HPV DNA test is commercially available and has been approved by the US Food and Drug Administration (FDA) and has a CE mark. It is a nucleic acid hybridisation assay with signal amplification that uses microplate chemiluminescence for the qualitative detection of 13 high-risk types (16/18/31/33/35/39/45/51/52/56/58/59/68) and 5 low-risk types (6/11/42/43/44). The test uses full genome probes to prevent false negatives caused by gene deletions. The specific HPV RNA probe is hybridized to the target DNA contained within infected specimens. The RNA:DNA hybrids are captured onto the surface of a microplate well coated with antibodies specific for RNA:DNA hybrids. The immobolised hybrids are then exposed to specific RNA:DNA alkaline phosphatase conjugated antibodies. Multiple conjugated antibodies bind to each hybrid resulting in substantial signal amplification, which is detected by using a chemiluminescent substrate. When the substrate is cleaved by the bound alkaline phosphatase, light is emitted and measured on a luminometer as relative light units (RLUs). The intensity of the light emitted indicates the presence or absence of target DNA in the specimen. An RLU measurement equal to or greater than the cutoff value denotes the presence of high or low-risk HPV DNA sequences in the specimen. An RLU measurement less than the cutoff value denotes the absence of the specific HPV DNA sequence or the levels of HPV DNA are below the detection limit of the assay. The details of the procedure can be found in Appendix III.

7.7 Integration assays

Two different assays were used to investigate viral integration in HPV 16 positive samples. E2 PCRs were performed to target and assess the physical state of the E2 ORF of the HPV genome. On a selection of samples a more labour intensive but much more specific assay, Detection of Integrated Papillomaviruses Sequences (DIPS), was performed.

7.7.1 E2 PCR

The E2 ORF was assessed by the E2 PCR assay described by Collins et al (2009). It involves a series of PCR reactions to amplify overlapping fragments that span across the E2 region to detect E2 disruption and potential integration. To control for presence of HPV a set of primers that amplify the E6 ORF are also used. Disruption to the E2 regions and potential integration was suggested if E6 primers produced an amplicon and one or more primer E2 ORF sets failed to produce an amplicon. 50 ng of DNA was added to the other reagents as shown in Table 7.3. The PCR conditions are shown in Table 7.12, and the primers in Table 7.13.

Reagents	Concentration	Volume (µl)
DNA	10 ng/µl	5
PCR Buffer (with 15mM $MgCl_2$)	10x	2
Forward Primer	10 µM	1
Reverse Primer	10 µM	1
HotstarTaq	5 U/µl	0.1
dNTP	2 mM	2
Water		6.9
Total volume		20

Table 7.3: Reagents for E2/E6 PCRs

7.7.1.1 L1 and E1 PCR

These assays were performed in the same way as E2 and E6 PCRs using the same combination of reagents (Table 7.3). The PCR conditions and primers can be found in Table 7.12 and Table 7.13.

7.7.2 Detection of Integrated Papillomaviruses Sequences

DIPS (Luft et al., 2001) was used to determine the integration status of HPV in the SuPerLy–HIM study. DIPS involves a single side specific ligation mediated PCR and amplifies a sequence of human genome that has integrated HPV alongside. Genomic DNA is first digested with a restriction enzyme. The digested DNA is then ligated to a double stranded adapter primer. A linear PCR uses HPV specific primers to amplify relevant fragments. This is followed by a nested PCR using HPV specific primers and a primer specific to the adapter (Figure 7.3). The nested PCR products are separated by gel electrophoresis and selected fragments are then purified and sequenced. A control primer that binds to a genomic locus on chromosome 21 ensures the assay is performing optimally.

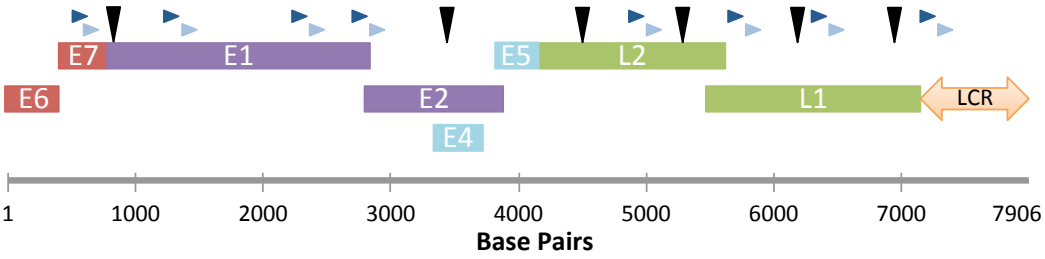

Figure 7.3: Schematic representation of the DIPS methods.

The black arrows represent restriction enzyme cut sites within the HPV16 genome. The blue arrows represent the primer locations for the linear (dark blue) and nested (light blue) PCRs.

7.7.2.1 DIPS procedure

Two aliquots of 1.2 µg DNA were digestin a reaction with restriction enzyme Sau3AI (New England Biolabs, Beverly, MA, USA), BSA, Buffer 1 and water (Table 7.4). HPV plasmid, SiHa, and water were used as controls. The reactions were incubated on a thermocycler at 37 °C for two hours and then incubated at 65 °C for 20 minutes to inactivate the enzyme.

Reagents	Concentration	Volume (µl)
DNA	1.2 µg	(up to 17.3)
Sau3AI	10 U	2.5
BSA	100x	0.5
Buffer 1	10x	5.0
Water		to 50
Total volume		50

Table 7.4: Reagents for the digestion reaction

The digestion adapter was pre-annealed by mixing the oligonucleotides AL1 and Sau3AI AS in 66 mM Tris HCl (pH 7.4)(Table 7.5). The adaptor mixture was heated to 90 °C for two minutes, slowly cooled overnight to 4 °C in a thermocycler, and stored at -20 °C.

Reagents	Concentration	Volume (µl)
AL1 primer	100 µM	25
Sau3AI AS primer	100 µM	25
Sterile TrisHCL	66 mM	50
Total volume	25 µM	100

Table 7.5: Reagents for making 25 µM Sau3AI adapter stock solution

The ligation reaction was prepared by mixing the digested DNA with ligase buffer, T4 Ligase (New England Biolabs, Beverly, MA, USA), Sau3AI adapter and water (Table 7.6). The samples were incubated at room temperature (20–25 °C) for two hours and then inactivated by heating at 65 °C for ten minutes.

Reagents	Concentration	Volume (µl)
Digested DNA		50
Ligase Buffer	10x	6
T4 Ligase	400 U/µl	1
Sau3AI Adapter	25 µM	1.2
Water		1.8
Total volume		60

Table 7.6: Reagents for ligation reaction

The primary linear PCR was set up using eight different HPV16 primers and the control primer. Nine mastermixes were made up using PCR buffer, dNTPs, HotstarTaq, water and one of the HPV specific or control primers (Table 7.7 and Table 7.13). The samples were placed in the thermocycler and programmed as described (Table 7.12).

Reagents	Concentration	Volume (µl)
Ligated DNA		3
PCR Buffer (with 15mM $MgCl_2$)	10x	2.5
dNTP	2 mM	2.5
HPV PCR 1 Primer	10 µM	0.5
HotstarTaq	1 U	0.125
Water		16.375
Total volume		25

Table 7.7: Reagents for linear PCR

For the nested PCR reaction the mastermix was made up in the main laboratory and transferred in a sealed PCR plate to the post-PCR laboratory in order to add the primary PCR product. The mastermix was made using PCR buffer, dNTPs, Hotstar Taq, AP1 primer, water and one of the secondary PCR 2 primers (Table 7.8, Table 7.12, and Table 7.13).

Reagents	Concentration	Volume (µl)
Linear PCR product		2
PCR Buffer (with 15mM $MgCl_2$)	10x	2.5
dNTP	2 mM	2.5
HPV PCR 2 Primer	10 µM	1
AP1 Primer	10 µM	1
HotstarTaq	1 U	0.125
Water		15.875
Total volume		25

Table 7.8: Reagents for nested PCR

For each sample, 5 µl of PCR product was mixed with 2.5 µl of Orange G on parafilm and loaded into a prepared 2% w/v agarose TBE gel with 2 µl/100ml of ethidium bromide (10mg/ml). The gel was run at 120 V for at least one hour or until adequate separation of the PCR products was achieved.

7.7.2.2 Sequencing

Amplicons that differed from the HPV 16 plasmid control were cut from the gel using a sterile disposable scalpel and a UV transilluminator. Suitable PPE was worn to protect the operator from the UV light. The gel fragments were placed into microcentrifuge tubes and purified using the illustra™ GFX™ PCR DNA and Gel Band Purification Kit (GE Healthcare Life Sciences, Buckinghamshire, UK). The fragments were first dissolved using heated buffers. The PCR product was then bound to a membrane and washed several times before finally being eluted into 15 µl of sterile water. This procedure was done for all fragments of interest as per manufacturer's instructions.

Once purified, 5 µl of the product was run on a 2% w/v agarose gel as before to confirm successful PCR product purification. The successfully purified PCR products were sent along with the relevant nested primer for Sanger sequencing at Source BioScience UK Ltd (Nottingham, UK) according to the requirements in Table 7.9.

Reagents	Concentration	Volume (µl)
PCR products	1 ng/µl/100 bp	5
Primers	3.2 µM	5

Table 7.9: The sequencing requirements for purified PCR products.

7.7.2.3 Data analysis

The chromatogram that was produced for each sample was analysed using a commercially available sequence editing application (4Peaks, Nucleobytes B.V., Netherlands) to ensure accuracy of the sequence. When the amplitude of the signal fell off and became equivalent to the background noise any further base calling was stopped and not considered for further data analyses. Sequences of poor quality where there was multiple overlapping sequences or high background noise were removed from subsequent analyses.

Sequence analysis was performed against human genomic sequences and the HPV16 reference sequence (GenBank ID NC001526.1) using the online National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST). To identify the exact region of the human genome where integration events occurred the University of California Santa Cruz (UCSC) Blast Like Alignment Tool (BLAT) was used.

7.8 Methylation assay

This technique involves treating methylated DNA with sodium bisulfite, which converts unmethylated cytosines into uracil. Methylated cytosines remain unchanged during the treatment. Once converted, the methylation profile of the DNA can be determined by PCR amplification followed by DNA sequencing.

7.8.1 Bisulfite Treatment

The treatment process was carried out using the EZ DNA Methylation™ Kit (Zymo Research Corporation, California, USA). Before starting the wash buffer was prepared by adding 24 ml of ethanol (100%) to 6 ml of M-Wash Buffer. The CT Conversion Reagent was prepared by adding 750 µl of sterile water and 185 µl of M-Dilution Buffer to one tube of CT conversion reagent. For each reaction 500 ng of DNA was made up with water to 32.5 µl and 7.5 µl of M-Dilution Buffer was added. CaSki and water were used for controls. The mixture was incubated at 42 °C for 30 minutes in a preheated block. To each sample 97.5 µl of prepared CT Conversion

Reagent was added. They were then incubated at 50 °C for 12–16 hours in the heated block under a blackout lid. Following incubation, the samples were cooled on ice for 10 minutes. A Zymo-Spin™ IC Column was labeled and prepared for each sample and control by adding 400 µl of M-Binding Buffer into the column and by placing in a collection tube. The samples were added to the corresponding columns and mixed by inversion before centrifuging at 13 000 rpm for 30 seconds. The flow-through was discarded and 100 µl of M-Wash Buffer was added followed by a further 30-second centrifuge. In to each column 200 µl of M-Desulphonation Buffer was added and left to incubate at room temperature for 15–20 minutes before another 30-second centrifuge. The samples were then washed twice by adding 200 µl of M-Wash Buffer, centrifuging for 30 seconds and repeating. Finally, the columns were placed into pre-labeled 1.5 ml microcentrifuge tubes, 10 µl of M-Elution Buffer was added and a 30-second centrifuge was performed to elute the treated DNA.

7.8.2 Pyrosequencing PCR

The BS treated DNA was first diluted 1:10. The PCRs were setup in 50 µl reactions and varied slightly depending on the primers used. The primer sequences and assay details were provided by Dr Bryant (Cardiff University, personal communication, March 2012). The locations of the CpG sites tested are shown in Table 7.10 and the details of the primers used are in Table 7.13. ZymoTaq™ PreMix (Zymo Research Corporation, California, USA), which included a hot start Taq polymerase, together with 2 µl of 1:10 BS treated DNA for each 50-µl PCR reaction. The reagents and PCR conditions are shown in Table 7.11 and Table 7.12, respectively.

Name of CpG site	Location of CpG sites in HPV16 genome (nt)
L1L2-1, L1L2-2, L1L2-3, L1L2-4	5615, 5609, 5606, 5600
E2-1, E2-2, E2-3, E2-4, E2-5, E2-6, E2-7, E2-8	3411, 3414, 3416, 3432, 3435, 3447, 3461, 3472

Table 7.10: Details of the CpG sites tested in the E2 and L1L2 regions of the HPV16 genome.

Reagents	Concentration	Volume (µl) for L1/L2	Volume (µl) for E2
Bisulfite treated DNA		2	2
Forward Primer	5 µM	2	2
Reverse Primer	5 µM	2	2
$MgCl_2$	25 mM	3	2
ZymoTaq premix		25	25
Water		16	17
Total volume		50	50

Table 7.11: Reagents for pyrosequencing PCR

7.8.3 Pyrosequencing procedure

Before pyrosequencing PCR products were checked using gel electrophoresis as previously described (see 7.7.2.1). Samples where no amplicon was produced (using CaSki as positive control) were not taken forward for sequencing. Samples with appropriate bands were diluted by adding 27 µl of PCR product to 13 µl of sterile water. The PyroMark Q96 Vacuum Prep Workstation was prepared with the appropriate solutions (1:10 diluted wash buffer, 70% ethanol, water (deionised and autoclaved), denaturation buffer) and the vacuum prep tool head was rinsed thoroughly with water for 20 seconds (Figure 7.4).

Details of the reaction and samples were input into the PyroMark CpG Software. Pyrosequencing PCR was performed using a biotin labelled primer. PCR products were immobilised by adding 1.75 µl of streptavidin sepharose bead suspension and 38.25 µl of PyroMark Binding Buffer per reaction. The mixes were then shaken on a shaking hot plate for at least 5 minutes (1400 rpm, 22 °C). Sequencing primers were made up by diluting (per reaction) 1.5 µl of 10 µM sequencing primer with 43.5 µl of PyroMark Annealing Buffer. 45 µl of sequencing primer mix was dispensed into each well of a pyrosequencing (PSQ) reaction plate and the plate was placed into the correct compartment of the vacuum workstation (Figure 7.4).

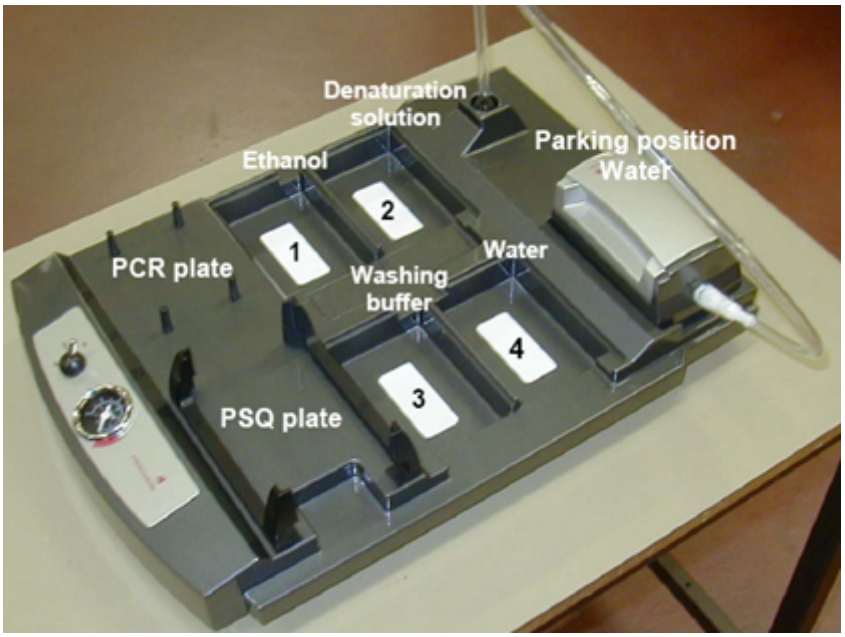

Figure 7.4: The layout of the PyroMark Q96 Vacuum Prep Workstation showing the correct reagent positions.

The PCR product/sepharose bead mix was removed from the hotplate and placed into the correct position on the vacuum workstation. Within 30 s of cessation of shaking, the vacuum was applied and the sepharose beads with immobilised PCR products were captured by slowly lowering the vacuum prep tool into the PCR plate. The vacuum prep tool was then placed into each the following solution trays (in order) for 5 seconds each; 70% v/v ethanol solution, PyroMark Denaturation Solution, 10% v/v PyroMark Washing Buffer. The prep tool was held vertically and any residual fluid was aspirated. The prep tool was lowered into the PSQ reaction plate containing sequencing primer and the vacuum switch was closed whilst hovering above the solution and then agitated to release the captured PCR products. The PSQ plate containing beads and sequencing primer was heated at 80 °C for two minutes then cooled to room temperature. The PyroMark Gold Q96 Reagent kit contained lyophilised enzyme and substrate pellets as well as dNTP mixes. Enzyme and substrate were reconstituted with the volume of water specified on the container 10 minutes prior to use. The PSQ96 Reagent Cartridge was filled using the volumes specified by the pyrosequencing software and the wells specified in Figure 7.5.

The cartridge and reaction plate were placed into the PyroMark Q96 ID Instrument and the run was started. At the end of the run, the data was analysed automatically and a pyrogram was produced for each pyrosequencing reaction (e.g. Figure 7.6).

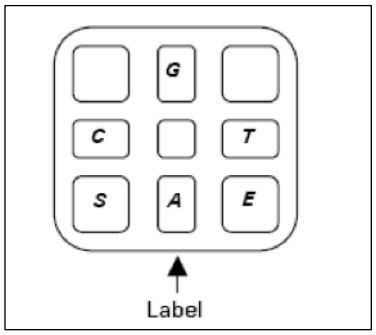

Figure 7.5: A diagram of the PSQ96 Reagent Cartridge detailing compartments of the cartridge where each reagent was loaded.

The cartridge label was positioned at the front, facing the operator. (S = substrate mix; E = enzyme mix; G = dGTP solution; T = dTTP solution; A = dATP solution; C = dCTP solution).

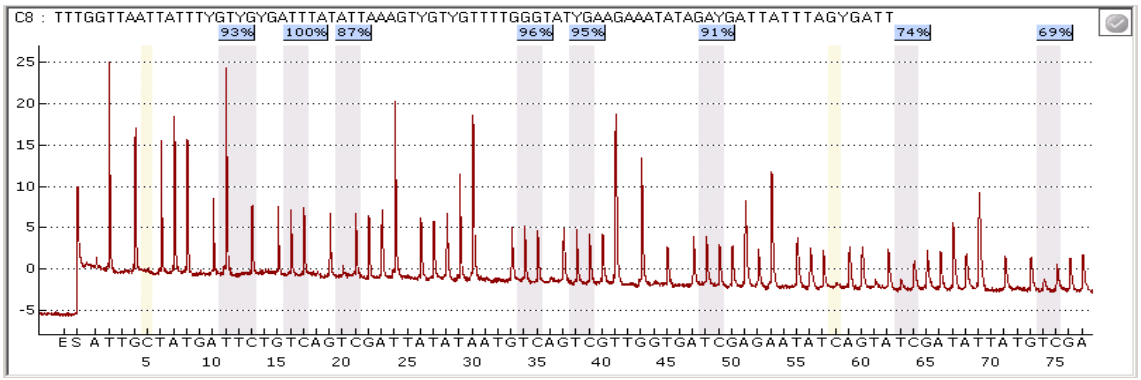

Figure 7.6: Example pyrogram for CaSki showing methylation at eight CpG sites in E2.

Along the X-axis is the dispensation order of the bases. The Y-axis is Relative Light Units (RLU). The higher the peak, the greater the number of nucleotides that are incorporated at that point of the sequence. The grey shaded areas represent the CpG sites. The heights of the C and T peaks are compared to determine the % methylation of that CpG, which is recorded in the coloured box above each CpG. The colour of the box indicates the data quality: blue, good; yellow, check; red, inadequate.

The internal controls within the software were used to select which samples could be included in subsequent analysis. The samples may fail any number of CpGs and the most common reason for failing was a low signal-to-noise ratio. Samples were repeated in duplicates to improve the chances of getting adequate data for all CpGs. At the end of each experiment the cartridge and vacuum workstation were cleaned thoroughly with water in order to prevent any blockages, which could have significant effects on subsequent runs.

7.9 PCR conditions and primers

7.9.1 PCR conditions

PCR	Initial denaturation		Amplification conditions			Final extension		
	Time(m)	Temp(°C)	Time(s)	Temp(°C)	Cycles	Time(s)	Temp(°C)	Cycles
B-globin	4	95	30	95	40	45	72	1
			30	55				
			30	72				
PapilloCheck	20	37	30	95	40	30	95	15
	15	95	25	55		45	72	
			45	72				
E2 / E6	15	95	30	95	40	60	72	10
			30	55/58*				
			60	72				
L1	15	95	30	94	40	60	72	10
			30	59				
			60	72				
E1	15	95	30	94	40	60	72	7
			30	60				
			180	72				
DIPS–Linear	15	95	30	94	40	60	72	7
			30	66				
			180	72				
DIPS–Nested	15	95	30	94	30	60	72	7
			30	66				
			180	72				
Pyro-sequencing	10	95	30	94	40	60	72	10
			45	48 / 54^				
			30	72				

Table 7.12: The PCR conditions

* The temperature for E6 and all the E2 primer sets except the E2 FL primers was 55 °C. ^ The temperature for E2 was 48 °C and for L1L2 was 54 °C.

7.9.2 Primers

B-globin primers	
BGPCO3	ACACAACTGTGTTCACTAGC
BGPC05	GAAACCCAAGAGTCTTCTCT
E2 primers	
E2 1 Forward	AGGACGTGGTCCAGATTAAG
E2 1 Reverse	TCAAACTGCACTTCCACTGT
E2 2 Forward	TAACTGCACCAACAGGATGT
E2 2 Reverse	GCCAAGTGCTGCCTAATAAT
E2 3 Forward	ATCTGTGTTTAGCAGCAACG
E2 3 Reverse	TAAATGCAGTGAGGATTGGA
E2 4 Forward	ACAGTGCTCCAATCCTCACT
E2 4 Reverse	TCACGTTGCCATTCACTATC
E2 5 Forward	GGCATTGGACAGGACATAAT
E2 5 Reverse	CAAAAGCACACAAAGCAAAG
E2 Full length (FL) Forward	TTAAGTTTGCACGAGGACGA
E2 Full length (FL) Reverse	CGCCAGTAATGTTGTGGATG
E6 primers	
E6 Forward	GAACAGCAATACAACAAACC
E6 Reverse	GATCTGCAACAAGACATACA
E1 primers	
E1 1 Forward	CTAGGAATTGTGTGCCCCATCTG
E1 1 Reverse	CTTTGTATCCATTCTGGCGTGTCT
E1 2 Forward	GATAGAGCCTCCAAAATTGCGT
E1 2 Reverse	ACGTTGGCAAAGAGTCTCCATC
L1 primers	
L1 Forward	TGTGCTGCCATATCTACTTCAGAAACTAC
L1 Reverse	TAGACCAAAATTCCAGTCCTCCAAA
DIPS – Adapter and adapter specific primers	
AL1 primer	GGGCCATCAGTCAGCAGTCGTAGCCCGGATCCAGACTTACACGTTG
AP1 primer	GGCCATCAGTCAGCAGTCGTAG
Sau3AI AS primer	PO_4–GATCCAACGTGTAAGTCTG–NH_2
DIPS – Primary PCR primers	
HPV PCR 1 primer 1(P1)	ACAAAGCACACACGTAGACATTCG
HPV PCR 1 primer 2(P2)	AGTAATAAATCAACGTGTTGCGATTG
HPV PCR 1 primer 3(P3)	TTTGGTTACAACCATTAGCAGATGC
HPV PCR 1 primer 4(P4)	GTGCCAACACTGGCTGTATCAAAG
HPV PCR 1 primer 5(P5)	TACCAATTCTACTGTACCTAATGCCAG
HPV PCR 1 primer 6(P6)	ACTTATTGGGGTCAGGTAAATGTATTC
HPV PCR 1 primer 7(P6)	AGTAGATATGGCAGCACATAATGAC
HPV PCR 1 primer 8(P8)	GTTGGCAAGCAGTGCAGGTCAG
Control PCR 1 primer	TTCTCTATGTGCGTTCTCTCCCTG
HPV PCR 1 16for4	GTTTGCACGAGGACGAGGAC
HPV PCR 1 16for5	CAGAGCCAGACACCGGAAAC

Table 7.13: Primer sequences

DIPS – Nested PCR primers	
HPV PCR 2 primer 1	CGTACTTTGGAAGACCTGTTAATGG
HPV PCR 2 primer 2	GGACTTACACCCAGTATAGCTGACAG
HPV PCR 2 primer 3	AATAGGTATGTTAGATGATGCTACAG
HPV PCR 2 primer 4	ACAAGCAATTGAACTGCAACTAACG
HPV PCR 2 primer 5	GAGGTTAATGCTGGCCTATGTAAAG
HPV PCR 2 primer 6	CCCTGTATTGTAATCCTGATACTTTAGG
HPV PCR 2 primer 7	TGCGTGTAGTATCAACAACAGTAAC
HPV PCR 2 primer 8	TTAAACCATAGTTGCTGACATAGAAC
Control PCR 2 primer	CAAACTCCAGGTCTCCAACCAG
HPV PCR 2 16for4	GACGAGGACAAGGAAAACGATGGAG
HPV PCR 2 16for5	GGAAACCCCTGCCACACCAC
Pyrosequencing primers	
E2 Forward	GTGAAATTATTAGGTAGTATTTGG
E2 Reverse	BTN–CAACAACTTAATAATATAACAAAAA
E2 Sequencing	GTGAAATTATTAGGTAGTA
L1L2 Forward	BTN–TTATTGTTGATGTAGGTGATTT
L1L2 Reverse	CCCAATAACCTCACTAAACAACC
L1L2 Sequencing	TAACCTCACTAAACAACCAA

Table 7.13 (continued): Primer sequences.

BTN, refers to the site of a biotin label

7.10 Statistical analysis

Statistical analysis was performed using Microsoft Excel (Microsoft, Washington, USA) and IBM SPSS Statistics (IBM Corporation, New York, USA). A number of definitions of statistical terms that are used in this thesis are given in Table 7.14.

Term	Definition
H_0 – null hypothesis	The default position; i.e., no relationship/difference between groups/means
H_1 – alternative hypothesis	The alternative position; i.e., a significant relationship/difference between groups/means
P-value	The Probability value; most of the time if $P<0.05$ reject the H_0 and accept the H_1
Type 1 error	Rejection of H_0 despite it being true
Type 2 error	Acceptance of H_0 despite it being false
Sensitivity	The proportion of people with a condition that test positive
Specificity	The proportion of people without a condition that test negative
Positive predictive value	The proportion of people that test positive and have the condition
Negative predictive value	The proportion of people that test negative and do not have the condition

Table 7.14: Definition of some statistical terms.

7.10.1 Statistical techniques

Parametric statistical techniques are considered more powerful, however, there are a number of assumptions that the data must first fulfil (Pallant, 2010). Two of the commonly required assumptions are that the data is normally distributed and the variances of any groups being studied are equal. There are a number of ways to test these assumptions, for this thesis Shapiro-Wilk's test was used to test normality and Levene's test was used to assess homogeneity of variances. In both cases a *P*-value >0.05 meant the assumptions had been met. In addition to using Shapiro-Wilk's test, the histograms for the data were also scrutinised because they can sometimes be a more reliable indicator of normality (Pallant, 2010). When the assumptions have not been met there are three options to consider. Firstly, the data (x) may be transformed, for example, by using $Log_{10}(x)$ or x^2, and then rerunning the tests. Alternatively, there are non-parametric techniques that have less stringent requirements, but consequently, are considered less accurate. Finally, there is the option accepting that there are violations and interpreting the results with caution.

7.10.1.1 T-tests and Mann-Whitney U-tests

The Student's T-test is a parametric test used to compare whether the means of two groups of independent data are equal. The T-test assumes that the data is normally distributed and has equal variances. When the assumptions were met a *P*-value ≤0.05 was used reject the H_0. A *P*-value of 0.05 means that there is a 5% probability of a Type I error occurring, which is the commonly accepted significance level used in clinical research.

Where the assumptions for the T-test were not met the non-parametric equivalent Mann-Whitney U-test was used. This test compares the median of two groups.

7.10.1.2 One-way ANOVA and Kruskall-Wallis tests

These tests are similar to the T-test and Mann-Whitney U-test but can be used to compare multiple groups. The same assumptions were required for the parametric ANOVA test and where these were not met the non-parametric alternative Kruskall-Wallis test was used. When a significant result was found further testing was required In order to identify between which specific groups it applied to. This involved performing Mann-Whitney U-tests with two groups at a time. However, when such multiple comparisons are made the chances of a Type I error occurring increases (Pallant, 2010). The Bonferroni correction can be used in these circumstances. The Bonferroni correction reduces the *P*-value to compensate for the multiple analyses by dividing the original *P*-value by the number of comparison made.

7.10.1.3 Chi-square tests and Fisher's Exact tests

This test was used to explore the relationship between groups of categorical data. It tests the H_0 that the observations based on two variables are not independent of each other. Yates' correction for continuity was used when Chi-square was calculated on a 2 by 2 table (contingency table) in order to compensate for the overestimation of the chi-square value (Pallant, 2010). One of the assumptions for chi-square tests was that the minimum value in any of the cells in the contingency table should be ≥5 (or at least 80% of cells have frequencies ≥5). If this assumption was not met then the Fisher's Exact Probability test, which is more accurate when samples sizes are small (Pallant, 2010).

7.10.1.4 Reproducibility and agreement

Two statistical techniques were used to test reproducibility. The intra-class correlation coefficient (ICC) described how strongly quantitative measurements resemble each other between different groups and the result was interpreted as described in the literature (Fermanian, 1984). The other technique involved plotting the mean differences between measurements on a Bland-Altman Plot, which is a well-recognised method for measuring agreement between two assays measured on a continuous scale (Bland and Altman, 1986). The limits of agreement were calculated (mean ± 2SD) together with the 95% confidence intervals.

For comparison of qualitative data Cohen's kappa coefficient was used. It is generally thought to be a more robust measure than simple percent agreement calculation since kappa takes into account the agreement occurring by chance (Vach, 2005).

7.11 Method development

During the course of this study a number of issues and challenges arose that meant methods had to be developed or adjusted. The most noteworthy of these are discussed here.

7.11.1 DNA extraction and the quality of the DNA

Ideally the same concentration of DNA should be put into assays to ensure consistency. In order to test the DNA concentration a NanoDrop instrument was used (see section 7.4.3). The concentration is calculated by the absorbance of light at different wavelengths. An indication

of the purity of DNA can also be obtained using this method. For the QIAamp MinElute extracted DNA the NanoDrop results revealed relatively low DNA concentrations (mean 18.5 ng/µl, range -3.7–74 ng/µl). The most notable finding, however, was the absorbance ratios that indicate the purity of the DNA, which can consequently affect concentration calculations. Typically the 260/280 and the 260/230 ratios for pure DNA should be approximately 1.8 and 2.0, respectively. In these samples there was a very varied 260/280 ratio (mean 2.4, median 2.5, range -2.0–5.4) and an equally varied 260/230 ratio (mean 0.73, median 0.69, range -59–80). When repeated, these findings were reproducible, including when water was used instead of AVE buffer for the 'blank' control measurement and when an alternative machine was used. Some reasons for abnormal ratios are given in Table 7.15 (Thermo Scientific, 2011).

Ratio Finding	Reason
High 260/280	Not usually indicative of an issue
Low 260/280	Residual phenol or other reagent in the extraction method
High 260/230	Blank measurement made on a dirty pedestal or inappropriate solution used for the blank measurement
Low 260/230	Carbohydrate carryover or residual phenol or guanidine (often used in column based kits
Negative ratios	Low DNA concentrations or inappropriate blank measurement

Table 7.15: Reasons for abnormal NanoDrop absorbance ratios.

It is also known that at low concentrations (<10 ng/µl) both of the ratios can be inaccurate (Thermo Scientific, 2012). Ultimately, the real test of DNA purity and concentration are the downstream PCR reactions. The B-globin PCR and amplification of the ADAT1 gene (DNA control for the PapilloCheck® assay), for example, worked in 92% and 99% of all samples, respectively (*n* = 534). This would imply that the DNA was of adequate purity and concentration for PCR. Despite this, efforts were made to try to improve the extraction procedure.

Firstly, additional wash steps were introduced in the sample processing in order to remove more contaminants. Furthermore, both ethanol precipitation and a number of clean-up kits were also used to improve the purity of the DNA. The only one that seemed to make a difference to the NanoDrop ratios was the clean-up kit. The ratios did improve a little, but it was at the expense of concentration. Another consideration was the time and costs required to make these small gains as well as the resulting reduction in the volume of DNA available for other assays. A decision was made to continue with the DNA extraction using the QIAamp

MinElute Kit as per study protocol and to use standard volumes of DNA product in any PCR reaction.

When more DNA was required for the DIPS and pyrosequencing assays the cell pellet was resuspended and the simpler Proteinase K extraction method was performed. The NanoDrop results for samples using this method were far more consistent in DNA concentration (mean 81.3 ng/µl, range 64.8–133 ng/µl), 260/280 ratio (mean 0.66, range 0.60–0.85), and 260/230 ratio (mean 0.18, range 0.16–0.25). These ratios were very low, consistent with salt contamination, a common feature of this type of extraction.

7.11.2 E6 and E2 PCRs

Early work using the E2 and E6 PCRs was carried out on a small range of clinical samples (n=28) in addition to control material. A significant number of clinical samples tested initially failed the E2 PCR and the E6 control PCR. In order to improve the success rate a number of steps were taken. The PCR conditions and reagents were revised, experiments were conducted with an increased DNA input and diluting the DNA also tested for potential sample inhibition. However, none of these steps made much improvement. The two modifications that made the greatest difference were the introduction of a linear reaction prior to E2 PCR using an upstream DIPS primer and the use of a hot-start Taq DNA polymerase. The linear PCR increases the amplification of the target region before narrowing down on a specific target sequence; thus, the second PCR should be more sensitive. HotstarTaq requires a 15-minute, 95 °C incubation step, which means that nonspecific amplification products are minimised and the yield of the specific PCR product is increased. An example of E6 PCR using these two methods is shown in Figure 7.7.

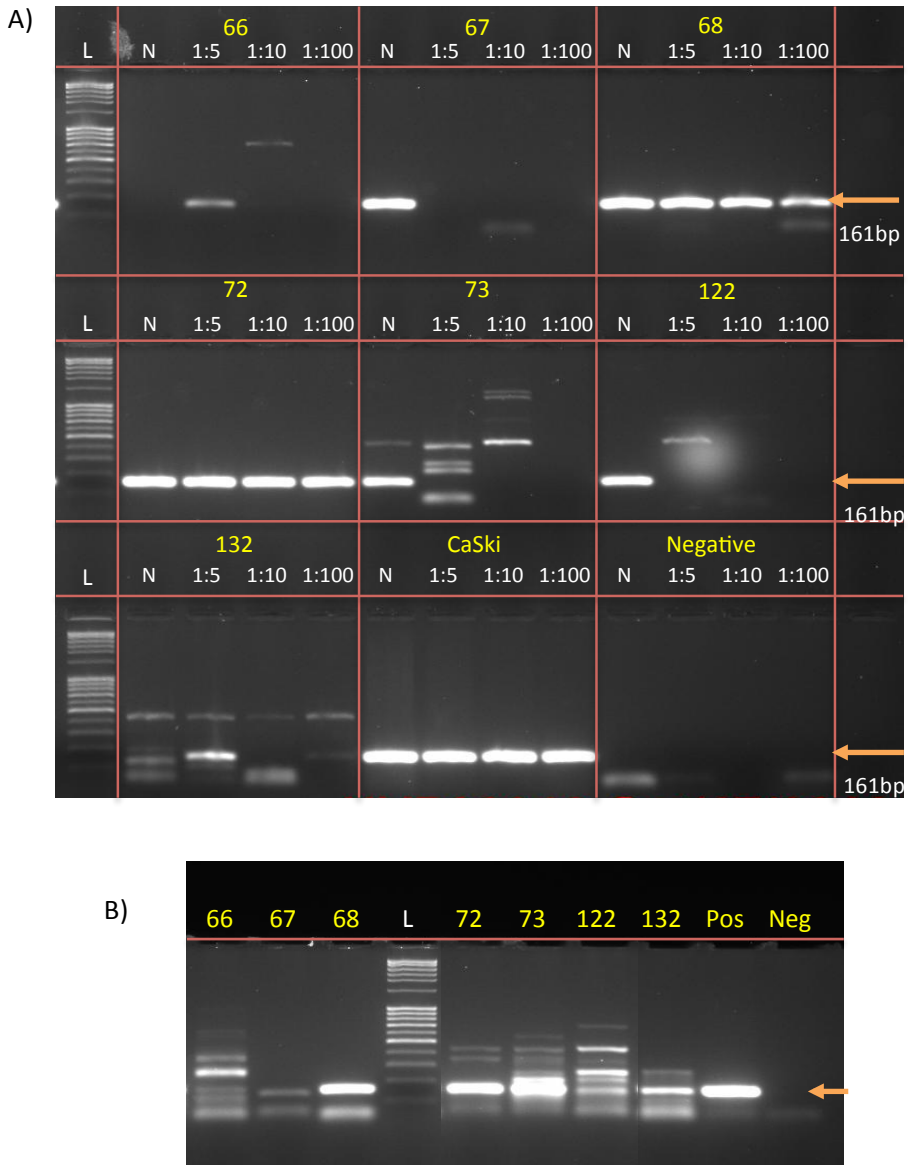

Figure 7.7: Example electrophoresis following E6 PCR using two different methods.

A – Dilution series (N=neat) using hot-start Taq method. All clinical samples shown here tested positive but not at all dilutions.

B – Linear PCR followed by standard Taq E6 PCR method. Although a correctly sized band can be seen for all samples it is slightly ambiguous for sample 66 and 122 where there are multiple bands present.

Orange arrows indicate the expected amplicon using E6 primers, 161 bp), L = 100bp DNA ladder.

After completing the comparisons on 22 samples (plus controls) for both E6 and E2-3 primer sets it was found some samples required dilution to 1:5 or 1:10 in order to test positive, whilst others would only test positive using neat DNA. Hence, it was decided to test E6 on all samples at 1:5 and if negative to retest using neat. The procedure was performed for E2 PCRs also. Both the linear PCR and hot-start Taq methods worked well, however, there were multiple bands on the gels. A selection of these bands were purified and sequenced in order to identify them. Reassuringly, they were consistent with the correct E6 region of HPV. However, multiple bands may be indicative of excess primers, excess Taq, high concentrations of DNA or low nonspecific

annealing temperatures. After further experimentation the assay was improved further with a combination of less HotStarTaq and less primer (Figure 7.8 and Figure 9.2).

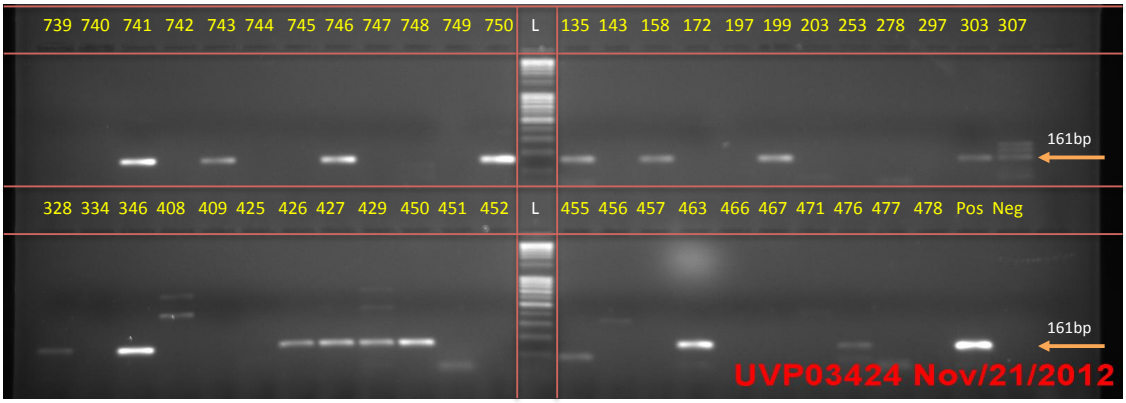

Figure 7.8: Example electrophoresis image following E6 PCR using optimum conditions.

The positive (pos) and negative (neg) controls are shown in the bottom right hand corner of the image. The samples are from the SuPerLy study. Samples 741, 743, 746, 750, 135, 158, 199, 303, 307, 328, 336, 426, 427, 429, 450, 463, and 476 have the expected 161 bp band for E6. The other samples are negative for HPV E6.

Chapter 8 – RESULTS AND DISCUSSIONS – SUPERLY STUDY

8.1 Introduction

This chapter reports the findings of the SuPerLy study. The characteristics of BD SurePath Plus™ and HPV typing in the context of a triage test for women with persistent low-grade cytological abnormalities are presented and discussed. Further HPV testing including type-specific PCRs are also detailed and an important finding relating to HPV testing and genomic disruption is considered.

The following hypotheses are covered in this chapter:

H_1 = BD SurePath Plus™ can predict the presence of high-grade cervical disease in women with persistent low-grade cytological abnormalities.

H_1 = HPV testing can predict the presence of high-grade cervical disease in women with persistent low-grade cytological abnormalities.

H_1 = BD SurePath Plus™ will predict with higher positive predictive value, but lower negative predictive value than HPV testing.

8.2 Study population

The number of women at each stage of the study is shown in the study flow chart (Figure 8.1).

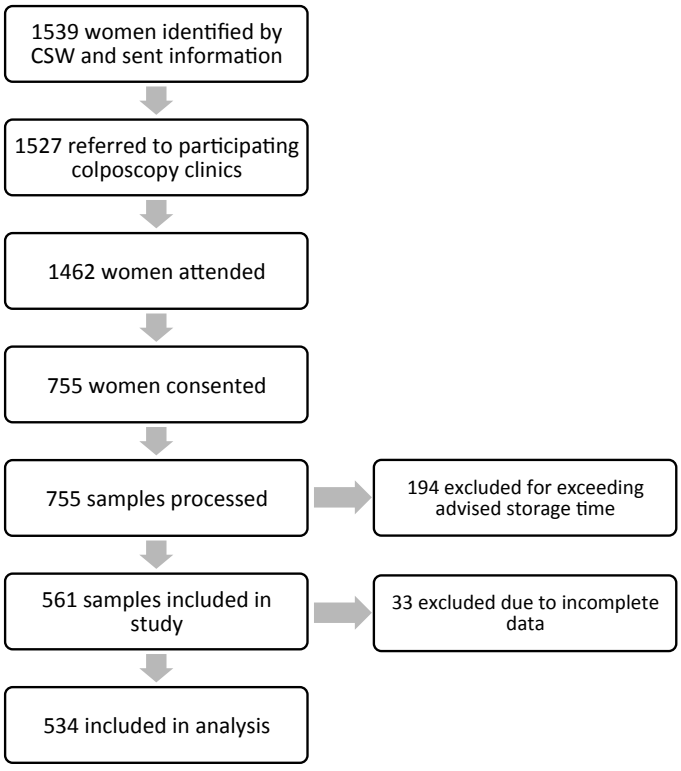

Figure 8.1: Study flow chart

Out of the eligible women that attended colposcopy clinic in Cardiff or Newport, 755 (52%) were consented and their samples were entered into the study. However, 194 samples had to be excluded because of a delay in reading the original smear meant that the samples had exceeded the manufacturer's recommended time that they should be kept in the SurePath media at room temperature. As a consequence of this the study had to be extended. The remaining 561 samples were stored according to protocol and suitable for SPP and HPV testing. There were 33 samples where the colposcopy and histology data was not available. Hence, data will be presented on a cohort of 534 women unless otherwise stated.

8.3 Referral cytology

All 534 women had low-grade cytology as per the inclusion criteria in the study protocol. The cytology result that triggered the referral to colposcopy was used in the analyses. The distribution of the various cytological grades is shown in Figure 8.2. The most common referral cytology was mild dyskaryosis (50%), followed by borderline changes (38%). The remaining 12% was made up of the other borderline categories.

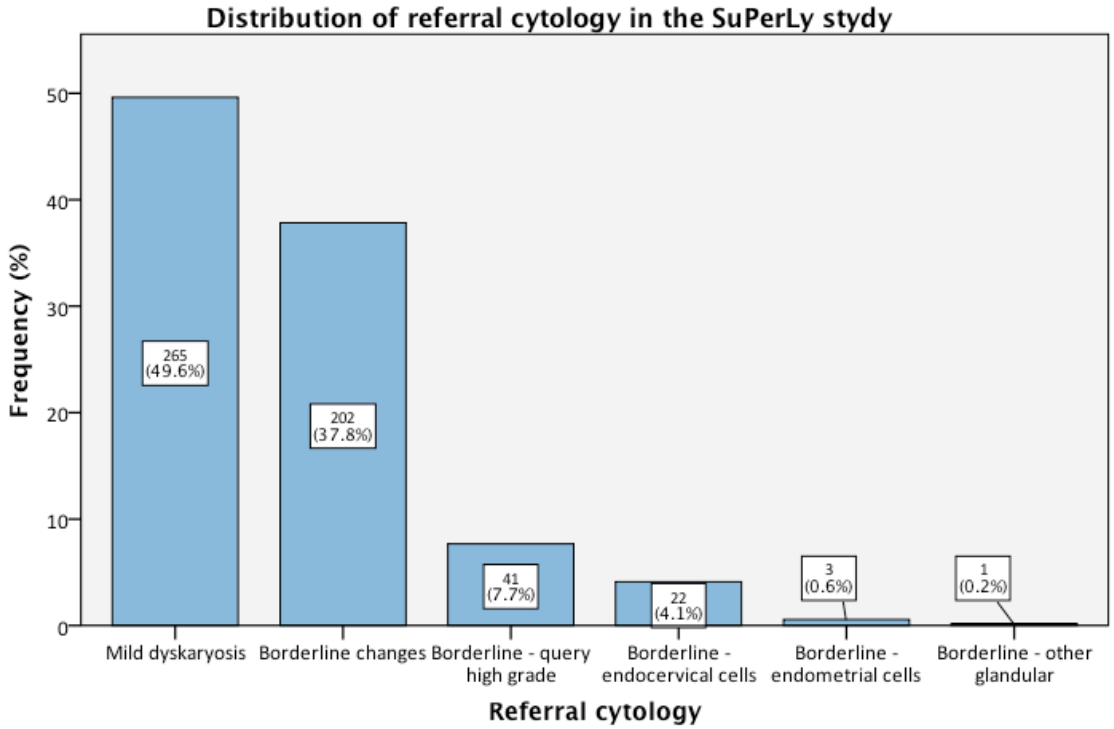

Figure 8.2: The distribution of referral cytology in the SuPerLy study.

8.4 Patient age

The mean and median ages of the study cohort were 33 and 28 years old respectively (range 20–66). The data was plotted in a histogram in order to appreciate the distribution (Figure 8.3). It is clear from the histogram that the cohort is skewed towards the 20–30 years olds.

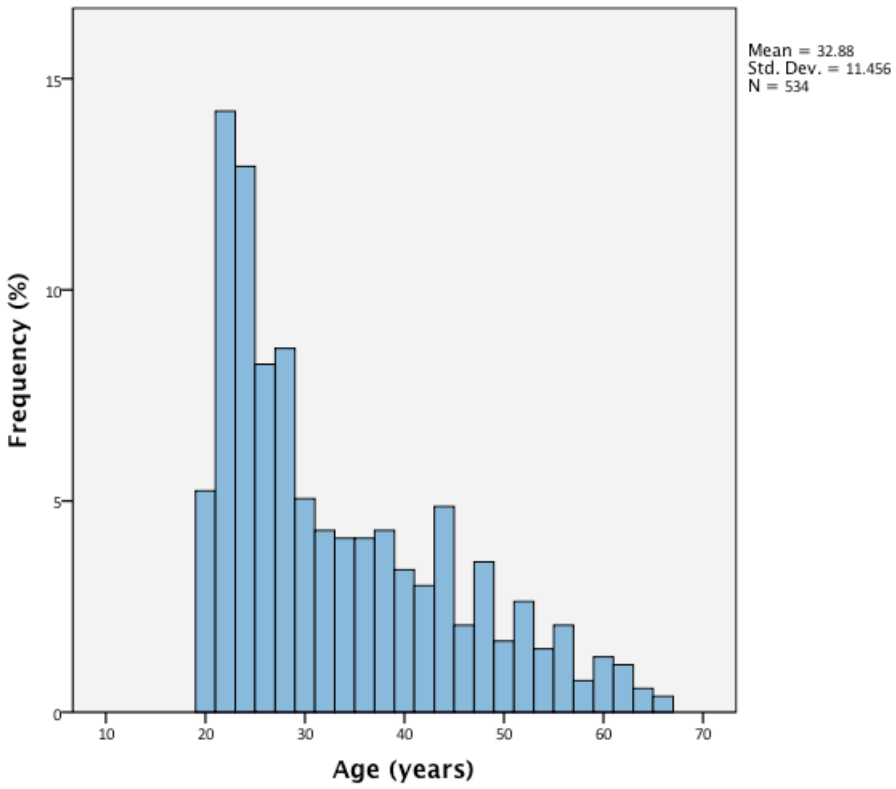

Figure 8.3: Histogram showing the ages of the SuPerLy study cohort.

8.5 DNA extraction

DNA extraction was performed on all 534 samples. There were two methods of confirming successful extraction: B-globin and the inbuilt PapilloCheck sample control. The B-globin PCRs showed successful DNA extraction in 489 (92%) of the samples; whereas the PapilloCheck control showed successful DNA extraction in 99% (526/534) of samples. All samples passed at least one of the DNA extraction controls.

8.6 Histological outcomes

The worst histology result within 26 weeks from consent was recorded. In the total study population 71 (13%) had a high-grade histology results (CIN2+). The different outcomes for the study population are shown in Figure 8.4.

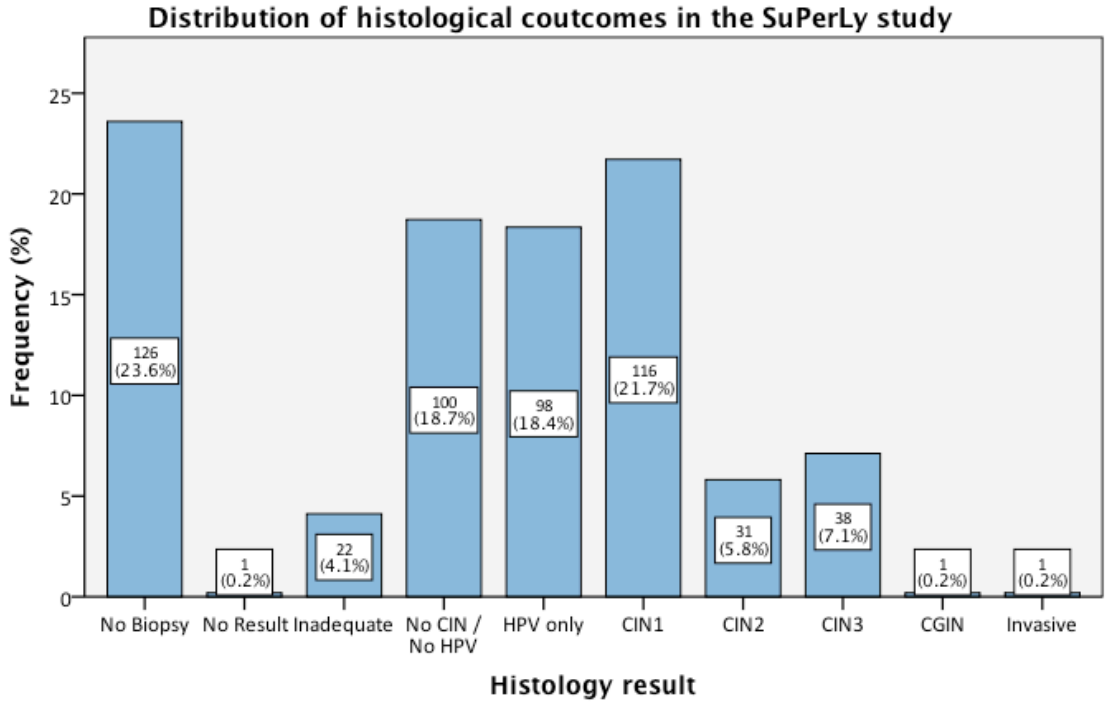

Figure 8.4: The histological outcomes in the SuPerLy study.

The worst histology result within 26 weeks from consent at first colposcopy is shown.

The colposcopy opinion was analysed to investigate the reason why no biopsy was taken in almost a quarter of the cohort. If a woman has an adequate colposcopy and normal TZ it is acceptable not to take a biopsy. In the design of the study it was considered inappropriate and unnecessary to biopsy every woman, although this has been done in other studies (Andersson et al., 2005, Kiatpongsan et al., 2006). Figure 8.5 shows the histological outcomes according to the colposcopic opinion. Of the 126 women that did not have a biopsy, 40 (32%) had a normal colposcopy opinion, another 40 (32%) had inflammatory / benign changes, 26 (9%) had low-grade, and the remaining 20 (16%) the data was not recorded. When the colposcopy opinion was recorded as normal (*n* = 45) and a biopsy was taken (*n* = 5) the result was either no CIN / no HPV or HPV only. However, when inflammatory / benign changes were noted (*n* = 90) there was a small but significant number of high-grade results (5 CIN2, 3 CIN3). Hence, for subsequent analyses of the women that had no biopsy taken, only those that had a normal colposcopy (*n* = 40) were considered to be disease-free at baseline. This is in keeping with other similar studies (ALTS Group, 2003, Cotton et al., 2010).

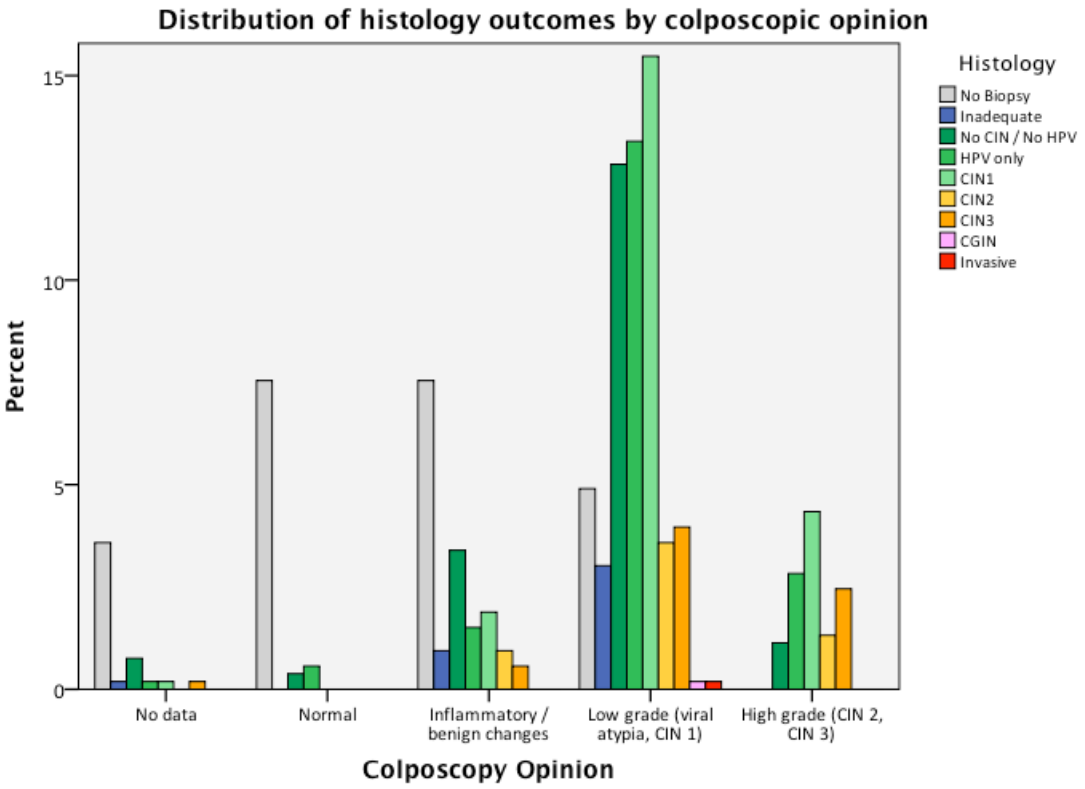

Figure 8.5: The histological outcomes for each colposcopic opinion.

The bars represent percentage of the total study population.

8.6.1 Referral cytology correlation with histology

The histology results were compared for the different cytological grades (Figure 8.6). The proportion of CIN2+ within mild dyskaryosis, borderline changes, borderline – endocervical cells and borderline – query high grade was 16%, 12%, 41%, and 29%, respectively. The odds ratio of CIN2+ in the presence of either borderline – query high-grade or borderline – endocervical cells result: OR = 2.9 (95 CI% [1.5–5.5], χ^2 = 9.6, P = 0.0019). If mild dyskaryosis was considered alone and compared with all other borderline results the histological outcomes were equivalent, OR = 0.95 (95% CI [0.63–1.47], χ^2 = 0, P = 1).

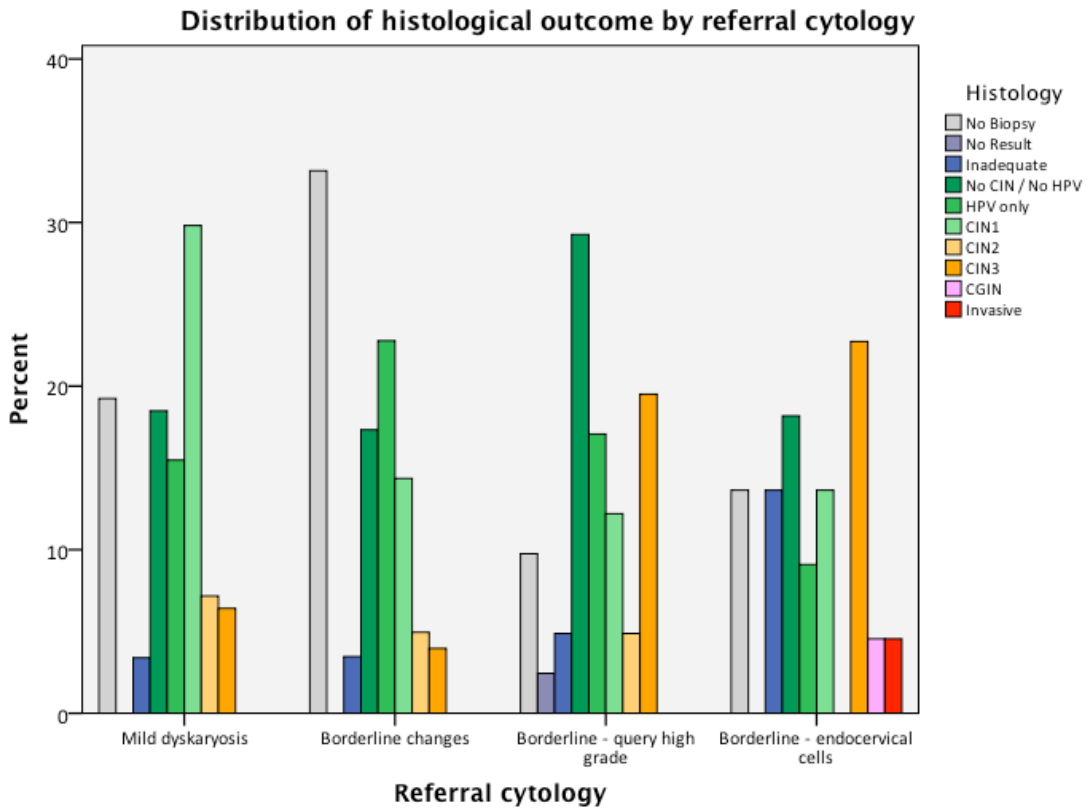

Figure 8.6: The histological outcomes for each referral cytology category.

The bars represent percentage of the total within each referral cytology category. The chart does not include borderline – endometrial cells (1 no biopsy and 2 HPV only) nor borderline – other glandular (1 no biopsy). No result indicates a biopsy was recorded as taken but no result could be found.

8.6.2 BD SurePath Plus™ correlation with histology

Due to the delay in closing the study and relocation of the cytological services during the study, BD SurePath Plus™ (SPP) data was available for 78% (417/534) of samples: 79% (328/417) were positive, 21% (88/417) were negative and there was one inadequate result.

The SPP result was compared across all histological grades (Figure 8.7). The positivity rate across all grades ranged from 72% in HPV only and negative biopsies up to 93% in CIN3 and 100% in invasive. A contingency table was constructed including the SPP result and the histological outcome (Table 8.1). The odds ratio showed some association with CIN2+, however, the CI were wide and the Chi-square was not significant; OR = 1.5 (95% CI [0.69–3.2]), χ^2 = 0.68, *P* = 0.41. The usefulness as a screening test was also calculated, sensitivity = 84% (95% CI [71–92%]), specificity = 22% (95% CI [17–27%]), PPV = 18% (95% CI [13–23%]), NPV = 87% (95% CI [77–94%]).

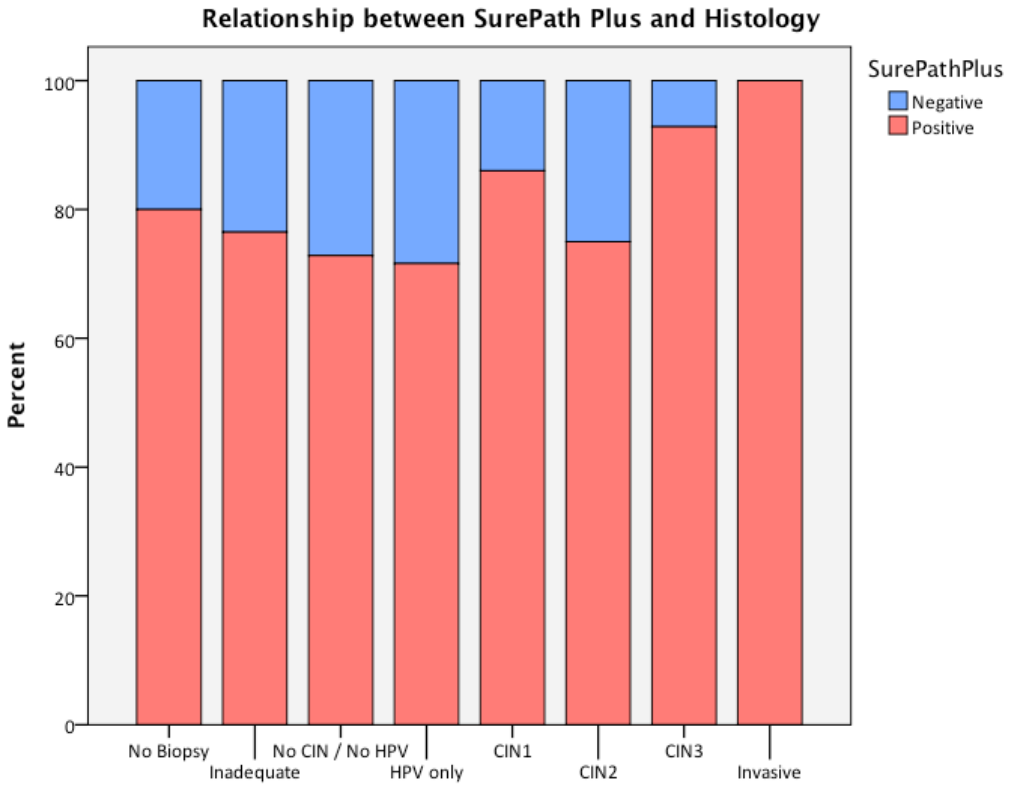

Figure 8.7: Relationship between BD SurePath Plus™ and histology.

	High-grade (CIN2+)		
	Absent	**Present**	**Totals**
BD SurePath Plus™ Positive	219	47	266
Negative	62	9	71
Totals	281	56	337

Table 8.1: Contingency table comparing SurePath Plus results and histological outcome.

8.6.3 HPV testing correlation with histology

HPV testing was done using both HC2 and PapilloCheck® as described in sections 7.6.1 and 7.6.2. HC2 testing was completed on 513 samples. A valid result was obtained for every sample tested and using the manufacturer's advised cutoff, 366 (71%) tested positive for hrHPV, while 153 (29%) tested negative. The HC2 result was compared to the histological outcome (Table 8.2). Further comparison of HC2 with PapilloCheck are given in section 8.8.

	High-grade (CIN2+)		
	Absent	**Present**	**Totals**
HC2 Positive	232	59	291
HC2 Negative	114	10	124
Totals	346	69	415

Table 8.2: Contingency table comparing HC2 result with histological outcome.

HC2 positive = hrHPV positive (i.e. one or more of HPV 16/18/31/33/35/39/45/51/52/56/58/59/68)

There was a significant association between a positive HC2 result and CIN2+, OR = 2.9 (95% CI [1.4–5.9%]), χ^2 = 8.5, *P* = 0.0036. The sensitivity of HC2 for identifying CIN2+ was 86% (95% CI [74–92%]), specificity = 33% (95% CI [28–38%]), PPV = 20% (95% CI [16–25%]), and NPV = 92% (95% CI [85–96%]).

PapilloCheck® was performed on all 534 samples: 526 (99%) of samples gave a valid result whereas 8 (1%) failed the in-built sample controls. Thirteen samples had failed initially and were consequently repeated; five of these passed on the repeat run. In this part of the data analysis the results for hrHPV types were pooled together to compare with histology. In addition, the types that were common to the HC2 test were also pooled together in order to compare PC to HC2 (Table 8.3). The ORs, sensitivities, specificities, PPVs, and NPVs are given for the PC hrHPV tests (18/13 types) in Table 8.4.

.

	High-grade (CIN2+)		
	Absent	**Present**	**Totals**
PC hrHPV Positive (18 types)	253	62	315
PC hrHPV Negative (18 types)	95	9	104
PC hrHPV Positive (13 types)	230	58	288
PC hrHPV Negative (13 types)	118	13	131
Totals	348	71	419

Table 8.3: Contingency table comparing PapilloCheck® hrHPV result with histological outcome.

The 13 types in the second part of the table refer to the HPV types that are used in the HC2 test.

HPV test	OR (95% CI)	X^2 (*P*-value)	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)
PC hrHPV (18 types)	2.6 (1.2–5.4)	6.0 (0.014)	87 (77–94)	27 (23–32)	20 (16–25)	91 (84–96)
PC hrHPV (13 types)	2.2 (1.2–4.2)	5.3 (0.02)	82 (70–90)	34 (29–39)	20 (16–25)	90 (83–94)

Table 8.4: Characteristics of PapilloCheck® as a screening test to identify high-grade disease.

8.7 Use of BD SurePath Plus™ and HPV testing as a triage screening test

In order to make a fair comparison, samples were only included here if data was complete for all categories: BD SurePath Plus™, HC2, PapilloCheck®, and histology (women with no biopsy and normal colposcopy opinion were included as mentioned previously). These results are summarised in Table 8.5.

Variable	Positive *n* (%)	Negative *n* (%)
SurePath Plus	261 (79)	71 (21)
HC2	228 (69)	104 (31)
PC hrHPV (18)	245 (74)	87 (26)
PC hrHPV (13)	224 (68)	108 (32)
Histology CIN2+	56 (17)	276 (83)
Histology CIN3+	32 (10)	300 (90)

Table 8.5: Comparison of the frequency distribution of the tests and histology in the SuPerLy study.

The extra CIN3+ histological category is necessary because CIN3 is the accepted surrogate marker for cervical cancer when any reduction in cervical cancer incidence is being investigated, which is the ultimate intention of any cervical screening programme. CIN2 is considered much less reproducible and less valid a diagnosis than CIN3 (Carreon et al., 2007). However, CIN2 is generally considered sufficient risk to warrant the same management as CIN3, hence, in the vast majority of studies it is used as the benchmark. Table 8.6 demonstrates the performance of SPP, HC2, and PC as a triage test for women with persistent low-grade cytology.

Variable	Ability to discriminate CIN2+				Ability to discriminate CIN3+			
	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
SPP	84 (71–92)	22 (18–28)	18 (14–23)	87 (77–94)	94 (78–99)	23 (18–28)	11 (8.0–16)	97 (89–100)
HC2	84 (71–92)	34 (29–40)	21 (16–27)	91 (84–96)	91 (74–98)	34 (28–39)	13 (8.8–18)	97 (91–99)
PC hrHPV (13)	78 (65–88)	35 (29–41)	20 (15–26)	89 (81–94)	84 (66–94)	34 (29–40)	12 (8.3–17)	95 (89–98)
PC hrHPV (18)	84 (71–92)	28 (23–34)	19 (15–25)	90 (81–95)	88 (70–96)	28 (23–33)	11 (7.9–16)	95 (88–99)

Table 8.6: Sensitivity, specificity, positive and negative predictive values for BD SurePath Plus™, HC2 and PapilloCheck®.

The best test for discriminating women with CIN2+ and CIN3+ appeared to be the well-established HC2. With such a high prevalence of BD SurePath Plus™ positive results, further analyses were performed in order to investigate and compare the performance of HC2 and PapilloCheck® in women that have a positive BD SurePath Plus™ result. The features of the various combinations are given in Table 8.7.

Combination with SPP	Ability to discriminate CIN2+				Ability to discriminate CIN3+			
	Sensitivity	Specificity	PPV	NPV	Sensitivity	Specificity	PPV	NPV
HC2	70	43	20	88	84	44	14	86
PC hrHPV (13)	68	46	20	88	81	46	14	96
PC hrHPV (18)	71	40	19	87	84	40	13	96

Table 8.7: Sensitivity, specificity, positive and negative predictive values for 3 HPV criteria, showing the effect of also requiring positivity by BD SurePath Plus™.

In all the combinations of tests the specificity was increased at the expense of sensitivity. The relative increase in specificity ranged from 11–13% for the HPV tests and for SPP it doubled. However, there was a relative loss in sensitivity of between 13–14%. The NPVs and PPVs remained largely unchanged. An NPV of 88% when HC2 was combined with SPP implied that even when SPP was negative and no HPV was found there was still a false negative rate of 12%.

One of the outcomes of a screening test is the identification of a high-risk group for the disease in question. However, another equally important outcome could be the creation of a very low-risk group, particularly when there is a relatively high prevalence of the disease. In the context of cervical screening this is important because if women can be allocated to a very low-risk group they could be discharged and potentially returned to normal recall in 3–5 years. For this low-risk group to be effective the rate of CIN2+ should be as low as possible and the group itself should be as large as possible. Table 8.8 shows the size of various groups using all combinations of tests and the proportion of that group that have CIN2+. The lowest risk group was when both BD SurePath Plus™ and HC2 or PapilloCheck® (including all 18 HPV types) were negative, 3% and 6% rate of CIN2+ respectively. However, the sizes of these groups were too small to be significant.

8.7.1 Comparison with other studies

Meta-analysis has revealed the majority of studies investigating the use of HPV testing as a triage test have used HC2 and furthermore most have only included women aged ≥30 years (in some cases ≥35 years)(Arbyn et al., 2012). The largest study in the UK investigating HPV testing in the triage of mild and borderline cytology used HPV PCR using GP5+/6+ consensus primers, followed by enzyme immunoassay (EIA) to amplify the L1 region of 14 hrHPV types (types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66 and 68) (Cotton et al., 2010). That study also recruited women from age 20 years from a UK cohort, hence, is an appropriate comparative study. In Table 8.8 the data from the SuPerLy study has been restricted to age ≥30 years in keeping with the body of literature. In Table 8.9 and Table 8.10 the data has been further divided into triage of mild dyskaryosis (equivalent to LSIL) and of borderline changes (equivalent to ASCUS), as per convention. In addition, for further comparison the data for HPV16 type-specific positive by PapilloCheck® (PC 16) was also included.

	n	As proportion of the whole group	CIN2+	As a proportion of CIN2+
HC2	332		56	
SPP –ve	71	0.21	9	0.13
SPP +ve	261	0.79	47	0.18
HPV –ve	104	0.31	9	0.09
HPV +ve	228	0.69	47	0.21
Both –ve	38	0.11	1	0.03
Either +ve	294	0.89	55	0.19
Either –ve	137	0.41	17	0.12
Both +ve	195	0.59	39	0.20
PC hrHPV (13)				
SPP –ve	71	0.21	9	0.13
SPP +ve	261	0.79	47	0.18
HPV –ve	108	0.33	12	0.11
HPV +ve	224	0.68	44	0.20
Both –ve	35	0.11	3	0.09
Either +ve	297	0.90	53	0.18
Either –ve	144	0.43	18	0.13
Both +ve	188	0.57	38	0.20
PC hrHPV (18)				
SPP –ve	71	0.214	9	0.127
SPP +ve	261	0.786	47	0.180
HPV –ve	87	0.262	9	0.103
HPV +ve	245	0.738	47	0.192
Both –ve	32	0.096	2	0.063
Either +ve	300	0.904	54	0.180
Either –ve	126	0.380	16	0.127
Both +ve	206	0.620	40	0.194

Table 8.8: Comparison of the proportion of CIN2+ in various risk groups in the SuPerLy study stratified according to the SPP and HPV test results. HC2 = Hybrid Capture 2, SPP = Sure Path Plus, PC hrHPV = Papillocheck using 13/18 hrHPVs.

Test (*n*)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
SurePath Plus (165)	86 (63–96)	23 (17–31)	14 (8.7–21)	92 (76–98)
HC2 (201)	92 (73–99)	45 (38–53)	20 (13–28)	98 (91–100)
PapilloCheck (201)	92 (73–99)	38 (31–46)	18 (12–26)	97 (89–99)
PC 16 (201)	31 (15–52)	75 (68–81)	15 (7.3–29)	88 (81–92)

Table 8.8: Comparison of test performance in the triage of low-grade cytological abnormalities in women ≥30 years.

Test (*n*)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
SurePath Plus (59)	50 (9.2–91)	31 (20–45)	5 (0.8–18)	89 (65–98)
HC2 (76)	83 (36–99)	51 (39–63)	13 (4.8–28)	97 (84–100)
PapilloCheck (77)	67 (24–94)	46 (35–59)	9.5 (3.1–24)	94 (79–99)
PC 16 (77)	17 (0.7–64)	80 (69–88)	6.7 (0.4–34)	92 (81–97)
Meta-analysis (Arbyn, 2012)	90 (88–92)	58 (54–63)		
TOMBOLA (1175)	70 (62–77)	71 (69–74)	25 (21–30)	95 (93–96)

Table 8.9: Comparison of test performance in the triage of borderline changes in women ≥30 years.

Test (*n*)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
SurePath Plus (99)	92 (60–100)	18 (10–29)	15 (8.0–26)	93 (64–100)
HC2 (100)	93 (64–99)	31 (22–42)	18 (10–29)	96 (80–100)
PapilloCheck (99)	100 (73–100)	29 (20–40)	19 (11–30)	100 (83–100)
PC 16 (99)	29 (9.5–58)	72 (61–81)	14 (4.7–34)	86 (75–93)
Meta-analysis (Arbyn, 2012)	95 (94–97)	28 (24–32)		
TOMBOLA (656)	75 (69–81)	47 (42–52)	39 (35–44)	81 (75–85)

Table 8.10: Comparison of test performance in the triage of mild dyskaryosis in women ≥30 years.

8.8 Further analysis of HPV testing results

The advantage of the PapilloCheck assay over the HC2 test is the provision of genotyping data. Hence, it was possible to analyse the type-specific HPV prevalence and investigate the frequency of infections with multiple types. In total 391 (73%, 95% CI [69–77%]) hrHPV positive cases and a further 38 (7.1%, 95% CI [5.2–9.6%]) cases with only lrHPV were identified. Thus, the overall HPV prevalence was 80% (95% CI [77–83%]). The HPV type-specific prevalence is presented in Figure 8.8 for the 534 eligible women from the SuPerLy study.

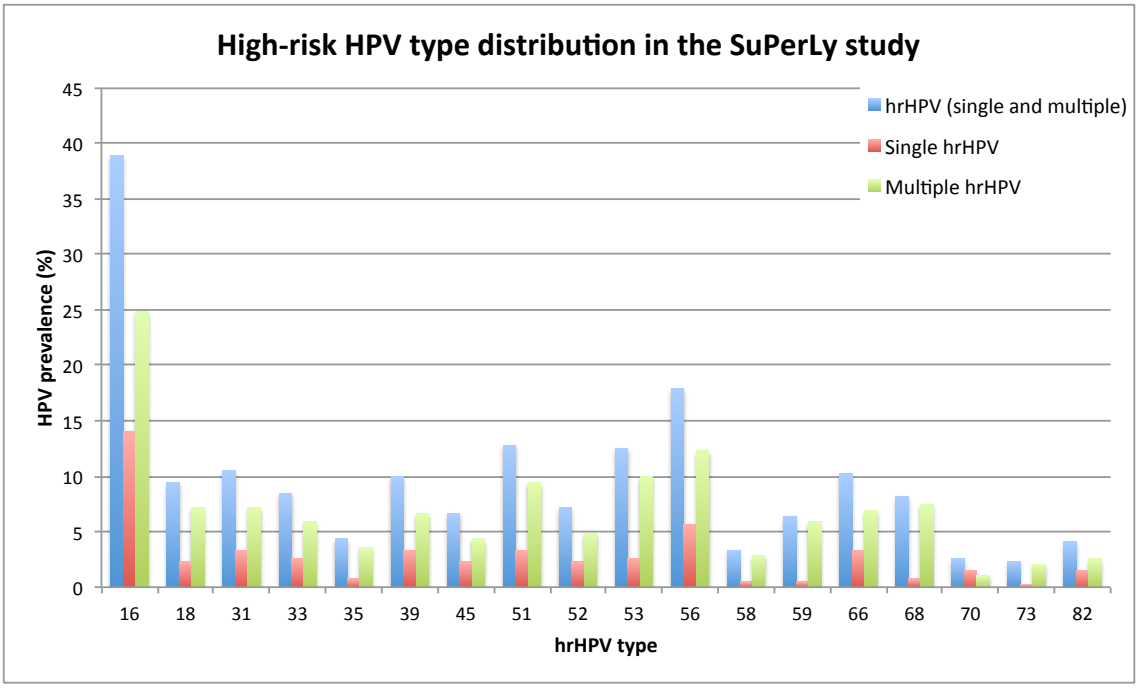

Figure 8.8: High-risk HPV type distribution in the SuPerLy study.

The total number of hrHPV cases with single and multiple hrHPV types for each genotype are shown as a percentage of the number of hrHPV positive cases (*n* = 391). The hrHPV prevalence of each type in order of predominance was as follows: HPV16 (*n* = 152, 38.9%), HPV56 (*n* = 70, 17.9%), HPV51 (*n* = 50, 12.8%), HPV53 (*n* = 49, 12.5%), HPV31 (*n* = 41, 10.5%), HPV66 (*n* = 40, 10.2%), HPV39 (*n* = 39, 10%), HPV18 (*n* = 37, 9.5%), HPV33 (*n* = 33, 8.4%), HPV68 (*n* = 32, 8.2%), HPV52 (*n* = 28, 7.2%), HPV45 (*n* = 26, 6.6%), HPV59 (*n* = 25, 6.4%), HPV35 (*n* = 17, 4.3%), HPV82 (*n* = 16, 4.1%), HPV58 (*n* = 13, 3.3%), HPV70 (*n* = 10, 2.6%), and HPV73 (*n* = 9, 2.3%).

PapilloCheck® detected a total of 687 hrHPV infections in 391 women. Of the hrHPV positive cases, 199 (51%, 95% CI [46–56%]) had a single hrHPV type infection and 192 (49%, 95% CI [44–54%]) had an infection with multiple types of hrHPV. The number of types ranged from 2-7 with a mean of 2.5. In both single and multiple type infections HPV16 was the predominant type (14% and 25%, respectively). There was also a high prevalence of HPV56 in both groups (5.6% and 12%, respectively).

The distribution of HPV was also analysed according to age group (Figure 8.9). Age appeared to have a strong association with the nature of hrHPV infection. Women aged 34 years or less were significantly more likely to have a multiple hrHPV type infection, $\chi^2 = 14.3$, $P = 0.0002$.

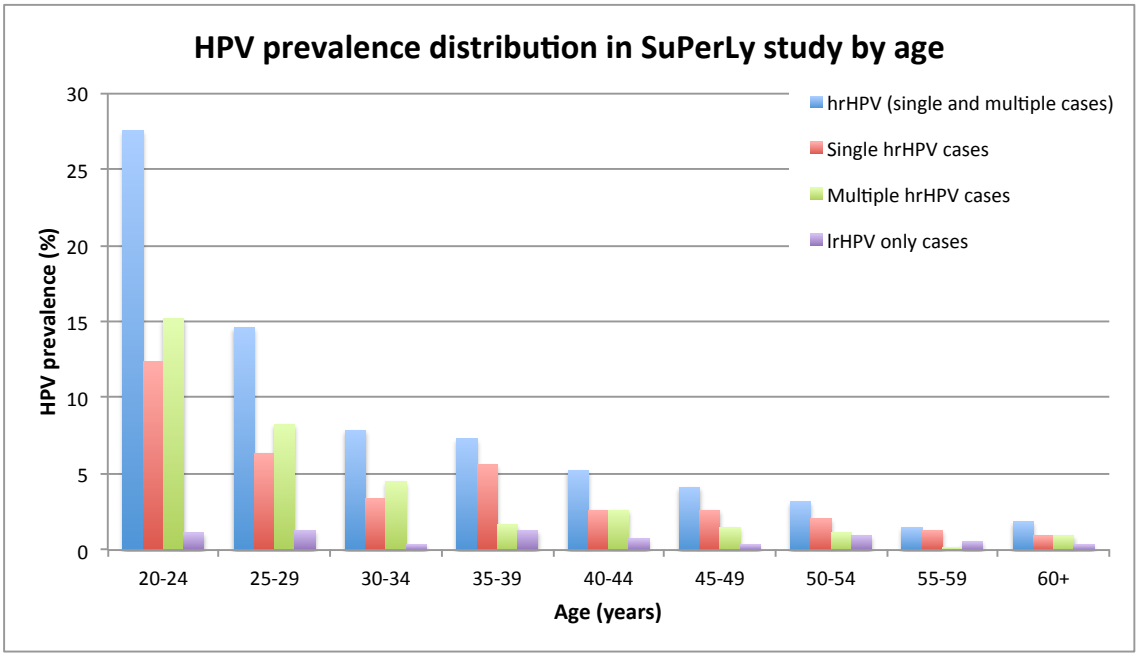

Figure 8.9: HPV prevalence in the SuPerLy study divided by age group.

The total number of hrHPV, single hrHPV, multiple hrHPV, and lrHPV only are shown as a percentage of the total study population (*n* = 534) and divided by the following age groups: 20–24 years (*n* = 173), 25–29 years (*n* = 102), 30–34 years (*n* = 60), 35-39 years (*n* = 58), 40–44 years (*n* = 47), 45–49 years (*n* = 32), 50–54 years (*n* = 29), 55–59 years (*n* = 17), 60+ years (*n* = 16).

The type-specific HPV prevalence in the different histological grades was analysed. In this analysis only single HPV infections were considered because when there is a multiple type HPV infection there was no way of knowing which of the HPV types were driving the disease without more detailed investigation with more clinical material. Furthermore, because within each genotype the numbers were quite small, the histological grades were grouped in the same way as before (see section 8.6.2)(Figure 8.10). Further statistical analyses was not performed for the same reason, however, it was noted that single infections with HPV16, HPV31, HPV33, HPV39, or HPV56 were much more associated with high-grade CIN.

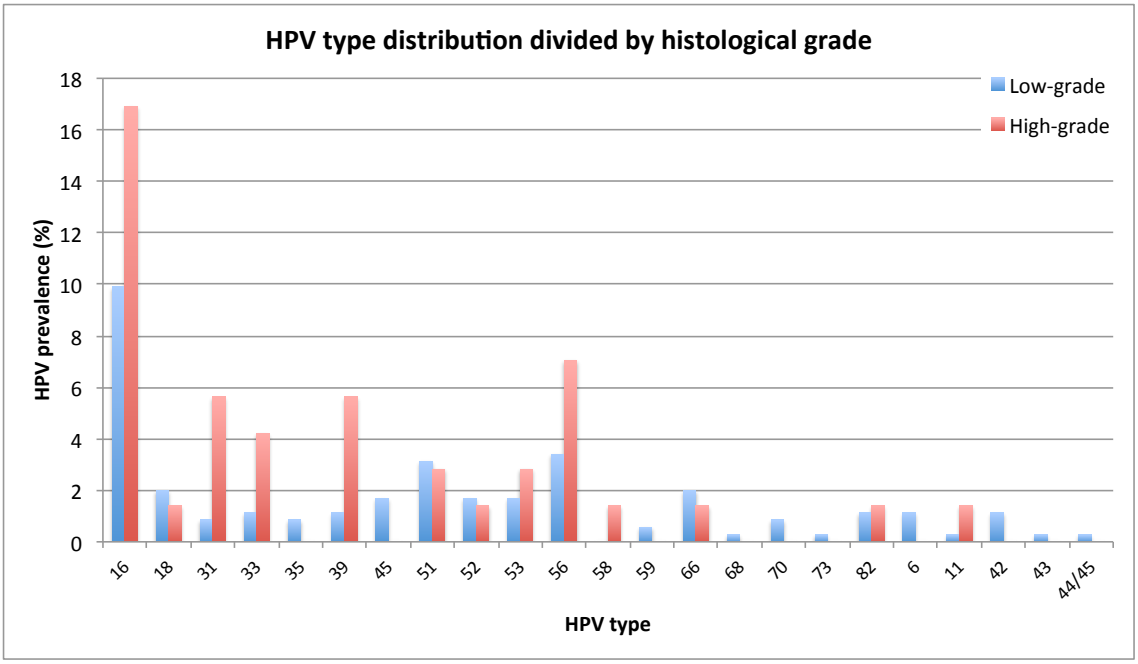

Figure 8.10: HPV type distribution classified by histological grade in cases with a single HPV infection. The total number of single HPV infections within each genotype was calculated as a percentage of the two histological grades. Low-grade histology (*n* = 353) included all biopsy results CIN1 or better and cases where there was no biopsy and a normal colposcopy opinion; high-grade histology (*n* = 71) included all biopsy results CIN2 or worse.

The risk of having high-grade histology was considered for cases with single type and multiple type hrHPV infections. High-grade disease was marginally more associated with single type HPV infections (OR = 1.2; 95% CI [0.69–2.13]), however, as the confidence intervals showed any difference was statistically inadequate.

8.8.1 Comparison of HC2 and PapilloCheck®

For a direct comparison of HC2 and PapilloCheck®, data was available for 511 samples. A contingency table was constructed in order to judge the agreement between the two tests (Table 8.11).

		HC2 Result		
		Negative	**Positive**	**Total**
PapilloCheck®	**Negative**	103	71	174
(13 hrHPV types)	**Positive**	48	289	337
	Total	151	360	511

Table 8.11: Contingency table comparing PapilloCheck® with HC2.
Only those PapilloCheck positive for one of the HPV types included in HC2 were included in this analysis.

HC2 identified an additional 23 HPV positive cases; however, of greater concern was the apparent disagreement between the two tests. This was confirmed by calculating a Cohen's

kappa value, kappa = 0.46 (95% CI [0.38–0.55]). Considering that the age made a large difference to the accuracy of the HPV tests for discriminating high-grade disease, the two tests were also compared with women aged 20–29 years excluded (Table 8.12). Although the kappa value was slightly improved, the difference was not significant. Interpretation of kappa values is not straightforward, however, the general consensus is that a kappa value between 0.41–0.60 would be classed as a moderate agreement (Landis and Koch, 1977).

		Agreement %	Kappa (95% CI)
All ages	**PC hrHPV (13) vs HC2**	77	0.46 (0.38–0.55)
(*n* = 511)	**PC hrHPV (18) vs HC2**	81	0.53 (0.44–0.61)
≥30 years	**PC hrHPV (13) vs HC2**	75	0.49 (0.38–0.60)
(*n* = 253)	**PC hrHPV (18) vs HC2**	78	0.54 (0.43–0.65)

Table 8.12: Agreement between PapiloCheck® and HC2.

8.8.1.1 HC2 and PapilloCheck® reproducibility

Both of these tests are commercially available and have been through the required validation to gain regulatory body approval. However, factors such as human error or contamination could be responsible for the disparate results seen. HC2 has been designed in a way that should reduce chances of error or contamination because it combines both DNA extraction and the hrHPV test into one assay. Furthermore, the steps involved are not complex, with minimal pipetting required, thus reducing the chance of manual errors. Owing to the costs of the HC2 test, and the requirement of additional clinical material, it was not possible to undertake repeat runs. However, it was reassuring that all the controls gave appropriate results.

For PapilloCheck®, the cost of testing was also a hindrance, however, a small number of samples were retested. The case where seven hrHPV types were found in one sample was repeated because this represented a more rigorous test. In addition, the preceding five samples were also repeated (Table 8.13).

Sample ID	First run	Second run
514	39	39
515	42	Negative
516	18	18
517	16	16
518	16, 51	16, 51
519	16, 31, 33, 39, 51, 53, 73, 42	16, 31, 33, 35, 39, 51, 53, 73, 42

Table 8.13: Reproducibility of the PapilloCheck® assay.

From the two separate PapilloCheck® runs it was found that four of the samples had 100% agreement. In the sample with seven hrHPV types (and one lrHPV type), only one of the seven was not identified in the second run. Furthermore, when the actual RLU measurement was analysed in both of these cases the value was very low when compared to those for the other HPV types. For example, in sample 519 the HPV35 was 40.2 RLU; whereas, HPV16, HPV31, and HPV33 were 1493.3, 180.5, and 446.2, respectively. It was likely that the two HPV types not repeated were at very low concentrations.

The PapilloCheck® assay in the HPV laboratory was also subject to a WHO HPV DNA proficiency study (Eklund et al., 2012). This contemporaneous test was performed by another member of the laboratory on a panel of 47 samples supplied by the WHO. The samples were made up of a variety of concentrations and included both single and multiple types. In order to be classed as proficient the test was required to detect 50 international units (IU) of HPV16 and HPV18 in 5µl, and 500 genome equivalents (GE) in 5 µl of the other 14 HPV types tested. In addition, it was also required that not more than one false positive was detected. In this dataset the results for the PapilloCheck® assay were considered proficient in 13 HPV types. The three types that did not meet the requirements were HPV16, HPV18 and HPV31. HPV18 and HPV31 were correctly identified as part of a multiple type infection, however, were missed when provided as single infection samples. For HPV16, detection was achieved at both 5 and 50 IU/5 µl, however, there was one false negative in a multiple type infection at 50 IU/5 µl.

Overall, there was some evidence of reproducibility, however, there was still concern over the reliability of the typing data. It was decided to reinforce the results by using a type-specific

PCR. This method was considered relatively inexpensive and amenable to high throughput; furthermore, only small amounts of DNA were required.

8.8.2 HPV type-specific PCRs

The E6 region was chosen as the target for PCR because it is rarely reported as a site of disruption (Wentzensen et al., 2004). It was not feasible to perform PCRs on all HPV types so the most common type, HPV16, was chosen. The method and method development for HPV16 E6 PCR is described in sections 6.7.1 and 6.11.2. Out of the 534 eligible women, 362 (68%, 95% CI [64–72%]) were positive for HPV16 E6 and 172 (32%, 95% CI [28–36%]) were negative. Of the positive results, 143 (40%) were only positive when 5 µl of neat DNA was used.

As the agreement between the E6 PCR results and the other HPV tests appeared to be low, a second type-specific PCR was chosen. HPV16 L1 PCR was performed on 519 samples; there was insufficient DNA in 15 samples. L1 PCR was positive in 293 (55%, 95% CI [52–61%]) and negative in the remaining 226 (42%, 95% CI [39–48%]) samples. The datasets available for all three tests were compared and tabulated (Table 8.14).

Test	*n*	Agreement %	Kappa (95% CI)
PC versus E6	526	66	0.36 (0.29–0.43)
L1 versus E6	519	71	0.40 (0.32–0.48)
L1 versus PC	511	57	0.26 (0.21–0.32)

Table 8.14: Comparison of HPV16 E6 and L1 PCR results with PapilloCheck®.

The L1 PCR identified less HPV16, however, it is also a potential site of disruption within the HPV genome, which may explain some of the difference. Importantly, E1, which is the target for the PapilloCheck® assay is also a recognised location for viral integration, and this has been described in low-grade lesions (Cricca et al., 2009). Furthermore, the target regions of HPV DNA for the L1 PCR and E6 PCR are a 133 bp and 120 bp long, respectively, compared to the target region for PapilloCheck® is 350 bp in length. This difference in length may be important, especially if the DNA had degraded during the collection and storage processes.

Finally, in order to assess the possibility of E1 disruption causing the PapilloCheck® assay to give a false negative result two primer sets that spanned the E1 ORF were used on all samples where there was sufficient DNA for analysis and HPV16 had been confirmed by E6 and L1 (*n* = 109). In addition, a random selection of PapilloCheck® positive samples (*n* = 22) were included (Table 8.15). The Kappa value for the agreement was only 0.38 (95% CI [0.21–0.56]).

		HPV16 E1 PCR		
		Negative	**Positive**	**Total**
PapilloCheck®	**Negative**	78	31	109
HPV16	**Positive**	3	19	22
	Total	81	50	131

Table 8.15: Comparison of HPV16 prevalence using PapilloCheck® and E1 PCR.

8.9 Discussion

The main aim of the SuPerLy study was to establish whether BD SurePath Plus™, a novel immunocytochemically-based test, could predict high-grade cervical abnormalities in women that had a persistent low-grade cervical abnormality. All women with persistent low-grade abnormalities are referred to colposcopy, however, only around one quarter of them have underlying high-grade disease (see section 3.3.3.1). Consequently, there is a need to improve the specificity of screening this group. The SuPerLy study recruited from a range of persistent low-grade cytological abnormalities following the consent of women at colposcopy clinic, with the range of ages sampled covering the entire screening population. The predominance of mild dyskaryosis and of women in the 20–30 year olds age group were both expected. It is important to consider that after recruitment for the SuPerLy study was completed, the Welsh Government announced the plan to increase the lower age limit for screening to 25, which was in line with England's policy (Welsh Government, 2013). Almost one third of the SuPerLy study population were aged 20–24 years. This may explain why the prevalence of high-grade disease was lower than expected (13% versus 25%). With this lower prevalence the study consequently had less power to detect any differences between the tests, however, there were a number of points for discussion.

8.9.1 BD SurePath Plus™

The positivity rate of BD SurePath Plus™ in this sample set was 79%. With such a high rate of positive tests and a low rate of high-grade disease, there was consequently a very low

specificity (22%). Perhaps, more surprising given this, was the number of false negative results. In this cohort a negative BD SurePath Plus™ result still gave a 13% chance of having high-grade disease. The premise of the BD SurePath Plus™ assay was that by staining two markers of cell proliferation, MCM2 and MCM7, HPV infections could be divided into two groups; those with higher proliferation and therefore high-risk of progression and those with no aberrant proliferation and a resulting low-risk of progression (Brake et al., 2003, Freeman et al., 1999, Henderson et al., 2011). Initial results from the forerunner to SPP, BD ProEx™ C, appeared to show an improved specificity (Kelly et al., 2006). However, as the authors pointed out, their study population had a high prevalence of CIN2+ due to their sample selection bias. Subsequent studies have shown improved sensitivity and specificity when BD ProEx™ C was used as a triage following primary hrHPV screening (Depuydt et al., 2011). However, there are also reports of false positive staining with this assay. In a study of 64 women, 84% (*n* = 27/32) of women that had a normal cytology result showed positive nuclear staining (Oberg et al., 2010). The implication of this finding was that normal proliferating cells exhibit sufficient MCM2 and TOP2A (the markers of BD ProEx™ C, see section 4.5.1) to give a positive result. The data presented in this thesis also imply that MCM2 and MCM7 are expressed sufficiently to give a false positive result.

One consideration is that the SuPerLy sample set includes women that have evidence of a persistent HPV infection. The majority of the women recruited in the study were based on the results of at least two cervical samples taken at least six months apart. It is not, therefore, very surprising that there were excess cell markers in the majority of samples, especially given that normal cells have been shown to exhibit the very similar markers (Oberg et al., 2010). Furthermore, in the SuPerLy study the 'gold standard' test for the identification of high-grade disease was a colposcopically guided biopsy; a small number of larger biopsies known as large loop excision of the transformation zone (LLETZ) were also performed. Colposcopy, however, is not precise and its sensitivity for CIN2–3 has been found to be as low as 56% (Massad and Collins, 2003). Colposcopically guided punch biopsies have also shown poor correlation (sensitivity as low as 46%) when a larger cone biopsy was performed as a control (Ang et al., 1995, Buxton et al., 1991, Skehan et al., 1990). The colposcopic opinion in the SuPerLy study supported this evidence with both false negative and false positive findings in the inflammatory/benign, low-grade and high-grade opinions. It is conceivable, therefore, that high-grade lesions could have been missed in our study. Two years of follow-up data will be available that should identify any missed lesions or lesions that were, perhaps, in the early

stages of transformation, with increased replication giving the positive SPP result, prior to exhibiting any physical changes at the surface.

8.10 HPV typing

The HPV prevalence in this study population with persistent low-grade cytological abnormalities was estimated using two commercially available tests. The comparison of HPV testing with other studies is not straightforward because of the unique SuPerLy study population tested. CSW guidelines call for the referral to colposcopy in women with persistent low-grade cytology; i.e., three borderline changes or two mild dyskaryosis. In most other studies examining HPV as a triage test they required one low-grade result (either borderline changes or mild dyskaryosis) to be included. Moreover, the other major difference is the age range of the vast majority of these studies. The standard inclusion criteria is 30–60 years, hence, comparison were made with the SuPerLy study's data by excluded women < 30 years. Increasing the age range made a substantial difference to all screening test characteristics and brought the HPV testing results to a similar level as in the meta-analysis (Arbyn et al., 2012). The TOMBOLA study (Cotton et al., 2010), which is the largest HPV triage study in the UK to date, was used to compare because it was the closest to the SuPerLy study in terms of classification systems, interpretation of cytology and the characteristics of the women. However, it found lower sensitivity and high specificity compared to both the SuPerLy study and the meta-analysis. There was a notably higher rate of CIN2+ combined with a lower rate of hrHPV positivity. It was possible that some of the difference may be explained by overcalling of low-grade abnormalities as CIN2 or CIN3, particularly since there was a concerning number of women with CIN2 that were hrHPV negative (using GP5+/6+ PCR EIA).

Additional HPV type-specific analyses using PapilloCheck® revealed a predominance of HPV16. The full type-specific prevalence was compared to meta-analysis of type distribution in low-grade cervical lesions (Clifford et al., 2005). The SuPerLy data correlated very well with the meta-analysis with the same four most common genotypes after HPV16, the only difference being that the positions of HPV56 and HPV31 were switched round compared to the meta-analysis. The type distribution beyond HPV16 was, however, quite different to the cross-sectional population based study of HPV prevalence in routine cervical samples in South Wales (Hibbitts et al., 2006). Monitoring of HPV type-specific prevalence will be most valuable, in light of the introduction of the HPV vaccine that protects against HPV16 and HPV18.

The two tests gave very similar prevalence figures, especially when the additional types that PapilloCheck® tests for were removed. However, on closer analysis there was considerable disagreement. Reproducibility was demonstrated to a small degree but not anywhere near the 10% sample repeats that others have performed (Hibbitts et al., 2006). The costs that would have been incurred for further repeat checks, in the case of this study, were inhibitory. Additionally, the amount of DNA required for multiple testing was also a limitation. There was some reassurance from the assays performance in the WHO tests; however, they did show some potential weaknesses. Moreover, accuracy and reproducibility can and should be assessed using control plasmids and cell lines initially, but the true test of an assay is using clinical material. Ideally clinical material should also be tested from a variety of different sources (e.g., LBC and biopsies). There is one study in the literature that has validated PapilloCheck® (Hesselink et al., 2010). They compared the assay with GP5+/6+ PCR EIA, which has previously been clinically validated and concluded that for the 14 hrHPV types that were common to both tests the PapilloCheck® assay was clinically compatible for the detection of CIN3+. They did not, however, address assay reproducibility in the study, but the authors did, at least, mention it as an outstanding requirement. Prerequisites for new HPV tests have been proposed (Meijer et al., 2009). They stipulate that intralaboratory reproducibility and interlaboratory agreement, as determined by the evaluation of at least 500 clinical samples, is required before a test is used for cervical screening. They also recommend that clinical validation of any new test be performed against HC2. The sensitivity for the detection of CIN2+ should be at least 90% of the sensitivity of HC2, whereas the specificity for the detection of CIN2+ should be at least 98% of that of HC2 in women aged ≥30 years. In the SuPerLy study both BD SurePath Plus™ and PapilloCheck® met the sensitivity criteria, however, they fell quite a long way short in terms of specificity for the detection of CIN2+.

Additional testing using HPV type-specific PCR were performed in order to investigate the possible reasons for HPV test disagreement. These demonstrated a higher HPV positivity, but also further disagreement. The nature of the PCR assays with small target regions on the HPV genome make them highly sensitive. Conversely, the commercial HPV assays used in the study have already been optimised in an attempt to reduce the number of false positive results that are associated with transient HPV infections. Moreover, the E1 type-specific PCRs in combination with the E2 status and integration data in this thesis have exposed potential flaws in HPV testing. The E1 region is the target of the PapilloCheck® assay and consequently it was hypothesised that disruption in E1 could result in a false negative result. There was certainly a trend towards agreement, however, the kappa was only 0.38. A total of 78 samples in the

study were positive for both HPV16 E6 and L1 but negative for E1 and PapilloCheck®. Furthermore, 48 of these samples were hrHPV positive by HC2. Integration in E1 is not uncommon (Wentzensen et al., 2004) and, although not proven, there is sufficient evidence presented here to warrant further investigation. It would also be prudent, as more research regarding site of viral disruption emerges, to consider other HPV tests that could be affected by integration.

8.10.1 Strengths and weaknesses

One of the strengths of the study was that it utilised standard screening cytological samples that meant the feasibility of introducing additional molecular tests could be more accurately assessed in the setting in which it would be used. However, central review of the cytology and histology would have strengthened the study further had the finances been available. The study also was able to compare two different commercially available HPV tests.

In the SuPerLy study we set out to recruit 600 women. Despite consenting more than this, the final number of according to protocol samples was lower (*n* = 561). This was mostly a consequence of carrying out the study in a 'real world' setting. Owing to the NHS' limited resources, the clinics were extremely busy, and with little or no slack built into the system, colposcopists regularly reported that they did not have time to explain and consent women to the study. There were also logistical issues, aside from the study, in the cytological laboratory that meant the cytology readers fell behind with their screening reads. This included the relocation of the laboratories during the study period. The unfortunate impact of this was that a number of samples were stored for over one month at room temperature, thus they exceeded the manufacturer's guidelines and were excluded from the study.

8.11 Conclusions

The novel test BD SurePath Plus™ showed no advantage over HPV testing in the triage of women with low-grade cytological abnormalities. It did correctly identify most of the women that had high-grade disease, however, it demonstrated very low specificity in this population. HPV testing using HC2 gave the most favourable results, especially when analysis was restricted to women ≥30 years. Agreement between HC2 and PapilloCheck® was moderate and further testing revealed that viral genomic disruption may result in false negative results. This could have significant consequences, especially if an affected women returns to routine screening intervals.

Chapter 9 – RESULTS AND DISCUSSION – INTEGRATION

9.1. Introduction

This chapter presents the results from the investigation of viral integration within the SuPerLy – HIM study. The clinical relevance of integration, in light of the E2 PCR and DIPS results, and their potential use as a screening biomarker are discussed. The main hypotheses tested in this chapter were:

H_1 = HPV16 E2 disruption is a marker of a transforming HPV infection and, therefore, an increased risk of having high-grade cervical disease.

H_1 = Viral integration is associated with a transforming HPV infection and, therefore, an increased risk of having high-grade cervical disease.

9.2. Viral Integration – E2

Optimisation of this method for assessing HPV16 E2 disruption was completed initially as previously described (see section 6.11.2). During this process E2 PCRs showed excellent reproducibility on CaSki cell line material. In a triplicate repeat of clinical material from six cases, four cases showed agreement using all primers sets, in one case there was a difference in the E2-4 primer set, and in the remaining case E2-1, E2-4, and E2-5 showed disagreement.

9.2.1. Study population

HPV 16 E2 PCRs were performed on 300 samples. Twenty-nine of these were performed before all HPV 16 assays had been performed and were excluded from analysis because they did not fulfill the criteria outlined in the methods (Figure 9.1). Of the 271 samples remaining, 161 (59%; 95% CI [53%–65%]) were shown to have at least one of the E2-1–E2-5 amplicons absent on gel electrophoresis and therefore disrupted, whereas 104 (38%; 95% CI [33–44%]) had all five present and were considered intact and in episomal form (Figure 9.2). In 6 (2%; 95% CI [1–5%]) samples the E2-FL primer produced an appropriately sized amplicon; however, not all the other smaller amplicons were present. Those samples appeared to be an anomaly; if the full-length primer worked then smaller regions of E2 should have been amplified also. One potential explanation is that there was some sequence variation where the primer targeted in the DNA templates for those cases. For the purposes of this analysis, those samples were included in the E2 intact group.

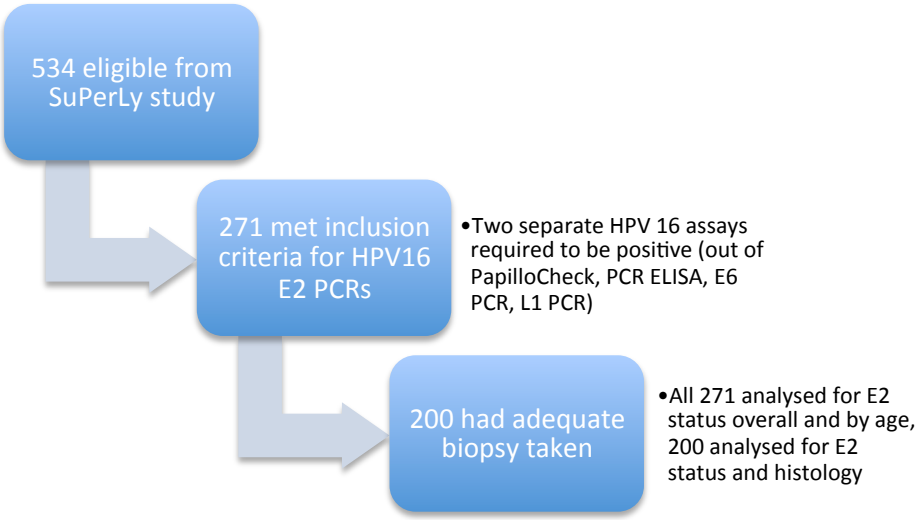

Figure 9.1: Sample flow for HPV16 E2 status analysis

In the 71 cases excluded in the final stage: 62 had no biopsy taken (55% disrupted, 45% intact) and 9 were inadequate (44% disrupted, 56% intact).

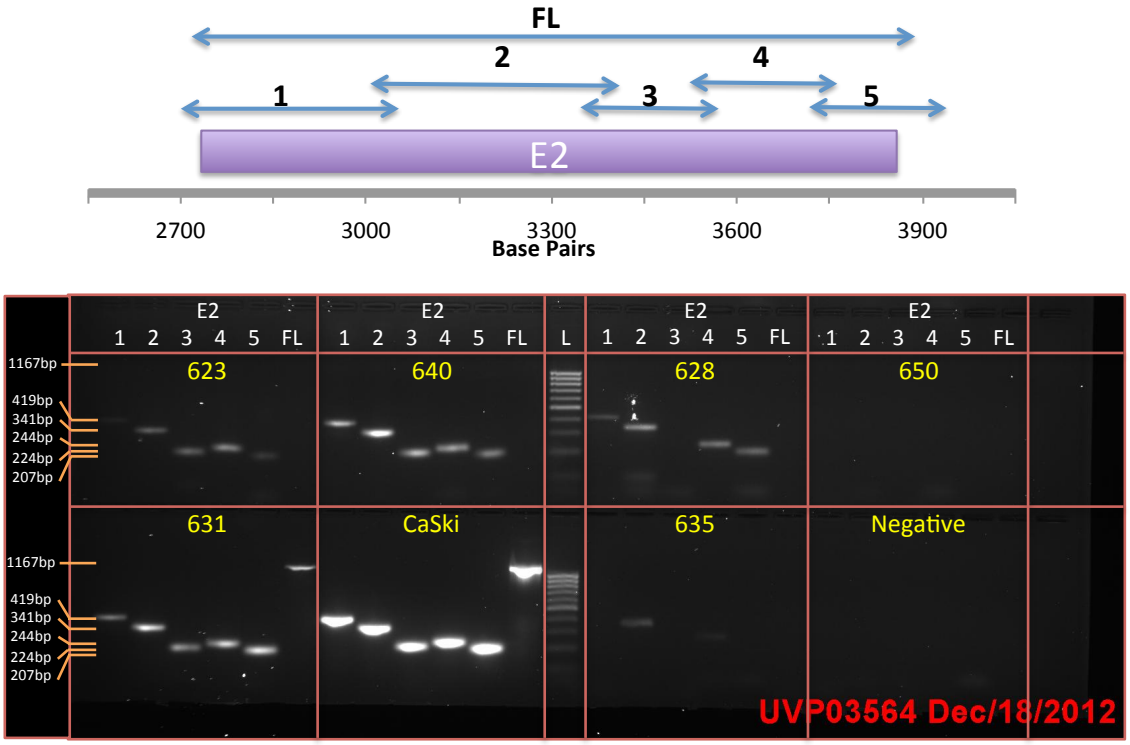

Figure 9.2: Example electrophoresis image showing HPV16 E2 PCR products and diagram showing the primer coverage.

There are six lanes per sample corresponding to the E2 primers (see lane headings E2 1-FL). The orange lines to the left hand side show the sizes of the bands. An example of fully intact E2 is shown for sample 631. Samples 623 and 640 demonstrated all the E2-1–E2-5 bands but not the E2-FL; these were also considered intact. An example of disrupted E2 is shown for sample samples 628, 635, and 650. All clinical samples tested positive for HPV 16 by two different assays. CaSki, positive control; Negative, no template DNA control; L, 100bp DNA ladder.

9.2.2. HPV 16 E2 status and histology

The E2 status was compared to the histology outcome following colposcopy. Figure 9.3 shows the association between E2 status and histology. In all grades of disease it was more common for E2 to be disrupted. In high-grade CIN, 22 out of 36 (61%; 95% CI [45–75%]) had disrupted E2; whereas, in low-grade CIN, 101 out of 164 (62%; 95% CI [54–69%]) had disrupted E2. A Chi-square test for independence confirmed there was no significant association between E2 status and histology (χ^2 (1) = 0.003; *P* = 1.00).

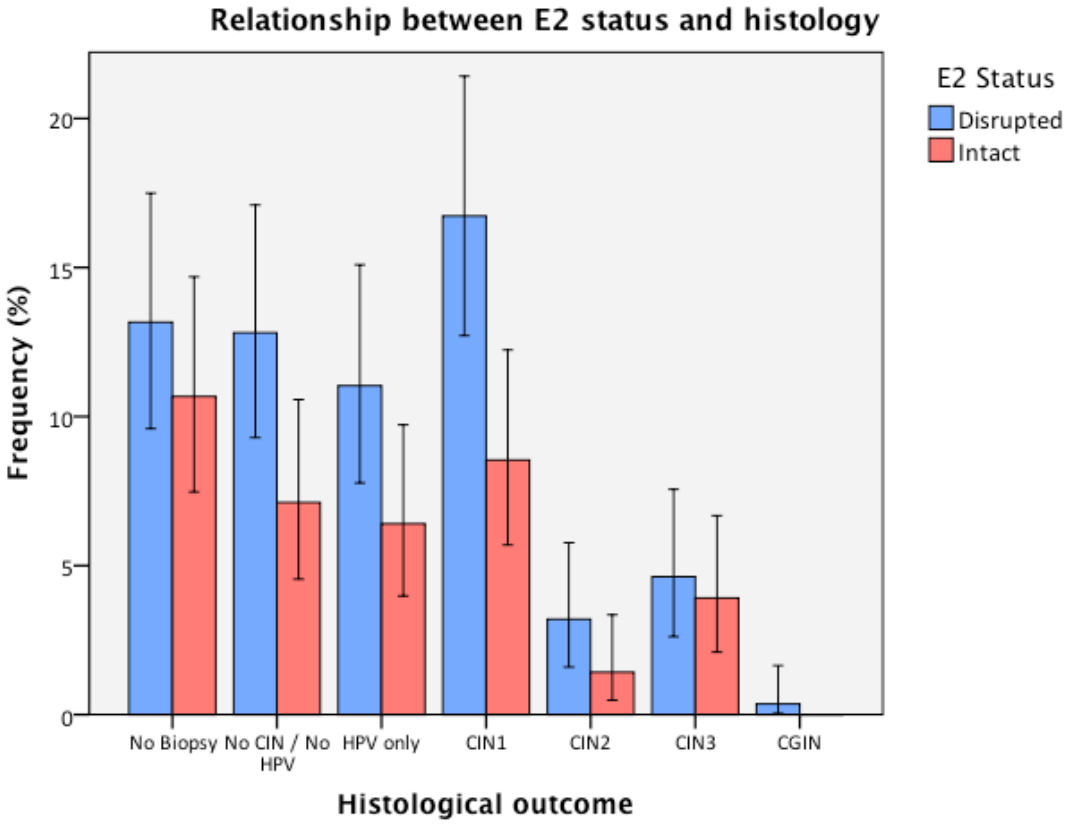

Figure 9.3: HPV16 E2 status according to histology grade.

9.2.3. HPV16 E2 status and age of patient

The E2 results were plotted in two groups to show the age distribution (Figure 9.4). The ages of the disrupted group appeared to be higher than the intact group. Before any correlation could be determined statistically, tests for normality and variance were performed. Levene's test showed equal variances between the groups (*P* = 0.88), however, a Shapiro-Wilk's test in combination with observing the histograms found that the ages were not normally distributed (*P* = 0.000). Hence, a Mann-Whitney U test was performed, which demonstrated the difference in the ages of the two groups E2 disrupted (*Md* = 30, *n* = 161) and E2 intact (*Md* = 26, *n* = 110) was significant, *U* = 7100, *z* = -2.77, *P* = 0.006.

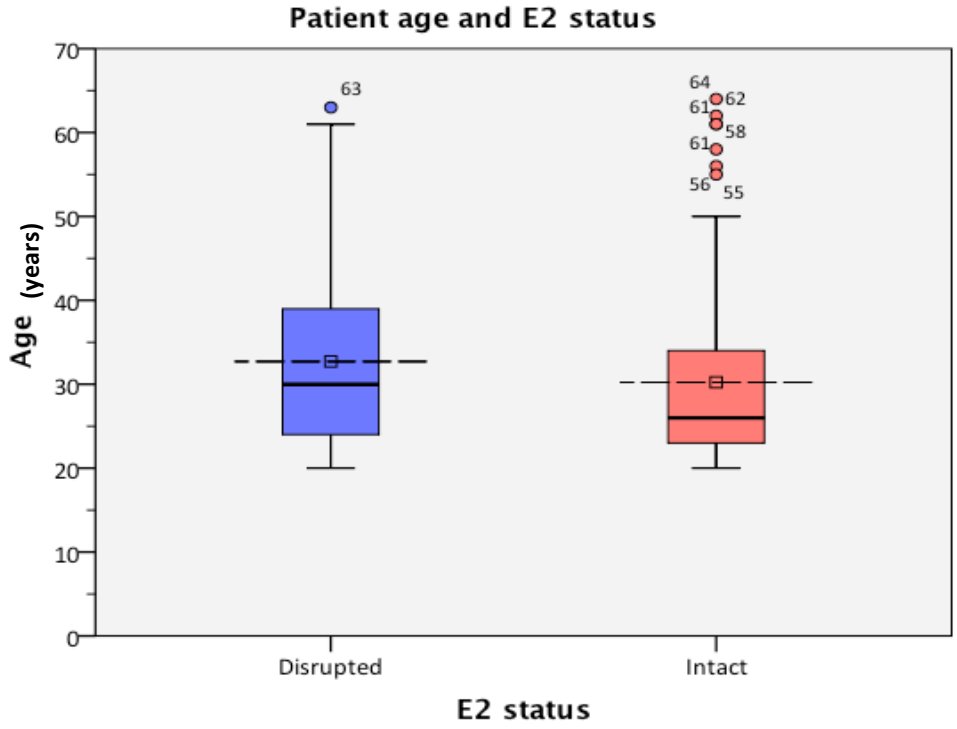

Figure 9.4: Box plot demonstrating the association between age and HPV16 E2 status.

The coloured rectangles represent the interquartile range (IQR), with the whiskers signifying the range of values up to 1.5 x IQR. Values outside this range (i.e., outliers) are signified by a matching coloured dot with the value alongside. The thick black line is the median value. The dashed line is the mean value.

9.2.4. HPV16 E2 status and referral cytology

E2 status was compared to referral cytology graphically (Figure 9.5) and statistically. Women with a borderline – query high-grade result had the highest proportion of E2 disruption (18/26, 69%; 95% CI [48%–85%]). However, a Chi-square test for independence indicated no significant association between E2 status and cytology of any grade, $\chi 2$ (4) = 2.21, P = 0.70. Chi-square assumptions were met.

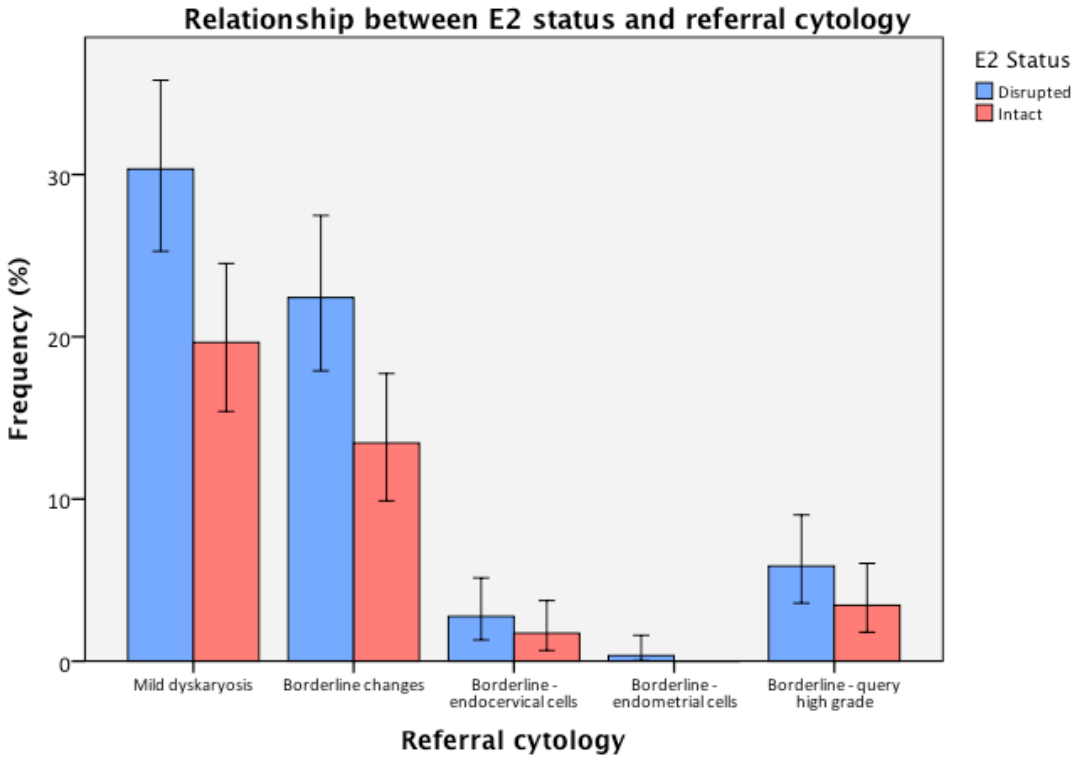

Figure 9.5: HPV16 E2 status according to referral cytology

9.2.5. HPV16 E2 status as a screening biomarker

The usefulness of E2 status as a screening biomarker was estimated using a contingency table (Table 9.1). The sensitivity of E2 status identifying women with HG CIN was 61% (95% CI [44–76%]) and the specificity was 38% (95% CI [31–46%]). The PPV was 18% (95% CI [12–26%]) and the NPV was 82% (95% CI [71–89%]).

	High-grade CIN		
	Absent	**Present**	**Totals**
E2 Disrupted	101	22	123
E2 Intact	63	14	77
Totals	164	36	200

Table 9.1: A contingency table showing how the status of HPV16 E2 relates to high-grade CIN

9.3. Viral integration – DIPS

9.3.1. Study population

DIPS PCRs were performed in duplicate on 45 samples (Figure 9.6), in order to assess HPV16 integration status. Initially 21 samples were selected from the first 350 eligible cases. Disrupted E2 was considered a surrogate marker of integration, and therefore, samples were initially selected based on their E2 status. Using approximately a 3:1 ratio, 15 cases with disrupted E2 and 6 cases of intact E2 were selected. Selection of cases and initial analysis of the data was performed whilst blinded to all clinical data including the histology result. Following Sau digestion and adapter ligation, ten HPV linear and nested primer sets were applied. An example of different PCR products obtained at the end of the DIPS process is shown in Figure 9.7.

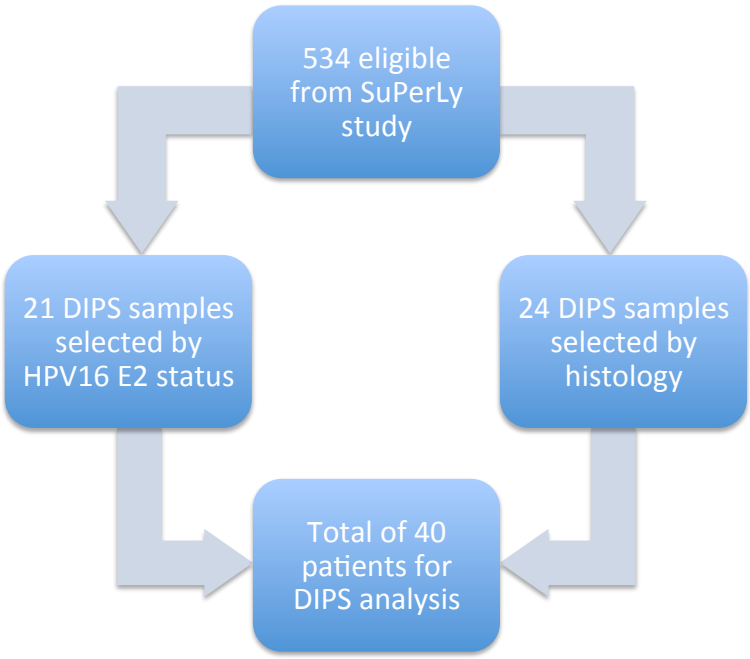

Figure 9.6: Sample flow for DIPS analysis.

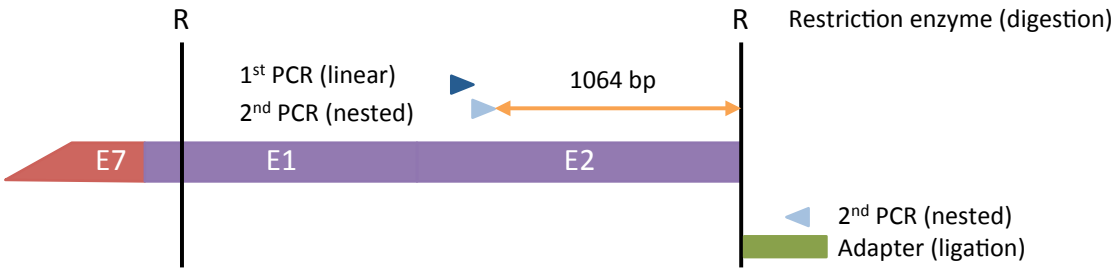

Figure 9.7a: Diagram showing DIPS assay using the P3 primer.

R indicates restriction enzyme cut site. The blue arrows indicate the linear and nested P3 primers. The orange arrow indicated the expected fragment size.

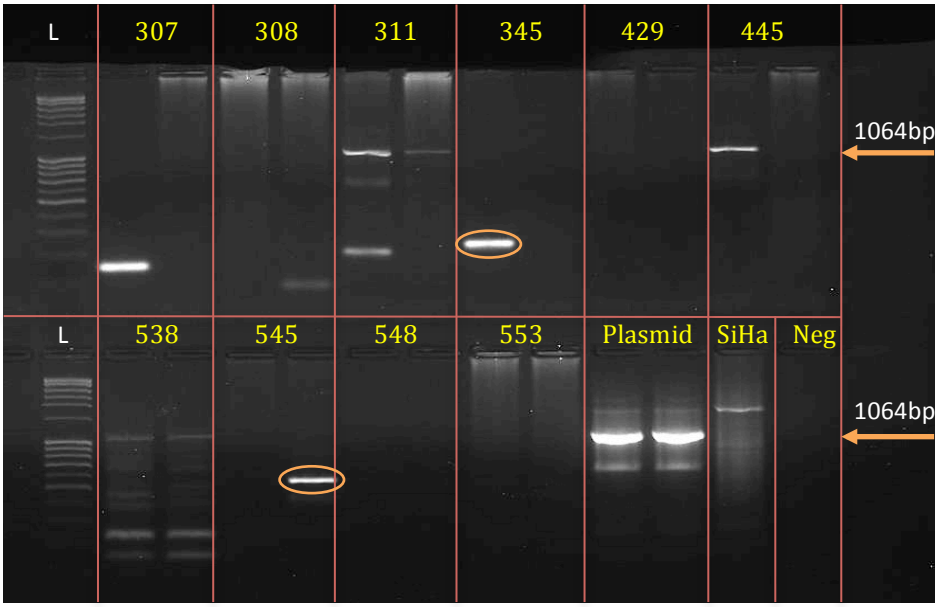

Figure 9.7b: Example electrophoresis image of PCR products obtained following DIPS using the P3 primer.

Duplicate runs for ten clinical samples are shown with HPV16 plasmid (episomal) and SiHa (integrated) as positive controls and water as DNA template negative control. The orange arrow indicates expected HPV fragment size (1064bp); present in samples 311, 445, 538. The bands within the orange ovals were sent for sequencing – both were HPV only. L, 100bp DNA ladder.

Out of the 21 cases analysed by DIPS, 13 cases had bands that differed from the plasmid control. The bands were excised and extracted from the gel as described in the methods (see section 7.7.2.2). Sequencing data for all bands excised aligned with HPV alone, using the BLAST and BLAT tools. Because no integration was identified in this selection and performing DIPS was not going to be possible on the entire sample set (mainly due to cost and the quantity of DNA required for the assay), it was decided to select subsequent cases based on the histology result. An example of the PCR products obtained following DIPS on the CIN3 sample set is shown in Figure 9.8.

The cases that had biopsy proven CIN3 were selected and DNA was extracted from the pellet. Only samples that were positive for HPV 16 by E6 PCR were included for analysis by DIPS. In this second sample set DIPS was performed on 24 cases (five of which had been through DIPS in the first sample set using their original SuPerLy study DNA). Two additional primers were used for the second set. Ten cases were identified as having different bands from the plasmid control. Of these nine cases had sequences that aligned to HPV DNA integrated with human DNA, and in the other case the DNA sequence aligned with HPV only.

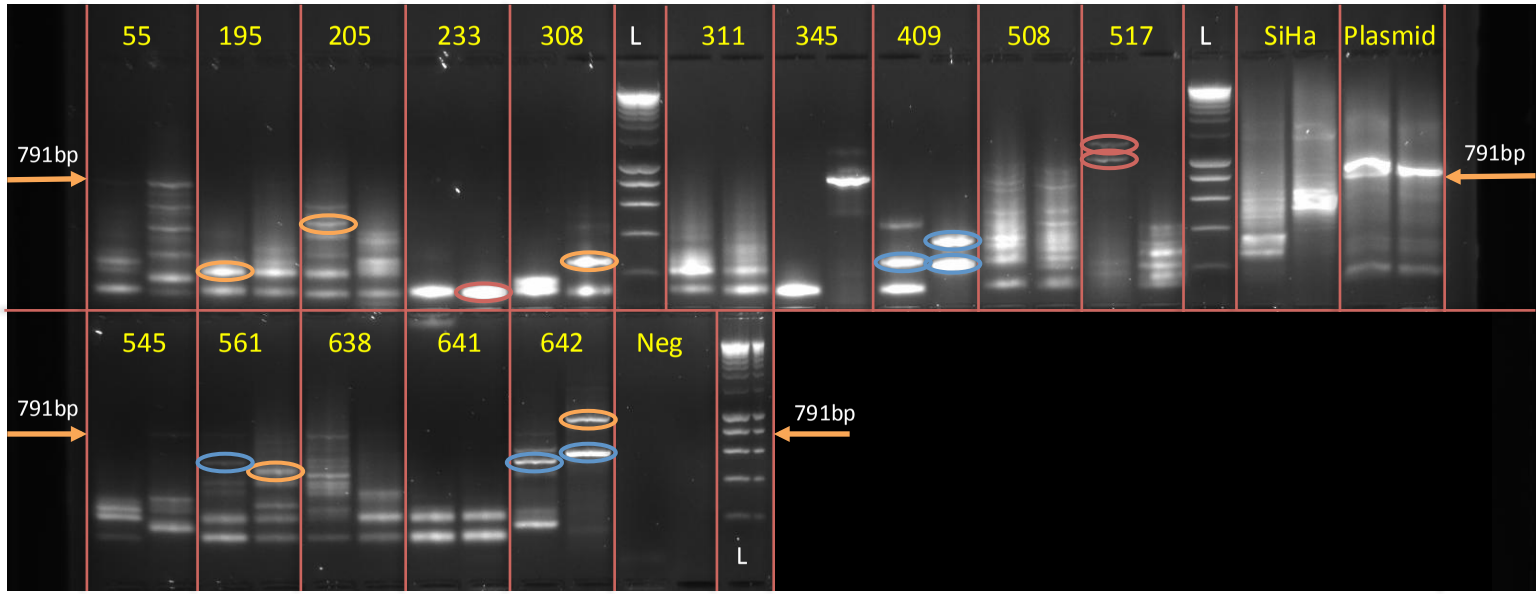

Figure 9.8: Electrophoresis image of PCR products following DIPS.

15 samples are shown with 3 controls (SiHa, Plasmid, Negative), all in duplicate repeats. The orange arrow indicates the expected size of fragment (i.e., from the nested primer to the next Sau site). The coloured ovals indicate which bands were extracted and sequenced. The results of the sequencing is indicated by the colour: orange oval, HPV only; blue oval, integration event; red oval, sequencing failed; L, wide-range DNA ladder.

9.3.2. Overall DIPS data

Overall, 9 (23%) of the women tested demonstrated viral integration in their cytology sample. In these 9 women, 24 separate integration events were identified. The mean number of events per patient was 3 (range 1–9). A summary of the integration events is shown in Table 9.2. For more detail regarding the site of integration see section 9.4.

Sample ID (Sequence ID)	Primer	Viral disruption (bp)[1]	HPV ORF	Chromosomal location[2]	Accession[3]	Match[4] (%)	Orientation
205 (9)	P4	3080	E2	Xq25	NT_011786.16	98	S
205 (10)	P4	3081	E2	1p13.3	NT_032977.9	100	AS
233 (11)	P4	3080	E2	9p21.1	NT_008470.19	100	AS
308 (43)	P7	6544	L1	10q26.3	NT_008818.16	100	S
409 (3)	P2	1789	E1	1q43	NT_032977.9	97	S
409 (24)	F4	2839	E1/E2	4p11	NT_016354.19	100	AS
409 (25)	F4	2830	E1/E2	15q26.3	NT_010194.17	100	S
409 (26)	F4	2831	E1/E2	1q31.2	NT_032977.9	99	AS
409 (44)	P7	6544	L1	8p23.2	NT_008046.16	100	AS
409 (45)	P7	6544	L1	14q23.1	NT_026437.12	99	S
508 (14)	P4	3080	E2	4q31.22	NT_016354.19	100	S
508 (15)	P4	3080	E2	5q33.1	NT_034772.6	100	AS
517 (5)	P2	1887	E1	1p21.3	NT_032977.9	100	S
517 (46)	P7	6544	L1	1q25.1	NT_032977.9	100	AS
545 (16)	P4	3080	E2	6p21.2	NT_025741.15	100	AS
561 (17)	P4	3080	E2	13q33.1	NT_009952.14	98	S
561 (18)	P4	3080	E2	14q13.3	NT_026437.12	100	S
561 (29)	F4	2839	E1/E2	RPT	—	98	AS
642 (19)	P4	3080	E2	2q36.3	NT_005403.17	100	AS
642 (31)	F4	3080	E2	13q32.1	NT_009952.14	100	AS
642 (33)	F4	3080	E2	7q36.1	NT_007933.15	100	AS
642 (39)	P5	5030	L2	16p13.13	NT_010498.15	100	AS
642 (40)	P5	4911	L2	5q31.1	NT_034772.6	99	S
642 (41)	P5	5073	L2	RPT	—	95-98	AS
642 (42)	P5	5086	L2	22q11.23	NT_011520.12	100	AS
642 (47)	P7	6544	L1	1q41	NT_032977.9	100	S
642 (48)	P7	6543	L1	3p21.1	NT_022517.18	100	AS

Table 9.2: Integration events identified using DIPS

[1]Base pair number of the last viral nucleotide before recombination to human sequence (according to GenBank accession number NC_001526). [2]Location of integration into human genome using UCSC database; RPT, repeat sequence with multiple hits. [3]EMBL Accession number from the NCBI database of the sequence with most likeness to human sequence data. [4]Percentage of agreement with NCBI database sequence. S, sense; AS, antisense.

9.3.3. Viral integration detected by DIPS and histology

The histological outcome for the 40 cases analysed by DIPS were as follows: 24 CIN3, 2 CIN1, 4 HPV, 7 No CIN / No HPV, and 3 that had not had a biopsy. After excluding the 3 that did not have a biopsy the groups were compared using a stacked bar chart (Figure 9.9) and Chi-Square tests. Integration was identified in 9 out of the 24 women (38%; 95% CI [20–59%]) with high-grade CIN (only women with CIN3 were tested) and no integration was found by DIPS in the 16 women with low-grade or negative histology (95% CI [0–24%]). This difference was significant (Fisher's Exact test; *P* = 0.02). The requirements for chi-square were met.

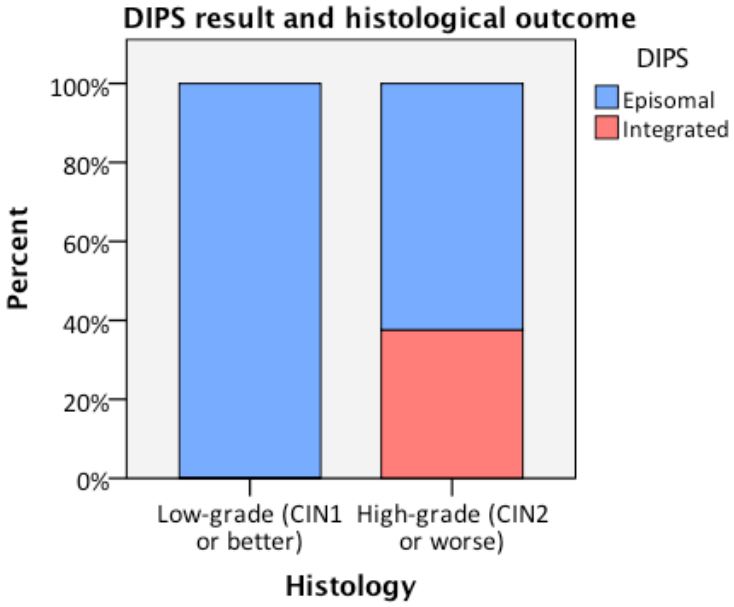

Figure 9.9: Comparison of DIPS result and histological outcome.
Low-grade includes CIN1, HPV, No CIN / No HPV. High-grade includes CIN2, CIN3, invasive.

9.3.4. Viral integration detected by DIPS and age

The ages of women were compared between the two groups: integration detected and no integration detected (considered episomal) (Figure 9.10). The episomal group appeared to be slightly older than the integrated group (Md = 28, *n* = 31; Md = 26, *n* = 9 respectively) and had a much greater spread of data (IQR = 14 and IQR = 5 respectively). A Mann-Whitney U test revealed, however, the difference in the ages was not significant, *U* = 115, *z* = -0.81, *P* = 0.42.

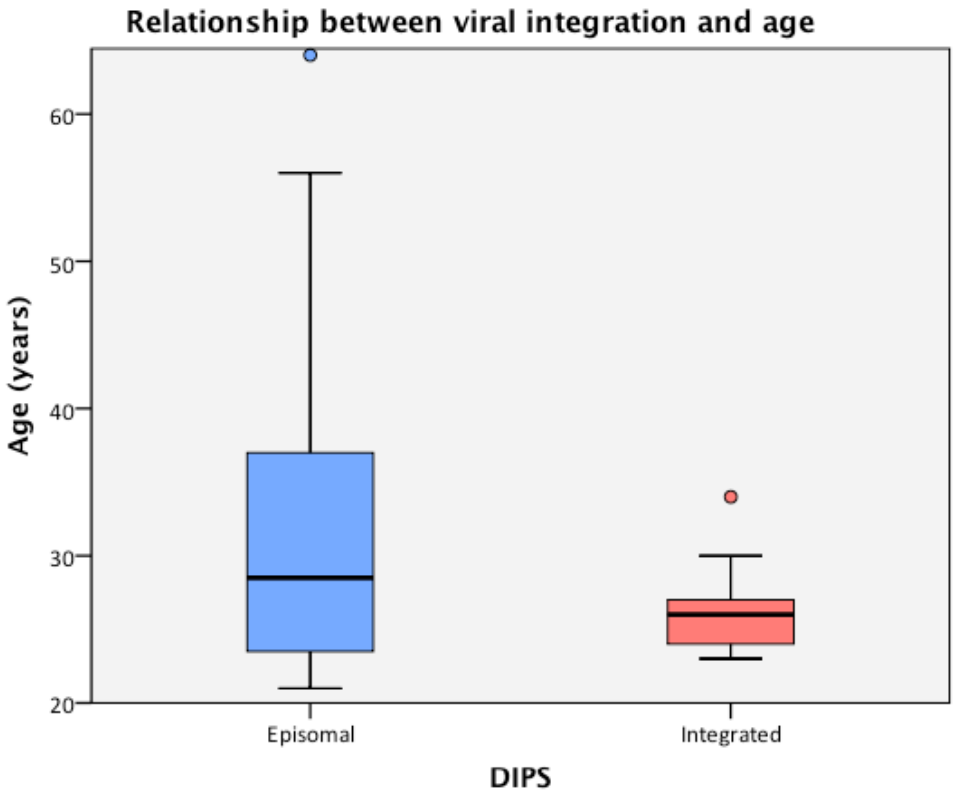

Figure 9.10: Box plot showing the difference in ages between women with episomal forms and women with integrated forms.

9.3.5. Viral integration detected by DIPS and referral cytology

Referral cytology and the DIPS result were compared to test for any association (Figure 9.11). The highest proportion of integration was seen in borderline – query high-grade group (*n* = 3 out of 7 [43%]). In the mild dyskaryosis and the borderline changes endocervical cells categories integration was seen in 4 out of 15 (27%) and 2 out of 5 (40%), respectively. No integration was seen in the borderline changes group. A Chi-square test for independence indicated the difference between the groups was not quite statistically significant, χ^2 (3) = 6.87, *P* = 0.08.

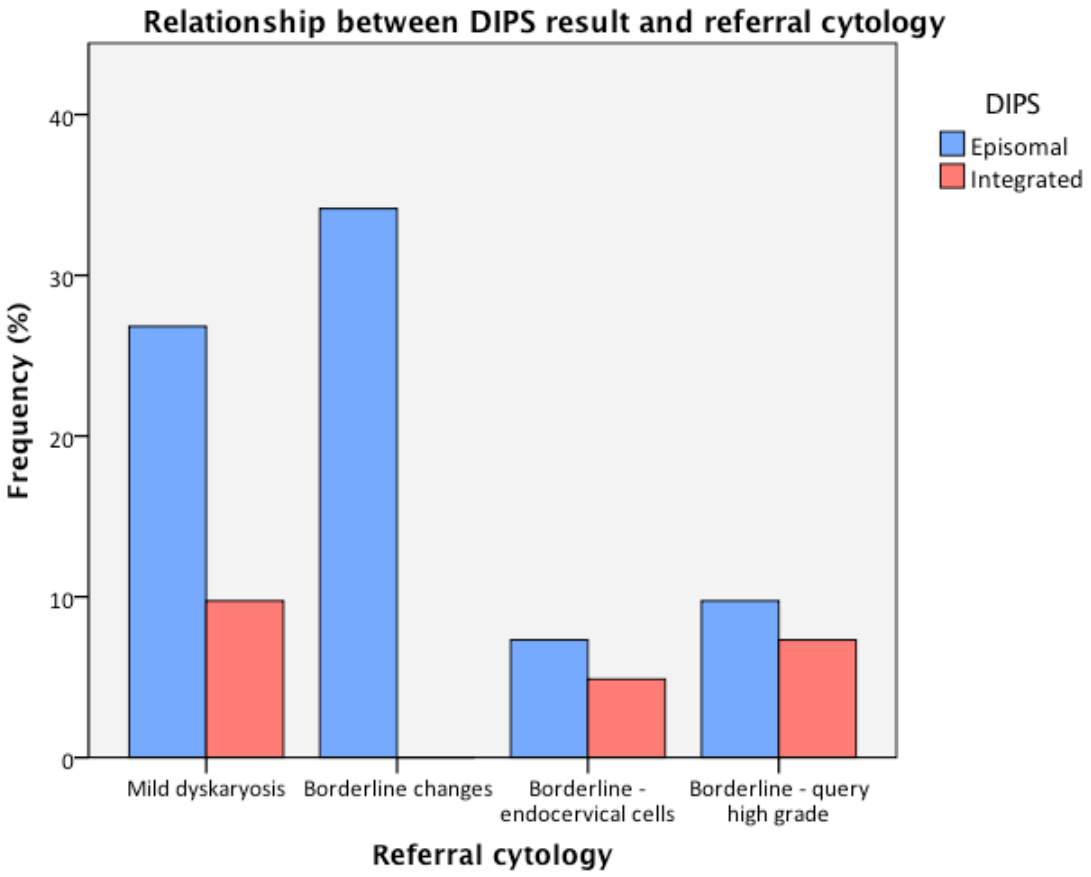

Figure 9.11: The relationship between viral integration and referral cytology.

9.3.6. Viral integration as a biomarker

The usefulness of viral integration as a screening test or biomarker was estimated using a contingency table (Table 9.3). The sensitivity of viral integration (as detected by DIPS) identifying women with CIN2+ was 38% (95% CI [20–59%]) and the specificity was 100% (95% CI [72–100%]). The PPV was 100% (95% CI [63–100%]) and the NPV was 46% (95% CI [28–66%]).

	High-grade CIN		
	Absent	**Present**	**Totals**
Integration detected	0	9	9
No integration detected	13	15	28
Totals	13	24	37

Table 9.3: A contingency table showing how the viral integration detected by DIPS relates to HG CIN. Samples that did not have a biopsy taken were excluded from this table.

9.4. Site of integration

Details of the site of integration for all the events identified by DIPS are shown in Table 9.4.

Sample ID (Sequence ID)	Viral disruption (bp)[1]	Chromosomal location[2]	Gene[3]	Fragile site[4]	RPT[5]	Viral cellular junction[6]
205 (9)	E2;3080	Xq25	GRIA3			GGATGT
205 (10)	E2;3081	1p13.3	NTNG1			GGATG
233 (11)	E2;3080	9p21.1	(TMEM215, 15)	FRA9C(C)		GATG
308 (43)	L1;6544	10q26.3		FRA10F(C)	SINE	GGCATC
409 (3)	E1;1789	1q43	PLD5			GAGCC
409 (24)	E1/E2;2839	4p11			SAT	TGACC
409 (25)	E1/E2;2830	15q26.3	(MEF2a, 24)		LINE	AGACC
409 (26)	E1/E2;2831	1q31.2	(RGS2, 17)	FRA1K(C)	SINE	AGACCT
409 (44)	L1;6544	8p23.2	CSMD1			CATC
409 (45)	L1;6544	14q23.1		FRA14B(C)	SINE	CATC
508 (14)	E2;3080	4q31.22				GATG
508 (15)	E2;3080	5q33.1	SLC36A3			GATG
517 (5)	E1;1887	1p21.3	(SNX7, 22)	FRA1M(R)	LINE	GATATAAA
517 (46)	L1;6544	1q25.1	(TNR, 3)	FRA1G(C)		CATC
545 (16)	E2;3080	6p21.2	KIF6		LINE	GATG
561 (17)	E2;3080	13q33.1			LTR	GATG
561 (18)	E2;3080	14q13.3	SLC25A21			GATG
561 (29)	E1/E2;2839	RPT			LINE	GACC
642 (19)	E2;3080	2q36.3	RHBDD1		DNA	GATG
642 (31)	E2;3080	13q32.1	(ABCC4, 6)	FRA13D(C)	LTR	GGATG
642 (33)	E2;3080	7q36.1	(ZNF783, 30)	FRA7I(C)		GATG
642 (39)	L2;5030	16p13.13	SNX29	FRA16A(R)	SINE/DNA	AG
642 (40)	L2;4911	5q31.1		FRA5C(C)		TCCA
642 (41)	L2;5073	RPT			SINE/DNA	GATC
642 (42)	L2;5086	22q11.23	GSBP11			A
642 (47)	L1;6544	1q41				CATC
642 (48)	L1;6543	3p21.1	CACNA2D3			CATCA

Table 9.4: Significant features of the integration sites.

[1]Base pair number and site within HPV16 genome (GenBank accession number NC_001526). [2]Location of integration into human genome using UCSC database; RPT, repeat sequence with multiple hits. [3]Human gene involved at site of integration; genes in parentheses were identified within 50Kbp of integration site (exact distance follows the gene [given in Kbp]). [4]Fragile site reported at site of integrant; R, rare; C, common (Debacker and Kooy, 2007). [5]Repeat elements within the human DNA identified; SINE, short interspersed nuclear element; SAT, satellite DNA made up of tandem repeats; LINE, long interspersed nuclear element; LTR, long transposed region; DNA, DNA transposon. [6]Overlapping sequence at viral-cellular junction.

The most common site of viral disruption in the 27 integrants described was within the E2 ORF. Including the 4 integrants where the sequence identified overlapped the E1 and E2 ORF, 15 (56%; 95% CI [37–72%]) of the samples involved the E2 ORF. A further 2 (7%; 95% CI [2–23%]) involved just the E1 ORF, 6 (22%; 95% CI [11–41%]) the L1 ORF, and 4 (15%; 95% CI [6–32%]) the L2 ORF. Out of the 9 women in whom integrants were found, 8 (89%; 95% CI [51–99%]) had at least one integrant disrupting the E2 ORF.

When viral disruption occurred in either E2 or L1, the break points were very consistent, with 3080bp and 6544bp respectively the most common locations. Integrants involving E1 or L2, however, had more varied break points.

9.4.1. Integration and correlation with tiling PCRs

In order to further assess the physical status of HPV in the samples that showed integration events, PCRs were performed for E1, E2, E6, and L1 (Table 9.5). The results of the PCRs were not entirely consistent with the DIPS findings, however, a finding of intact tiling PCRs does not rule out the possibility of a mixture of integrants and episomes.

Sample ID	Break point in DIPS	E1 PCR	E2 PCR	E6 PCR	L1 PCR
205	E2	Intact	Intact	Intact	Intact
233	E2	Disrupted	Disrupted	Intact	Intact
308	L1	Intact	Disrupted	Intact	Intact
409	E1/E2, L1	Disrupted	Intact	Intact	Intact
508	E2	Intact	Intact	Intact	Intact
517	E1, L1	Intact	Intact	Intact	Intact
545	E2	Intact	Disrupted	Intact	Intact
561	E1/E2	Intact	Disrupted	Intact	Intact
642	E2, L1, L2	Intact	Intact	Intact	Intact

Table 9.5: HPV tiling PCRs' correlation with integration events.

9.4.2. Site of integration into human genome

The sites of integration were spread throughout 15 out of 23 (65%) chromosomes. In this study all the integration sites found were unique for each woman. The integration events are shown in Figure 9.13 alongside all the sites found in a systematic review on the topic (Wentzensen et al., 2004). When data from this study was compared to Wentzensen et al. (2004), half of the

integration sites were in the same chromosome bands as previously reported; whereas the other half were in new sites. The most common chromosome involved in this study was chromosome 1; 6 (22%) of the integrants contained DNA sequences that strongly matched to bands from chromosome 1. Figure 9.12 is a chromagram from one of the samples demonstrating integration between HPV (E1/E2) and human DNA (1q31.2).

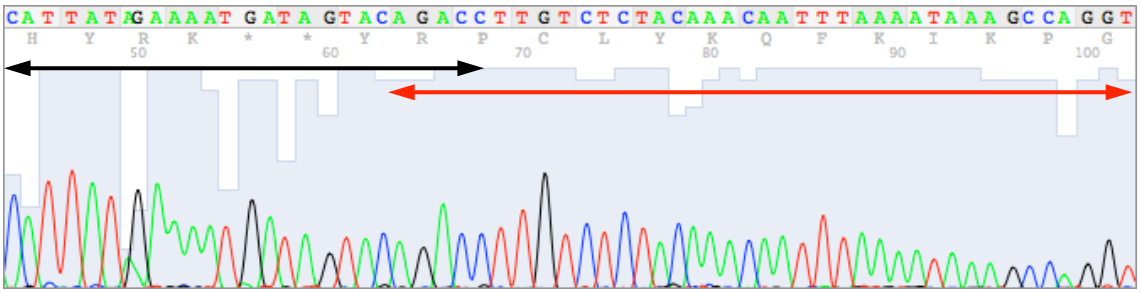

Figure 9.12: Chromagram from sample ID 405(26).

The black arrow indicates DNA sequence that aligns to 2805–2830bp (E1/E2) of the HPV16 genome. The red arrow indicates DNA sequence that aligns to 192269139–192269091bp (1q31.2).

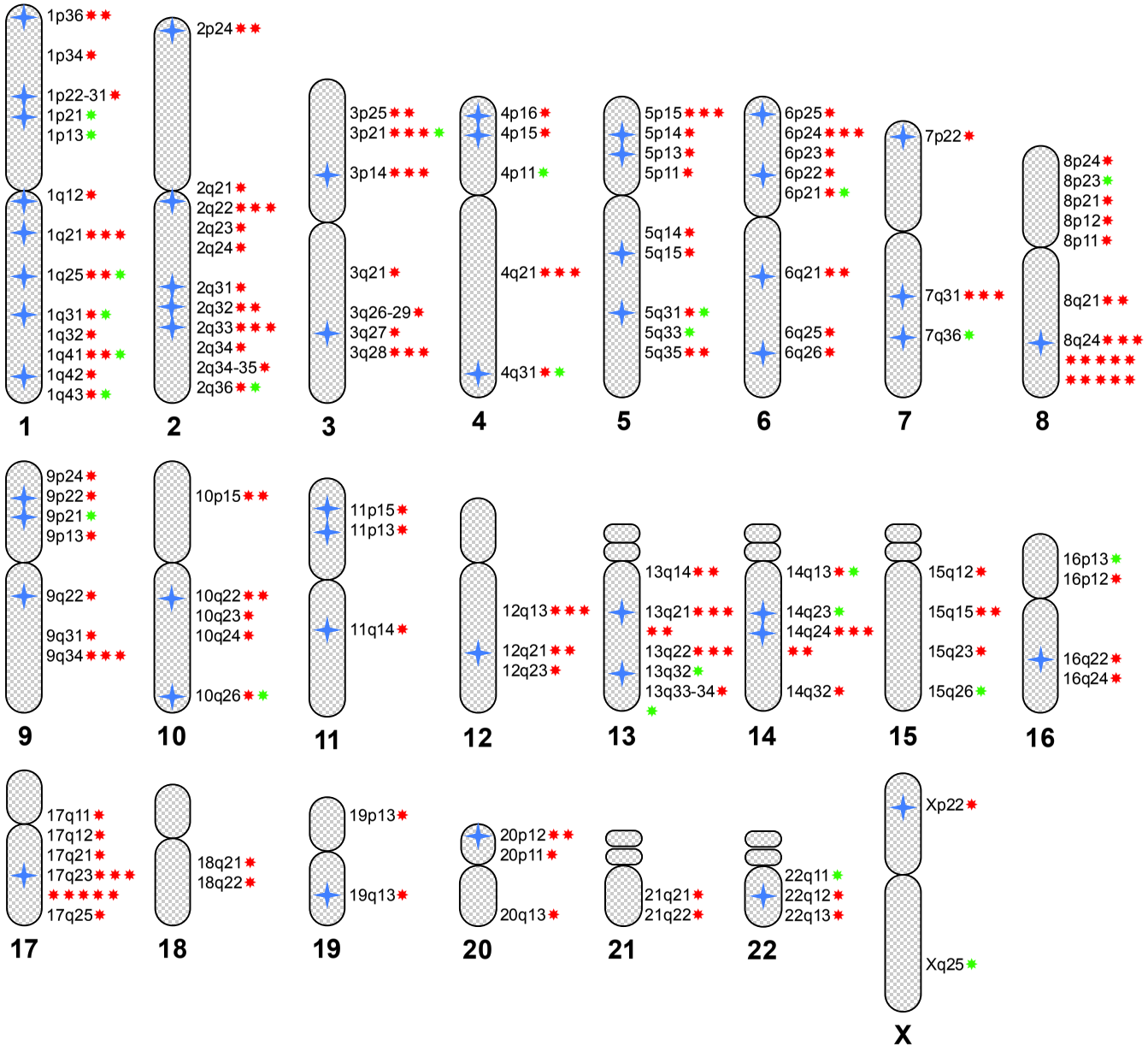

Figure 9.13: An overview of the integration sites within the human genome.

Red stars indicate integration sites identified in (Wentzensen et al., 2004). Green stars indicate integration events identified in this thesis. Blue crosses represent common fragile sites as listed in the NCBI map viewer.

9.4.2.1. Integration and mechanism

Integration was detected at known fragile sites in 10 samples (37%; 95% CI [22–56%]; 8 were within common fragile sites (CFSs), 2 were within rare fragile sites. Only half of the fragile sites have previously been associated with integration events in a systematic review (Wentzensen et al., 2004). Two of the CFSs have been molecularly mapped: FRA7I and FRA10F (Ma et al., 2012). A break in FRA7I has been associated with carcinogenesis in breast cancer (Ciullo et al., 2002). FRA10F contains a cancer-associated CFS gene known as *FATS* (Fragile-site Associated Tumour Suppressor), which plays a role in regulating DNA damage checkpoints and suppressing tumourigenesis (Li et al., 2010, Ma et al., 2012).

In 11 (41%) samples, HPV DNA was integrated directly into the ORF of human genes (Table 9.4). In a further 7 (26%) samples, integration occurred within 50Kbp of human genes. Some of the genes concerned had previously been associated with cancer (Table 9.6).

Gene	Association with cancer
Neptin G1 (NTNG1)	Colon cancer (Yi et al., 2011)
Glutamate receptor (GRIA3)	Pancreatic cancer (Ripka et al., 2010)
Regulator of G-protein signaling 2 (RGS2)	Prostate, colorectal, and breast cancers (Cao et al., 2006, Jiang et al., 2010, Smalley et al., 2007)
CUB and Sushi multiple domains (CSMD1)	Breast, lung, skin, head and neck cancers (Kamal et al., 2010), (Ma et al., 2009)
Rhomboid-related protein (RHBDD1)	Modulates apoptotic activity (Wang et al., 2008)
Calcium channel, voltage-depenedent, alpha2/delta subunit 3 (CACNA2D3)	Renal and gastric cancers (Hanke et al., 2001, Wanajo et al., 2008)

Table 9.6: Genes identified at or very close (within 50 Kbp) to site of integration and their association with cancer.

A significant proportion (*n* = 13, 48%) of human DNA aligned in this study was found to contain repeat elements. The most common was the short interspersed nuclear element (Table 9.4).

9.5. Discussion

9.5.1. HPV16 E2 status

One way of investigating viral integration is by using E2 PCR to assess the integrity of the E2 gene, which is a common site of integration. Whether E2 disruption is one of the first events to enable the HPV virus to persist and evade the host immune system is still not clear. Several

studies have shown evidence of early E2 disruption and viral integration in low-grade precancerous lesions (Li et al., 2008, Huang et al., 2008, Collins et al., 2009, Gallo et al., 2003). However, many studies have reported that integration only occurs in high-grade lesions and cancer (Daniel et al., 1995, Hudelist et al., 2004). Evidence suggests that the degree of E2 disruption, and therefore, integration correlates with the grade of cervical lesion (Li et al., 2008). The status of the E2 gene has been suggested as a possible biomarker of disease progression. One of the hypotheses of this study was that integration could predict high-grade cervical disease in women with persistent low-grade cytology.

The data presented in this thesis suggests that E2 disruption is a common and early event. Furthermore, it appears just as likely to occur in low-grade disease as it does in high-grade disease. This is consistent with a longitudinal study that recruited 15–19 year olds and followed them for up to nine years (Collins et al., 2009). The incomplete cytological or histological data in the study make direct comparisons difficult. Nevertheless, they also showed early disruption to the E2 gene, in HPV infections, at baseline and during follow-up (total E2 disruption 56% (36/64)). Huang *et al.* (2008) also reported high rates of integration in CIN1 (83%), but showed no significant difference between the different grades of disease. In two other studies that used E2 PCR to assess E2 status in cervical samples and snap-frozen biopsies from women with high-grade CIN, 61-67% were found to have disrupted E2, which correlate very closely to the findings of this study (Alazawi et al., 2004, Tonon et al., 2001).

The women within the SuPerLy study were selected on the basis of repeated low-grade cytological abnormalities. Normally the host clears HPV infection within months. However, this cohort of women will most likely have had the same HPV infection persisting for between 12-18 months and probably in many cases a lot longer. The E2 data within this study suggests that E2 disruption could be a sign of persistent HPV infection rather than a sign of high-grade disease. Changes at the molecular level appear to precede morphological changes at the epithelial surface. The physical changes in the epithelium that result from persistent HPV occur over several months to years and it is widely believed that increased expression of E6 and E7 is one of the requirements for the infection enduring for so long (Doorbar et al., 2012). When E2, a regulator of E6 and E7 expression, is disrupted then this prerequisite is achieved, the virus persists and epithelial restructuring becomes evident. This theory can explain why the majority of women in this study showed evidence of E2 disruption.

The E2 disruption model, however, does not explain how the infection persisted where E2 was found to be intact. One explanation is that E2 disrupted DNA coexisted with E2 intact DNA and

because the assay was PCR-based, it would not have required much intact E2 for sufficient amplification to occur. It is also conceivable that intact E2 DNA may be found in superficial layers long after disruption of DNA has occurred in the basal layers of epithelium where viral transcription occurs. The fact that cervical sampling devices mainly scrape the surface of the cervix would support this notion (Dey et al., 1996). It is quite plausible that there is more than one focus of infection within the relatively large TZ, perhaps at different stages of the life cycle. Moreover, an infected partner may continue to introduce viral episomes from their own productive infection. In the same vein, multi-type infections are commonly discovered on HPV testing (49% in the SuPerLy study). It is not known to what extent different HPV types may interact within the host and affect any screening investigations.

As can be seen in the DIPS data presented in this thesis and from previous studies (Thorland et al., 2003, Wentzensen et al., 2004) viral disruption may occur at other sites in the HPV genome. This disruption could represent another mechanism that causes deregulation of the oncogenes E6 and E7 or a different process of carcinogenesis altogether. It holds, therefore, that E2 disruption is not the only way of promoting prolonged viral infection, further genomic instability and carcinogenesis.

There is evidence that in approximately 12.5% of cervical cancers, only transcripts from episomal HPV are present (Klaes et al., 1999). This would suggest that there might be an episomal-driven carcinogenesis with quantitative deregulation without the need for integration or physical E2 disruption (Pett and Coleman, 2007). This is an important consideration should detection of integration ever be incorporated into the screening programme.

9.5.2. DIPS

Another way for investigating integration status in DNA from cervical samples is the DIPS PCR assays. In the SuPerLy – HIM study, integration of HPV16 into human DNA was identified using DIPS in 23% of samples tested. Notably, integration events were only detected in women with CIN3 on histology. These findings would point toward integration being a late event found in precancerous high-grade lesions. Several studies support this view (Klaes et al., 1999, Hafner et al., 2008, Melsheimer et al., 2004, Matovina et al., 2009, Hopman et al., 2004, Hudelist et al., 2004), however, others have published evidence that integration occurs earlier, even in low-grade lesions (Li et al., 2012, Huang et al., 2008). Some of this discordance can be put down to the different methods of identifying integration (e.g. DIPS, APOT, RSPCR, E2 PCR, qPCR) and

the variety of cervical samples (e.g., smears, swabs, and biopsies) fixed in a variety of media (e.g., LBC, paraffin, liquid nitrogen) that these methods have been applied to. One of the major differences between DIPS and APOT is that, DIPS identifies any HPV integration within the DNA, whereas APOT will only find transcriptionally active integrants. It is possible that as the viral genome becomes unstable multiple integration events occur, however, only a very small number of these are selected and transcription occurs involving the integrated DNA. It has been postulated that the non-transcriptionally active integrants may remain in a latent state until activated by a cellular selection process (Pett and Coleman, 2007).

9.5.3. Use as a biomarker

In this study HPV16 E2 status did not appear to be useful at discriminating different disease grades but instead appeared to be more indicative of a persistent infection. There is still a potential that it could be used as a biomarker to separate out a transient infection from a persistent infection. This would be more appropriate in a primary screening setting with HPV as the first line test and a subsequent E2 status test if the woman was HPV positive. Disrupted E2 could indicate a persistent infection and potential high-grade disease, whereas, an intact E2 may indicate a transient infection is more likely and a repeat HPV test could be done some time later. There are many flaws to this proposed use as a screening biomarker. The main one is the relatively frequent finding of intact E2 in high-grade lesions and cancer, where either episomes co-exist in the background of integrants or an alternative episomal driven carcinogenesis is responsible. Unless there is a way to identify these cases at the same time then the future of an E2 biomarker is likely to be very limited. Furthermore, the assay would need to be adapted for different HPV types due to sequence variation.

Identification of integration events using DIPS did show much better discrimination between disease grades. The results showed a relatively high specificity and PPV so that if it were used as a biomarker then a positive result would be strongly associated with a high-grade lesion. However, with such a low sensitivity and NPV, a negative DIPS result would not give much reassurance that there was only low-grade disease. However, if it was combined with another test that had a high sensitivity and NPV; e.g., an hrHPV test, it could potentially be very useful.

9.5.4. Site of integration

When any one study looking at HPV integration in cervical cancer is examined in isolation it would appear that integration is entirely random. However, when studies are combined, for

example in Figure 9.13, some patterns do emerge. The most common site for integration is 8q24, which corresponds with the location of the *MYC* oncogene. *MYC* has been associated with a number of different cellular pathways influencing cell proliferation, differentiation, genomic stability, and tumourigenesis (Meyer and Penn, 2008). This integration site was not found in this study; however, it is usually only found in cervical cancer and cancer cell lines (Wentzensen et al., 2004). Despite that, studies investigating the potential of *MYC*-copy number as a biomarker for high-grade CIN found that it was less sensitive but more specific compared to cytology (Obermann et al., 2013).

There were several cancer-associated genes found at the site of integration in this study (Table 9.5). Interestingly, the six genes identified were found in three different women (two each). Several studies report only one integration event per individual (Ziegert et al., 2003). However, in this study multiple integration events were found in the majority (67%) of women. Another recent study supports this finding; it reported multiple integration events (using an expanded DIPS method) in 55% of samples (Li et al., 2013). Thus, multiple integration events may represent early clonal events with only one or two integrants involving cancer-associated genes that offer a selective growth advantage resulting in progression to cancer. A study supporting this notion found that only one out of three integrants identified (by DIPS) was transcriptionally active; furthermore the study showed loss of gene function as a result of insertional mutagenesis (Schmitz et al., 2012). The cancer-associated genes found in this study have not previously been reported as integration sites for HPV and they may represent novel mechanisms for cervical carcinogenesis. In order to assess their role further it would be important to establish if the integrants were transcriptionally active using APOT.

CFSs have previously been linked with tumour-associated viruses (Thorland et al., 2000, Wilke et al., 1996); moreover, they are frequently reported as sites of integration in high-grade cervical disease and cervical cancer (38%–55%)(Dall et al., 2008, Thorland et al., 2003, Wentzensen et al., 2004, Yu et al., 2005). The data in this thesis has reinforced the finding, demonstrating integration into CFSs in 56% of women. It is still not clear, however, whether CFSs are targeted by viral integration because of a selection advantage resulting from physical or functional alterations, or simply because their characteristic instability makes them more susceptible.

Identification of repeating sequences at the site of integration has not been reported in many studies. In this study, 48% of integration events involved repetitive elements, although, if analysed for each individual woman it was 67%. Repeating elements have been reported in

45% of the human genome. Hence, the finding in this study is not particularly significant. However, another study using the DIPS method found that 92% of integration events were found in repetitive elements when analysing the viral-cellular junctions (Li et al., 2013).

Viral disruption was predominantly seen in E2 and E1 ORFs (63%), a discovery that echoes the findings of several integration studies (Luft et al., 2001, Matovina et al., 2009, Ziegert et al., 2003). It has not been commonly reported that HPV has recurrent sites of disruption. Another study using the DIPS method (with an additional 12 primer sets) found disruption repeated between 1–6 times at certain nucleotides (Li et al., 2013). However, none of those nucleotides were found in this study. The most common site in this study was at nucleotide 3080. The sequence just preceding this site is GATG and commonly there was a 4bp overlap at the junction with human DNA. This sequence is quite similar to the sequence that is cut by Sau restriction enzyme. There is, therefore, a possibility that the enzyme may be cutting the viral and human DNA at inappropriate points (this is known as star activity). However, there was plenty of evidence to suggest the results were genuine. No star activity was found in women with low-grade histology at all. The correct adapter sequence was found on the end of the majority of integration events within this study. When the adapter sequence was not found the chromagram showed that the DNA signal was petering out. Furthermore, the tiling HPV PCRs corresponded in most cases with the site of viral disruption identified by DIPS (Table 9.6). Hence, these repeated sites may represent fragile sites within the HPV genome. In order to be even more confident the viral-cellular junction could be tested with flanking PCR primers confirming integration. In addition, it would be useful to perform APOT, although that would not be possible without obtaining a new sample in order to extract quality RNA.

9.5.5. Strengths and weaknesses

The physical state of HPV16 was assessed for women with low-grade cytological abnormalities. The integration assays were applied to LBC samples that were obtained in a real world setting. The two assays were assessed and the resulting analysis will be extremely useful when planning larger clinical trials. Two years of follow-up data will be also available and it will be very interesting to see whether the HPV16 E2 status at baseline may have been a predictor of subsequent high-grade disease.

The E2 PCR method is not able to identify disruption elsewhere in the viral genome and, furthermore, it does not recognise when a full-length, head-to-tail tandem repeat integration occurs (as in the case of CaSki). When no bands are produced in this PCR it is difficult to know

for sure whether it is because there is disruption, if the DNA was at too low concentrations or the PCR failed for another reason. Similar limitations apply to DIPS where false negative results are also possible. Using consistent concentrations, performing repeats and use of controls does, however, provide reassurance in the validity of the data.

One of the inherent problems with DIPS is the use of agarose gel, which is known to have poor resolution. The observer selects the bands based on size so that any bands that are not of the expected size are extracted and sent for sequencing. Ideally every band would be sent for sequencing as even if a correctly sized band is produced it is still possible that there is integration within it. However, sequencing is an expensive process and success also relies on certain DNA concentrations. Despite these limitations, DIPS is good at providing an accurate determination of the site of integration, which can be a problem for other assays.

Another consideration is the practicalities of the tests particularly if they were to be used in a screening setting. E2 PCRs are straightforward, inexpensive and amenable to high-throughput; on the other hand, DIPS is costly, laborious and time-consuming. The DIPS method, in its current format, would be impractical in a screening setting unless it was a second or third test for a small subset of the population. Both tests are also type specific, so development of assays for other high-risk types would be necessary. Detection of integration using real-time PCR technique may offer a more accurate, quantitative and potentially high-throughput method.

9.6. Conclusion

The frequent finding of E2 disruption in this study suggested that it was an early event and more a marker of persistent infection rather than a transforming infection. E2 PCRs offer a limited assessment of integration and there are a number of issues with the assay itself. On the other hand, the results using the DIPS methodology demonstrated a highly specific test for high-grade cervical disease. Recurring and novel features of the integration sites were found. However, in order to be used as a screening biomarker further testing in a greater number of samples is required; furthermore the development of less labour intensive techniques should be a priority.

Chapter 10 – RESULTS AND DISCUSSION – METHYLATION

10.1. Introduction

This chapter presents the results from the investigation of HPV16 DNA methylation within the SuPerLy–HIM study. The reproducibility of the method is considered before detailing the methylation results. The main hypothesis tested in this chapter was:

H_1 = Hypermethylation within the viral genome correlates with high-grade cervical disease development.

10.2. Study population

In order to investigate viral DNA methylation a small number of samples from the SuPerLy study were selected for BS conversion and pyrosequencing. The sample flow for this stage of the study is shown in Figure 10.1.

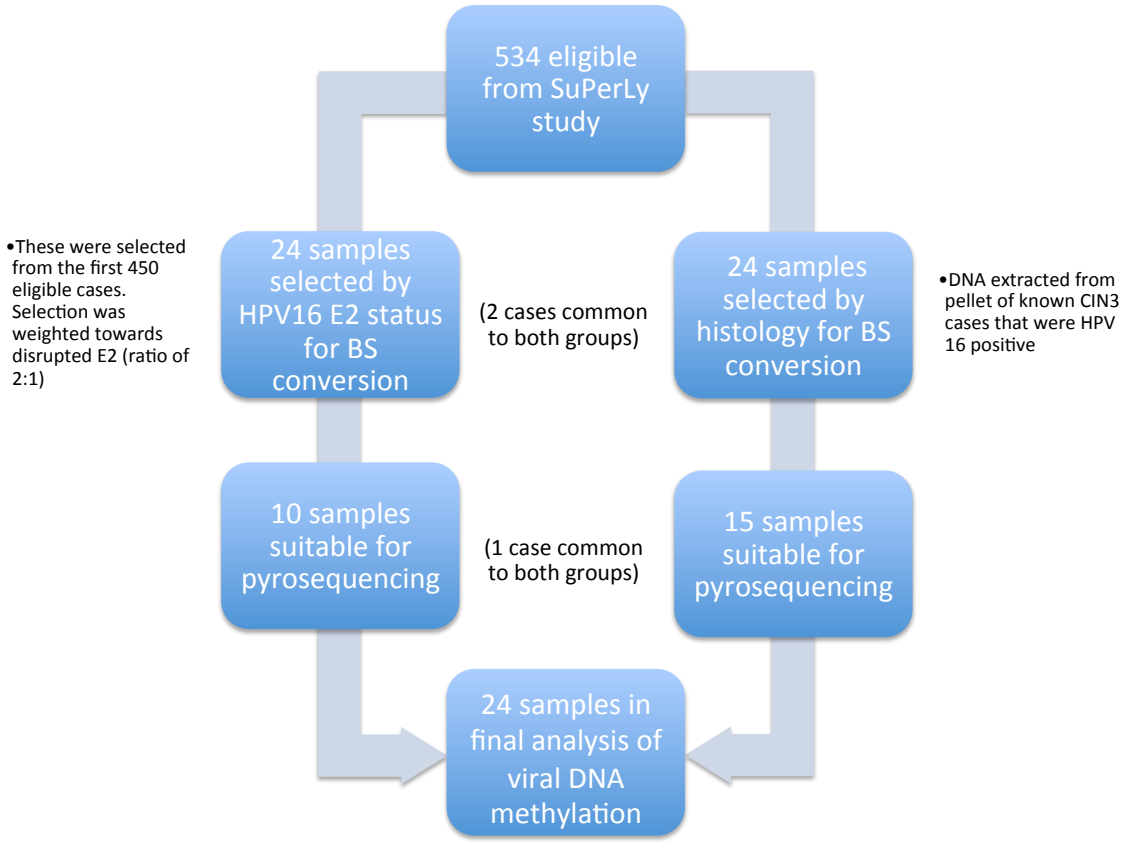

Figure 10.1: Sample flow for viral DNA methylation analysis

It was not possible to select the same samples that were chosen for DIPS analysis in the first sample set (shown on the left side of Figure 9.6) due to the amount of DNA required for both assays. The second sample set in both DIPS and BS conversion was from the same group of patients, in which DNA was extracted from the cell pellet (shown on the right sides of Figure

9.6 and Figure 10.1). After BS conversion the PCR was performed on all converted samples in duplicate. Only samples that produced the correct size band by gel electrophoresis were used in subsequent pyrosequencing reactions (Figure 10.2).

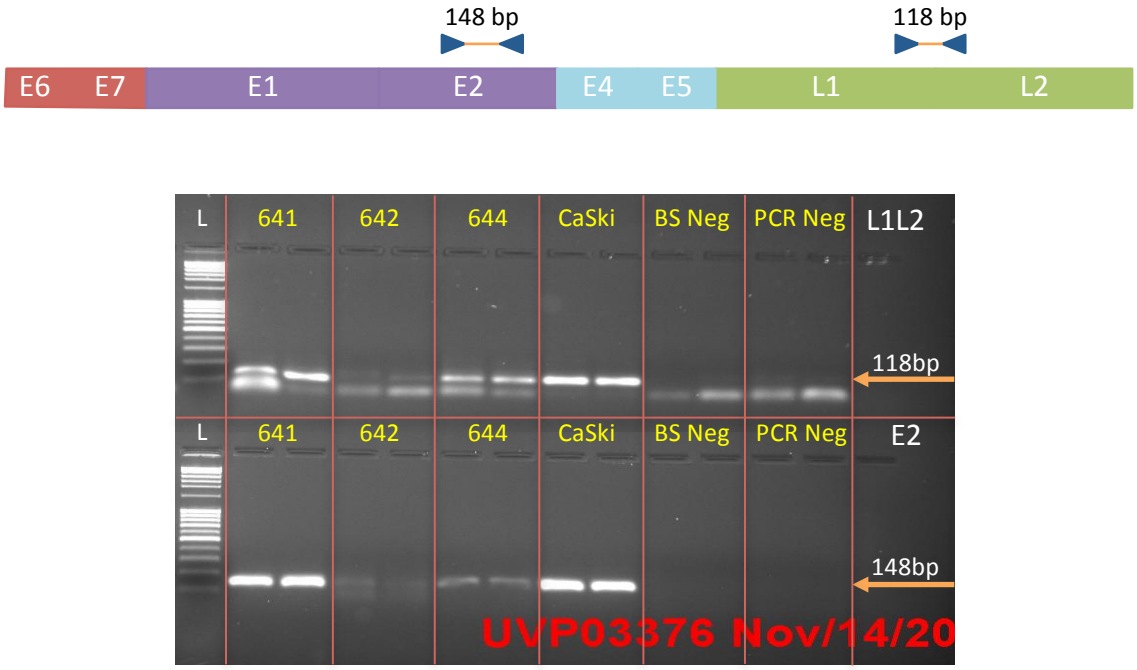

Figure 10.2: Example electrophoresis image showing PCR product after BS conversion.

PCR products are shown in duplicate repeats for three different clinical samples, CaSki (positive control) and two negative controls (BS neg = BS treated water followed by PCR; PCR neg = water only as template in PCR). The top half of gel corresponds to the L1L2 primer set as shown in the diagram by the blue triangles. The bottom half corresponds to the E2 primer set. Samples 641, 642 and 644 all demonstrate the correct sized fragment, however, the bands for sample 642 appear much weaker.

10.3. Initial analyses of data

When the samples that failed the PCR were analysed further, four were excluded from the analysis because they were only HPV16 positive in one HPV16 assay (E6) in the original study and when E6 PCR was repeated with the proteinase K extracted DNA all four were negative. The remaining 17 (34%) samples that failed the methylation PCRs (both E2 and L1L2) also had E6 PCR repeated to confirm the presence of HPV 16 and all were positive; furthermore, HPV16 was identified in at least one other assay for all of these samples. However, 14 (82%) demonstrated disruption of the E2 ORF when they were tested with the six E2 primer sets. Whereas, out of the 25 that passed the methylation PCRs, only 6 (24%) had E2 disruption. A Chi-square test (with Yates Continuity Correction) confirmed that E2 disruption had a significant affect on the outcome of methylation PCR, χ^2 (1) = 11.57, $P = 0.0007$, *phi* = 0.57 (large effect size).

Overall, of the 24 individual samples that underwent pyrosequencing, 14 (58%) gave adequate data for all CpGs, 7 (29%) failed in one or more of the CpG sites in the E2 region, 1 (4%) failed in all CpGs in the L1L2 region and 2 (8%) failed completely in both regions. Samples that failed pyrosequencing correlated with having a weak band following PCR of BS treated DNA. In total there were 21 cases to analyse for the CpGs in the L1L2 region and 17 for the E2 region. The methylation data is shown in full for the L1L2 region in Figure 10.3 and for the E2 region in Figure 10.4.

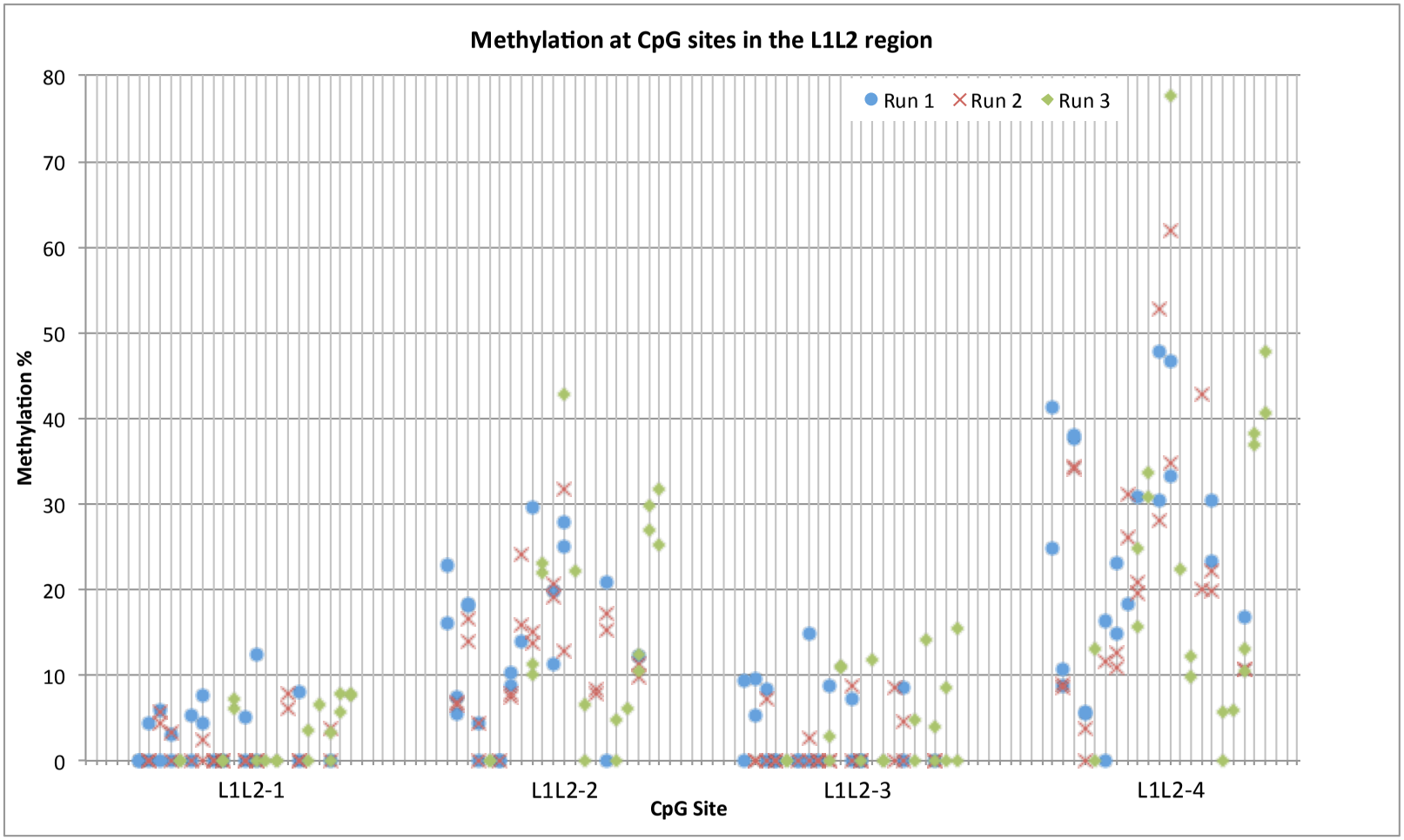

Figure 10.3: Methylation at CpG sites in the L1L2 region. Runs 1 and 2 represent data from the DNA extracted using Proteinase K. Run 3 represents data using DNA extracted using Qiagen MinElute Kit. For each run up to 2 successful results are shown. All results on the same vertical gridline are from the same case.

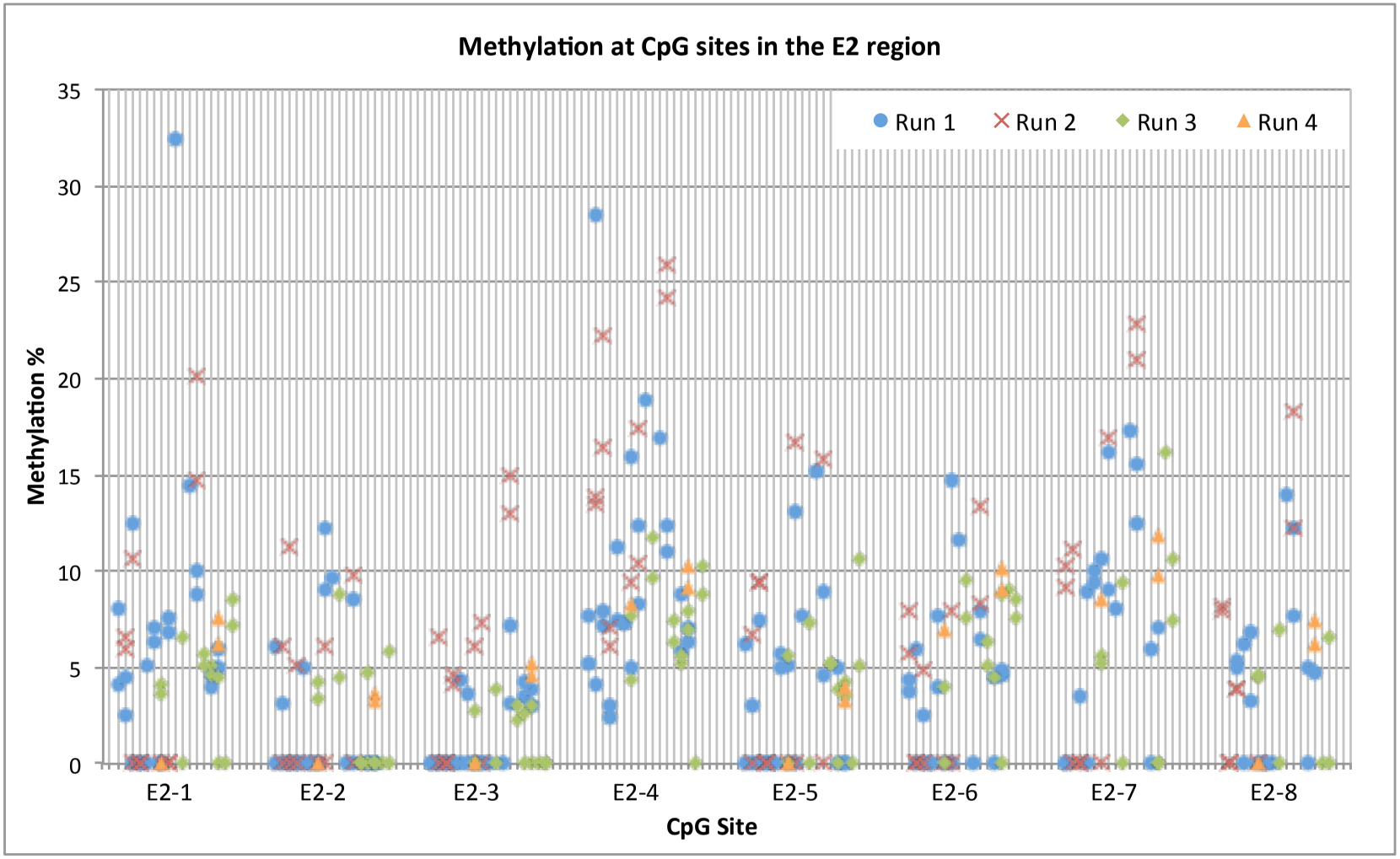

Figure 10.4: Methylation at CpG sites in the E2 region. Runs 1 and 2 represent data from the DNA extracted using Proteinase K. Run 3 represents data using DNA extracted using Qiagen MinElute Kit. Run 4 was only successful in two samples due to a mechanical problem with the pyrosequencer. For each run up to 2 successful results are shown. All results on the same vertical gridline are from the same case.

By comparing the methylation result for each duplicate sample the repeatability for each CpG region was assessed (Figure 10.5 and Figure 10.6).

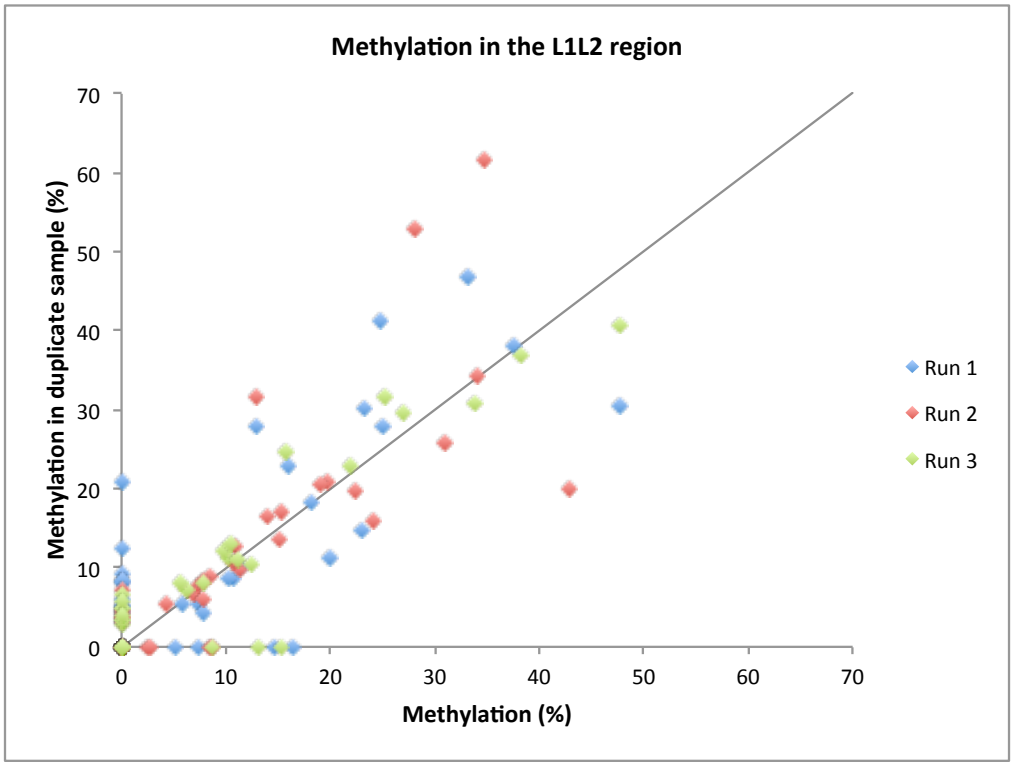

Figure 10.5.: Methylation measured in duplicate at CpG sites in the L1L2 region.
The line represents the line of equality. Runs are described in Figure 10.3.

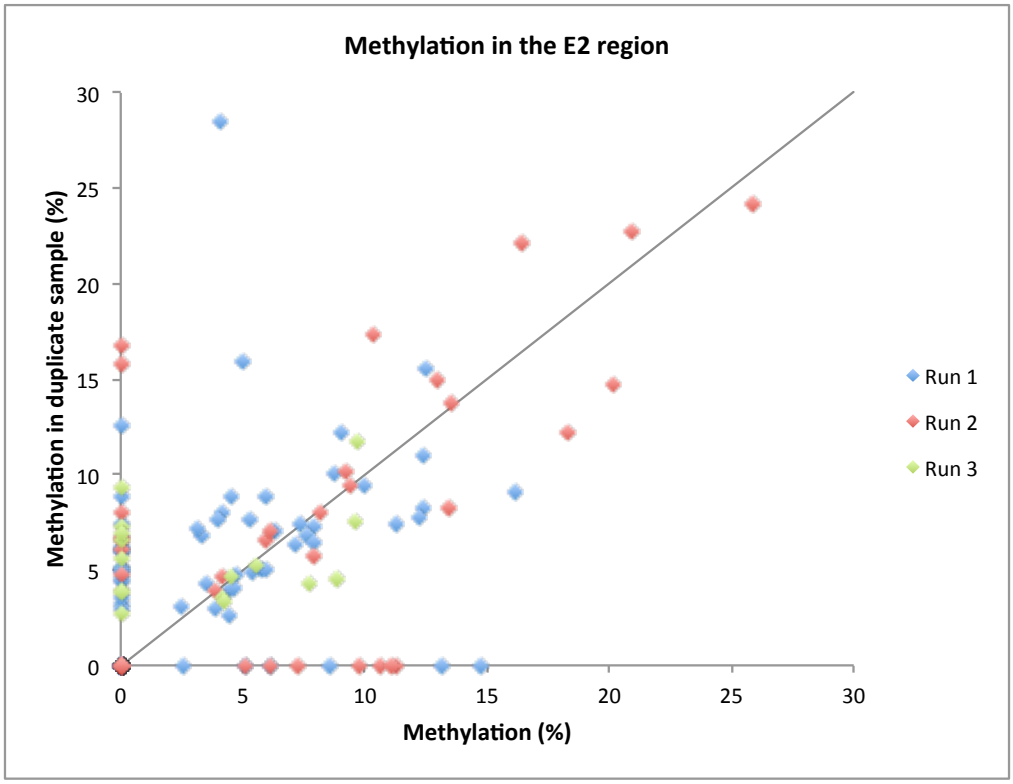

Figure 10.6: Methylation measured in duplicate at CpG sites in the E2 region.
The line represents the line of equality. Runs are described in Figure 10.4.

The reproducibility of the methylation results obtained for clinical material was compared to the reproducibility of results for the control CaSki cell line DNA (Figure 10.7 and Figure 10.8).

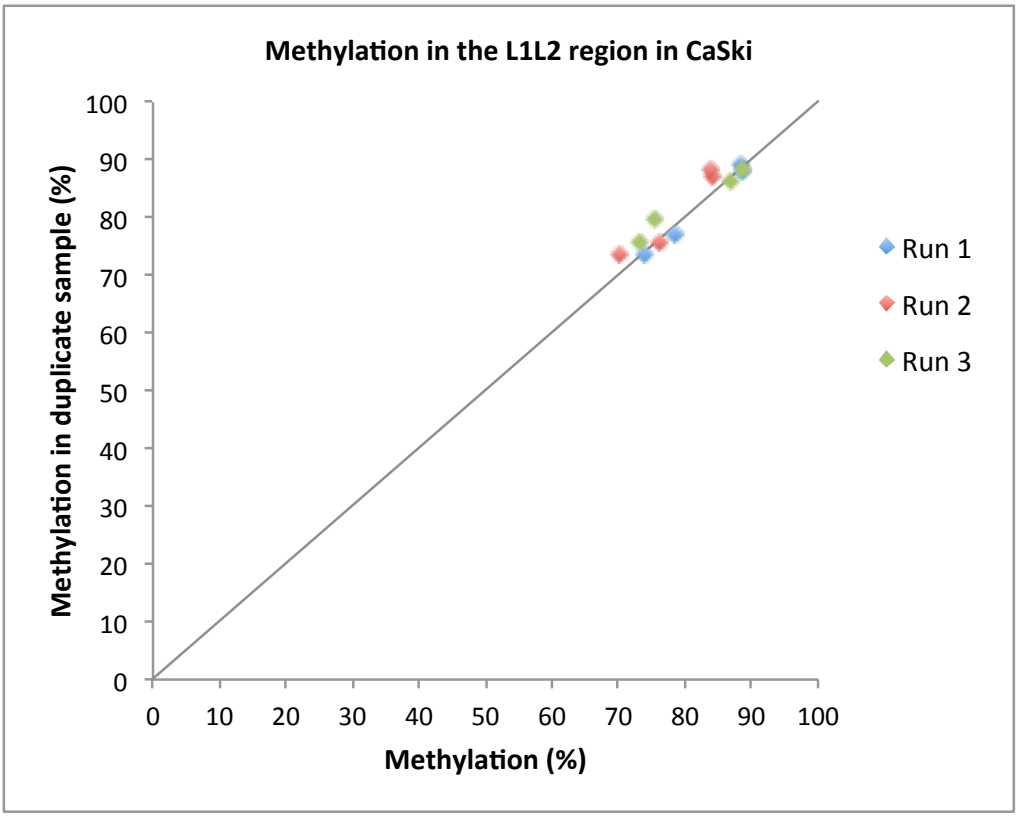

Figure 10.7: Methylation measured in duplicate at CpG sites in the L1L2 region of CaSki DNA.
The line represents the line of equality. Runs are described in Figure 10.3.

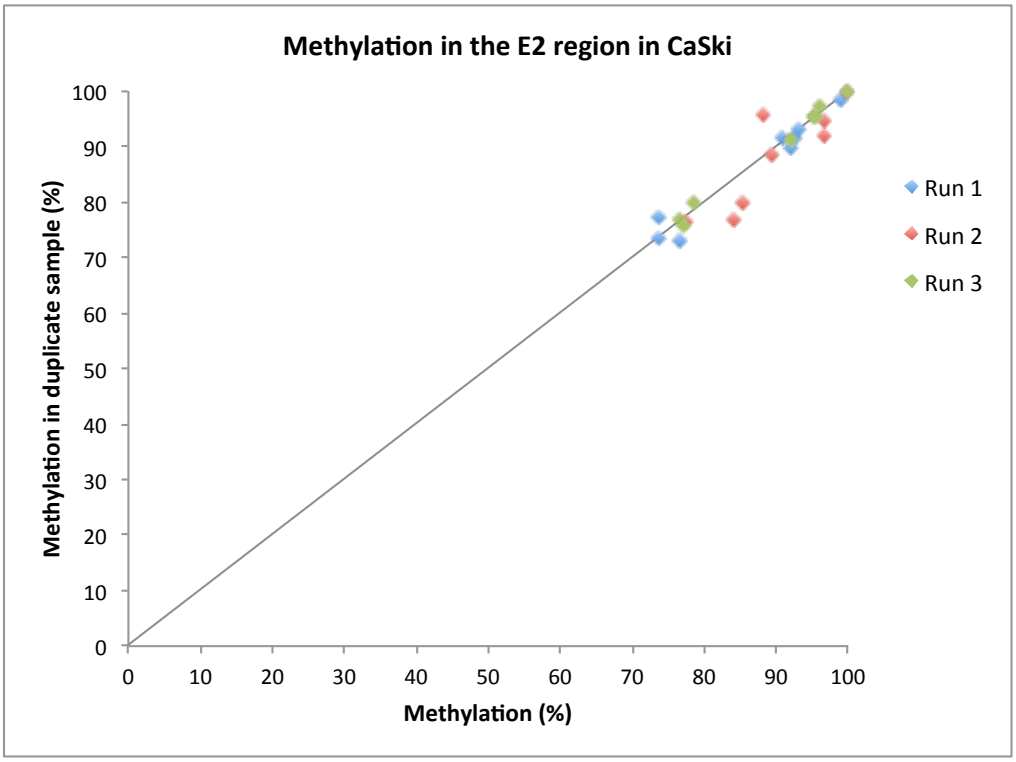

Figure 10.8: Methylation measured in duplicate at CpG sites in the E2 region of CaSki DNA.
The line represents the line of equality. Runs are described in Figure 10.4.

In the L1L2 region paired methylation measurements were available for 13 samples, whereas, in the E2 region there were only eight samples. These results were obtained from different runs performed on different days. In order to assess inter-run reproducibility the intraclass correlation coefficient was calculated: for L1L2 it was 0.82 (95% CI [0.76–0.88]; for E2 it was 0.62 (95% CI [0.43–0.76]). The agreement between runs was moderate to good (Fermanian, 1984). The agreement between runs was also examined using Bland-Altman plots (Figure 10.9 and Figure 10.10).

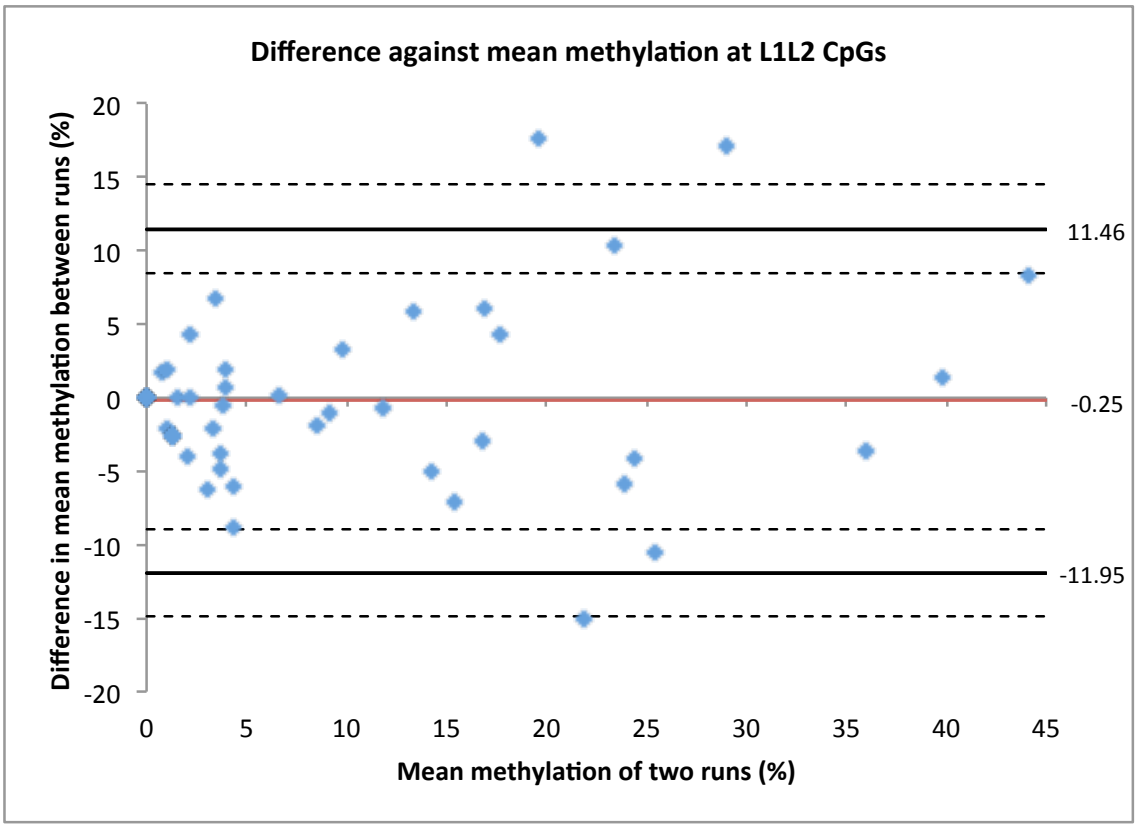

Figure 10.9: Bland-Altman plot comparing mean methylation at L1L2 CpGs from two separate runs.
The mean difference between the two runs is shown by the red line, the limits of agreements by the solid black lines, and the 95% CIs for the limits by the dashed lines.

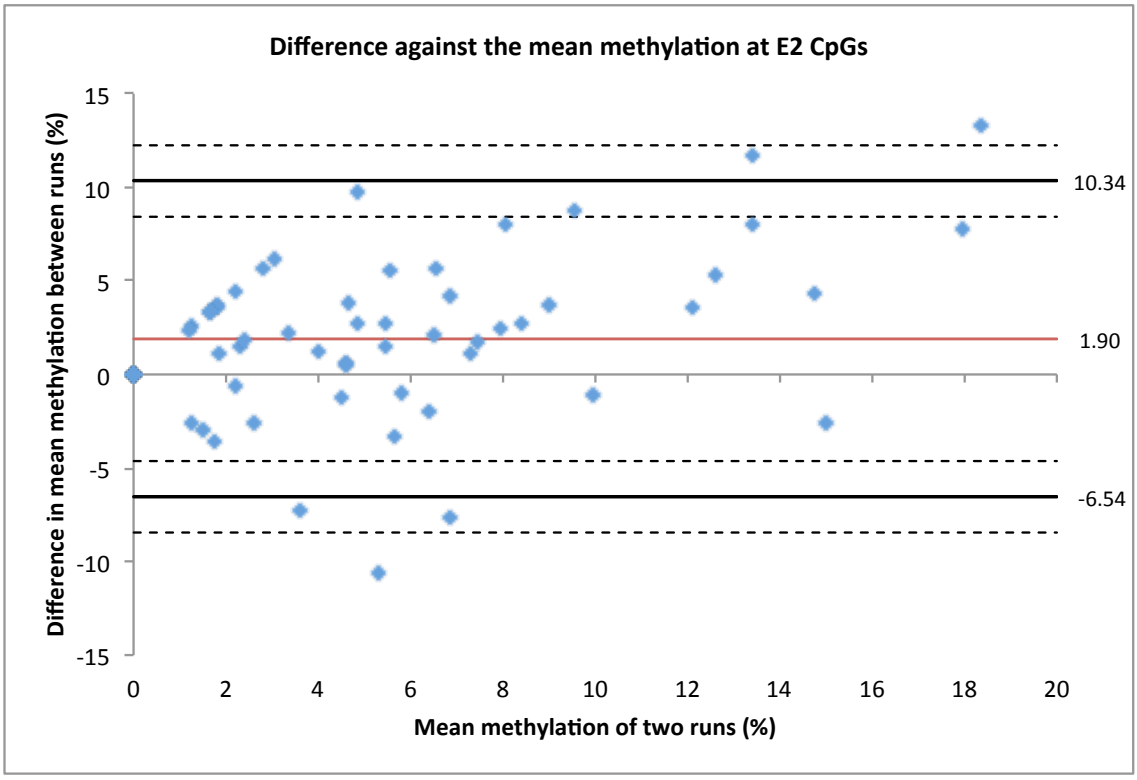

Figure 10.10: Bland-Altman plot comparing mean methylation at E2 CpGs from two separate runs.

The mean difference between the two runs is shown by the red line, the limits of agreements by the solid black lines, and the 95% CIs for the limits by the dashed lines.

The Bland-Altman analysis gave relatively wide limits of agreement; for L1L2, -11.95 to 11.46; and for E2, -6.54 to 10.34. Despite the fact that most of the differences lie within these limits it is not yet clear whether they represent a clinically relevant difference that would affect the use methylation as a screening biomarker.

10.3.1. Viral DNA methylation and histology

The mean methylation at individual CpG sites within each region was compared across disease grades. Within the eight CpGs of the HPV 16 E2 region, higher mean methylation was seen in high-grade disease in E2-1, E2-4, E2-7, and E2-8 (Figure 10.11). However, Mann-Whitney U tests showed no significant differences for the different histology results (Table 10.1).

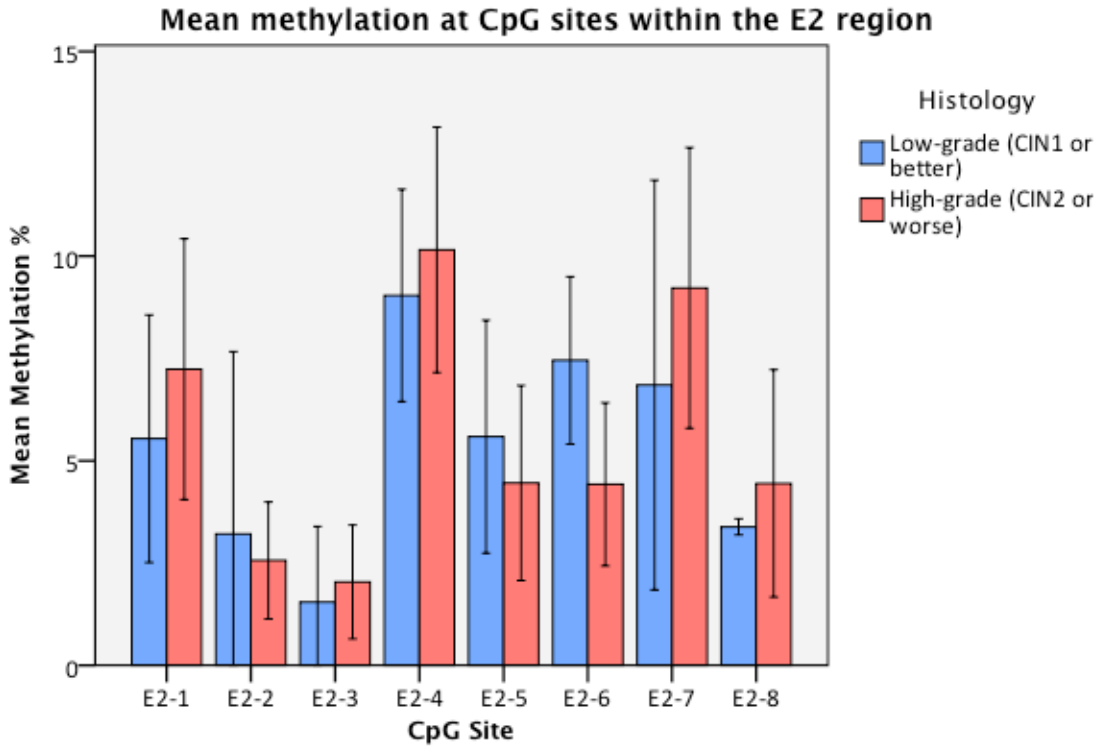

Figure 10.11: Mean methylation at CpG sites within the E2 region of HPV 16 divided by histological grade.

Number of cases for low-grade and high-grade histology for each CpG site respectively: E2-1,*n* = 3, *n* = 12; E2-2–E2-5, *n* = 3, *n* = 14; E2-6, *n* = 3, *n* = 13; E2-7–E2-8, *n* = 2, *n* = 13. Error bars represent 95% CIs.

	HPV 16 CpG Site							
	E2-1	E2-2	E2-3	E2-4	E2-5	E2-6	E2-7	E2-8
Mann-Whitney *U*-value	19	0.94	0.34	0.75	0.75	0.40	0.39	0.90
***P*-value**	1.00	0.77	0.95	1.00	0.36	0.30	0.69	0.93

Table 10.1: Statistical analyses of mean methylation at each E2 CpG site with histological grade.

Non-parametric tests were performed as not all assumptions were met.

In the L1L2 region, the mean methylation for the four CpG sites was also compared between the different histological grades (Figure 10.12). The mean methylation at L1L2-2 and L1L2-4 appeared to be much higher in high-grade (*M* = 13.1 and *M* = 24.4) compared to low-grade histology (*M* = 7.0 and *M* = 13.1, respectively). However, this difference was not significant using Mann-Whitney U tests (Table 10.2).

.

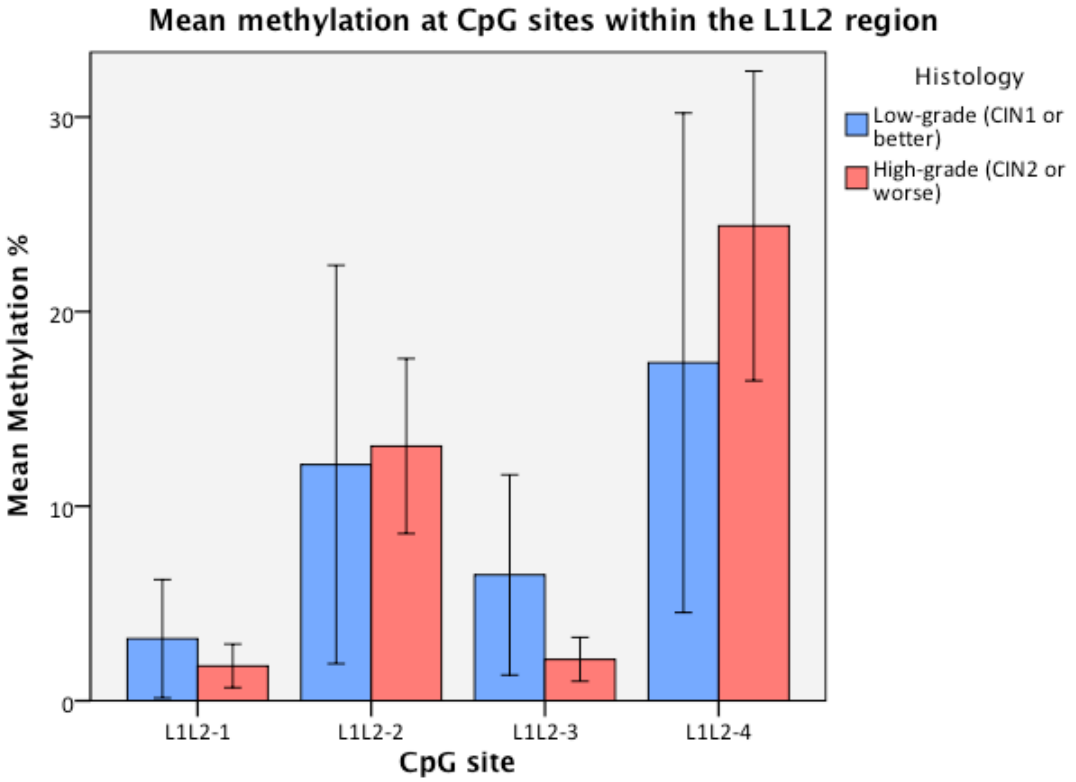

Figure 10.12: Mean methylation at CpG sites within the L1L2 region of HPV 16 divided by histological grade.

n = 7 low-grade and *n* = 14 high-grade histology for each CpG site. Error bars represent 95% CIs.

	HPV 16 CpG Site			
	L1L2-1	L1L2-2	L1L2-3	L1L2-4
Mann-Whitney *U*-value	43	53	30	64
***P*-value**	0.69	0.80	0.17	0.29

Table 10.2: Statistical analyses of mean methylation at each L1L2 CpG site with histological grade.

10.3.2. Viral DNA methylation and referral cytology

The greatest mean methylation was seen in borderline high grade or borderline endocervical cells for each CpG in both of the HPV regions tested (E2, Figure 10.13; L1L2, Figure 10.14).

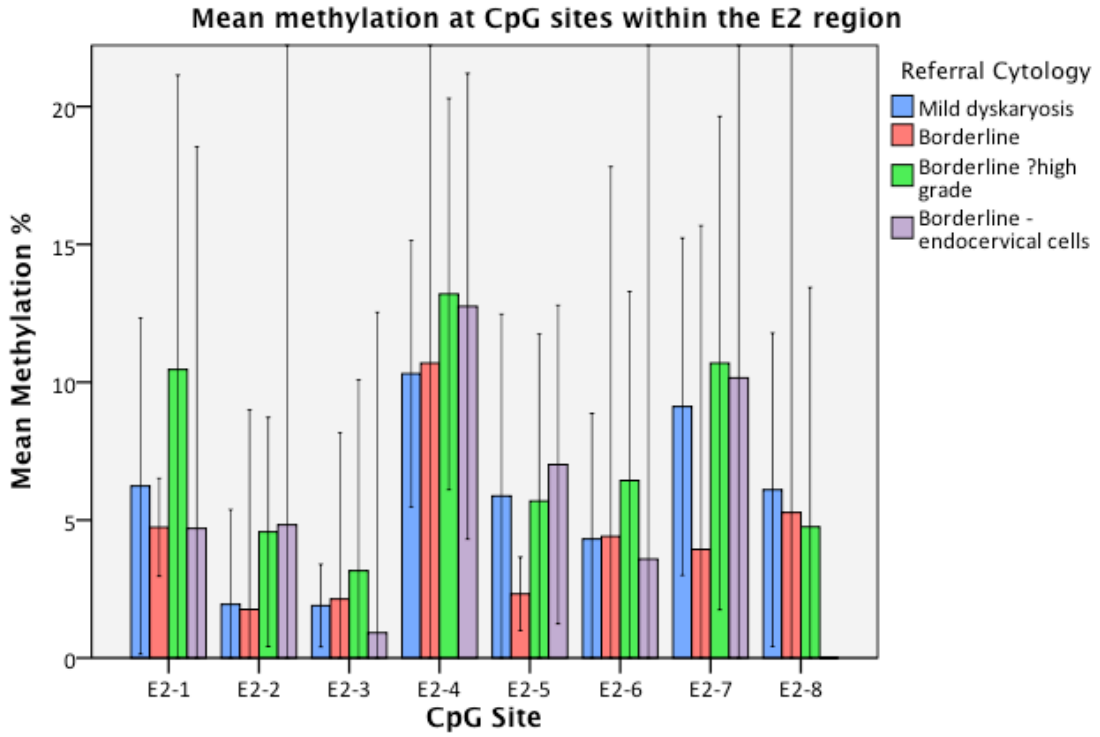

Figure 10.13: Mean methylation at CpG sites within the E2 region divided by referral cytology.

The number of cases analysed for each CpG site in order of referral cytology shown in the legend were: E2-1, *n* = 6, 2, 5, 2; E2-2–E2-5, 8, 2, 5, 2; E2-6, 8, 2, 4, 2; E2-7–E2-8, 7, 2, 4, 2. Error bars represent 95% CIs.

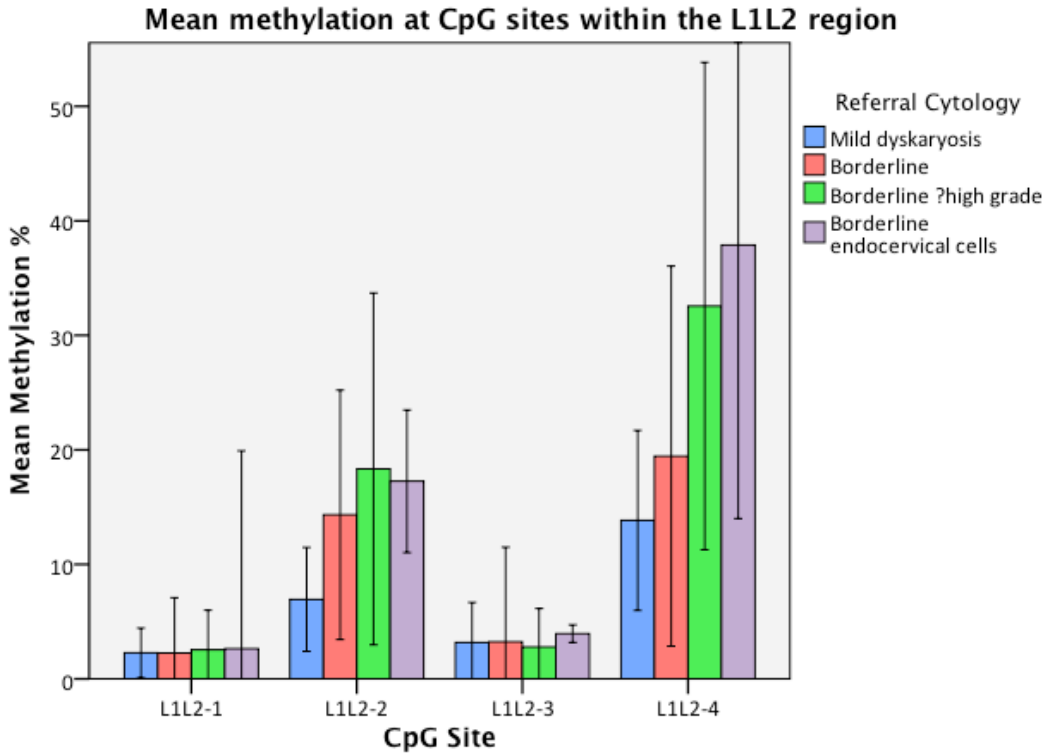

Figure 10.14: Mean methylation at CpG sites within the L1L2 region divided by referral cytology.

The number of cases analysed for each CpG site in order of referral cytology shown in the legend were: *n* = 10, 4, 5, 2. Error bars represent 95% CIs.

The differences in mean methylation were compared statistically using a Kruskall-Wallis test (Table 10.3). There were statistically significant differences in the mean methylation of different cytological grades at the E2-2 and L1L2-4 CpG sites. In order to identify which groups were significantly different follow-up Mann-Whitney U tests were performed between pairs of groups at these two CpG sites. Both at E2-2 and L1L2-4, mean methylation in women with borderline high grade was significantly greater than in women with mild dyskaryosis, $U = 4$, $P = 0.02$; $U = 6$, $P = 0.03$, respectively. However, to control for Type 1 errors the Bonferroni adjustment was applied to the alpha value (four comparisons were made; $P = 0.05/4 = 0.013$). Consequently, the differences at E2-2 and L1L2-4 were not significant at the adjusted alpha level.

	HPV 16 CpG Site											
	E2-1	E2-2	E2-3	E2-4	E2-5	E2-6	E2-7	E2-8	L1L2-1	L1L2-2	L1L2-3	L1L2-4
***H*-value**	2.27	8.91	1.35	3.02	3.81	1.90	2.58	3.84	0.97	5.37	0.68	8.86
df	3	3	3	3	3	3	3	3	3	3	3	3
***P*-value**	0.52	**0.03**	0.72	0.39	0.28	0.59	0.46	0.28	0.81	0.15	0.88	**0.03**

Table 10.3: Statistical analyses of viral methylation at each CpG site for different cytological grades. Kruskall-Wallis test was used because the assumptions were not met for parametric tests.

10.4. Relationship between viral integration and methylation

The integration and methylation data were combined to investigate any associations between the two. The mean methylation was analysed at each CpG site tested within the E2 region (Figure 10.15) and the L1L2 region (Figure 10.16). Overall, integration was associated with hypermethylation. However, a similar pattern that was observed with high-grade histology was seen in the L1L2 CpG sites, with a combination of hypomethylation at L1L2-1 and L1L2-3, and hypermethylation at L1L2-2 and L1L2-4. Mann-Whitney U tests showed that the hypermethylation found in integrated samples was not statistically significant (Table 10.4).

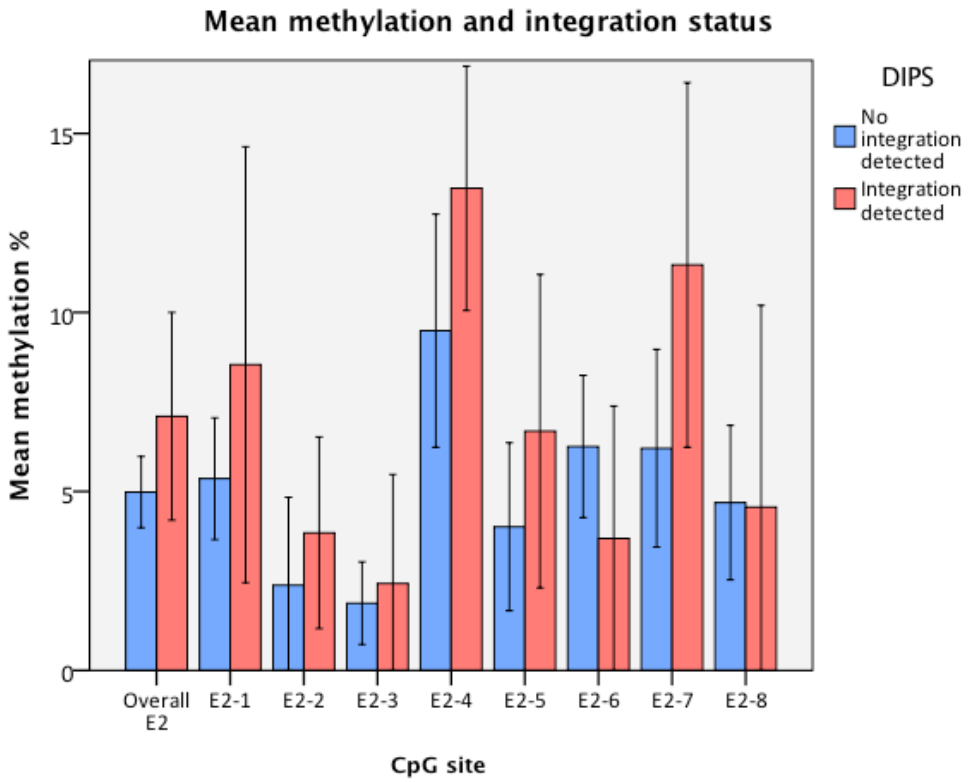

Figure 10.15: Mean methylation at CpG sites within the E2 region divided by integration status.

Number of cases in the no integration detected and integration detected groups for CpG sites; E2-1, 8, 7; E2-2–E2-5, 8, 9; E2-6, 7, 9; E2-7–E2-8, 6, 9, respectively. Error bars represent 95% CIs.

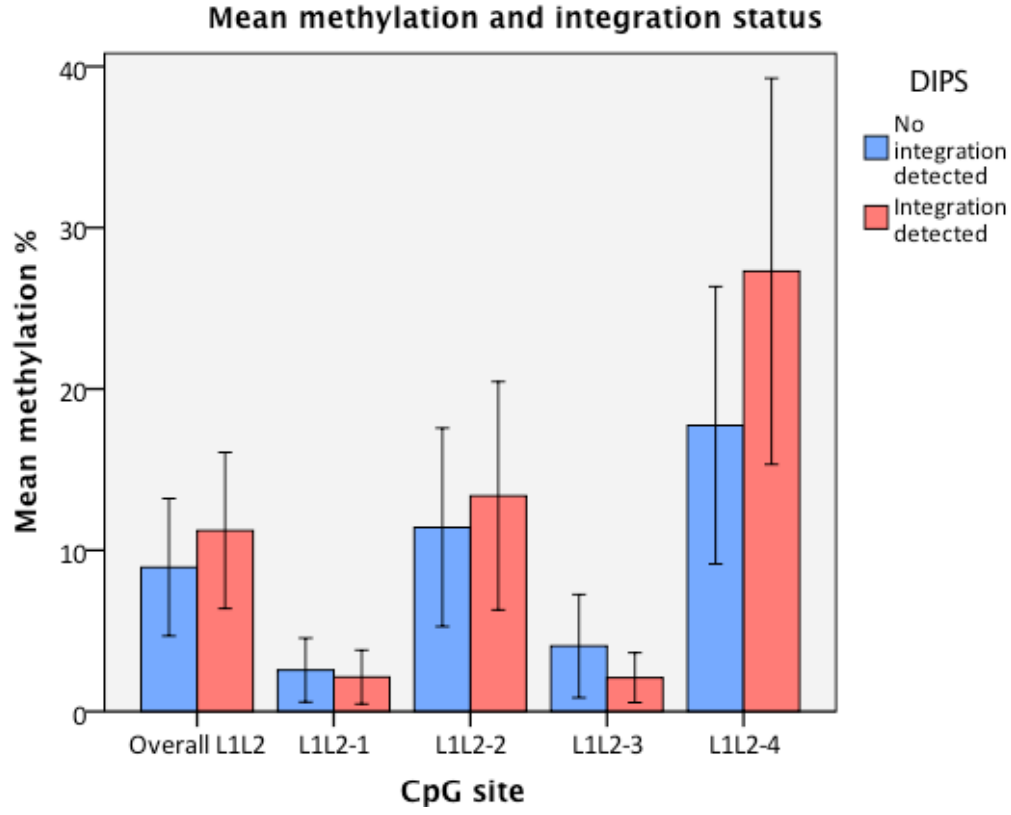

Figure 10.16: Mean methylation at CpG sites within the L1L2 region divided by integration status.

For all CpG sites; no integration detected, *n* = 12; integration detected, *n* = 9. Error bars represent 95% CIs.

	HPV 16 CpG Site											
	E2-1	E2-2	E2-3	E2-4	E2-5	E2-6	E2-7	E2-8	L1L2-1	L1L2-2	L1L2-3	L1L2-4
***U*-value**	30	43	32	48	39	23	38	16	62	56	36	73
***P*-value**	0.87	0.54	0.74	0.28	0.81	0.41	0.22	0.22	0.60	0.92	0.22	0.19

Table 10.4: Statistical analyses of mean methylation at CpG sites within E2 and L1L2 with integration identified by DIPS.

Mann-Whitney U tests were used because the assumptions were not met for parametric tests.

10.4.1. Site of integration and methylation

Analyses of the relationship between the site of integration (see 9.4.1) and methylation were performed. There was some difficulty in comparing groups because there were multiple integration events in most of the women. Moreover, only one woman had an integrant exclusively involving L1. Integration in the E2 ORF appeared to be associated with hypermethylation, in general, but given the small number of cases formal comparisons were not made.

10.4.2. Viral DNA methylation and E2 status

The E2 status (see 9.2.1) was compared to mean methylation at the CpG sites in the E2 region (Figure 10.17) and L1L2 region (Figure 10.18).

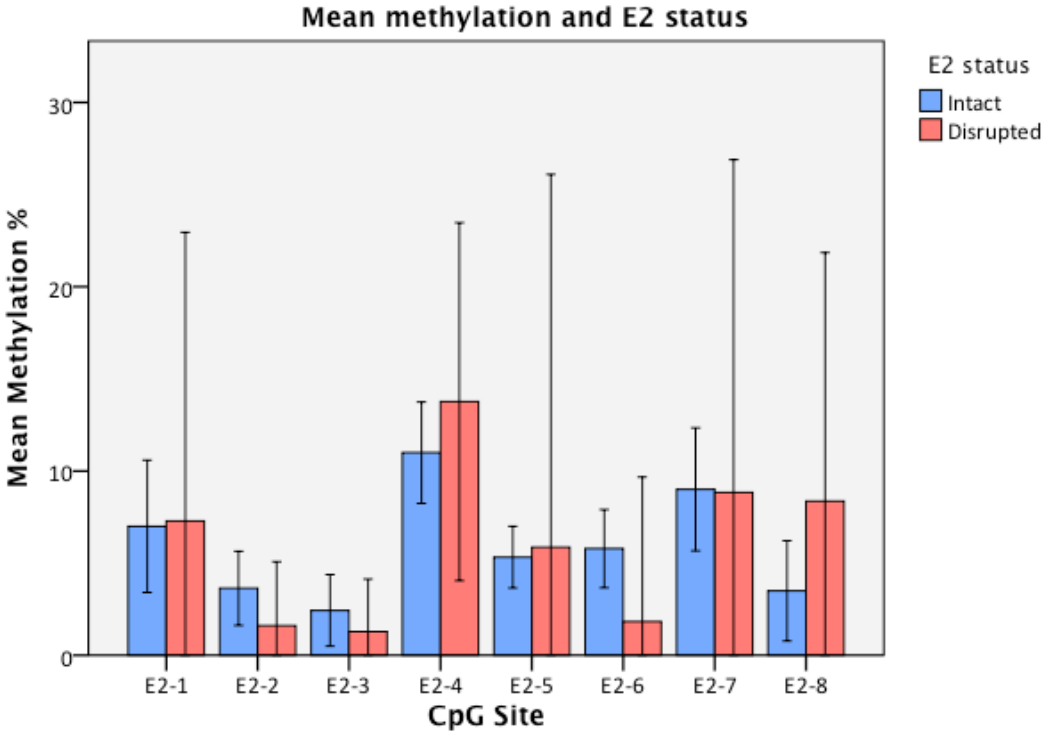

Figure 10.17: Mean methylation at CpG sites within the E2 region divided by E2 status.

Number of cases in the E2 intact and E2 disrupted groups for CpG sites; E2-1, 12, 3; E2-2–E2-5, 13, 4; E2-6, 12, 4; E2-7–E2-8, 11, 4, respectively. Error bars represent 95% CIs.

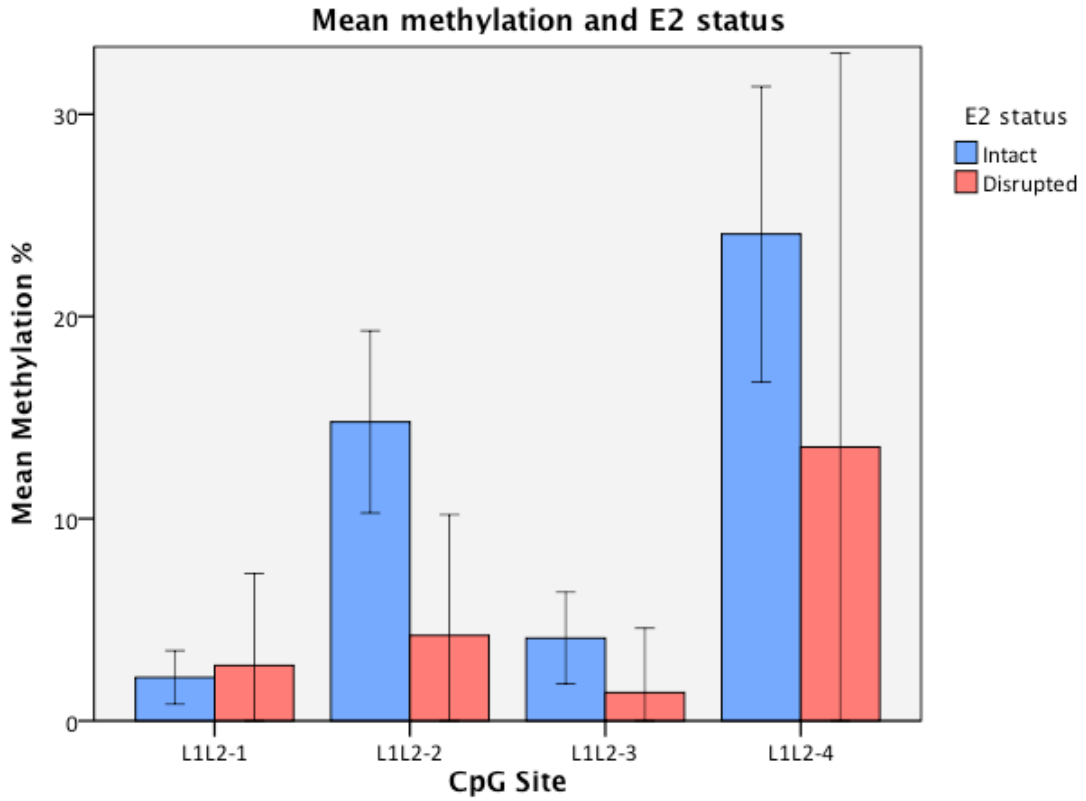

Figure 10.18: Mean methylation at CpG sites within the L1L2 region divided by E2 status.
For all CpG sites; E2 intact, *n* = 17; E2 disrupted, *n* = 4. Error bars represent 95% CIs.

E2 disruption appeared to be associated with a mixture of hypermethylation and hypomethylation within the CpG sites in the E2 region. Interestingly, the pattern that had been seen for histology and integration in the L1L2 region was not present when analysing the E2 status within that region. Instead, there appeared to be hypomethylation at all CpG sites. Mann-Whitney U tests were used and identified statistically significant hypomethylation at E2-6 and L1L2-2 (Table 10.5).

	HPV 16 CpG Site											
	E2-1	E2-2	E2-3	E2-4	E2-5	E2-6	E2-7	E2-8	L1L2-1	L1L2-2	L1L2-3	L1L2-4
***U*-value**	19	34	29	21	36	43	31	9	27	59	46	50
***P*-value**	1.00	0.41	0.79	0.62	0.30	**0.02**	0.28	0.10	0.57	**0.02**	0.32	0.17

Table 10.5: Statistical analyses of mean methylation at CpG sites within E2 and L1L2 with E2 status.
Mann-Whitney U tests were used because the assumptions were not met for parametric tests. Corrected Bonferroni *P*-value = 0.025.

10.5. Discussion

HPV DNA methylation has recently emerged as a potential biomarker for use in cervical screening (Clarke et al., 2012). However, studies have been mainly exploratory to date with the optimal method, most discriminating CpG sites and accurate algorithms for analyses yet to be established. Viral methylation in 24 samples was studied using bisulfite treatment followed by pyrosequencing. The nature of this study was mainly exploratory and, importantly, was designed to utilise clinical material (LBC) that had been collected as part of the national cervical screening programme in South Wales.

10.5.1. Reproducibility

Reproducibility is an essential quality of any assay to be used on clinical material. Before clinical material was investigated in this study cell line DNA was used to confirm that the assay was working. Both with intra-run and inter-run results excellent reproducibility was demonstrated when investigating methylation in CaSki DNA. The reproducibility of the assay on clinical material, however, was less convincing. Similar, high degrees of variation have been seen in other studies (Brandsma et al., 2009, Clarke et al., 2012). However, most studies in the literature do not give their data on intra- or inter-run reproducibility. Furthermore, in most cases error bars and confidence intervals are not given.

One of the most noticeable observations in this study was the number of CpGs that recorded 0% (no methylation) in one sample and up to 20% methylation in the duplicate sample within the same run. Although most results fell within the 95% limits of agreement, those limits were relatively wide and a 10% change in percentage methylation may become significant if potential cutoffs were to be generated in the context of a screening test. There are a number of possibilities that may explain this lack of reproducibility. Firstly, there may have been human error at some point in the process. PCR-based assays are known to be very sensitive to small variations in reagents, and contaminants or inhibitory salts can also affect the reaction. Standard operating procedures were followed and should have reduced this risk and, furthermore, control material was used in every part of the process for each run. Secondly, the differences may be due to the quality of the DNA in the samples. Cell line DNA worked very well and maintained excellent reproducibility (even down to 1/1000th dilution). The clinical samples in this study had been collected by exfoliative cervical sampling, stored in transport media for a few weeks (at room temperature), were then processed in the laboratory and finally had their DNA extracted. It is known that the transport media, in which the samples from this study were stored, is associated with RNA and DNA degradation over time (Powell et

al., 2006). This may also explain why a number of samples failed the PCR following the BS treatment. The CaSki samples went through all the same steps in the assay but were not exposed to long periods in transport media.

The third possible explanation is that the variability is due to quantity of HPV DNA present in each specimen. The exfoliated epithelial cells that are collected during the cervical sampling process represent varied cell types. The amount of cells and types of cells would certainly be affected by the technique of the operator. Detection of methylated CpG sites may, therefore, be diluted by the presence of multifocal and heterogeneous populations (Clarke et al., 2012). Although, this fact should have been minimised by using standard concentrations and ensuring that the sample was homogenised, as far as possible, before using in a PCR. Low DNA concentrations may also significantly affect the BS treatment and PCR reaction. This study found that pyrosequencing was more likely to fail when the PCR product showed a weak band on gel electrophoresis (Figure 10.2). Finally, it is conceivable that methylation is not constant in all cells and that, in the same way, as integrants can co-exist with episomes methylation may vary across cells. It maybe the case that when hypermethylation is found consistently at a CpG site it is more significant than when there is variation on repeat testing, implying a selection advantage and progression towards cancer.

The large variety of sampling techniques, population studied, methylation assays and the multitude of CpG sites available in HPV for analysis can make comparisons between studies problematic. Initial comparisons were made to recent work completed by Dr Dean Bryant, who developed the assays used in this study. He found similar degrees of variation between samples, especially in LBC material, and the methylation values found in women with severe dyskaryosis were similar to those found in women with CIN2+ in this study (Figure 10.19)(Bryant, 2012).

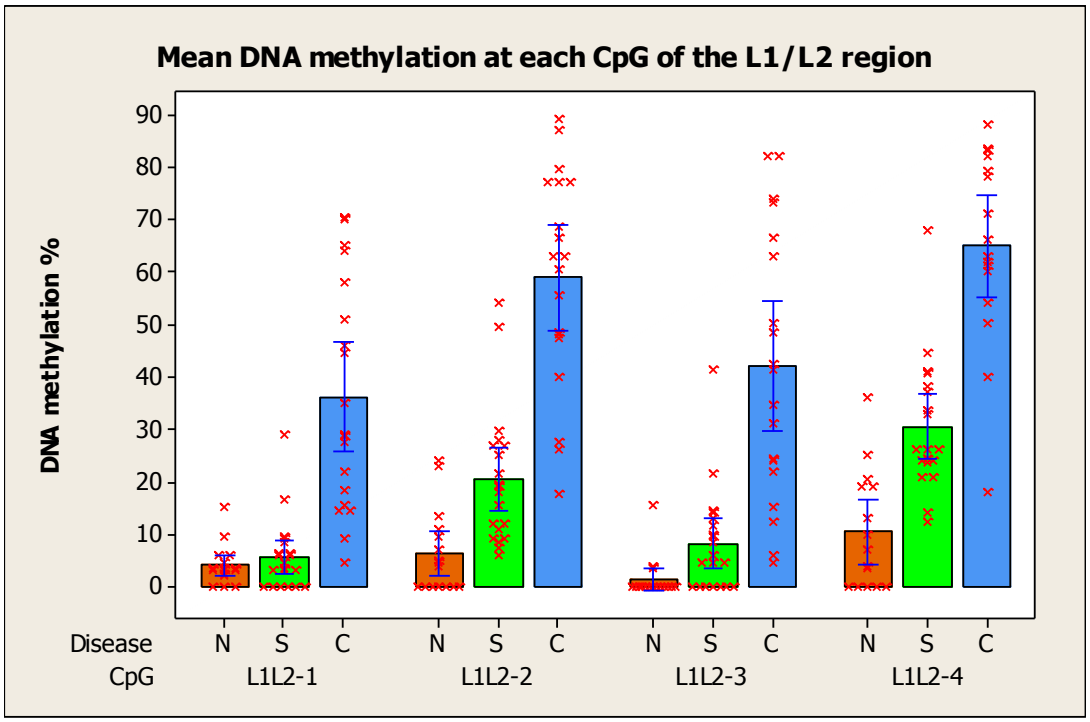

Figure 10.19: Variation in CpG methylation with disease grade for the L1L2 region.

Used with permission from (Bryant, 2012). DNA was extracted from HPV16 LBC samples with normal (N) cytology (*n* = 20) or severe (S) dyskaryosis (*n* = 20), and from fixed blocks of tissue from cervical cancers (C)(*n* = 27). Crosses indicate individual sample values. Bars and error bars represent the mean with 95% CIs.

10.5.2. Methylation and disease grade

It has been reported previously that low or no methylation is associated with normal or low-grade CIN and hypermethylation is associated with high-grade CIN (Brandsma et al., 2009). In this study hypermethylation was, in general, associated with high-grade CIN. None of the differences in mean methylation between histological grades, at each CpG site, reached statistical significance, although due to the relatively small sample size, this finding was not surprising. Furthermore, it was also apparent that hypomethylation at certain individual CpGs may also be a significant marker of high-grade disease. In particular, this study demonstrated an alternating pattern of hypomethylation followed by hypermethylation in the mean methylation at four CpGs in the L1L2 region. A similar "N" shaped pattern was seen in CaSki and also in other studies involving the same CpGs (Bryant, 2012, Sun et al., 2011).

Other studies have also found increased methylation in L1 and/or L2 CpGs in high-grade disease (Fernandez et al., 2009, Kalantari et al., 2010, Mirabello et al., 2013, Sun et al., 2011). Sun *et al* used cervical lavages samples and used a similar process of BS treatment followed by pyrosequencing. That study found very similar levels of methylation in CIN3 to this study; whereas, in low-grade disease groups and CIN2 they found less methylation, which

consequently showed statistical significance. In the same study, however, error bars were not shown and confidence intervals were not given.

In order for methylation to be effective as a screening test it not only has to be reproducible, but also it needs to reliably predict women with high-grade disease. From the data shown in this thesis, along with the published data, it appears most unlikely that there is any one CpG that could be used as a biomarker. One of the ways of overcoming some of the variability in methylation results for the various CpGs in HPV is to devise a model or algorithm where there are a number of features that correlate with high-grade disease. One potential algorithm has been developed and was recently validated on a very similar clinical cohort of women from Cardiff and Newport with low-grade cytology enrolled in the CRISP study (Mirabello et al., 2013, Lorincz et al., 2013). The only significant difference in this cohort was that the women were enrolled following their first smear showing low-grade abnormalities, rather than with persistent low-grade smears. During the development process CpG sites within the L1 region were found to be the most informative in predicting high-grade CIN (Mirabello et al., 2013). The devised classifier score involved the proportion of methylated L2 CpG sites combined with the mean methylation at two CpG sites within L1 (not tested in this study). The area under the curve for the classifier in the CRISP study cohort was 0.74, although it had broad CIs due to a relatively small sample size (n = 73). It is possible that an algorithm like this offers the most potential for use in screening programmes in the future. Currently methods have only been described for assessing methylation in HPV 16, so assays would need to be developed and evaluated for additional HR HPV types, if this was to be used as a screening test.

10.5.3. Methylation and cytology

The majority of CpG sites showed hypermethylation in borderline high-grade and borderline endocervical cells compared to mild dyskaryosis or borderline cytology. Two of the CpG sites were not quite significant at the Bonferroni adjusted P-value, E2-2 and L1L2-4. This finding fits with clinical practice where a borderline high-grade cytology result is considered higher risk than mild dyskaryosis or any of the other borderline categories.

10.5.4. Methylation and integration

One of the important findings in the context of integration was the finding that samples with HPV 16 E2 disruption were far less likely to successfully produce an appropriate PCR product following BS treatment (P = 0.0007). Given the frequency of E2 disruption identified by this

study and others (Collins et al., 2009), this may pose problems in the use of any methylation as a screening biomarker. Interestingly, the PCR for the L1L2 region also failed in the samples with E2 disruption. This may be because of further disruption in the ORF of the late genes or because of a reduced viral load in the samples caused by reduced viral replication. The results from this study also found that E2 disruption was associated mostly with hypomethylation, in particular at the E2-6 and L1L2-2 CpG sites. Viral oncogenes target DNA methyltransferases (Burgers et al., 2007), which may explain alterations in methylation prior to integrant selection. Studies have also shown that hypermethylation may control replication control and lead to oncogenesis (Vinokurova and von Knebel Doeberitz, 2011). In this study and others (Kalantari et al., 2008a, Kalantari et al., 2008b) samples with integrants were shown to be more associated with hypermethylation. However, it must be noted that pyrosequencing is unable to differentiate between samples with just hypermethylated genomes and those with a mixture of hypomethylated and hypermethylated genomes.

10.5.5. Strengths and weaknesses

The main strength of this study has been the broadly successful use of a high-throughput quantitative assay to assess the methylation of viral DNA in LBC samples collected within a national screening programme. Furthermore, very few studies have investigated the reproducibility of the assay using LBC material. Another strength is that both integration and methylation studies were performed on the same sample set.

However, the sample set may well have caused some of the problems in the mostly non-significant findings of the data. The study population consisted of women with persistent cytological abnormality, albeit low-grade. Hence, when comparing the methylation for variations of low-grade cytology there was little difference. Furthermore, persistent infection may be associated with hypermethylation, thus, diluting the difference in methylation associated with a transforming infection and high-grade disease.

One of the weaknesses with the assay employed in this study is that it is currently type-specific. If the assay were to be used to identify screening biomarkers it would need to be expanded to include other HPV types. One of the alternatives is to measure DNA methylation in the host cells and research in this area is on-going. Although the study would have been stronger with a larger sample size, the exploratory data generated is essential in order to explore potential pitfalls, generate hypotheses and plan for future studies.

10.6. Conclusions

Measuring the methylation of viral DNA using pyrosequencing offers the benefits of a quantitative, high throughput assay. Hypermethylation appears to be associated with high-grade disease; however, the degree of hypermethylation can vary considerably at different CpG sites within the HPV genome. Some potential problems with the reproducibility of the assay have been highlighted regarding its use in LBC samples and in samples where viral disruption may be present. When these issues are addressed viral DNA methylation will be a strong candidate as a screening biomarker.

Chapter 11 – GENERAL DISCUSSION AND CONCLUSIONS

11.1 Introduction

In this chapter the original aims and hypotheses are reviewed together with the key findings of thesis. The implications are considered in the context of cervical screening and clinical relevance. Areas for future research are also highlighted and conclusions are given.

11.2 General discussion

The discovery of HPV as a causative agent of cervical cancer and its precursor CIN has enabled significant progress in their prevention over the last few decades. Primary prevention with the HPV vaccine has been rolled out in the majority of developed countries, however, it will be more than a decade before there is any impact on the prevention of cervical cancer. Moreover, there are still at least two generations of women that have not received the vaccine and require secondary prevention in the form of cervical screening. Furthermore, in vaccinated women, the vaccine is not expected to prevent all cervical cancer and CIN. Cervical screening has been very successful at reducing rates of cervical cancer and, importantly, at identifying cancers at an earlier stage (Peto et al., 2004b). Within current cervical screening programmes, there are certain scenarios where the management is not straightforward and there is a danger of over treating women. Women with persistent low-grade cytological abnormalities are one such group. Transient HPV infection is common and is associated with low-grade abnormalities in the cervix. In this group a better screening test or additional triage test is required to improve the discrimination of high-grade cervical disease.

Many studies have been performed trying to improve on cytological screening with HPV testing or other biomarkers. The aim of this thesis was to investigate some novel potential biomarkers and compare and contrast with HPV testing.

11.3 Hypothesis 1: BD SurePath Plus™ can predict the presence of high-grade cervical disease in women with persistent low-grade cytological abnormalities.

Using the novel biomarker BD SurePath Plus™, within a prospective observational study, tested this hypothesis. The results showed that although the test could identify the majority of women with high-grade cervical disease, it was also positive in a considerable number of women with low-grade or no cervical disease. Some of the limitations of the test and the

reliance on colposcopy and a punch biopsy in some cases as the only indicator of high-grade disease have been discussed already (see section 8.9.1). Moreover, the final conclusion on the use of BD SurePath Plus™ to predict high-grade cervical disease cannot be made until the full study is complete with the two-year follow-up data. Data from two studies in the USA using BD SurePath Plus™ showed promise when presented in abstract form at national conferences (Whitehead et al., 2012, Whitehead et al., 2011) and the publication of the full reports will make a very useful comparison with the data presented here. Based on the evidence in this thesis, BD SurePath Plus™ was better at predicting the absence, by a negative result, of high-grade disease in women with persistent low-grade cytological abnormalities, than it was at predicting the presence of disease. The PPV for HG disease was 18%, with CI of 14-23%. The hypothesis was therefore rejected.

11.4 Hypothesis 2: HPV testing can predict the presence of high-grade cervical disease in women with persistent low-grade cytological abnormalities.

To test this hypothesis two commercially available HPV tests were used within the same prospective, observational study. The results showed that a positive PapilloCheck® or HC2 test was associated with high-grade cervical disease (P = 0.014 and P = 0.004). However, there were a number of false positive tests, thus weakening their use in the context of women with persistent low-grade cytological abnormalities. The effect of the age of the cohort on HPV testing and the associated higher prevalence of multiple type hrHPV infections were found to be significant factors when determining the tests' ability to predict high-grade disease. hrHPV testing, particularly in the ≥30 years age group was shown to be a predictor of CIN2+ and the hypothesis was, therefore, accepted.

11.5 Hypothesis 3: BD SurePath Plus™ will predict with higher positive predictive value, but lower negative predictive value than HPV testing.

BD SurePath Plus™ was the weakest performer in all parameters when compared with the two HPV tests in this study. The high positivity rate severely restricted its ability to maintain a high positive predictive value. It is likely that some more cases of high-grade disease may be identified at subsequent colposcopy in women that had a positive BD SurePath Plus™ result but had a low-grade / normal biopsy or no biopsy. However, it is highly improbable that these numbers will be significant enough to produce PPV and NPV values that would outperform those of HPV testing.

11.6 Hypothesis 4: E2 disruption is a marker of a transforming HPV infection and, therefore, increases the risk of having high-grade cervical disease.

To test this hypothesis E2 PCRs were performed on all HPV16 positive cases in the SuPerLy–HIM study. E2 was disrupted in the majority of women and no association was shown with grade of disease, hence the hypothesis was rejected. Moreover, it was postulated that E2 disruption might be a marker of persistent infection rather than a transforming one. E2 disruption can result in deregulation of the HPV oncogenes E6 and E7 (Doorbar et al., 2012), however, it is not considered the only mechanism (Pett and Coleman, 2007). The evidence in this thesis does not support the use of E2 PCR as a screening tool, rather it appeared to ask more questions than it answered. Without further testing it would not be possible to ascertain from a negative result whether any one or all of the 6 PCRs had simply failed; furthermore, if the PCR showed intact E2 there could still have been some disrupted viral genes in the background of episomes.

11.7 Hypothesis 5: Viral integration is associated with a transforming HPV infection and, therefore, an increased risk of having high-grade cervical disease.

In order to test this hypothesis DIPS analyses were performed on a selection of samples with different grades of cervical disease. It must be noted, however, that the selection was biased toward samples with CIN3, in order to maximise positive results. Indeed, integration events were only discovered in women with CIN3 ($n = 9$, $P = 0.02$), and therefore the findings would strongly support the hypothesis. The assay was labour intensive and not really conducive to a screening test. An integration assay that is rapid and amenable for high throughput would be more appropriate. In addition to finding a number of known sites of integration, several new integration sites were discovered in this thesis. However, in order to be even more confident in this finding a PCR using primers flanking the viral-cellular junction would be required. The finding of multiple integration events per sample fits with the notion that there may be many integrants in a sample but only one offers a growth advantage and consequently is the integrant that gets transcribed (Pett and Coleman, 2007).

11.8 Hypothesis 6: Hypermethylation within the viral genome correlates with high-grade cervical disease.

Methylation of viral DNA was studied in a selection of women and hypermethylation was associated with high-grade disease, however, it was not statistically significant. The method of

measuring the viral DNA methylation using bisulfite treatment followed by pyrosequencing was relatively straightforward. However, successfully getting samples through both steps of the process proved to be problematic. Further analysis revealed that the method may not be suitable for some samples with disruption of the viral genome. Samples with disruption of E2 were significantly associated with failure of the assay in both the L1L2 and E2 regions (P = 0.0007). The advantages of this assay include its quantitative nature and its capacity for high throughput. However, with such variety of viral DNA methylation shown at different CpGs in this work and as reported by others, it is likely that specific algorithms made up of different combinations of hypomethylation and hypermethylation at various CpG sites will be most useful. One such algorithm has already shown potential in predicting high-grade CIN in women with low-grade cytological abnormalities (Lorincz et al., 2013). However, in order to be more applicable in the screening environment the assay would have to cover the majority of HPV types. Another way around that issue is to test the methylation of human DNA. Promising human CpG sites have already been linked with the development of high-grade cervical disease (Hesselink et al., 2011, Overmeer et al., 2011) and work is being carried out in the HPV laboratory to develop novel assays for this purpose.

11.9 HPV testing

Cervical screening using HPV testing has only recently started to be included in certain settings. In England, HPV testing has already been introduced as a triage test of cervical samples with low-grade cytological abnormalities. The combination of two commercially available HPV tests with further HPV type-specific testing has not been performed in a persistent low-grade cytological setting previously. Disagreement between HPV tests is a concern and warrants further consideration, especially since a negative HPV test in most screening algorithms mean that the woman is returned to normal recall (3 years until the next screening test). The data presented here appeared to show that either the commercial assays were giving false negatives or the type-specific PCRs were giving false positives. It is important to remember that the commercial assays have been developed with clinical application in mind so that a loss in sensitivity is exchanged for an improvement in specificity. They ideally would only identify HPV infections that were associated with disease as opposed to those associated with a transient infection. The clinical relevance is the most important aspect of any screening test and although more cases may be identified via type-specific PCR they may well be less clinically relevant.

This thesis does highlight issues regarding HPV testing that relies on intact DNA in certain HPV regions. The PapilloCheck assay targets an undisclosed region in the E1 ORF. In order to ascertain the E1 genomic status in the samples an E1 assay was designed in our HPV laboratory consisting of two primers sets. Both the E1 PCRs and integration data in this thesis, along with other published data regarding integration (Wentzensen et al., 2004) question whether E1 is a suitable target for a HPV test. Our evidence is far from conclusive but a study to look at E1 and PapilloCheck in more detail is needed.

11.10 Development of biomarkers and reproducibility of assays

A recommended framework for developing biomarkers has previously been proposed (Arbyn et al., 2009b). This thesis presents findings both in the form of original data and in the discussion of other published studies that would support the framework. Novel HPV tests and other biomarkers are being continually developed. Many of these tests have been used in prospective screening studies before full and proper clinical validation. Reproducibility of an assay is a fundamental quality of a successful screening test, however, it is rarely included in published studies. This thesis found a potential issue within the reproducibility of the pyrosequencing methylation assay. The evidence suggested that the variation might have been due to degradation of DNA caused by the use of storage media. A well-structured study using multiple cervical samples from same group of women exposed to a variety of external factors is warranted.

11.11 Future work

Aside from the studies already proposed more research into the interplay between viral genomic disruption and HPV testing is needed. A sample with integration or disruption within a HPV typing assay's target region would most likely result in a false negative result. Although this occurrence may be relatively rare, it could have serious consequences if it resulted in a missed opportunity to diagnose a cancer still in its earliest stages. Now that HPV testing is becoming an established part of the screening programme it is entirely appropriate that these potential issues are addressed.

11.12 Conclusions

This study successfully assessed a novel biomarker in the important clinical context of persistent low-grade cytological abnormalities. Although there had been reasonable evidence to suggest that the biomarker would be a good discriminator of high-grade disease, following the clinical study it was found to be inferior to HPV testing. The study did demonstrate that an immunocytochemically-based test was feasible and the same format may be optimised with alternative more specific markers in the future.

The age of patients with persistent low-grade cytological abnormalities was a significant confounding factor. A high degree of false positive and negative results in the 20–30 year olds found in this study could have implications for the implementation of HPV triage in the UK. Although the commencement of screening age is going to be aligned to 25 across the UK there is likely to be a large number of colposcopy referrals with the introduction of an HPV triage test in the 25–30 year olds.

Viral integration and viral DNA methylation were both associated with high-grade disease, however, more work is required in both areas to both understand the HPV biology, their interaction with each other, and the optimal method for identifying them as biomarkers. New sites of viral integration were found in the study and the predilection of common fragile sites and repeat sequences as sites of integration was also reinforced.

More consideration of the reproducibility of biomarker assays should be given and journal editors and reviewers should be more stringent on this issue. Furthermore, the possibility of integration in a target region of an HPV test resulting in a false negative result appears to have been largely overlooked. Should this finding be repeated in other studies an alternative viral target may be required to ensure a high-grade cervical lesion or early stage cervical cancer is not missed, which is ultimately the aim of the cervical screening programme.

References

AGOFF, S. N., LIN, P., MORIHARA, J., MAO, C., KIVIAT, N. B. & KOUTSKY, L. A. 2003. p16(INK4a) expression correlates with degree of cervical neoplasia: a comparison with Ki-67 expression and detection of high-risk HPV types. *Mod Pathol,* 16**,** 665-73.

AI, W., NARAHARI, J. & ROMAN, A. 2000. Yin yang 1 negatively regulates the differentiation-specific E1 promoter of human papillomavirus type 6. *J Virol,* 74**,** 5198-205.

ALAZAWI, W., PETT, M., STRAUSS, S., MOSELEY, R., GRAY, J., STANLEY, M. & COLEMAN, N. 2004. Genomic imbalances in 70 snap-frozen cervical squamous intraepithelial lesions: associations with lesion grade, state of the HPV16 E2 gene and clinical outcome. *British Journal of Cancer,* 91**,** 2063-70.

ALTS GROUP 2000. Human papillomavirus testing for triage of women with cytologic evidence of low-grade squamous intraepithelial lesions: baseline data from a randomized trial. The Atypical Squamous Cells of Undetermined Significance/Low-Grade Squamous Intraepithelial Lesions Triage Study (ALTS) Group. *Journal of the National Cancer Institute,* 92**,** 397-402.

ALTS GROUP 2003. Results of a randomized trial on the management of cytology interpretations of atypical squamous cells of undetermined significance. *American Journal of Obstetrics & Gynecology,* 188**,** 1383-92.

ANDERSSON, S., DILLNER, L., ELFGREN, K., MINTS, M., PERSSON, M. & RYLANDER, E. 2005. A comparison of the human papillomavirus test and Papanicolaou smear as a second screening method for women with minor cytological abnormalities. *Acta Obstetricia et Gynecologica Scandinavica,* 84**,** 996-1000.

ANDERSSON-ELLSTROM, A., HAGMAR, B. M., JOHANSSON, B., KALANTARI, M., WARLEBY, B. & FORSSMAN, L. 1996. Human papillomavirus deoxyribonucleic acid in cervix only detected in girls after coitus. *Int J STD AIDS,* 7**,** 333-6.

ANG, M. S., KAUFMAN, R. H., ADAM, E., RIDDLE, G., IRWIN, J. F., REEVES, K. O. & SIMON, T. L. 1995. Colposcopically directed biopsy and loop excision of the transformation zone. Comparison of histologic findings. *J Reprod Med,* 40**,** 167-70.

ANTEQUERA, F. & BIRD, A. 1993. Number of CpG islands and genes in human and mouse. *Proceedings of the National Academy of Sciences of the United States of America,* 90**,** 11995-9.

APPLEBY, P., BERAL, V., BERRINGTON DE GONZALEZ, A., COLIN, D., FRANCESCHI, S., GOODILL, A., GREEN, J., PETO, J., PLUMMER, M. & SWEETLAND, S. 2006. 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. *International Journal of Cancer,* 118**,** 1481-95.

ARBYN, M., KYRGIOU, M., SIMOENS, C., RAIFU, A. O., KOLIOPOULOS, G., MARTIN-HIRSCH, P., PRENDIVILLE, W. & PARASKEVAIDIS, E. 2008. Perinatal mortality and other severe adverse pregnancy outcomes associated with treatment of cervical intraepithelial neoplasia: meta-analysis. *BMJ,* 337**,** a1284.

ARBYN, M., MARTIN-HIRSCH, P., BUNTINX, F., VAN RANST, M., PARASKEVAIDIS, E. & DILLNER, J. 2009a. Triage of women with equivocal or low-grade cervical cytology results: a meta-analysis of the HPV test positivity rate. *J Cell Mol Med,* 13**,** 648-59.

ARBYN, M., MARTIN-HIRSCH, P. & WENTZENSEN, N. 2010. Human papillomavirus-based triage of women showing a cervical cytology result of borderline or mild dyskaryosis. *Bjog,* 117**,** 641-4.

ARBYN, M., RONCO, G., ANTTILA, A., MEIJER, C. J., POLJAK, M., OGILVIE, G., KOLIOPOULOS, G., NAUCLER, P., SANKARANARAYANAN, R. & PETO, J. 2012. Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer. *Vaccine,* 30 Suppl 5**,** F88-99.

ARBYN, M., RONCO, G., CUZICK, J., WENTZENSEN, N. & CASTLE, P. E. 2009b. How to evaluate emerging technologies in cervical cancer screening? *International Journal of Cancer,* 125**,** 2489-96.

ARIAS-PULIDO, H., PEYTON, C. L., JOSTE, N. E., VARGAS, H. & WHEELER, C. M. 2006. Human papillomavirus type 16 integration in cervical carcinoma in situ and in invasive cervical cancer. *Journal of clinical microbiology,* 44**,** 1755-62.

ASHRAFI, G. H., HAGHSHENAS, M., MARCHETTI, B. & CAMPO, M. S. 2006. E5 protein of human papillomavirus 16 downregulates HLA class I and interacts with the heavy chain via its first hydrophobic domain. *International Journal of Cancer,* 119**,** 2105-12.

BAILIS, J. M., LUCHE, D. D., HUNTER, T. & FORSBURG, S. L. 2008. Minichromosome maintenance proteins interact with checkpoint and recombination proteins to promote s-phase genome stability. *Mol Cell Biol,* 28**,** 1724-38.

BASEMAN, J. G. & KOUTSKY, L. A. 2005. The epidemiology of human papillomavirus infections. *J Clin Virol,* 32 Suppl 1**,** S16-24.

BERKHOF, J., BULKMANS, N. W., BLEEKER, M. C., BULK, S., SNIJDERS, P. J., VOORHORST, F. J. & MEIJER, C. J. 2006. Human papillomavirus type-specific 18-month risk of high-grade cervical intraepithelial neoplasia in women with a normal or borderline/mildly dyskaryotic smear. *Cancer Epidemiology, Biomarkers & Prevention,* 15**,** 1268-73.

BERNARD, H. U., BURK, R. D., CHEN, Z., VAN DOORSLAER, K., HAUSEN, H. & DE VILLIERS, E. M. 2010. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology,* 401**,** 70-9.

BIOMARKERS DEFINITIONS WORKING GROUP 2001. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther,* 69**,** 89-95.

BIPAT, S., GLAS, A. S., VAN DER VELDEN, J., ZWINDERMAN, A. H., BOSSUYT, P. M. & STOKER, J. 2003. Computed tomography and magnetic resonance imaging in staging of uterine cervical carcinoma: a systematic review. *Gynecologic Oncology,* 91**,** 59-66.

BIRCHMEIER, C., MEYER, D. & RIETHMACHER, D. 1995. Factors controlling growth, motility, and morphogenesis of normal and malignant epithelial cells. *Int Rev Cytol,* 160**,** 221-66.

BLAND, J. M. & ALTMAN, D. G. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet,* 1**,** 307-10.

BOSCH, F. X. & DE SANJOSE, S. 2003. Chapter 1: Human papillomavirus and cervical cancer--burden and assessment of causality. *Journal of the National Cancer Institute. Monographs***,** 3-13.

BOSCH, F. X., MANOS, M. M., MUNOZ, N., SHERMAN, M., JANSEN, A. M., PETO, J., SCHIFFMAN, M. H., MORENO, V., KURMAN, R. & SHAH, K. V. 1995. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. International biological study on cervical cancer (IBSCC) Study Group. *Journal of the National Cancer Institute,* 87**,** 796-802.

BOUVARD, V., BAAN, R., STRAIF, K., GROSSE, Y., SECRETAN, B., EL GHISSASSI, F., BENBRAHIM-TALLAA, L., GUHA, N., FREEMAN, C., GALICHET, L. & COGLIANO, V. 2009. A review of human carcinogens--Part B: biological agents. *Lancet Oncol,* 10**,** 321-2.

BOUVARD, V., STOREY, A., PIM, D. & BANKS, L. 1994. Characterization of the human papillomavirus E2 protein: evidence of trans-activation and trans-repression in cervical keratinocytes. *The EMBO journal,* 13**,** 5451-9.

BOYER, S. N., WAZER, D. E. & BAND, V. 1996a. E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway. *Cancer Res,* 56**,** 4620-4.

BOYER, S. N., WAZER, D. E. & BAND, V. 1996b. E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway. *Cancer Research,* 56**,** 4620-4.

BRAKE, T., CONNOR, J. P., PETEREIT, D. G. & LAMBERT, P. F. 2003. Comparative analysis of cervical cancer in women and in a human papillomavirus-transgenic mouse model: identification of minichromosome maintenance protein 7 as an informative biomarker for human cervical cancer. *Cancer Res,* 63**,** 8173-80.

BRAKE, T. & LAMBERT, P. F. 2005. Estrogen contributes to the onset, persistence, and malignant progression of cervical cancer in a human papillomavirus-transgenic mouse model. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 2490-5.

BRANDSMA, J. L., SUN, Y., LIZARDI, P. M., TUCK, D. P., ZELTERMAN, D., HAINES, G. K., 3RD, MARTEL, M., HARIGOPAL, M., SCHOFIELD, K. & NEAPOLITANO, M. 2009. Distinct human papillomavirus type 16 methylomes in cervical cells at different stages of premalignancy. *Virology,* 389**,** 100-7.

BREITBURD, F., KIRNBAUER, R., HUBBERT, N. L., NONNENMACHER, B., TRIN-DINH-DESMARQUET, C., ORTH, G., SCHILLER, J. T. & LOWY, D. R. 1995. Immunization with viruslike particles from cottontail rabbit papillomavirus (CRPV) can protect against experimental CRPV infection. *J Virol,* 69**,** 3959-63.

BROWN, D. R., KITCHIN, D., QADADRI, B., NEPTUNE, N., BATTEIGER, T. & ERMEL, A. 2006. The human papillomavirus type 11 E1--E4 protein is a transglutaminase 3 substrate and induces abnormalities of the cornified cell envelope. *Virology,* 345**,** 290-8.

BRUNI, L., DIAZ, M., CASTELLSAGUE, X., FERRER, E., BOSCH, F. X. & DE SANJOSE, S. 2010. Cervical human papillomavirus prevalence in 5 continents: meta-analysis of 1 million women with normal cytological findings. *Journal of Infectious Diseases,* 202**,** 1789-99.

BRYANT, D. 2012. *Methylation of Human Papillomavirus DNA: Biological Significance and Clinical Utility.* PhD, Cardiff University.

BULKMANS, N. W., BERKHOF, J., ROZENDAAL, L., VAN KEMENADE, F. J., BOEKE, A. J., BULK, S., VOORHORST, F. J., VERHEIJEN, R. H., VAN GRONINGEN, K., BOON, M. E., RUITINGA, W., VAN BALLEGOOIJEN, M., SNIJDERS, P. J. & MEIJER, C. J. 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.

BURGERS, W. A., BLANCHON, L., PRADHAN, S., DE LAUNOIT, Y., KOUZARIDES, T. & FUKS, F. 2007. Viral oncoproteins target the DNA methyltransferases. *Oncogene,* 26**,** 1650-5.

BUXTON, E. J., LUESLEY, D. M., SHAFI, M. I. & ROLLASON, M. 1991. Colposcopically directed punch biopsy: a potentially misleading investigation. *Br J Obstet Gynaecol,* 98**,** 1273-6.

CALLEJA-MACIAS, I. E., KALANTARI, M. & BERNARD, H. U. 2009. Cholinergic signaling through nicotinic acetylcholine receptors stimulates the proliferation of cervical cancer cells: an explanation for the molecular role of tobacco smoking in cervical carcinogenesis? *International Journal of Cancer,* 124**,** 1090-6.

CANCER RESEARCH UK 2013. *CancerStats Key Facts - Cervical cancer*, Cancer Research UK.

CAO, X., QIN, J., XIE, Y., KHAN, O., DOWD, F., SCOFIELD, M., LIN, M. F. & TU, Y. 2006. Regulator of G-protein signaling 2 (RGS2) inhibits androgen-independent activation of androgen receptor in prostate cancer cells. *Oncogene,* 25**,** 3719-34.

CARREON, J. D., SHERMAN, M. E., GUILLEN, D., SOLOMON, D., HERRERO, R., JERONIMO, J., WACHOLDER, S., RODRIGUEZ, A. C., MORALES, J., HUTCHINSON, M., BURK, R. D. & SCHIFFMAN, M. 2007. CIN2 is a much less reproducible and less valid diagnosis than CIN3: results from a histological review of population-based cervical samples. *International Journal of Gynecological Pathology,* 26**,** 441-6.

CASTLE, P. E., STOLER, M. H., WRIGHT, T. C., JR., SHARMA, A., WRIGHT, T. L. & BEHRENS, C. M. 2011. Performance of carcinogenic human papillomavirus (HPV) testing and HPV16 or HPV18 genotyping for cervical cancer screening of women aged 25 years and older: a subanalysis of the ATHENA study. *Lancet Oncol,* 12**,** 880-90.

CERVICAL SCREENING WALES 2007. KC53/61/65 Statistical Report – Adroddiad Ystadegol 2006/07. Cardiff, UK: Cervical Screening Wales.

CERVICAL SCREENING WALES. 2012a. *Cervical Screening Programme, Wales: 2011/2012* [Online]. Available: http://www.screeningservices.org.uk/csw/prof/reports/KC53-61-65_11-12.pdf [Accessed 28/10/2012].

CERVICAL SCREENING WALES 2012b. Cervical Screening Wales Quality Manual. Section 2R – Core Reference. Version 1.0 revision 5. Cardiff, UK: Cervical Screening Wales.

CHAMPOUX, J. J. 2001. DNA topoisomerases: structure, function, and mechanism. *Annu Rev Biochem,* 70**,** 369-413.

CHAUDHURI, B., XU, H., TODOROV, I., DUTTA, A. & YATES, J. L. 2001. Human DNA replication initiation factors, ORC and MCM, associate with oriP of Epstein-Barr virus. *Proceedings of the National Academy of Sciences of the United States of America,* 98**,** 10085-9.

CHEN, Y., MILLER, C., MOSHER, R., ZHAO, X., DEEDS, J., MORRISSEY, M., BRYANT, B., YANG, D., MEYER, R., CRONIN, F., GOSTOUT, B. S., SMITH-MCCUNE, K. & SCHLEGEL, R. 2003. Identification of cervical cancer markers by cDNA and tissue microarrays. *Cancer Res,* 63**,** 1927-35.

CHENG, S., SCHMIDT-GRIMMINGER, D. C., MURANT, T., BROKER, T. R. & CHOW, L. T. 1995. Differentiation-dependent up-regulation of the human papillomavirus E7 gene reactivates cellular DNA replication in suprabasal differentiated keratinocytes. *Genes & development,* 9**,** 2335-49.

CHO, N. H., KIM, Y. T. & KIM, J. W. 2002. Alteration of cell cycle in cervical tumor associated with human papillomavirus: cyclin-dependent kinase inhibitors. *Yonsei Med J,* 43**,** 722-8.

CHOO, K. B., PAN, C. C. & HAN, S. H. 1987. Integration of human papillomavirus type 16 into cellular DNA of cervical carcinoma: preferential deletion of the E2 gene and invariable retention of the long control region and the E6/E7 open reading frames. *Virology,* 161**,** 259-61.

CIULLO, M., DEBILY, M. A., ROZIER, L., AUTIERO, M., BILLAULT, A., MAYAU, V., EL MARHOMY, S., GUARDIOLA, J., BERNHEIM, A., COULLIN, P., PIATIER-TONNEAU, D. & DEBATISSE, M. 2002. Initiation of the breakage-fusion-bridge mechanism through common fragile site activation in human breast cancer cells: the model of PIP gene duplication from a break at FRA7I. *Hum Mol Genet,* 11**,** 2887-94.

CLARKE, M. A., WENTZENSEN, N., MIRABELLO, L., GHOSH, A., WACHOLDER, S., HARARI, A., LORINCZ, A., SCHIFFMAN, M. & BURK, R. D. 2012. Human papillomavirus DNA methylation as a potential biomarker for cervical cancer. *Cancer Epidemiology, Biomarkers & Prevention,* 21**,** 2125-37.

CLIFFORD, G. M., RANA, R. K., FRANCESCHI, S., SMITH, J. S., GOUGH, G. & PIMENTA, J. M. 2005. Human papillomavirus genotype distribution in low-grade cervical lesions: comparison by geographic region and with cervical cancer. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology,* 14**,** 1157-64.

COLLINS, S., ROLLASON, T. P., YOUNG, L. S. & WOODMAN, C. B. 2010. Cigarette smoking is an independent risk factor for cervical intraepithelial neoplasia in young women: a longitudinal study. *Eur J Cancer,* 46**,** 405-11.

COLLINS, S. I., CONSTANDINOU-WILLIAMS, C., WEN, K., YOUNG, L. S., ROBERTS, S., MURRAY, P. G. & WOODMAN, C. B. 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.

CONWAY, M. J., ALAM, S., CHRISTENSEN, N. D. & MEYERS, C. 2009. Overlapping and independent structural roles for human papillomavirus type 16 L2 conserved cysteines. *Virology,* 393**,** 295-303.

CORDEN, S. A., SANT-CASSIA, L. J., EASTON, A. J. & MORRIS, A. G. 1999. The integration of HPV-18 DNA in cervical carcinoma. *Mol Pathol,* 52**,** 275-82.

COTTON, S., SHARP, L., LITTLE, J., CRUICKSHANK, M., SETH, R., SMART, L., DUNCAN, I., HARRILD, K., NEAL, K. & WAUGH, N. 2010. The role of human papillomavirus testing in the management of women with low-grade abnormalities: multicentre randomised controlled trial. *Bjog,* 117**,** 645-59.

CRICCA, M., VENTUROLI, S., LEO, E., COSTA, S., MUSIANI, M. & ZERBINI, M. 2009. Disruption of HPV 16 E1 and E2 genes in precancerous cervical lesions. *Journal of Virological Methods,* 158**,** 180-3.

CULLEN, A. P., REID, R., CAMPION, M. & LORINCZ, A. T. 1991. Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasm. *J Virol,* 65**,** 606-12.

CUSCHIERI, K. & WENTZENSEN, N. 2008. Human papillomavirus mRNA and p16 detection as biomarkers for the improved diagnosis of cervical neoplasia. *Cancer Epidemiology, Biomarkers & Prevention,* 17**,** 2536-45.

CUZICK, J., BERGERON, C., VON KNEBEL DOEBERITZ, M., GRAVITT, P., JERONIMO, J., LORINCZ, A. T., MEIJER, C., SANKARANARAYANAN, R., SNIJDERS, P. & SZAREWSKI, A. 2012. New technologies and procedures for cervical cancer screening. *Vaccine,* 30 Suppl 5**,** F107-16.

CUZICK, J., CLAVEL, C., PETRY, K. U., MEIJER, C. J., HOYER, H., RATNAM, S., SZAREWSKI, A., BIREMBAUT, P., KULASINGAM, S., SASIENI, P. & IFTNER, T. 2006. Overview of the European and North American studies on HPV testing in primary cervical cancer screening. *Int J Cancer,* 119**,** 1095-101.

DALL, K. L., SCARPINI, C. G., ROBERTS, I., WINDER, D. M., STANLEY, M. A., MURALIDHAR, B., HERDMAN, M. T., PETT, M. R. & COLEMAN, N. 2008. Characterization of Naturally Occurring HPV16 Integration Sites Isolated from Cervical Keratinocytes under Noncompetitive Conditions. *Cancer Research,* 68**,** 8249-8259.

DANIEL, B., MUKHERJEE, G., SESHADRI, L., VALLIKAD, E. & KRISHNA, S. 1995. Changes in the physical state and expression of human papillomavirus type 16 in the progression of cervical intraepithelial neoplasia lesions analysed by PCR. *The Journal of general virology,* 76 (Pt 10)**,** 2589-93.

DAO, L. D., DUFFY, A., VAN TINE, B. A., WU, S. Y., CHIANG, C. M., BROKER, T. R. & CHOW, L. T. 2006. Dynamic localization of the human papillomavirus type 11 origin binding protein E2 through mitosis while in association with the spindle apparatus. *J Virol,* 80**,** 4792-800.

DAVY, C. E., AYUB, M., JACKSON, D. J., DAS, P., MCINTOSH, P. & DOORBAR, J. 2006. HPV16 E1--E4 protein is phosphorylated by Cdk2/cyclin A and relocalizes this complex to the cytoplasm. *Virology,* 349**,** 230-44.

DAVY, C. E., JACKSON, D. J., RAJ, K., PEH, W. L., SOUTHERN, S. A., DAS, P., SORATHIA, R., LASKEY, P., MIDDLETON, K., NAKAHARA, T., WANG, Q., MASTERSON, P. J., LAMBERT, P. F., CUTHILL, S., MILLAR, J. B. & DOORBAR, J. 2005. Human papillomavirus type 16 E1 E4-induced G2 arrest is associated with cytoplasmic retention of active Cdk1/cyclin B1 complexes. *J Virol,* 79**,** 3998-4011.

DE SANJOSE, S., QUINT, W. G., ALEMANY, L., GERAETS, D. T., KLAUSTERMEIER, J. E., LLOVERAS, B., TOUS, S., FELIX, A., BRAVO, L. E., SHIN, H. R., VALLEJOS, C. S., DE RUIZ, P. A., LIMA, M. A., GUIMERA, N., CLAVERO, O., ALEJO, M., LLOMBART-BOSCH, A., CHENG-YANG, C., TATTI, S. A., KASAMATSU, E., ILJAZOVIC, E., ODIDA, M., PRADO, R., SEOUD, M., GRCE, M., USUBUTUN, A., JAIN, A., SUAREZ, G. A., LOMBARDI, L. E., BANJO, A., MENENDEZ, C., DOMINGO, E. J., VELASCO, J., NESSA, A., CHICHAREON, S. C., QIAO, Y. L., LERMA, E., GARLAND, S. M., SASAGAWA, T., FERRERA, A., HAMMOUDA, D., MARIANI, L., PELAYO, A., STEINER, I., OLIVA, E., MEIJER, C. J., AL-JASSAR, W. F., CRUZ, E., WRIGHT, T. C., PURAS, A., LLAVE, C. L., TZARDI, M., AGORASTOS, T., GARCIA-BARRIOLA, V., CLAVEL, C., ORDI, J., ANDUJAR, M., CASTELLSAGUE, X., SANCHEZ, G. I., NOWAKOWSKI, A. M., BORNSTEIN, J., MUNOZ, N. & BOSCH, F. X. 2010. Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol,* 11**,** 1048-56.

DE VILLIERS, E.-M., FAUQUET, C., BROKER, T. R., BERNARD, H.-U. & ZUR HAUSEN, H. 2004. Classification of papillomaviruses. *Virology,* 324**,** 17-27.

DEBACKER, K. & KOOY, R. F. 2007. Fragile sites and human disease. *Hum Mol Genet,* 16 Spec No. 2**,** R150-8.

DEMERET, C., GOYAT, S., YANIV, M. & THIERRY, F. 1998. The human papillomavirus type 18 (HPV18) replication protein E1 is a transcriptional activator when interacting with HPV18 E2. *Virology,* 242**,** 378-86.

DENNY, L. A., FRANCESCHI, S., DE SANJOSE, S., HEARD, I., MOSCICKI, A. B. & PALEFSKY, J. 2012. Human papillomavirus, human immunodeficiency virus and immunosuppression. *Vaccine,* 30 Suppl 5**,** F168-74.

DENTON, K. J., BERGERON, C., KLEMENT, P., TRUNK, M. J., KELLER, T. & RIDDER, R. 2010. The sensitivity and specificity of p16(INK4a) cytology vs HPV testing for detecting high-grade cervical disease in the triage of ASC-US and LSIL pap cytology results. *Am J Clin Pathol,* 134**,** 12-21.

DEPARTMENT OF HEALTH. 2011. *Improving Outcomes: A Strategy for Cancer* [Online]. Available: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_123371 [Accessed 28/10/2012].

DEPUYDT, C. E., MAKAR, A. P., RUYMBEKE, M. J., BENOY, I. H., VEREECKEN, A. J. & BOGERS, J. J. 2011. BD-ProExC as adjunct molecular marker for improved detection of CIN2+ after HPV primary screening. *Cancer Epidemiology, Biomarkers & Prevention,* 20**,** 628-37.

DESAINTES, C. & DEMERET, C. 1996. Control of papillomavirus DNA replication and transcription. *Seminars in cancer biology,* 7**,** 339-47.

DEY, P., COLLINS, S., DESAI, M. & WOODMAN, C. 1996. Adequacy of cervical cytology sampling with the Cervex brush and the Aylesbury spatula: a population based randomised controlled trial. *BMJ,* 313**,** 721-3.

DILLNER, J., KJAER, S. K., WHEELER, C. M., SIGURDSSON, K., IVERSEN, O. E., HERNANDEZ-AVILA, M., PEREZ, G., BROWN, D. R., KOUTSKY, L. A., TAY, E. H., GARCIA, P., AULT, K. A., GARLAND, S. M., LEODOLTER, S., OLSSON, S. E., TANG, G. W., FERRIS, D. G., PAAVONEN, J., LEHTINEN, M., STEBEN, M., BOSCH, F. X., JOURA, E. A., MAJEWSKI, S., MUNOZ, N., MYERS, E. R., VILLA, L. L., TADDEO, F. J., ROBERTS, C., TADESSE, A., BRYAN, J. T., MAANSSON, R., LU, S., VUOCOLO, S., HESLEY, T. M., BARR, E. & HAUPT, R. 2010. Four year efficacy of prophylactic human papillomavirus quadrivalent vaccine against low grade cervical, vulvar, and vaginal intraepithelial neoplasia and anogenital warts: randomised controlled trial. *BMJ,* 341**,** c3493.

DOORBAR, J. 2005. The papillomavirus life cycle. *Journal of Clinical Virology,* 32, Supplement**,** 7-15.

DOORBAR, J. 2006. Molecular biology of human papillomavirus infection and cervical cancer. *Clinical science,* 110**,** 525-41.

DOORBAR, J., CAMPBELL, D., GRAND, R. J. & GALLIMORE, P. H. 1986. Identification of the human papilloma virus-1a E4 gene products. *The EMBO journal,* 5**,** 355-62.

DOORBAR, J., ELY, S., STERLING, J., MCLEAN, C. & CRAWFORD, L. 1991. Specific interaction between HPV-16 E1-E4 and cytokeratins results in collapse of the epithelial cell intermediate filament network. *Nature,* 352**,** 824-7.

DOORBAR, J., QUINT, W., BANKS, L., BRAVO, I. G., STOLER, M., BROKER, T. R. & STANLEY, M. A. 2012. The biology and life-cycle of human papillomaviruses. *Vaccine,* 30 Suppl 5**,** F55-70.

DOSTATNI, N., LAMBERT, P. F., SOUSA, R., HAM, J., HOWLEY, P. M. & YANIV, M. 1991. The functional BPV-1 E2 trans-activating protein can act as a repressor by preventing formation of the initiation complex. *Genes & development,* 5**,** 1657-71.

DUENSING, S. & MUNGER, K. 2004. Mechanisms of genomic instability in human cancer: insights from studies with human papillomavirus oncoproteins. *International Journal of Cancer,* 109**,** 157-62.

DUPONT, J. M., TOST, J., JAMMES, H. & GUT, I. G. 2004. De novo quantitative bisulfite sequencing using the pyrosequencing technology. *Anal Biochem,* 333**,** 119-27.

EKLUND, C., WALLIN, K. L., FORSLUND, O. & DILLNER, J. 2012. Technical Report on the HPV LabNet 2011 HPV DNA Genotyping Proficiency Panel. Geneva, Switzerland.

ESTELLER, M., FRAGA, M. F., GUO, M., GARCIA-FONCILLAS, J., HEDENFALK, I., GODWIN, A. K., TROJAN, J., VAURS-BARRIERE, C., BIGNON, Y. J., RAMUS, S., BENITEZ, J., CALDES, T., AKIYAMA, Y., YUASA, Y., LAUNONEN, V., CANAL, M. J., RODRIGUEZ, R., CAPELLA, G., PEINADO, M. A., BORG, A., AALTONEN, L. A., PONDER, B. A., BAYLIN, S. B. & HERMAN,

J. G. 2001. DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. *Hum Mol Genet,* 10**,** 3001-7.

FEHRMANN, F., KLUMPP, D. J. & LAIMINS, L. A. 2003. Human papillomavirus type 31 E5 protein supports cell cycle progression and activates late viral functions upon epithelial differentiation. *J Virol,* 77**,** 2819-31.

FEINBERG, A. P. & TYCKO, B. 2004. The history of cancer epigenetics. *Nature reviews. Cancer,* 4**,** 143-53.

FERLAY J, S. H., BRAY F, FORMAN D, MATHERS C AND PARKIN DM. 2010. *GLOBOCAN 2008, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10* [Online]. Lyon, France: International Agency for Research on Cancer. Available: http://globocan.iarc.fr [Accessed 18/01/2013].

FERMANIAN, J. 1984. [Measuring agreement between 2 observers: a quantitative case]. *Rev Epidemiol Sante Publique,* 32**,** 408-13.

FERNANDEZ, A. F., ROSALES, C., LOPEZ-NIEVA, P., GRANA, O., BALLESTAR, E., ROPERO, S., ESPADA, J., MELO, S. A., LUJAMBIO, A., FRAGA, M. F., PINO, I., JAVIERRE, B., CARMONA, F. J., ACQUADRO, F., STEENBERGEN, R. D., SNIJDERS, P. J., MEIJER, C. J., PINEAU, P., DEJEAN, A., LLOVERAS, B., CAPELLA, G., QUER, J., BUTI, M., ESTEBAN, J. I., ALLENDE, H., RODRIGUEZ-FRIAS, F., CASTELLSAGUE, X., MINAROVITS, J., PONCE, J., CAPELLO, D., GAIDANO, G., CIGUDOSA, J. C., GOMEZ-LOPEZ, G., PISANO, D. G., VALENCIA, A., PIRIS, M. A., BOSCH, F. X., CAHIR-MCFARLAND, E., KIEFF, E. & ESTELLER, M. 2009. The dynamic DNA methylomes of double-stranded DNA viruses associated with human cancer. *Genome Res,* 19**,** 438-51.

FLORES, E. R., ALLEN-HOFFMANN, B. L., LEE, D. & LAMBERT, P. F. 2000. The human papillomavirus type 16 E7 oncogene is required for the productive stage of the viral life cycle. *J Virol,* 74**,** 6622-31.

FLORES, E. R., ALLEN-HOFFMANN, B. L., LEE, D., SATTLER, C. A. & LAMBERT, P. F. 1999. Establishment of the human papillomavirus type 16 (HPV-16) life cycle in an immortalized human foreskin keratinocyte cell line. *Virology,* 262**,** 344-54.

FLORES, E. R. & LAMBERT, P. F. 1997. Evidence for a switch in the mode of human papillomavirus type 16 DNA replication during the viral life cycle. *J Virol,* 71**,** 7167-79.

FRAGA, M. F. & ESTELLER, M. 2002. DNA methylation: a profile of methods and applications. *Biotechniques,* 33**,** 632, 634, 636-49.

FRANCESCHI, S., HERRERO, R., CLIFFORD, G. M., SNIJDERS, P. J., ARSLAN, A., ANH, P. T., BOSCH, F. X., FERRECCIO, C., HIEU, N. T., LAZCANO-PONCE, E., MATOS, E., MOLANO, M., QIAO, Y. L., RAJKUMAR, R., RONCO, G., DE SANJOSE, S., SHIN, H. R., SUKVIRACH, S., THOMAS, J. O., MEIJER, C. J. & MUNOZ, N. 2006. Variations in the age-specific curves of human papillomavirus prevalence in women worldwide. *International Journal of Cancer,* 119**,** 2677-84.

FREEMAN, A., MORRIS, L. S., MILLS, A. D., STOEBER, K., LASKEY, R. A., WILLIAMS, G. H. & COLEMAN, N. 1999. Minichromosome maintenance proteins as biological markers of dysplasia and malignancy. *Clinical Cancer Research,* 5**,** 2121-32.

FROMMER, M., MCDONALD, L. E., MILLAR, D. S., COLLIS, C. M., WATT, F., GRIGG, G. W., MOLLOY, P. L. & PAUL, C. L. 1992. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proceedings of the National Academy of Sciences of the United States of America,* 89**,** 1827-31.

FU, L., VAN DOORSLAER, K., CHEN, Z., RISTRIANI, T., MASSON, M., TRAVE, G. & BURK, R. D. 2010. Degradation of p53 by human Alphapapillomavirus E6 proteins shows a stronger correlation with phylogeny than oncogenicity. *PLoS ONE,* 5.

FUTURE II STUDY GROUP 2007. Quadrivalent vaccine against human papillomavirus to prevent high-grade cervical lesions. *New England Journal of Medicine,* 356**,** 1915-27.

GALLO, G., BIBBO, M., BAGELLA, L., ZAMPARELLI, A., SANSEVERINO, F., GIOVAGNOLI, M. R., VECCHIONE, A. & GIORDANO, A. 2003. Study of viral integration of HPV-16 in young patients with LSIL. *Journal of clinical pathology,* 56**,** 532-6.

GARIGLIO, P., GUTIERREZ, J., CORTES, E. & VAZQUEZ, J. 2009. The role of retinoid deficiency and estrogens as cofactors in cervical cancer. *Archives of Medical Research,* 40**,** 449-65.

GARLAND, S. M., HERNANDEZ-AVILA, M., WHEELER, C. M., PEREZ, G., HARPER, D. M., LEODOLTER, S., TANG, G. W., FERRIS, D. G., STEBEN, M., BRYAN, J., TADDEO, F. J., RAILKAR, R., ESSER, M. T., SINGS, H. L., NELSON, M., BOSLEGO, J., SATTLER, C., BARR, E. & KOUTSKY, L. A. 2007. Quadrivalent vaccine against human papillomavirus to prevent anogenital diseases. *New England Journal of Medicine,* 356**,** 1928-43.

GARTEL, A. L. & RADHAKRISHNAN, S. K. 2005. Lost in transcription: p21 repression, mechanisms, and consequences. *Cancer Res,* 65**,** 3980-5.

GIROGLOU, T., FLORIN, L., SCHAFER, F., STREECK, R. E. & SAPP, M. 2001. Human papillomavirus infection requires cell surface heparan sulfate. *Journal of virology,* 75**,** 1565-70.

GLOSS, B. & BERNARD, H. U. 1990. The E6/E7 promoter of human papillomavirus type 16 is activated in the absence of E2 proteins by a sequence-aberrant Sp1 distal element. *J Virol,* 64**,** 5577-84.

GUAN, P., HOWELL-JONES, R., LI, N., BRUNI, L., DE SANJOSE, S., FRANCESCHI, S. & CLIFFORD, G. M. 2012. Human papillomavirus types in 115,789 HPV-positive women: a meta-analysis from cervical infection to cancer. *International Journal of Cancer,* 131**,** 2349-59.

HAFNER, N., DRIESCH, C., GAJDA, M., JANSEN, L., KIRCHMAYR, R., RUNNEBAUM, I. B. & DURST, M. 2008. Integration of the HPV16 genome does not invariably result in high levels of viral oncogene transcripts. *Oncogene,* 27**,** 1610-7.

HAGENSEE, M. E., OLSON, N. H., BAKER, T. S. & GALLOWAY, D. A. 1994. Three-dimensional structure of vaccinia virus-produced human papillomavirus type 1 capsids. *J Virol,* 68**,** 4503-5.

HANKE, S., BUGERT, P., CHUDEK, J. & KOVACS, G. 2001. Cloning a calcium channel alpha2delta-3 subunit gene from a putative tumor suppressor gene region at chromosome 3p21.1 in conventional renal cell carcinoma. *Gene,* 264**,** 69-75.

HARWOOD, C. A. & PROBY, C. M. 2002. Human papillomaviruses and non-melanoma skin cancer. *Current opinion in infectious diseases,* 15**,** 101-14.

HEALTH AND SOCIAL CARE INFORMATION CENTRE. 2012. *Cervical Screening Programme, England 2011-12* [Online]. Available: http://www.ic.nhs.uk/webfiles/publications/008_Screening/cervscreen1112/Cervical_Screening_Bulletin_1112.pdf [Accessed 28/10/2012].

HENDERSON, D., HALL, L., PRPIC, N., HESSLING, J., PARKER, M., SAMPSON, S., SIMKINS, S., BROUGH, G., DIXON, E., LENZ, K., KNAPP, S., MURPHY, P., TAYLOR, A., FISCHER, T. & MALINOWSKI, D. P. 2011. The selection and characterization of antibodies to minichromosome maintenance proteins that highlight cervical dysplasia. *J Immunol Methods,* 370**,** 1-13.

HEROD, J. 2010. Cervical Cancer Standards of Care. *In:* ACHESON, N. & LUESLEY, D. (eds.) *Gynaecological Oncology for the MRCOG and Beyond - Second Edition.* London: RCOG.

HERRICK, J., CONTI, C., TEISSIER, S., THIERRY, F., COUTURIER, J., SASTRE-GARAU, X., FAVRE, M., ORTH, G. & BENSIMON, A. 2005. Genomic organization of amplified MYC genes suggests distinct mechanisms of amplification in tumorigenesis. *Cancer Res,* 65**,** 1174-9.

HESSELINK, A. T., HEIDEMAN, D. A., BERKHOF, J., TOPAL, F., POL, R. P., MEIJER, C. J. & SNIJDERS, P. J. 2010. Comparison of the clinical performance of PapilloCheck human papillomavirus detection with that of the GP5+/6+-PCR-enzyme immunoassay in population-based cervical screening. *Journal of clinical microbiology,* 48**,** 797-801.

HESSELINK, A. T., HEIDEMAN, D. A., STEENBERGEN, R. D., COUPE, V. M., OVERMEER, R. M., RIJKAART, D., BERKHOF, J., MEIJER, C. J. & SNIJDERS, P. J. 2011. Combined promoter methylation analysis of CADM1 and MAL: an objective triage tool for high-risk human papillomavirus DNA-positive women. *Clinical Cancer Research,* 17**,** 2459-65.

HIBBITTS, S., RIECK, G. C., HART, K., POWELL, N. G., BEUKENHOLDT, R., DALLIMORE, N., MCREA, J., HAUKE, A., TRISTRAM, A. & FIANDER, A. N. 2006. Human papillomavirus infection: an anonymous prevalence study in South Wales, UK. *Br J Cancer,* 95**,** 226-32.

HINES, C. S., MEGHOO, C., SHETTY, S., BIBURGER, M., BRENOWITZ, M. & HEGDE, R. S. 1998. DNA structure and flexibility in the sequence-specific binding of papillomavirus E2 proteins. *J Mol Biol,* 276**,** 809-18.

HOFFMANN, R., HIRT, B., BECHTOLD, V., BEARD, P. & RAJ, K. 2006. Different modes of human papillomavirus DNA replication during maintenance. *J Virol,* 80**,** 4431-9.

HOLMGREN, S. C., PATTERSON, N. A., OZBUN, M. A. & LAMBERT, P. F. 2005. The minor capsid protein L2 contributes to two steps in the human papillomavirus type 31 life cycle. *J Virol,* 79**,** 3938-48.

HONEYCUTT, K. A., CHEN, Z., KOSTER, M. I., MIERS, M., NUCHTERN, J., HICKS, J., ROOP, D. R. & SHOHET, J. M. 2006. Deregulated minichromosomal maintenance protein MCM7 contributes to oncogene driven tumorigenesis. *Oncogene,* 25**,** 4027-32.

HOPMAN, A. H., SMEDTS, F., DIGNEF, W., UMMELEN, M., SONKE, G., MRAVUNAC, M., VOOIJS, G. P., SPEEL, E. J. & RAMAEKERS, F. C. 2004. Transition of high-grade cervical intraepithelial neoplasia to micro-invasive carcinoma is characterized by integration of HPV 16/18 and numerical chromosome abnormalities. *The Journal of pathology,* 202**,** 23-33.

HOPMAN, E. H., ROZENDAAL, L., VOORHORST, F. J., WALBOOMERS, J. M., KENEMANS, P. & HELMERHORST, T. J. 2000. High risk human papillomavirus in women with normal cervical cytology prior to the development of abnormal cytology and colposcopy. *Bjog,* 107**,** 600-4.

HOWIE, H. L., KATZENELLENBOGEN, R. A. & GALLOWAY, D. A. 2009. Papillomavirus E6 proteins. *Virology,* 384**,** 324-34.

HRICAK, H., GATSONIS, C., CHI, D. S., AMENDOLA, M. A., BRANDT, K., SCHWARTZ, L. H., KOELLIKER, S., SIEGELMAN, E. S., BROWN, J. J., MCGHEE, R. B., JR., IYER, R., VITELLAS, K. M., SNYDER, B., LONG, H. J., 3RD, FIORICA, J. V. & MITCHELL, D. G. 2005. Role of imaging in pretreatment evaluation of early invasive cervical cancer: results of the intergroup study American College of Radiology Imaging Network 6651-Gynecologic Oncology Group 183. *J Clin Oncol,* 23**,** 9329-37.

HUANG, L. W., CHAO, S. L. & LEE, B. H. 2008. Integration of human papillomavirus type-16 and type-18 is a very early event in cervical carcinogenesis. *Journal of clinical pathology,* 61**,** 627-31.

HUDELIST, G., MANAVI, M., PISCHINGER, K. I., WATKINS-RIEDEL, T., SINGER, C. F., KUBISTA, E. & CZERWENKA, K. F. 2004. Physical state and expression of HPV DNA in benign and dysplastic cervical tissue: different levels of viral integration are correlated with lesion grade. *Gynecologic Oncology,* 92**,** 873-80.

INTERNATIONAL AGENCY FOR RESEARCH ON CANCER 2005. *Cervix Cancer Screening. IARC Handbooks of Cancer Prevention,* Lyon, France, IARC Press.

ISD SCOTLAND. 2012. *Scottish Cervical Screening Programme Statistics 2011/12* [Online]. Available: http://www.isdscotland.org/Health-Topics/Cancer/Publications/2012-08-28/2012-08-28-Cervical-Screening-Report.pdf?71010988951 [Accessed 28/10/2012].

ISMAIL, S. M., COLCLOUGH, A. B., DINNEN, J. S., EAKINS, D., EVANS, D. M., GRADWELL, E., O'SULLIVAN, J. P., SUMMERELL, J. M. & NEWCOMBE, R. G. 1989. Observer variation in histopathological diagnosis and grading of cervical intraepithelial neoplasia. *BMJ,* 298**,** 707-10.

JEON, S., ALLEN-HOFFMANN, B. L. & LAMBERT, P. F. 1995. Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells. *J Virol,* 69**,** 2989-97.

JIANG, Z., WANG, Z., XU, Y., WANG, B., HUANG, W. & CAI, S. 2010. Analysis of RGS2 expression and prognostic significance in stage II and III colorectal cancer. *Biosci Rep,* 30**,** 383-90.

JOHANSSON, C., SOMBERG, M., LI, X., BACKSTROM WINQUIST, E., FAY, J., RYAN, F., PIM, D., BANKS, L. & SCHWARTZ, S. 2012. HPV-16 E2 contributes to induction of HPV-16 late gene expression by inhibiting early polyadenylation. *The EMBO journal,* 31**,** 3212-27.

JONES, M. H., SINGER, A. & JENKINS, D. 1996. The mildly abnormal cervical smear: patient anxiety and choice of management. *J R Soc Med,* 89**,** 257-60.

JORDAN, J., ARBYN, M., MARTIN-HIRSCH, P., SCHENCK, U., BALDAUF, J. J., DA SILVA, D., ANTTILA, A., NIEMINEN, P. & PRENDIVILLE, W. 2008. European guidelines for quality assurance in cervical cancer screening: recommendations for clinical management of abnormal cervical cytology, part 1. *Cytopathology : official journal of the British Society for Clinical Cytology,* 19**,** 342-54.

JOURA, E. A., LEODOLTER, S., HERNANDEZ-AVILA, M., WHEELER, C. M., PEREZ, G., KOUTSKY, L. A., GARLAND, S. M., HARPER, D. M., TANG, G. W., FERRIS, D. G., STEBEN, M., JONES, R. W., BRYAN, J., TADDEO, F. J., BAUTISTA, O. M., ESSER, M. T., SINGS, H. L., NELSON, M., BOSLEGO, J. W., SATTLER, C., BARR, E. & PAAVONEN, J. 2007. Efficacy of a quadrivalent prophylactic human papillomavirus (types 6, 11, 16, and 18) L1 virus-like-particle vaccine against high-grade vulval and vaginal lesions: a combined analysis of three randomised clinical trials. *Lancet,* 369**,** 1693-702.

KAHN, J. A. & BURK, R. D. 2007. Papillomavirus vaccines in perspective. *Lancet,* 369**,** 2135-7.

KALANTARI, M., CHASE, D. M., TEWARI, K. S. & BERNARD, H. U. 2010. Recombination of human papillomavirus-16 and host DNA in exfoliated cervical cells: a pilot study of L1 gene methylation and chromosomal integration as biomarkers of carcinogenic progression. *Journal of medical virology,* 82**,** 311-20.

KALANTARI, M., LEE, D., CALLEJA-MACIAS, I. E., LAMBERT, P. F. & BERNARD, H. U. 2008a. Effects of cellular differentiation, chromosomal integration and 5-aza-2'-deoxycytidine treatment on human papillomavirus-16 DNA methylation in cultured cell lines. *Virology,* 374**,** 292-303.

KALANTARI, M., VILLA, L. L., CALLEJA-MACIAS, I. E. & BERNARD, H. U. 2008b. Human papillomavirus-16 and -18 in penile carcinomas: DNA methylation, chromosomal recombination and genomic variation. *International Journal of Cancer,* 123**,** 1832-40.

KAMAL, M., SHAABAN, A. M., ZHANG, L., WALKER, C., GRAY, S., THAKKER, N., TOOMES, C., SPEIRS, V. & BELL, S. M. 2010. Loss of CSMD1 expression is associated with high tumour grade and poor survival in invasive ductal breast carcinoma. *Breast cancer research and treatment,* 121**,** 555-63.

KAMPER, N., DAY, P. M., NOWAK, T., SELINKA, H. C., FLORIN, L., BOLSCHER, J., HILBIG, L., SCHILLER, J. T. & SAPP, M. 2006. A membrane-destabilizing peptide in capsid protein L2 is required for egress of papillomavirus genomes from endosomes. *Journal of virology,* 80**,** 759-68.

KANODIA, S., FAHEY, L. M. & KAST, W. M. 2007. Mechanisms used by human papillomaviruses to escape the host immune response. *Curr Cancer Drug Targets,* 7**,** 79-89.

KEARSEY, S. E. & LABIB, K. 1998. MCM proteins: evolution, properties, and role in DNA replication. *Biochimica et biophysica acta,* 1398**,** 113-36.

KELLY, D., KINCAID, E., FANSLER, Z., ROSENTHAL, D. L. & CLARK, D. P. 2006. Detection of cervical high-grade squamous intraepithelial lesions from cytologic samples using a novel immunocytochemical assay (ProEx C). *Cancer,* 108**,** 494-500.

KELLY, R. S., PATNICK, J., KITCHENER, H. C. & MOSS, S. M. 2011. HPV testing as a triage for borderline or mild dyskaryosis on cervical cytology: results from the Sentinel Sites study. *British Journal of Cancer,* 105**,** 983-8.

KHAN, M. J., CASTLE, P. E., LORINCZ, A. T., WACHOLDER, S., SHERMAN, M., SCOTT, D. R., RUSH, B. B., GLASS, A. G. & SCHIFFMAN, M. 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. *Journal of the National Cancer Institute,* 97**,** 1072-9.

KHLEIF, S. N., DEGREGORI, J., YEE, C. L., OTTERSON, G. A., KAYE, F. J., NEVINS, J. R. & HOWLEY, P. M. 1996. Inhibition of cyclin D-CDK4/CDK6 activity is associated with an E2F-

mediated induction of cyclin kinase inhibitor activity. *Proceedings of the National Academy of Sciences of the United States of America,* 93**,** 4350-4.

KIATPONGSAN, S., NIRUTHISARD, S., MUTIRANGURA, A., TRIVIJITSILP, P., VASURATNA, A., CHAITHONGWONGWATTHANA, S. & LERTKHACHONSUK, R. 2006. Role of human papillomavirus DNA testing in management of women with atypical squamous cells of undetermined significance. *Int J Gynecol Cancer,* 16**,** 262-5.

KIM, K., ANGELETTI, P. C., HASSEBROEK, E. C. & LAMBERT, P. F. 2005. Identification of cis-acting elements that mediate the replication and maintenance of human papillomavirus type 16 genomes in Saccharomyces cerevisiae. *J Virol,* 79**,** 5933-42.

KIM, K. & LAMBERT, P. F. 2002. E1 protein of bovine papillomavirus 1 is not required for the maintenance of viral plasmid DNA replication. *Virology,* 293**,** 10-4.

KINES, R. C., THOMPSON, C. D., LOWY, D. R., SCHILLER, J. T. & DAY, P. M. 2009. The initial steps leading to papillomavirus infection occur on the basement membrane prior to cell surface binding. *Proceedings of the National Academy of Sciences of the United States of America,* 106**,** 20458-63.

KITCHENER, H. C., ALMONTE, M., THOMSON, C., WHEELER, P., SARGENT, A., STOYKOVA, B., GILHAM, C., BAYSSON, H., ROBERTS, C., DOWIE, R., DESAI, M., MATHER, J., BAILEY, A., TURNER, A., MOSS, S. & PETO, J. 2009. HPV testing in combination with liquid-based cytology in primary cervical screening (ARTISTIC): a randomised controlled trial. *Lancet Oncol,* 10**,** 672-82.

KITCHENER, H. C., GILHAM, C., SARGENT, A., BAILEY, A., ALBROW, R., ROBERTS, C., DESAI, M., MATHER, J., TURNER, A., MOSS, S. & PETO, J. 2011. A comparison of HPV DNA testing and liquid based cytology over three rounds of primary cervical screening: extended follow up in the ARTISTIC trial. *Eur J Cancer,* 47**,** 864-71.

KIVIAT, N. B., CRITCHLOW, C. W. & KURMAN, R. J. 1992. Reassessment of the morphological continuum of cervical intraepithelial lesions: does it reflect different stages in the progression to cervical carcinoma? *IARC Sci Publ***,** 59-66.

KIYONO, T., FOSTER, S. A., KOOP, J. I., MCDOUGALL, J. K., GALLOWAY, D. A. & KLINGELHUTZ, A. J. 1998. Both Rb/p16INK4a inactivation and telomerase activity are required to immortalize human epithelial cells. *Nature,* 396**,** 84-8.

KJAER, S. K., FREDERIKSEN, K., MUNK, C. & IFTNER, T. 2010. Long-term absolute risk of cervical intraepithelial neoplasia grade 3 or worse following human papillomavirus infection: role of persistence. *Journal of the National Cancer Institute,* 102**,** 1478-88.

KJAER, S. K., SIGURDSSON, K., IVERSEN, O. E., HERNANDEZ-AVILA, M., WHEELER, C. M., PEREZ, G., BROWN, D. R., KOUTSKY, L. A., TAY, E. H., GARCIA, P., AULT, K. A., GARLAND, S. M., LEODOLTER, S., OLSSON, S. E., TANG, G. W., FERRIS, D. G., PAAVONEN, J., LEHTINEN, M., STEBEN, M., BOSCH, F. X., DILLNER, J., JOURA, E. A., MAJEWSKI, S., MUNOZ, N., MYERS, E. R., VILLA, L. L., TADDEO, F. J., ROBERTS, C., TADESSE, A., BRYAN, J., MAANSSON, R., LU, S., VUOCOLO, S., HESLEY, T. M., SAAH, A., BARR, E. & HAUPT, R. M. 2009. A pooled analysis of continued prophylactic efficacy of quadrivalent human papillomavirus (Types 6/11/16/18) vaccine against high-grade cervical and external genital lesions. *Cancer Prev Res,* 2**,** 868-78.

KLAES, R., FRIEDRICH, T., SPITKOVSKY, D., RIDDER, R., RUDY, W., PETRY, U., DALLENBACH-HELLWEG, G., SCHMIDT, D. & VON KNEBEL DOEBERITZ, M. 2001. Overexpression of p16(INK4A) as a specific marker for dysplastic and neoplastic epithelial cells of the cervix uteri. *International Journal of Cancer,* 92**,** 276-84.

KLAES, R., WOERNER, S. M., RIDDER, R., WENTZENSEN, N., DUERST, M., SCHNEIDER, A., LOTZ, B., MELSHEIMER, P. & VON KNEBEL DOEBERITZ, M. 1999. Detection of high-risk cervical intraepithelial neoplasia and cervical cancer by amplification of transcripts derived from integrated papillomavirus oncogenes. *Cancer Res,* 59**,** 6132-6.

KLINGELHUTZ, A. J., FOSTER, S. A. & MCDOUGALL, J. K. 1996. Telomerase activation by the E6 gene product of human papillomavirus type 16. *Nature,* 380**,** 79-82.

KOCH, A., HANSEN, S. V., NIELSEN, N. M., PALEFSKY, J. & MELBYE, M. 1997. HPV detection in children prior to sexual debut. *International Journal of Cancer,* 73**,** 621-4.

KOUSHIK, A. & FRANCO, E. L. 2006. Epidemiology and the role of human papillomaviruses. *In:* JORDAN, J. & SINGER, A. (eds.) *The Cervix - Second Edition.* Oxford, UK: Blackwell Publishing Ltd.

KOUTSKY, L. 1997. Epidemiology of genital human papillomavirus infection. *Am J Med,* 102**,** 3-8.

KRAWCZYK, E., SUPRYNOWICZ, F. A., LIU, X., DAI, Y., HARTMANN, D. P., HANOVER, J. & SCHLEGEL, R. 2008. Koilocytosis: a cooperative interaction between the human papillomavirus E5 and E6 oncoproteins. *Am J Pathol,* 173**,** 682-8.

KUKIMOTO, I., AIHARA, S., YOSHIIKE, K. & KANDA, T. 1998. Human papillomavirus oncoprotein E6 binds to the C-terminal region of human minichromosome maintenance 7 protein. *Biochemical and biophysical research communications,* 249**,** 258-62.

LAMBERT, P. F., MONK, B. C. & HOWLEY, P. M. 1990. Phenotypic analysis of bovine papillomavirus type 1 E2 repressor mutants. *J Virol,* 64**,** 950-6.

LANDIS, J. R. & KOCH, G. G. 1977. The measurement of observer agreement for categorical data. *Biometrics,* 33**,** 159-74.

LANG, A. J., MIRSKI, S. E., CUMMINGS, H. J., YU, Q., GERLACH, J. H. & COLE, S. P. 1998. Structural organization of the human TOP2A and TOP2B genes. *Gene,* 221**,** 255-66.

LASKEY, R. A. & MADINE, M. A. 2003. A rotary pumping model for helicase function of MCM proteins at a distance from replication forks. *EMBO Rep,* 4**,** 26-30.

LEA, J. S., COLEMAN, R., KURIEN, A., SCHORGE, J. O., MILLER, D. S., MINNA, J. D. & MULLER, C. Y. 2004. Aberrant p16 methylation is a biomarker for tobacco exposure in cervical squamous cell carcinogenesis. *American Journal of Obstetrics & Gynecology,* 190**,** 674-9.

LEE, D., KIM, H. Z., JEONG, K. W., SHIM, Y. S., HORIKAWA, I., BARRETT, J. C. & CHOE, J. 2002. Human papillomavirus E2 down-regulates the human telomerase reverse transcriptase promoter. *The Journal of biological chemistry,* 277**,** 27748-56.

LEHTINEN, M., PAAVONEN, J., WHEELER, C. M., JAISAMRARN, U., GARLAND, S. M., CASTELLSAGUE, X., SKINNER, S. R., APTER, D., NAUD, P., SALMERON, J., CHOW, S. N., KITCHENER, H., TEIXEIRA, J. C., HEDRICK, J., LIMSON, G., SZAREWSKI, A., ROMANOWSKI, B., AOKI, F. Y., SCHWARZ, T. F., POPPE, W. A., DE CARVALHO, N. S., GERMAR, M. J., PETERS, K., MINDEL, A., DE SUTTER, P., BOSCH, F. X., DAVID, M. P., DESCAMPS, D., STRUYF, F. & DUBIN, G. 2012. Overall efficacy of HPV-16/18 AS04-adjuvanted vaccine against grade 3 or greater cervical intraepithelial neoplasia: 4-year end-of-study analysis of the randomised, double-blind PATRICIA trial. *Lancet Oncol,* 13**,** 89-99.

LEVI, F., LUCCHINI, F., NEGRI, E., FRANCESCHI, S. & LA VECCHIA, C. 2000. Cervical cancer mortality in young women in Europe: patterns and trends. *Eur J Cancer,* 36**,** 2266-71.

LEY, C., BAUER, H. M., REINGOLD, A., SCHIFFMAN, M. H., CHAMBERS, J. C., TASHIRO, C. J. & MANOS, M. M. 1991. Determinants of genital human papillomavirus infection in young women. *Journal of the National Cancer Institute,* 83**,** 997-1003.

LI, H., YANG, Y., ZHANG, R., CAI, Y., YANG, X., WANG, Z., LI, Y., CHENG, X., YE, X., XIANG, Y. & ZHU, B. 2013. Preferential sites for the integration and disruption of human papillomavirus 16 in cervical lesions. *J Clin Virol*.

LI, K., JIN, X., FANG, Y., WANG, C., GONG, M., CHEN, P., LIU, J., DENG, D. & AI, J. 2012. Correlation between physical status of human papilloma virus and cervical carcinogenesis. *J Huazhong Univ Sci Technolog Med Sci,* 32**,** 97-102.

LI, W., WANG, W., SI, M., HAN, L., GAO, Q., LUO, A., LI, Y., LU, Y., WANG, S. & MA, D. 2008. The physical state of HPV16 infection and its clinical significance in cancer precursor lesion and cervical carcinoma. *Journal of cancer research and clinical oncology,* 134**,** 1355-61.

LI, Z., ZHANG, Q., MAO, J. H., WEISE, A., MRASEK, K., FAN, X., ZHANG, X., LIEHR, T., LU, K. H., BALMAIN, A. & CAI, W. W. 2010. An HDAC1-binding domain within FATS bridges p21 turnover to radiation-induced tumorigenesis. *Oncogene,* 29**,** 2659-71.

LIU, J. S., KUO, S. R., BROKER, T. R. & CHOW, L. T. 1995. The functions of human papillomavirus type 11 E1, E2, and E2C proteins in cell-free DNA replication. *The Journal of biological chemistry,* 270**,** 27283-91.

LIU, Y., CHERRY, J. J., DINEEN, J. V., ANDROPHY, E. J. & BALEJA, J. D. 2009. Determinants of stability for the E6 protein of papillomavirus type 16. *J Mol Biol,* 386**,** 1123-37.

LOBATO, S., TAFURI, A., FERNANDES, P. A., CALIARI, M. V., SILVA, M. X., XAVIER, M. A. & VAGO, A. R. 2012. Minichromosome maintenance 7 protein is a reliable biological marker for human cervical progressive disease. *J Gynecol Oncol,* 23**,** 11-5.

LONGO, M. C., BERNINGER, M. S. & HARTLEY, J. L. 1990. Use of uracil DNA glycosylase to control carry-over contamination in polymerase chain reactions. *Gene,* 93**,** 125-8.

LONGWORTH, M. S. & LAIMINS, L. A. 2004. The binding of histone deacetylases and the integrity of zinc finger-like motifs of the E7 protein are essential for the life cycle of human papillomavirus type 31. *J Virol,* 78**,** 3533-41.

LORINCZ, A. T. 2011. The Promise and the Problems of Epigenetics Biomarkers in Cancer. *Expert Opin Med Diagn,* 5**,** 375-379.

LORINCZ, A. T., BRENTNALL, A. R., VASILJEVIC, N., SCIBIOR-BENTKOWSKA, D., CASTANON, A., FIANDER, A., POWELL, N., TRISTRAM, A., CUZICK, J. & SASIENI, P. 2013. HPV16 L1 and L2 DNA methylation predicts high grade cervical intraepithelial neoplasia in women with mildly abnormal cervical cytology. *International Journal of Cancer*.

LUESLEY, D. & LEESON, S. (eds.) 2010. *Colposcopy and Programme Management. Guidelines for the NHS Cervical Screening Programme. NHSCSP Publication no. 20. 2nd Edition,* Sheffield: NHS Cancer Screening Programmes.

LUFT, F., KLAES, R., NEES, M., DÜRST, M., HEILMANN, V., MELSHEIMER, P. & VON KNEBEL DOEBERITZ, M. 2001. Detection of integrated papillomavirus sequences by ligation-mediated PCR (DIPS-PCR) and molecular characterization in cervical cancer cells. *International Journal of Cancer,* 92**,** 9-17.

MA, C., QUESNELLE, K. M., SPARANO, A., RAO, S., PARK, M. S., COHEN, M. A., WANG, Y., SAMANTA, M., KUMAR, M. S., AZIZ, M. U., NAYLOR, T. L., WEBER, B. L., FAKHARZADEH, S. S., WEINSTEIN, G. S., VACHANI, A., FELDMAN, M. D. & BROSE, M. S. 2009. Characterization CSMD1 in a large set of primary lung, head and neck, breast and skin cancer tissues. *Cancer Biol Ther,* 8**,** 907-16.

MA, K., QIU, L., MRASEK, K., ZHANG, J., LIEHR, T., QUINTANA, L. G. & LI, Z. 2012. Common fragile sites: genomic hotspots of DNA damage and carcinogenesis. *Int J Mol Sci,* 13**,** 11974-99.

MARKOWITZ, L. E., TSU, V., DEEKS, S. L., CUBIE, H., WANG, S. A., VICARI, A. S. & BROTHERTON, J. M. 2012. Human papillomavirus vaccine introduction--the first five years. *Vaccine,* 30 Suppl 5**,** F139-48.

MARTIN-HIRSCH, P. L., KOLIOPOULOS, G. & PARASKEVAIDIS, E. 2002. Is it now time to evaluate the true accuracy of cervical cytology screening? A review of the literature. *Eur J Gynaecol Oncol,* 23**,** 363-5.

MASSAD, L. S. & COLLINS, Y. C. 2003. Strength of correlations between colposcopic impression and biopsy histology. *Gynecologic Oncology,* 89**,** 424-8.

MATOVINA, M., SABOL, I., GRUBISIC, G., GASPEROV, N. M. & GRCE, M. 2009. Identification of human papillomavirus type 16 integration sites in high-grade precancerous cervical lesions. *Gynecologic Oncology,* 113**,** 120-127.

MCBRIDE, A. A. 2008. Replication and partitioning of papillomavirus genomes. *Adv Virus Res,* 72**,** 155-205.

MCCREDIE, M. R., PAUL, C., SHARPLES, K. J., BARANYAI, J., MEDLEY, G., SKEGG, D. C. & JONES, R. W. 2010. Consequences in women of participating in a study of the natural history of cervical intraepithelial neoplasia 3. *Aust N Z J Obstet Gynaecol,* 50**,** 363-70.

MCCREDIE, M. R., SHARPLES, K. J., PAUL, C., BARANYAI, J., MEDLEY, G., JONES, R. W. & SKEGG, D. C. 2008. Natural history of cervical neoplasia and risk of invasive cancer in women with cervical intraepithelial neoplasia 3: a retrospective cohort study. *Lancet Oncol,* 9**,** 425-34.

MCINDOE, W. A., MCLEAN, M. R., JONES, R. W. & MULLINS, P. R. 1984. The invasive potential of carcinoma in situ of the cervix. *Obstetrics & Gynecology,* 64**,** 451-8.

MCINTOSH, P. B., MARTIN, S. R., JACKSON, D. J., KHAN, J., ISAACSON, E. R., CALDER, L., RAJ, K., GRIFFIN, H. M., WANG, Q., LASKEY, P., ECCLESTON, J. F. & DOORBAR, J. 2008. Structural analysis reveals an amyloid form of the human papillomavirus type 16 E1--E4 protein and provides a molecular basis for its accumulation. *J Virol,* 82**,** 8196-203.

MEIJER, C. J., BERKHOF, J., CASTLE, P. E., HESSELINK, A. T., FRANCO, E. L., RONCO, G., ARBYN, M., BOSCH, F. X., CUZICK, J., DILLNER, J., HEIDEMAN, D. A. & SNIJDERS, P. J. 2009. Guidelines for human papillomavirus DNA test requirements for primary cervical cancer screening in women 30 years and older. *International Journal of Cancer,* 124**,** 516-20.

MELNIKOW, J., NUOVO, J., WILLAN, A. R., CHAN, B. K. & HOWELL, L. P. 1998. Natural history of cervical squamous intraepithelial lesions: a meta-analysis. *Obstetrics & Gynecology,* 92**,** 727-35.

MELSHEIMER, P., VINOKUROVA, S., WENTZENSEN, N., BASTERT, G. & VON KNEBEL DOEBERITZ, M. 2004. DNA aneuploidy and integration of human papillomavirus type 16 e6/e7 oncogenes in intraepithelial neoplasia and invasive squamous cell carcinoma of the cervix uteri. *Clinical Cancer Research,* 10**,** 3059-63.

MEYER, N. & PENN, L. Z. 2008. Reflecting on 25 years with MYC. *Nature reviews. Cancer,* 8**,** 976-90.

MIDDLETON, K., PEH, W., SOUTHERN, S., GRIFFIN, H., SOTLAR, K., NAKAHARA, T., EL-SHERIF, A., MORRIS, L., SETH, R., HIBMA, M., JENKINS, D., LAMBERT, P., COLEMAN, N. & DOORBAR, J. 2003. Organization of human papillomavirus productive cycle during neoplastic progression provides a basis for selection of diagnostic markers. *Journal of virology,* 77**,** 10186-201.

MILLIGAN, S. G., VEERAPRADITSIN, T., AHAMET, B., MOLE, S. & GRAHAM, S. V. 2007. Analysis of novel human papillomavirus type 16 late mRNAs in differentiated W12 cervical epithelial cells. *Virology,* 360**,** 172-81.

MIRABELLO, L., SCHIFFMAN, M., GHOSH, A., RODRIGUEZ, A. C., VASILJEVIC, N., WENTZENSEN, N., HERRERO, R., HILDESHEIM, A., WACHOLDER, S., SCIBIOR-BENTKOWSKA, D., BURK, R. D. & LORINCZ, A. T. 2013. Elevated methylation of HPV16 DNA is associated with the development of high grade cervical intraepithelial neoplasia. *International Journal of Cancer,* 132**,** 1412-22.

MODIS, Y., TRUS, B. L. & HARRISON, S. C. 2002. Atomic model of the papillomavirus capsid. *The EMBO journal,* 21**,** 4754-62.

MONNIER-BENOIT, S., MAUNY, F., RIETHMULLER, D., GUERRINI, J. S., CAPILNA, M., FELIX, S., SEILLES, E., MOUGIN, C. & PRETET, J. L. 2006. Immunohistochemical analysis of CD4+ and CD8+ T-cell subsets in high risk human papillomavirus-associated pre-malignant and malignant lesions of the uterine cervix. *Gynecologic Oncology,* 102**,** 22-31.

MOSCICKI, A. B., ELLENBERG, J. H., FARHAT, S. & XU, J. 2004. Persistence of human papillomavirus infection in HIV-infected and -uninfected adolescent girls: risk factors and differences, by phylogenetic type. *Journal of Infectious Diseases,* 190**,** 37-45.

MOSCICKI, A. B., HILLS, N., SHIBOSKI, S., POWELL, K., JAY, N., HANSON, E., MILLER, S., CLAYTON, L., FARHAT, S., BROERING, J., DARRAGH, T. & PALEFSKY, J. 2001. Risks for incident human papillomavirus infection and low-grade squamous intraepithelial lesion development in young females. *JAMA,* 285**,** 2995-3002.

MOSCUFO, N., SVERDRUP, F., BREIDING, D. E. & ANDROPHY, E. J. 1999. Two distinct regions of the BPV1 E1 replication protein interact with the activation domain of E2. *Virus research,* 65**,** 141-54.

MUKHERJEE, G., MURALIDHAR, B., BAFNA, U. D., LASKEY, R. A. & COLEMAN, N. 2007. MCM immunocytochemistry as a first line cervical screening test in developing countries: a prospective cohort study in a regional cancer centre in India. *British Journal of Cancer,* 96**,** 1107-11.

MUNGER, K. & HOWLEY, P. M. 2002. Human papillomavirus immortalization and transformation functions. *Virus research,* 89**,** 213-28.

MUNOZ, N., BOSCH, F. X., DE SANJOSE, S., TAFUR, L., IZARZUGAZA, I., GILI, M., VILADIU, P., NAVARRO, C., MARTOS, C., ASCUNCE, N. & ET AL. 1992. The causal link between human papillomavirus and invasive cervical cancer: a population-based case-control study in Colombia and Spain. *International Journal of Cancer,* 52**,** 743-9.

MUNOZ, N., KJAER, S. K., SIGURDSSON, K., IVERSEN, O. E., HERNANDEZ-AVILA, M., WHEELER, C. M., PEREZ, G., BROWN, D. R., KOUTSKY, L. A., TAY, E. H., GARCIA, P. J., AULT, K. A., GARLAND, S. M., LEODOLTER, S., OLSSON, S. E., TANG, G. W., FERRIS, D. G., PAAVONEN, J., STEBEN, M., BOSCH, F. X., DILLNER, J., HUH, W. K., JOURA, E. A., KURMAN, R. J., MAJEWSKI, S., MYERS, E. R., VILLA, L. L., TADDEO, F. J., ROBERTS, C., TADESSE, A., BRYAN, J. T., LUPINACCI, L. C., GIACOLETTI, K. E., SINGS, H. L., JAMES, M. K., HESLEY, T. M., BARR, E. & HAUPT, R. M. 2010. Impact of human papillomavirus (HPV)-6/11/16/18 vaccine on all HPV-associated genital diseases in young women. *Journal of the National Cancer Institute,* 102**,** 325-39.

MUSAHL, C., HOLTHOFF, H. P., LESCH, R. & KNIPPERS, R. 1998. Stability of the replicative Mcm3 protein in proliferating and differentiating human cells. *Experimental cell research,* 241**,** 260-4.

NAKAHARA, T., PEH, W. L., DOORBAR, J., LEE, D. & LAMBERT, P. F. 2005. Human papillomavirus type 16 E1circumflexE4 contributes to multiple facets of the papillomavirus life cycle. *J Virol,* 79**,** 13150-65.

NAUCLER, P., RYD, W., TORNBERG, S., STRAND, A., WADELL, G., ELFGREN, K., RADBERG, T., STRANDER, B., FORSLUND, O., HANSSON, B. G., HAGMAR, B., JOHANSSON, B., RYLANDER, E. & DILLNER, J. 2009. Efficacy of HPV DNA testing with cytology triage and/or repeat HPV DNA testing in primary cervical cancer screening. *Journal of the National Cancer Institute,* 101**,** 88-99.

NAUCLER, P., RYD, W., TORNBERG, S., STRAND, A., WADELL, G., ELFGREN, K., RADBERG, T., STRANDER, B., JOHANSSON, B., FORSLUND, O., HANSSON, B. G., RYLANDER, E. & DILLNER, J. 2007. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. *New England Journal of Medicine,* 357**,** 1589-97.

NHS CERVICAL SCREENING PROGRAMMES 2010. Colposcopy and programme management: guidelines for the NHS cervical screening programme. Second ed.: NHSCSP.

NHSCSP 2013. Achievable standards, Benchmarks for reporting, and Criteria for evaluating cervical cytopathology – 3rd edition. NHSCSP Publication No. 1. Sheffield: NHS Cancer Screening Programmes.

NICE 2003. Guidance on the use of liquid-based cytology for cervical screening. Technology Appraisal 69. London: National Insitute for Health and Clinical Excellence.

NICHOLLS, P. K., KLAUNBERG, B. A., MOORE, R. A., SANTOS, E. B., PARRY, N. R., GOUGH, G. W. & STANLEY, M. A. 1999. Naturally occurring, nonregressing canine oral papillomavirus infection: host immunity, virus characterization, and experimental infection. *Virology,* 265**,** 365-74.

NICHOLLS, P. K., MOORE, P. F., ANDERSON, D. M., MOORE, R. A., PARRY, N. R., GOUGH, G. W. & STANLEY, M. A. 2001. Regression of canine oral papillomas is associated with infiltration of CD4+ and CD8+ lymphocytes. *Virology,* 283**,** 31-9.

O'CONNOR, M. J. 2000. Targeting of transcriptional cofactors by the HPV E6 protein: another tale of David and Goliath. *Trends Microbiol,* 8**,** 45-7.

O'CONNOR, M. J., STUNKEL, W., KOH, C. H., ZIMMERMANN, H. & BERNARD, H. U. 2000. The differentiation-specific factor CDP/Cut represses transcription and replication of human papillomaviruses through a conserved silencing element. *J Virol,* 74**,** 401-10.

O'CONNOR, M. J., TAN, S. H., TAN, C. H. & BERNARD, H. U. 1996. YY1 represses human papillomavirus type 16 transcription by quenching AP-1 activity. *J Virol,* 70**,** 6529-39.

OBERG, T. N., KIPP, B. R., VRANA, J. A., BARTHOLET, M. K., FALES, C. J., GARCIA, R., MCDONALD, A. N., ROSAS, B. L., HENRY, M. R. & CLAYTON, A. C. 2010. Comparison of p16INK4a and

ProEx C immunostaining on cervical ThinPrep cytology and biopsy specimens. *Diagn Cytopathol,* 38**,** 564-72.

OBERMANN, E. C., PRINCE, S. S., BARASCUD, A., GRILLI, B., HERZOG, M., KAUP, D., CATHOMAS, G., FREY TIRRI, B., ZLOBEC, I., WIGHT, E. & BUBENDORF, L. 2013. Prediction of outcome in patients with low-grade squamous intraepithelial lesions by fluorescence in situ hybridization analysis of human papillomavirus, TERC, and MYC. *Cancer Cytopathol*.

OH, S. T., KYO, S. & LAIMINS, L. A. 2001. Telomerase activation by human papillomavirus type 16 E6 protein: induction of human telomerase reverse transcriptase expression through Myc and GC-rich Sp1 binding sites. *J Virol,* 75**,** 5559-66.

ORIGONI, M., CARMINATI, G., ROLLA, S., CLEMENTI, M., SIDERI, M., SANDRI, M. T. & CANDIANI, M. 2012. Human papillomavirus viral load expressed as relative light units (RLU) correlates with the presence and grade of preneoplastic lesions of the uterine cervix in atypical squamous cells of undetermined significance (ASCUS) cytology. *Eur J Clin Microbiol Infect Dis,* 31**,** 2401-6.

OSTER, A. G. 1993. Natural history of cervical intraepithelial neoplasia: a critical review. *International Journal of Gynecological Pathology,* 12**,** 186-92.

OVERMEER, R. M., LOUWERS, J. A., MEIJER, C. J., VAN KEMENADE, F. J., HESSELINK, A. T., DAALMEIJER, N. F., WILTING, S. M., HEIDEMAN, D. A., VERHEIJEN, R. H., ZAAL, A., VAN BAAL, W. M., BERKHOF, J., SNIJDERS, P. J. & STEENBERGEN, R. D. 2011. Combined CADM1 and MAL promoter methylation analysis to detect (pre-)malignant cervical lesions in high-risk HPV-positive women. *International Journal of Cancer,* 129**,** 2218-25.

OZBUN, M. A. & MEYERS, C. 1998. Human papillomavirus type 31b E1 and E2 transcript expression correlates with vegetative viral genome amplification. *Virology,* 248**,** 218-30.

PAAVONEN, J., JENKINS, D., BOSCH, F. X., NAUD, P., SALMERON, J., WHEELER, C. M., CHOW, S. N., APTER, D. L., KITCHENER, H. C., CASTELLSAGUE, X., DE CARVALHO, N. S., SKINNER, S. R., HARPER, D. M., HEDRICK, J. A., JAISAMRARN, U., LIMSON, G. A., DIONNE, M., QUINT, W., SPIESSENS, B., PEETERS, P., STRUYF, F., WIETING, S. L., LEHTINEN, M. O. & DUBIN, G. 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.

PAAVONEN, J., NAUD, P., SALMERON, J., WHEELER, C. M., CHOW, S. N., APTER, D., KITCHENER, H., CASTELLSAGUE, X., TEIXEIRA, J. C., SKINNER, S. R., HEDRICK, J., JAISAMRARN, U., LIMSON, G., GARLAND, S., SZAREWSKI, A., ROMANOWSKI, B., AOKI, F. Y., SCHWARZ, T. F., POPPE, W. A., BOSCH, F. X., JENKINS, D., HARDT, K., ZAHAF, T., DESCAMPS, D., STRUYF, F., LEHTINEN, M. & DUBIN, G. 2009. Efficacy of human papillomavirus (HPV)-16/18 AS04-adjuvanted vaccine against cervical infection and precancer caused by oncogenic HPV types (PATRICIA): final analysis of a double-blind, randomised study in young women. *Lancet,* 374**,** 301-14.

PALLANT, J. 2010. *SPSS Survival Manual – 4th edition,* Maidenhead, UK, McGraw-Hill.

PAPANICOLAOU, G. N. 1946. Diagnostic value of exfoliated cells from cancerous tissues. *J Am Med Assoc,* 131**,** 372-8.

PERES, J. 2011. For cancers caused by HPV, two vaccines were just the beginning. *Journal of the National Cancer Institute,* 103**,** 360-2.

PETER, M., ROSTY, C., COUTURIER, J., RADVANYI, F., TESHIMA, H. & SASTRE-GARAU, X. 2006. MYC activation associated with the integration of HPV DNA at the MYC locus in genital tumors. *Oncogene,* 25**,** 5985-93.

PETO, J., GILHAM, C., DEACON, J., TAYLOR, C., EVANS, C., BINNS, W., HAYWOOD, M., ELANKO, N., COLEMAN, D., YULE, R. & DESAI, M. 2004a. Cervical HPV infection and neoplasia in a large population-based prospective study: the Manchester cohort. *British Journal of Cancer,* 91**,** 942-53.

PETO, J., GILHAM, C., FLETCHER, O. & MATTHEWS, F. E. 2004b. The cervical cancer epidemic that screening has prevented in the UK. *Lancet,* 364**,** 249-56.

PETT, M. & COLEMAN, N. 2007. Integration of high-risk human papillomavirus: a key event in cervical carcinogenesis? *The Journal of pathology,* 212**,** 356-67.

PFISTER, H. 2003. Chapter 8: Human papillomavirus and skin cancer. *Journal of the National Cancer Institute. Monographs***,** 52-6.

PIM, D. & BANKS, L. 2010. Interaction of viral oncoproteins with cellular target molecules: infection with high-risk vs low-risk human papillomaviruses. *APMIS,* 118**,** 471-93.

PIRAMI, L., GIACHE, V. & BECCIOLINI, A. 1997. Analysis of HPV16, 18, 31, and 35 DNA in pre-invasive and invasive lesions of the uterine cervix. *Journal of clinical pathology,* 50**,** 600-4.

POLJAK, M., CUZICK, J., KOCJAN, B. J., IFTNER, T., DILLNER, J. & ARBYN, M. 2012. Nucleic acid tests for the detection of alpha human papillomaviruses. *Vaccine,* 30 Suppl 5**,** F100-6.

PORTELA, A. & ESTELLER, M. 2010. Epigenetic modifications and human disease. *Nat Biotechnol,* 28**,** 1057-68.

POWELL, N., SMITH, K. & FIANDER, A. 2006. Recovery of human papillomavirus nucleic acids from liquid-based cytology media. *Journal of Virological Methods,* 137**,** 58-62.

PRINSEN, C. F., FLES, R., WIJNEN-DUBBERS, C. W., DE VALK, H. A., KLAASSEN, C. H., MRAVUNAC, M., HORREVORTS, A. M. & THUNNISSEN, F. B. 2007. Baseline human papillomavirus status of women with abnormal smears in cervical screening: a 5-year follow-up study in The Netherlands. *Bjog,* 114**,** 951-7.

PUNTMANN, V. O. 2009. How-to guide on biomarkers: biomarker definitions, validation and applications with examples from cardiovascular disease. *Postgrad Med J,* 85**,** 538-45.

QUINN, M., BABB, P., JONES, J. & 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.

RAYBOULD, R., FIANDER, A. & HIBBITTS, S. 2011. Human Papillomavirus Integration and its Role in Cervical Malignant Progression. *The Open Clinical Cancer Journal,* 5**,** 1–7.

REBOLJ, M., BONDE, J., NJOR, S. H. & LYNGE, E. 2011. Human papillomavirus testing in primary cervical screening and the cut-off level for hybrid capture 2 tests: systematic review. *BMJ,* 342.

RICHARDSON, H., KELSALL, G., TELLIER, P., VOYER, H., ABRAHAMOWICZ, M., FERENCZY, A., COUTLEE, F. & FRANCO, E. L. 2003. The natural history of type-specific human papillomavirus infections in female university students. *Cancer Epidemiology, Biomarkers & Prevention,* 12**,** 485-90.

RICHART, R. M. 1973. Cervical intraepithelial neoplasia. *Pathol Annu,* 8**,** 301-28.

RIPKA, S., RIEDEL, J., NEESSE, A., GRIESMANN, H., BUCHHOLZ, M., ELLENRIEDER, V., MOELLER, F., BARTH, P., GRESS, T. M. & MICHL, P. 2010. Glutamate receptor GRIA3--target of CUX1 and mediator of tumor progression in pancreatic cancer. *Neoplasia,* 12**,** 659-67.

ROBERTSON, K. D. 2005. DNA methylation and human disease. *Nat Rev Genet,* 6**,** 597-610.

ROBERTSON, K. D. & JONES, P. A. 2000. DNA methylation: past, present and future directions. *Carcinogenesis,* 21**,** 461-7.

ROBLES, S. C., WHITE, F. & PERUGA, A. 1996. Trends in cervical cancer mortality in the Americas. *Bull Pan Am Health Organ,* 30**,** 290-301.

ROMANCZUK, H. & HOWLEY, P. M. 1992. Disruption of either the E1 or the E2 regulatory gene of human papillomavirus type 16 increases viral immortalization capacity. *Proceedings of the National Academy of Sciences of the United States of America,* 89**,** 3159-63.

ROMANCZUK, H., THIERRY, F. & HOWLEY, P. M. 1990. Mutational analysis of cis elements involved in E2 modulation of human papillomavirus type 16 P97 and type 18 P105 promoters. *J Virol,* 64**,** 2849-59.

RONAGHI, M. 2001. Pyrosequencing sheds light on DNA sequencing. *Genome research,* 11**,** 3-11.

RONCO, G., CUZICK, J., PIEROTTI, P., CARIAGGI, M. P., DALLA PALMA, P., NALDONI, C., GHIRINGHELLO, B., GIORGI-ROSSI, P., MINUCCI, D., PARISIO, F., POJER, A., SCHIBONI, M. L., SINTONI, C., ZORZI, M., SEGNAN, N. & CONFORTINI, M. 2007. Accuracy of liquid

based versus conventional cytology: overall results of new technologies for cervical cancer screening: randomised controlled trial. *BMJ,* 335**,** 28.

RONCO, G., SEGNAN, N., GIORGI-ROSSI, P., ZAPPA, M., CASADEI, G. P., CAROZZI, F., DALLA PALMA, P., DEL MISTRO, A., FOLICALDI, S., GILLIO-TOS, A., NARDO, G., NALDONI, C., SCHINCAGLIA, P., ZORZI, M., CONFORTINI, M. & CUZICK, J. 2006. Human papillomavirus testing and liquid-based cytology: results at recruitment from the new technologies for cervical cancer randomized controlled trial. *Journal of the National Cancer Institute,* 98**,** 765-74.

ROSENBERGER, S., DE-CASTRO ARCE, J., LANGBEIN, L., STEENBERGEN, R. D. & ROSL, F. 2010. Alternative splicing of human papillomavirus type-16 E6/E6* early mRNA is coupled to EGF signaling via Erk1/2 activation. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 7006-11.

RUDLINGER, R., SMITH, I. W., BUNNEY, M. H. & HUNTER, J. A. 1986. Human papillomavirus infections in a group of renal transplant recipients. *British Journal of Dermatology,* 115**,** 681-92.

SANO, T., OYAMA, T., KASHIWABARA, K., FUKUDA, T. & NAKAJIMA, T. 1998. Expression status of p16 protein is associated with human papillomavirus oncogenic potential in cervical and genital lesions. *Am J Pathol,* 153**,** 1741-8.

SARAFI, T. R. & MCBRIDE, A. A. 1995. Domains of the BPV-1 E1 replication protein required for origin-specific DNA binding and interaction with the E2 transactivator. *Virology,* 211**,** 385-96.

SASIENI, P. & ADAMS, J. 1999. Effect of screening on cervical cancer mortality in England and Wales: analysis of trends with an age period cohort model. *BMJ,* 318**,** 1244-5.

SASIENI, P., ADAMS, J. & CUZICK, J. 2003. Benefit of cervical screening at different ages: evidence from the UK audit of screening histories. *British Journal of Cancer,* 89**,** 88-93.

SASIENI, P., CASTANON, A. & CUZICK, J. 2009a. Effectiveness of cervical screening with age: population based case-control study of prospectively recorded data. *BMJ,* 339**,** b2968.

SASIENI, P., CASTANON, A. & PARKIN, D. M. 2009b. How many cervical cancers are prevented by treatment of screen-detected disease in young women? *International Journal of Cancer,* 124**,** 461-4.

SASIENI, P. D., CUZICK, J. & LYNCH-FARMERY, E. 1996. Estimating the efficacy of screening by auditing smear histories of women with and without cervical cancer. The National Co-ordinating Network for Cervical Screening Working Group. *British Journal of Cancer,* 73**,** 1001-5.

SCHEFFNER, M., WERNESS, B. A., HUIBREGTSE, J. M., LEVINE, A. J. & HOWLEY, P. M. 1990. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell,* 63**,** 1129-36.

SCHELHAAS, M., SHAH, B., HOLZER, M., BLATTMANN, P., KUHLING, L., DAY, P. M., SCHILLER, J. T. & HELENIUS, A. 2012. Entry of human papillomavirus type 16 by actin-dependent, clathrin- and lipid raft-independent endocytosis. *PLoS pathogens,* 8**,** e1002657.

SCHIFFMAN, M. & CASTLE, P. E. 2003. Human papillomavirus: epidemiology and public health. *Archives of pathology & laboratory medicine,* 127**,** 930-4.

SCHIFFMAN, M. H. & BRINTON, L. A. 1995. The epidemiology of cervical carcinogenesis. *Cancer,* 76**,** 1888-901.

SCHILLER, J. T., DAY, P. M. & KINES, R. C. 2010. Current understanding of the mechanism of HPV infection. *Gynecologic Oncology,* 118**,** S12-7.

SCHMIDT, D., BERGERON, C., DENTON, K. J. & RIDDER, R. 2011. p16/ki-67 dual-stain cytology in the triage of ASCUS and LSIL papanicolaou cytology: results from the European equivocal or mildly abnormal Papanicolaou cytology study. *Cancer Cytopathol,* 119**,** 158-66.

SCHMITZ, M., DRIESCH, C., BEER-GRONDKE, K., JANSEN, L., RUNNEBAUM, I. B. & DURST, M. 2012. Loss of gene function as a consequence of human papillomavirus DNA integration. *International Journal of Cancer,* 131**,** E593-602.

SCHNEIDER, V. 2003. Symposium part 2: Should the Bethesda System terminology be used in diagnostic surgical pathology?: Counterpoint. *International journal of gynecological pathology : official journal of the International Society of Gynecological Pathologists,* 22**,** 13-17.

SCOTTISH GOVERNMENT. 2012. *Cervical screening* [Online]. Available: http://www.scotland.gov.uk/News/Releases/2012/12/cervical-screening11122012 [Accessed 11/02/2013].

SEDMAN, J. & STENLUND, A. 1998. The papillomavirus E1 protein forms a DNA-dependent hexameric complex with ATPase and DNA helicase activities. *J Virol,* 72**,** 6893-7.

SELINKA, H. C., FLORIN, L., PATEL, H. D., FREITAG, K., SCHMIDTKE, M., MAKAROV, V. A. & SAPP, M. 2007. Inhibition of transfer to secondary receptors by heparan sulfate-binding drug or antibody induces noninfectious uptake of human papillomavirus. *J Virol,* 81**,** 10970-80.

SELINKA, H. C., GIROGLOU, T. & SAPP, M. 2002. Analysis of the infectious entry pathway of human papillomavirus type 33 pseudovirions. *Virology,* 299**,** 279-287.

SELLORS, J. W., KARWALAJTYS, T. L., KACZOROWSKI, J., MAHONY, J. B., LYTWYN, A., CHONG, S., SPARROW, J. & LORINCZ, A. 2003. Incidence, clearance and predictors of human papillomavirus infection in women. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne,* 168**,** 421-5.

SEN, E., ALAM, S. & MEYERS, C. 2004. Genetic and biochemical analysis of cis regulatory elements within the keratinocyte enhancer region of the human papillomavirus type 31 upstream regulatory region during different stages of the viral life cycle. *J Virol,* 78**,** 612-29.

SERRANO, M., HANNON, G. J. & BEACH, D. 1993. A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. *Nature,* 366**,** 704-7.

SIDDIQUI, M. T., HORNAMAN, K., COHEN, C. & NASSAR, A. 2008. ProEx C immunocytochemistry and high-risk human papillomavirus DNA testing in papanicolaou tests with atypical squamous cell (ASC-US) cytology: correlation study with histologic biopsy. *Archives of pathology & laboratory medicine,* 132**,** 1648-52.

SINGER, A. & JORDAN, J. 2006. The functional anatomy of the cervix, the cervical epithelium and the stroma. *In:* JORDAN, J. & SINGER, A. (eds.) *The Cervix - Second Edition.* Oxford, UK: Blackwell Publishing Ltd.

SKEHAN, M., SOUTTER, W. P., LIM, K., KRAUSZ, T. & PRYSE-DAVIES, J. 1990. Reliability of colposcopy and directed punch biopsy. *Br J Obstet Gynaecol,* 97**,** 811-6.

SMALLEY, M. J., IRAVANI, M., LEAO, M., GRIGORIADIS, A., KENDRICK, H., DEXTER, T., FENWICK, K., REGAN, J. L., BRITT, K., MCDONALD, S., LORD, C. J., MACKAY, A. & ASHWORTH, A. 2007. Regulator of G-protein signalling 2 mRNA is differentially expressed in mammary epithelial subpopulations and over-expressed in the majority of breast cancers. *Breast cancer research : BCR,* 9**,** R85.

SMITH, J. H. 2012. ABC3 Part I: a review of the guidelines for terminology, classification and management of cervical cytology in England. *Cytopathology : official journal of the British Society for Clinical Cytology,* 23**,** 353-9.

SMITH, J. H. & DESAI, M. S. 2007. European guidelines on cervical screening. *Cytopathology : official journal of the British Society for Clinical Cytology,* 18**,** 211-2.

SMITH, J. S., GREEN, J., BERRINGTON DE GONZALEZ, A., APPLEBY, P., PETO, J., PLUMMER, M., FRANCESCHI, S. & BERAL, V. 2003. Cervical cancer and use of hormonal contraceptives: a systematic review. *Lancet,* 361**,** 1159-67.

SOLOMON, D., DAVEY, D., KURMAN, R., MORIARTY, A., O'CONNOR, D., PREY, M., RAAB, S., SHERMAN, M., WILBUR, D., WRIGHT, T., JR. & YOUNG, N. 2002. The 2001 Bethesda System: terminology for reporting results of cervical cytology. *JAMA,* 287**,** 2114-9.

SOLOMON, D., PAPILLO, J. L. & DAVEY, D. D. 2009. Statement on human papillomavirus DNA test utilization. *Archives of pathology & laboratory medicine,* 133**,** 1276-7.

SPINK, K. M. & LAIMINS, L. A. 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, Supplement 1**,** S16-S22.

STANLEY, M. 2010. Pathology and epidemiology of HPV infection in females. *Gynecologic Oncology,* 117**,** S5-10.

STANLEY, M. A. 2012. Epithelial cell responses to infection with human papillomavirus. *Clin Microbiol Rev,* 25**,** 215-22.

STEENBERGEN, R. D., KRAMER, D., BRAAKHUIS, B. J., STERN, P. L., VERHEIJEN, R. H., MEIJER, C. J. & SNIJDERS, P. J. 2004. TSLC1 gene silencing in cervical cancer cell lines and cervical neoplasia. *Journal of the National Cancer Institute,* 96**,** 294-305.

STEGER, G. & CORBACH, S. 1997. Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein. *J Virol,* 71**,** 50-8.

STENLUND, A. 2003. Initiation of DNA replication: lessons from viral initiator proteins. *Nat Rev Mol Cell Biol,* 4**,** 777-85.

STRAIGHT, S. W., HINKLE, P. M., JEWERS, R. J. & MCCANCE, D. J. 1993. The E5 oncoprotein of human papillomavirus type 16 transforms fibroblasts and effects the downregulation of the epidermal growth factor receptor in keratinocytes. *J Virol,* 67**,** 4521-32.

SUN, C., REIMERS, L. L. & BURK, R. D. 2011. Methylation of HPV16 genome CpG sites is associated with cervix precancer and cancer. *Gynecologic Oncology,* 121**,** 59-63.

SUTCLIFFE, J. G. 1979. Complete nucleotide sequence of the Escherichia coli plasmid pBR322. *Cold Spring Harb Symp Quant Biol,* 43 Pt 1**,** 77-90.

SUZICH, J. A., GHIM, S. J., PALMER-HILL, F. J., WHITE, W. I., TAMURA, J. K., BELL, J. A., NEWSOME, J. A., JENSON, A. B. & SCHLEGEL, R. 1995. Systemic immunization with papillomavirus L1 protein completely prevents the development of viral mucosal papillomas. *Proceedings of the National Academy of Sciences of the United States of America,* 92**,** 11553-7.

SZAREWSKI, A., AMBROISINE, L., CADMAN, L., AUSTIN, J., HO, L., TERRY, G., LIDDLE, S., DINA, R., MCCARTHY, J., BUCKLEY, H., BERGERON, C., SOUTTER, P., LYONS, D. & CUZICK, J. 2008. Comparison of predictors for high-grade cervical intraepithelial neoplasia in women with abnormal smears. *Cancer Epidemiology, Biomarkers & Prevention,* 17**,** 3033-42.

SZAREWSKI, A. & CUZICK, J. 1998. Smoking and cervical neoplasia: a review of the evidence. *J Epidemiol Biostat,* 3**,** 229–256.

SZAREWSKI, A., POPPE, W. A., SKINNER, S. R., WHEELER, C. M., PAAVONEN, J., NAUD, P., SALMERON, J., CHOW, S. N., APTER, D., KITCHENER, H., CASTELLSAGUE, X., TEIXEIRA, J. C., HEDRICK, J., JAISAMRARN, U., LIMSON, G., GARLAND, S., ROMANOWSKI, B., AOKI, F. Y., SCHWARZ, T. F., BOSCH, F. X., HARPER, D. M., HARDT, K., ZAHAF, T., DESCAMPS, D., STRUYF, F., LEHTINEN, M. & DUBIN, G. 2012. Efficacy of the human papillomavirus (HPV)-16/18 AS04-adjuvanted vaccine in women aged 15-25 years with and without serological evidence of previous exposure to HPV-16/18. *International Journal of Cancer,* 131**,** 106-16.

TAKAI, D. & JONES, P. A. 2002. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proceedings of the National Academy of Sciences of the United States of America,* 99**,** 3740-5.

TAMBOURET, R. H., MISDRAJI, J. & WILBUR, D. C. 2008. Longitudinal clinical evaluation of a novel antibody cocktail for detection of high-grade squamous intraepithelial lesions on cervical cytology specimens. *Archives of pathology & laboratory medicine,* 132**,** 918-25.

TAN, S. H., LEONG, L. E., WALKER, P. A. & BERNARD, H. U. 1994. The human papillomavirus type 16 E2 transcription factor binds with low cooperativity to two flanking sites and represses the E6 promoter through displacement of Sp1 and TFIID. *J Virol,* 68**,** 6411-20.

THERMO SCIENTIFIC 2011. T042–Technical Bulletin NanoDrop Spectrophotometers. Assessment of Nucleic Acid Purity. *Thermo Fisher Scientific.* Wilmington, Delaware, USA.

THERMO SCIENTIFIC 2012. T123–Technical Bulletin. Interpretation of Nucleic Acid 260/280 Ratios. *Thermo Fisher Scientific.* Wilmington, Delaware, USA.

THIERRY, F. 2009. Transcriptional regulation of the papillomavirus oncogenes by cellular and viral transcription factors in cervical carcinoma. *Virology,* 384**,** 375-9.

THOMAS, M. & BANKS, L. 1999. Human papillomavirus (HPV) E6 interactions with Bak are conserved amongst E6 proteins from high and low risk HPV types. *The Journal of general virology,* 80 (Pt 6)**,** 1513-7.

THORLAND, E. C., MYERS, S. L., GOSTOUT, B. S. & SMITH, D. I. 2003. Common fragile sites are preferential targets for HPV16 integrations in cervical tumors. *Oncogene,* 22**,** 1225-37.

THORLAND, E. C., MYERS, S. L., PERSING, D. H., SARKAR, G., MCGOVERN, R. M., GOSTOUT, B. S. & SMITH, D. I. 2000. Human papillomavirus type 16 integrations in cervical tumors frequently occur in common fragile sites. *Cancer Res,* 60**,** 5916-21.

TONON, S. A., PICCONI, M. A., BOS, P. D., ZINOVICH, J. B., GALUPPO, J., ALONIO, L. V. & TEYSSIE, A. R. 2001. Physical status of the E2 human papilloma virus 16 viral gene in cervical preneoplastic and neoplastic lesions. *J Clin Virol,* 21**,** 129-34.

TSOUMPOU, I., ARBYN, M., KYRGIOU, M., WENTZENSEN, N., KOLIOPOULOS, G., MARTIN-HIRSCH, P., MALAMOU-MITSI, V. & PARASKEVAIDIS, E. 2009. p16(INK4a) immunostaining in cytological and histological specimens from the uterine cervix: a systematic review and meta-analysis. *Cancer Treat Rev,* 35**,** 210-20.

UK NATIONAL SCREENING COMMITTEE 2012. UK National Screening Committee recommendations update – Autumn 2012. London.

VACH, W. 2005. The dependence of Cohen's kappa on the prevalence does not matter. *J Clin Epidemiol,* 58**,** 655-61.

VALENCIA, C., BONILLA-DELGADO, J., OKTABA, K., OCADIZ-DELGADO, R., GARIGLIO, P. & COVARRUBIAS, L. 2008. Human papillomavirus E6/E7 oncogenes promote mouse ear regeneration by increasing the rate of wound re-epithelization and epidermal growth. *J Invest Dermatol,* 128**,** 2894-903.

VELDMAN, T., HORIKAWA, I., BARRETT, J. C. & SCHLEGEL, R. 2001. Transcriptional activation of the telomerase hTERT gene by human papillomavirus type 16 E6 oncoprotein. *J Virol,* 75**,** 4467-72.

VINOKUROVA, S. & VON KNEBEL DOEBERITZ, M. 2011. Differential methylation of the HPV 16 upstream regulatory region during epithelial differentiation and neoplastic transformation. *PLoS ONE,* 6**,** e24451.

WALBOOMERS, J. M., JACOBS, M. V., MANOS, M. M., BOSCH, F. X., KUMMER, J. A., SHAH, K. V., SNIJDERS, P. J., PETO, J., MEIJER, C. J. & MUNOZ, N. 1999. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *The Journal of pathology,* 189**,** 12-9.

WANAJO, A., SASAKI, A., NAGASAKI, H., SHIMADA, S., OTSUBO, T., OWAKI, S., SHIMIZU, Y., EISHI, Y., KOJIMA, K., NAKAJIMA, Y., KAWANO, T., YUASA, Y. & AKIYAMA, Y. 2008. Methylation of the calcium channel-related gene, CACNA2D3, is frequent and a poor prognostic factor in gastric cancer. *Gastroenterology,* 135**,** 580-90.

WANG, Q., GRIFFIN, H., SOUTHERN, S., JACKSON, D., MARTIN, A., MCINTOSH, P., DAVY, C., MASTERSON, P. J., WALKER, P. A., LASKEY, P., OMARY, M. B. & DOORBAR, J. 2004. Functional analysis of the human papillomavirus type 16 E1=E4 protein provides a mechanism for in vivo and in vitro keratin filament reorganization. *J Virol,* 78**,** 821-33.

WANG, Y., GUAN, X., FOK, K. L., LI, S., ZHANG, X., MIAO, S., ZONG, S., KOIDE, S. S., CHAN, H. C. & WANG, L. 2008. A novel member of the Rhomboid family, RHBDD1, regulates BIK-mediated apoptosis. *Cell Mol Life Sci,* 65**,** 3822-9.

WELSH GOVERNMENT. 2013. *Changes to cervical screening in Wales anounced* [Online]. Available: http://www.wales.nhs.uk/news/25662 [Accessed 11/02/2013].

WENTZENSEN, N., VINOKUROVA, S. & VON KNEBEL DOEBERITZ, M. 2004. Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract. *Cancer Res,* 64**,** 3878-84.

WHITEHEAD, C., JUDGE, K., ROSENTHAL, D., STEPHENSON, B., TAYLOR, A. & MALINOWSKI, D. 2012. Increased Cervical Disease Detection using SurePath Plus which Combines

Biomarker Immunocytochemistry within a LBC Pap. *Journal of the American Society of Cytopathology,* 1**,** S46-S47.

WHITEHEAD, C. M., HE, Q., ALLEN, L., WILLSE, K., NELSON, R., HUDSON, R., DOOBAY, H., PURNELL, D., SIMPSON, L., TAYLOR, A., FISCHER, T. J. & MALINOWSKI, D. 2011. Improving Identification of CIN2+ Disease by Combining MCM2 and MCM7 Biomarker Over-Expression with a Pap Counterstain on a Standard Cervical LBC Specimen. *United States & Canadian Academy of Pathology Annual Meeting.* San Antonio, Texas.

WILGENBURG, B. J., BUDGEON, L. R., LANG, C. M., GRIFFITH, J. W. & CHRISTENSEN, N. D. 2005. Characterization of immune responses during regression of rabbit oral papillomavirus infections. *Comp Med,* 55**,** 431-9.

WILKE, C. M., HALL, B. K., HOGE, A., PARADEE, W., SMITH, D. I. & GLOVER, T. W. 1996. FRA3B extends over a broad region and contains a spontaneous HPV16 integration site: direct evidence for the coincidence of viral integration sites and fragile sites. *Hum Mol Genet,* 5**,** 187-95.

WILLIAMS, G. H., ROMANOWSKI, P., MORRIS, L., MADINE, M., MILLS, A. D., STOEBER, K., MARR, J., LASKEY, R. A. & COLEMAN, N. 1998. Improved cervical smear assessment using antibodies against proteins that regulate DNA replication. *Proceedings of the National Academy of Sciences of the United States of America,* 95**,** 14932-7.

WILSON, J. M. G. & JUNGNER, G. 1968. *Principles and practice of screening for disease* [Online]. Geneva: WHO. Available: http://www.who.int/bulletin/volumes/86/4/07-050112BP.pdf [Accessed 10/02/2013].

WILSON, V. G., WEST, M., WOYTEK, K. & RANGASAMY, D. 2002. Papillomavirus E1 proteins: form, function, and features. *Virus genes,* 24**,** 275-90.

WINER, R. L., LEE, S. K., HUGHES, J. P., ADAM, D. E., KIVIAT, N. B. & KOUTSKY, L. A. 2003. Genital human papillomavirus infection: incidence and risk factors in a cohort of female university students. *Am J Epidemiol,* 157**,** 218-26.

WOODMAN, C. B., COLLINS, S., WINTER, H., BAILEY, A., ELLIS, J., PRIOR, P., YATES, M., ROLLASON, T. P. & YOUNG, L. S. 2001. Natural history of cervical human papillomavirus infection in young women: a longitudinal cohort study. *Lancet,* 357**,** 1831-6.

WRIGHT, T. C., JR., STOLER, M. H., SHARMA, A., ZHANG, G., BEHRENS, C. & WRIGHT, T. L. 2011. Evaluation of HPV-16 and HPV-18 genotyping for the triage of women with high-risk HPV+ cytology-negative results. *Am J Clin Pathol,* 136**,** 578-86.

YANG, B. H., BRAY, F. I., PARKIN, D. M., SELLORS, J. W. & ZHANG, Z. F. 2004. Cervical cancer as a priority for prevention in different world regions: an evaluation using years of life lost. *International Journal of Cancer,* 109**,** 418-24.

YI, J. M., DHIR, M., VAN NESTE, L., DOWNING, S. R., JESCHKE, J., GLÖCKNER, S. C., DE FREITAS CALMON, M., HOOKER, C. M., FUNES, J. M., BOSHOFF, C., SMITS, K. M., VAN ENGELAND, M., WEIJENBERG, M. P., IACOBUZIO-DONAHUE, C. A., HERMAN, J. G., SCHUEBEL, K. E., BAYLIN, S. B. & AHUJA, N. 2011. Genomic and Epigenomic Integration Identifies a Prognostic Signature in Colon Cancer. *Clinical Cancer Research,* 17**,** 1535-1545.

YOU, K., LIANG, X., QIN, F., GUO, Y. & GENG, L. 2007. High-risk human papillomavirus DNA testing and high-grade cervical intraepithelial lesions. *Aust N Z J Obstet Gynaecol,* 47**,** 141-4.

YU, T., FERBER, M. J., CHEUNG, T. H., CHUNG, T. K., WONG, Y. F. & SMITH, D. I. 2005. The role of viral integration in the development of cervical cancer. *Cancer genetics and cytogenetics,* 158**,** 27-34.

YUGAWA, T. & KIYONO, T. 2009. Molecular mechanisms of cervical carcinogenesis by high-risk human papillomaviruses: novel functions of E6 and E7 oncoproteins. *Rev Med Virol,* 19**,** 97-113.

ZANIER, K., OULD M'HAMED OULD SIDI, A., BOULADE-LADAME, C., RYBIN, V., CHAPPELLE, A., ATKINSON, A., KIEFFER, B. & TRAVE, G. 2012. Solution structure analysis of the HPV16 E6 oncoprotein reveals a self-association mechanism required for E6-mediated degradation of p53. *Structure,* 20**,** 604-17.

ZHANG, B., SPANDAU, D. F. & ROMAN, A. 2002. E5 protein of human papillomavirus type 16 protects human foreskin keratinocytes from UV B-irradiation-induced apoptosis. *J Virol,* 76**,** 220-31.

ZHAO, K. N., HENGST, K., LIU, W. J., LIU, Y. H., LIU, X. S., MCMILLAN, N. A. & FRAZER, I. H. 2000. BPV1 E2 protein enhances packaging of full-length plasmid DNA in BPV1 pseudovirions. *Virology,* 272**,** 382-93.

ZHAO, X., RUSH, M. & SCHWARTZ, S. 2004. Identification of an hnRNP A1-dependent splicing silencer in the human papillomavirus type 16 L1 coding region that prevents premature expression of the late L1 gene. *J Virol,* 78**,** 10888-905.

ZIEGERT, C., WENTZENSEN, N., VINOKUROVA, S., KISSELJOV, F., EINENKEL, J., HOECKEL, M. & VON KNEBEL DOEBERITZ, M. 2003. A comprehensive analysis of HPV integration loci in anogenital lesions combining transcript and genome-based amplification techniques. *Oncogene,* 22**,** 3977-84.

ZUR HAUSEN, H. 1976. Condylomata acuminata and human genital cancer. *Cancer Res,* 36**,** 794.

ZUR HAUSEN, H. 1977. Human papillomaviruses and their possible role in squamous cell carcinomas. *Current topics in microbiology and immunology,* 78**,** 1-30.

ZUR HAUSEN, H., MEINHOF, W., SCHEIBER, W. & BORNKAMM, G. W. 1974. Attempts to detect virus-secific DNA in human tumors. I. Nucleic acid hybridizations with complementary RNA of human wart virus. *International Journal of Cancer,* 13**,** 650-6.

APPENDIX I – REVIEW ARTICLE ACCEPTED FOR PUBLICATION

This article has been accepted by the editor (after peer review) and is awaiting publication in The Obstetrician and Gynaecologist.

Non-cervical human papillomavirus related disease

Jonathan Lippiatt, Ned Powell, Amanda Tristram

Key content:
• Human papillomavirus (HPV) is associated with neoplastic disease at sites other than the cervix
• Cancer caused by HPV may behave differently to other cancers at the same site
• The presence of HPV alone is not sufficient to demonstrate causal association
• HPV vaccination has the potential to impact on disease burden beyond cervical cancer

Learning objectives:
• To understand the spectrum of disease caused by HPV
• To review the management of HPV-associated disease in pregnancy
• To be aware that future research into HPV related disease may alter management

Ethical issues:
• What advice should women be given regarding their sexual partners?
• Should boys be offered HPV vaccination?

Introduction

The recent Nobel Prize winner in Physiology or Medicine Harald zur Hausen first hypothesized a role for human papillomavirus (HPV) in the development of cervical cancer.[1] Since the 1970s HPV has been proposed as a causative factor in a variety of benign and malignant diseases. HPV is a non-enveloped double-stranded DNA virus that infects the epithelial basal layer. The majority of HPV infections occur without symptoms and are cleared by the host over 8-12 months. However, infection may persist resulting in intraepithelial neoplasia and, over time, progression to invasive carcinoma.

One hundred and twenty different types of HPV affecting humans have now been described.[2] HPV has been identified in various benign and malignant lesions within anogenital sites, aerodigestive tract, skin and conjunctiva.[3] There are extensive data in the literature regarding HPV and cervical disease. Walboomers *et al*[4] reported that the worldwide prevalence of HPV in cervical carcinomas is 99.7%. The burden of HPV in non-cervical disease has been increasingly recognised and this review will summarise the evidence available.

HPV is species and tissue specific (see table 1). There are approximately 40 types that affect human mucocutaneous tissue, such as are found in the anal, genital and oral tract. In addition they can be divided between those of malignant potential (high risk) and those associated with benign conditions (low risk). Although the focus of this article is on HPV associated neoplastic disease, there is reference to benign conditions affecting the anogenital site with particular emphasis on their management in pregnancy.

Anogenital

Vulval:
In the UK, in 2007, 1120 women were diagnosed with vulval cancer. This accounts for 6% of gynaecological cancers in the UK.[5] The incidence of both vulval cancer and its precursor vulval intraepithelial neoplasia (VIN) are increasing especially in women below age 50 years [6-8]. In the UK, the proportion of vulval cancer diagnosed under the age of 50 rose from 6% in 1975 to 14% in 2007.[5] Vulval squamous cell carcinoma (VSCC) accounts for over 90% of vulval cancers. There appear to be at least two distinct pathways for the development of vulval cancer and VIN (see figure 1). Usual type VIN is most common and is associated with HPV, younger patients, multifocal lesions, and other anogenital intraepithelial neoplasia.[9] It generally gives rise to warty/basaloid VSCC. Differentiated VIN is usually found adjacent to invasive keratinizing VSCC. It is typically HPV negative and is seen most frequently in older women with other epithelial disorders such as lichen sclerosus or lichen simplex chronicus. [10]
Most of the studies looking at HPV prevalence in vulval cancer and VIN have been small and results vary considerably. A recent meta-analysis[11] reported HPV prevalence in vulval cancer, high-grade VIN and low-grade VIN to be 40.4%, 85.3% and 67.8%, respectively. HPV 16 is by far the most commonly found type accounting for 67.5% of VIN and 32.2% of vulval cancer. The study demonstrates the paucity of data regarding vulval neoplasia. HPV positivity in VSCC varies widely between populations, from 34.7% in some European countries to 63.2% in the US.[12]

The symptoms associated with VIN and vulval cancer can be quite distressing and embarrassing.[13] However, the diagnosis of vulval cancer is often delayed. A Dutch study[12] found that out of all the gynaecological malignancies, vulval cancer had the longest delay in diagnosis. The most common presenting symptom of VIN is pruritus, however, this is also a common symptom found in many benign conditions, and therefore, may not be recognised as serious. Clinical examination is essential and in order to make an accurate diagnosis a biopsy is always recommended.[14] VIN lesions, in particular, can vary in size, shape, regularity, pigmentation and location on the vulva and without a biopsy misdiagnosis is common.

Vaginal:
Primary cancer of the vagina is rare. The commonest causes of squamous cell vaginal cancer are HPV and irradiation. There were 243 cases of vaginal cancer in the UK in 2007.[15] There are a limited number of studies looking at HPV prevalence in vaginal cancer and its precursor, vaginal intraepithelial neoplasia (VAIN). The combined overall HPV prevalence from a total of 15 studies has been reported as 100% in VAIN 1, 90.1% in VAIN 2/3 and 69.9% in vaginal carcinomas[16] HPV 16 is responsible for the majority of

the HPV related vaginal disease. VAIN may be responsible for abnormal cervical smears. Careful vaginal examination is, therefore, mandatory when a cervical cause for the abnormal smear cannot be found. Vaginal lesions may also coexist with cervical or vulval lesions and may not, therefore, be recorded as a separate entity, leading to underreporting.

Anal:
The incidence of anal cancer is approaching 1000 new diagnoses each year in the UK.[17]Although a relatively uncommon malignancy, several studies have shown an increase in incidence over the last few decades. [18 19] Both anal cancer and anal intraepithelial neoplasia have been linked to HPV infection. Overall HPV prevalence has been reported as 91.5%, 93.9% and 84.3% in AIN 1, AIN 2/3 anal carcinomas respectively.[11] HPV 16 is responsible for 68-76% of anal cancers.[3 11 20] However, in immunocompromised individuals with anal cancers or AIN 2/3, several studies have shown proportionally higher rates of infection with other genotypes, and they are often multi-type infections.[3 11 20] In the absence of any effective prevention or screening, the current trends of increasing incidence appear unlikely to change until the effects of HPV prophylactic vaccination become apparent in 10-30 years.

Penile:
Penile cancer is rare in the UK, with approximately 400 cases diagnosed each year.[21] It appears that, like vulval neoplasia, the prevalence of HPV in penile cancer varies with histological type. [3 22] The review by Backes et al[22] reports the overall prevalence of HPV in penile cancer to be 47.9%. Basaloid and warty subtypes are far more consistently associated with HPV infection than verrucous penile cancers (66.3% versus 22.4%).[22] HPV 16 is again the most prevalent HPV type. These histological types have also been associated with a precancerous stage called penile intraepithelial neoplasia, which has a HPV prevalence of 75-100%.[23 24]

Aerodigestive tract

Oropharyngeal cancer:
HPV is associated with the development of a subset of head and neck cancers arising in the oropharynx, which includes the tonsils, base of tongue and soft palate. The UK incidence of oropharyngeal cancer in 2007 was 1063.[25] In the majority of head and neck cancers, the UK incidence is stable or decreasing, whereas the incidence of oropharyngeal cancer, and tonsillar cancer in particular, is increasing.[26] This trend has also been demonstrated in many other countries including the USA, Sweden and Greece and is known to affect younger patients with less exposure to the potentially carcinogenic effects of tobacco and alcohol, relative to HPV-negative patients.[27-30] In Scotland, incidence rates for OPC are increasing faster than rates for any other cancer.[31] The hypothesis that HPV is the fundamental reason for this is supported by a recent UK study. [29] It reports an increase in HPV positive cases from 15% to 57% between 1988 and 2009.

The natural history of HPV infection in the oral cavity is currently poorly defined.[32] Oral HPV infection is believed to be sexually acquired.[27 33] Case-control studies have shown an association between HPV-positive OPC and sexual behaviours including a high lifetime number of oral sex or vaginal sex partners, early age at first intercourse and

infrequent use of condoms.[27] However questions regarding the duration of infection, degree of oral transmission, and whether productive viral infections are established in the oropharynx, have yet to be answered.

There is accumulating evidence from prospective studies that HPV-positive OPC responds better to treatment, including chemotherapy and radiotherapy, than HPV-negative OPC, and that patients with HPV- positive disease have excellent long-term survival rates.[34] Clinical trials to investigate de-escalation of treatment for HPV cases are now underway in the UK and USA.[35]

Others:
There are only few studies looking at HPV prevalence in the oesophagus, larynx and lung. The findings are inconsistent and a causative role for HPV in the development of these cancers cannot yet be reliably demonstrated.[36]

Skin and conjunctiva

Skin:
The skin of both healthy populations and immunocompromised patients harbours a very large spectrum of HPV genotypes. Genus β papillomaviruses, in particular, are involved in cutaneous lesions in humans.[37] Recently, Karagas et al[38] showed that people with squamous cell carcinoma, but not basal cell carcinoma, were far more likely to have each of the β HPV types compared to people in the control group. In addition, they found the more HPV types present increased the probability of having squamous cell carcinoma. However, the pathogenesis in relation to HPV still remains unclear. Further molecular and epidemiological data is required.

Conjunctiva:
The evidence for a role of HPV in ocular disease is controversial. HPV has been implicated in lid warts, conjunctival and lacrimal sac papillomas, conjunctival and corneal dysplasia as well as squamous cell carcinoma of the conjunctiva. However the major studies examining the role of HPV infection in ocular disease have mixed findings. In a recent review, Hughes et al[39] found that 80% of conjunctival papillomas were HPV positive. Thirty-two per cent of dysplasias, including conjunctival intraepithelial neoplasia and carcinoma in situ, were HPV positive. HPV 16 and 18 were most prevalent, although HPV 6 and 11 were also identified. Thirty-one per cent of conjunctival squamous carcinomas were HPV positive and this was almost exclusively HPV 16 and 18.

HPV disease in pregnancy

Genital warts:
The innate immune defences normally activated in pregnancy can contribute to increased clinical severity of some maternal infections.[40] Often women will see deterioration in symptoms during pregnancy. The warts may become more prolific and more friable and much harder to treat. The treatment of genital warts in pregnancy is also a challenge because the treatment options are more limited. (see box 1)
Transmission of HPV to the neonate is a rare but associated risk. A small number of neonates with HPV infection at delivery will go onto to have Recurrent Respiratory

Papillomatosis (RRP), which can lead to airway obstruction, repeated surgery and, rarely, malignancy. Quantifying the risk of RRP is very difficult because the condition is so uncommon. A Danish study[41] found, over a 20-year period the overall rate of RRP was 3.5 per 1,000,000 person-years. It also reported that seven of every 1000 births with a maternal history of genital warts resulted in RPP.[41] Caesarean delivery has so far not been shown to reduce the risk and should only be performed for genital warts if the birth canal is obstructed or if vaginal delivery would result in excessive bleeding.[41]

Box 1 – Management of genital warts in pregnancy

- Warts often grow in size and number or may appear for the first time during pregnancy
- The aims of treatment are: symptom relief and minimising the number of lesions present at delivery
- Treatment options include: laser, diathermy, surgical excision, trichloroacetic acid
- Most topical treatments should be avoided due to possible teratogenic effects
- Only in severe cases should caesarean delivery be considered
- There is an extremely small risk of their child developing recurrent respiratory papillomatosis

Anogenital intraepithelial neoplasia:

The NHS Cervical Screening Programme has guidance for managing cervical intraepithelial neoplasia (CIN) in pregnancy.[42] (see box 2) The primary aim of management in pregnancy is to identify invasive disease. Cervical carcinomas can present in pregnancy and enquiring about the patient's smear history and a speculum examination of the cervix should be routinely performed following an episode of antepartum haemorrhage. There are currently no guidelines for the management of vulval, vaginal or anal intraepithelial neoplasia in pregnancy.

Box 2 – NHSCSP Guidelines in Pregnancy

- If a woman has been called for routine screening during pregnancy it should be deferred
- If a previous test was abnormal and in the interim the woman becomes pregnant, then the test should not be delayed but should be taken in mid-trimester unless there is a clinical contraindication
- If a pregnant woman requires colposcopy or cytology after treatment (or follow up of untreated CIN 1), her assessment may be delayed until after delivery. Unless there is an obstetric contraindication, however, assessment should not be delayed if a first follow up cytology or colposcopy is required following treatment for cervical glandular intra-epithelial neoplasia, or treatment for CIN 2/3 with involved or uncertain margin status
- A woman who meets the criteria for colposcopy should be examined in the colposcopy clinic even if she is pregnant
- If there is a suspicion of invasive disease then a biopsy is indicated. Punch biopsies suggesting only CIN cannot reliably exclude invasion. Therefore, cone, wedge and diathermy loop biopsies are preferred. Beware of the increased risk of haemorrhage

HPV disease in males / implications for sexual partners

With the introduction of the HPV vaccine, knowledge of the virus and its link with cervical cancer is improving in the general population. The vast majority of people, however, would not be aware of the risk of non-cervical HPV related disease. It is very important, therefore, for healthcare workers to inform their patients about the risks to both the patients and their sexual partners. The uptake of cervical screening is currently only 79% so every opportunity should be taken to check a woman's cervical smear history.[43] Female sexual partners should be strongly encouraged to attend routine cervical screening and both male and female partners should be told to report any unusual symptoms or lesions in the anogenital area. All should also be aware of and report persistent or unusual oropharnygeal symptoms. The ubiquitous nature of the infection should be emphasised, as should the fact that sexually active adults will have most likely been exposed to HPV already.

HPV vaccine discussion

The full impact of the HPV vaccine on non-cervical disease is not yet known, but the uptake of the vaccine and the choice of vaccine are likely to have the greatest significance. There are a few studies that have looked at the incidence of anogenital intraepithelial neoplasia post vaccination.[44-46] Kjaer and colleagues[46] have shown that after 42 months the vaccine was 100% effective against HPV 6/11/16/18-related high-grade vulval and vaginal lesions. Similar efficacy has been shown for low-grade cervical, vulval, and vaginal intraepithelial neoplasia.[45] Smith et al[47] has predicted that the first-generation prophylactic HPV vaccines have the potential to prevent around 45% of all HPV associated cancers, including approximately 70% of invasive cervical cancers and 25% of non-cervical cancers.[47] This was based on recent systematic reviews and worldwide cancer statistics produced by the International Agency for Research on Cancer (IARC).[48] One of the problems in predicting the effect of the vaccine is a possible overestimation in disease attributed to HPV, as the presence of HPV DNA in tumour tissue does not necessarily indicate a causal relationship (see box 3). Where polymerase chain reaction (PCR) detection of HPV DNA is used alone, it is reasonable to assume that there would be false positives and therefore overestimation of the aetiological significance. Smeets et al[40] demonstrated this by comparing HPV DNA PCR detection alone with more detailed analysis using detection of oncogene expression. More evidence is required to establish the full burden of HPV disease (see table 2).

Many reviews on this topic have shown variations in HPV prevalence in different populations. When combining this data care must be taken in interpreting it especially if it will influence health policy in a particular country. Recent UK-based studies have shown the importance of population specific data[49] and have indicated that the potential influence of the HPV vaccine may actually be underestimated in the UK. There is on-going debate of whether boys should be vaccinated as well as girls. The arguments for and against have been summarised by Hibbitts[50] and Cuschieri[51] (see box 4)

Box 3 – Establishing causal links between HPV and cancer

Certain criteria must be met to confirm a causative link between HPV and a cancer. The basis for this are Koch postulates published in 1890[52] (for bacterial disease). These postulates can be adapted for HPV and cancer[53]:

1. Is HPV infection found in affected patients?
2. Are viral genes found in cancer cells?
3. Can the cancer be linked to the presence of an active viral gene product e.g. an oncogene?
4. Does prevention of the infection stop the cancer e.g. by vaccination?

Box 4 – Should we vaccinate boys?

Arguments for:

- Herd immunity obtained by vaccinating only women is likely to be insufficient to eradicate the targeted HPV types
- Increased protection from non-cervical HPV disease, particularly in men who have sex with men risk groups
- Reduced disease burden on the NHS, particularly if a quadrivalent vaccine is used

Arguments against:

- Increased costs of extending the vaccination programme to prevent relatively few non-cervical malignancies
- Targeted screening instead of widespread vaccination may be more appropriate for high risk groups

Conclusion

HPV causes disease in many non-cervical sites, the full extent of which has not been fully established. Our traditional anatomical approach to the management of cancers may be replaced by a system related to disease aetiology that will allow improved management strategies.

The management of HPV associated disease in pregnancy necessitates a modified approach. The treatment as well as the counselling given should be individualised to the patient.

HPV vaccination has been a huge step forward in cancer prevention, but there are many social, economic and ethical questions still to be answered.

Table 1: Common association of HPV and disease.
Diseases commonly associated with individual HPV types

HPV Type	Disease
1,2	Verrucca vulgaris
1,2 & 4	Plantar warts
3, 10	Flat cutaneous warts
5, 8	Carcinogenesis in epidermodysplasia verruciformis
6, 11	Anogenital warts, Respiratory papillomatosis
16, 18	Anogenital neoplasia, oropharyngeal cancers

Table 2: Summary of level of evidence for role of human papillomavirus (HPV) in carcinogenesis

Level of evidence	HPV type	Site
Sufficient	16	Cervix, vulva[b], vagina, penis[b], anus, oral cavity, oropharynx
	18, 31, 33, 35, 39, 45, 51, 56, 58, 59, 66	Cervix
	5, 8	Skin in patients with epidermodysplasia verruciformis[s]
Limited	16	Larynx, periungal[s]
	18	Vulva[b], vagina, penis[b], anus, oral cavity, larynx
	6,11	Larynx, vulva[v], penis[v], anus[v]
	Beta genus types	Skin[s]
	All types	Conjunctiva
Inadequate		Oesophagus, lung, colon, ovary, breast, prostate, urinary bladder, nasal and sinus

Adapted from International Association for Research on Cancer Monograph volume 90
[b] Basaloid and warty tumours [s] Squamous carcinoma [v]Verruccous carcinoma

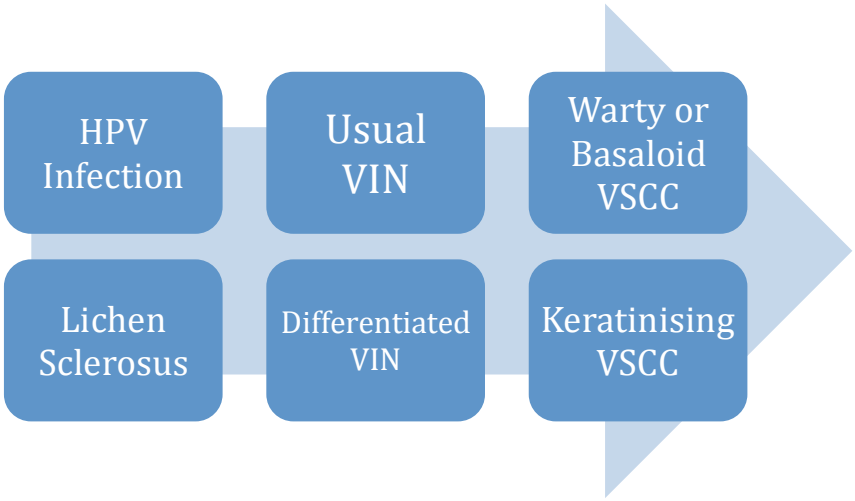

Figure 1: Diagram showing the two aetiologies of vulval cancer

References

1. zur Hausen H. Condylomata acuminata and human genital cancer. *Cancer research* 1976;36(2 pt 2):794.
2. Bernard HU, Burk RD, Chen Z, van Doorslaer K, Hausen H, de Villiers EM. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* 2010;401(1):70-9.
3. International Agency for Research on Cancer. *IARC Monographs on the Evaluation of Carinogenic Risks to Humans: Human Papillomaviruses*. Lyon, France, 2007.
4. Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA,Shah KV, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999;189(1):12-9.
5. Cancer Research UK. Vulval Cancer - UK incidence statistics. London: Cancer Research UK, 2009.
6. Judson PL, Habermann EB, Baxter NN, Durham SB, Virnig BA. Trends in the Incidence of Invasive and In Situ Vulvar Carcinoma. *Obstet Gynecol* 2006;107(5):1018-22.
7. Joura E, Lösch A, Haider-Angeler M, Breitenecker G, Leodolter S. Trends in vulvar neoplasia. Increasing incidence of vulvar intraepithelial neoplasia and squamous cell carcinoma of the vulva in young women. *J Reprod Med* 2000;45(8):613-5.
8. Jones R, Baranyai J, Stables S. Trends in squamous cell carcinoma of the vulva: the influence of vulvar intraepithelial neoplasia. *Obstet Gynecol* 1997;90(3):448-52.
9. McCluggage W. Recent developments in vulvovaginal pathology. *Histopathology* 2009;54(2):156-73.
10. van Seters M, ten Kate F, van Beurden M, Verheijen R, Meijer C, Burger M, et al. In the absence of (early) invasive carcinoma, vulvar intraepithelial neoplasia associated with lichen sclerosus is mainly of undifferentiated type: new insights in histology and aetiology. *J Clin Pathol* 2007;60(5):504-9.
11. De Vuyst H, Clifford GM, Nascimento MC, Madeleine MM, Franceschi S. Prevalence and type distribution of human papillomavirus in carcinoma and intraepithelial neoplasia of the vulva, vagina and anus: a meta-analysis. *Int J Cancer* 2009;124(7):1626-36.
12. van de Nieuwenhof HP, van Kempen LCLT, de Hullu JA, Bekkers RLM, Bulten J, Melchers WJG, et al. The Etiologic Role of HPV in Vulvar Squamous Cell Carcinoma Fine Tuned. *Cancer Epidemiology Biomarkers & Prevention* 2009;18(7):2061-67.
13. Shylasree TS, Karanjgaokar V, Tristram A, Wilkes AR, MacLean AB, Fiander AN. Contribution of demographic, psychological and disease-related factors to quality of life in women with high-grade vulval intraepithelial neoplasia. *Gynecol Oncol* 2008;110(2):185-9.
14. Royal College of Obstetricians and Gynaecologists. Management of Vulval Cancer - Working Party Report, 2006.
15. Cancer Research UK. Vaginal Cancer - UK incidence statistics. London: CRUK, 2009.
16. Baldwin P, van der Burg S, Boswell C, Offringa R, Hickling J, Dobson J, et al. Vaccinia-expressed human papillomavirus 16 and 18 e6 and e7 as a therapeutic vaccination for vulval and vaginal intraepithelial neoplasia. *Clin Cancer Res* 2003;9(14):5205-13.
17. Cancer Research UK. Risks and causes of anal cancer.

18. Frisch M, Melbye M, Møller H. Trends in incidence of anal cancer in Denmark. *BMJ* 1993;306(6875):419-22.
19. Robinson D, Coupland V, Moller H. An analysis of temporal and generational trends in the incidence of anal and other HPV-related cancers in Southeast England. *Br J Cancer* 2009;100(3):527-31.
20. Wong AK, Chan RC, Aggarwal N, Singh MK, Nichols WS, Bose S. Human papillomavirus genotypes in anal intraepithelial neoplasia and anal carcinoma as detected in tissue biopsies. *Mod Pathol* 2009;23(1):144-50.
21. Cancer Research UK. Penis cancer - A to Z.
22. Backes D, Kurman R, Pimenta J, Smith J. Systematic review of human papillomavirus prevalence in invasive penile cancer. *Cancer Causes and Control* 2009;20(4):449-57.
23. Aynaud O, Ionesco M, Barrasso R. Penile intraepithelial neoplasia. Specific clinical features correlate with histologic and virologic findings. *Cancer* 1994;74(6):1762-7.
24. Rubin MA, Kleter B, Zhou M, Ayala G, Cubilla AL, Quint WG, et al. Detection and typing of human papillomavirus DNA in penile carcinoma: evidence for multiple independent pathways of penile carcinogenesis. *The American journal of pathology* 2001;159(4):1211-8.
25. Cancer Research UK. Oral cancer - UK incidence statistics.
26. Evans M, Powell NG. The changing aetiology of head and neck cancer: the role of human papillomavirus. *Clin Oncol (R Coll Radiol)* 2010;22(7):538-46.
27. D'Souza G, Kreimer AR, Viscidi R, Pawlita M, Fakhry C, Koch WM, et al. Case-control study of human papillomavirus and oropharyngeal cancer. *N Engl J Med* 2007;356(19):1944-56.
28. Kreimer AR, Clifford GM, Boyle P, Franceschi S. Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2005;14(2):467-75.
29. Nasman A, Attner P, Hammarstedt L, Du J, Eriksson M, Giraud G, et al. Incidence of human papillomavirus (HPV) positive tonsillar carcinoma in Stockholm, Sweden: an epidemic of viral-induced carcinoma? *Int J Cancer* 2009;125(2):362-6.
30. Romanitan M, Nasman A, Ramqvist T, Dahlstrand H, Polykretis L, Vogiatzis P, et al. Human papillomavirus frequency in oral and oropharyngeal cancer in Greece. *Anticancer Res* 2008;28(4B):2077-80.
31. Junor EJ, Kerr GR, Brewster DH. Oropharyngeal cancer. Fastest increasing cancer in Scotland, especially in men. *Bmj* 2010;340:c2512.
32. Adelstein DJ, Ridge JA, Gillison ML, Chaturvedi AK, D'Souza G, Gravitt PE, et al. Head and neck squamous cell cancer and the human papillomavirus: summary of a National Cancer Institute State of the Science Meeting, November 9-10, 2008, Washington, D.C. *Head Neck* 2009;31(11):1393-422.
33. Kreimer AR, Alberg AJ, Daniel R, Gravitt PE, Viscidi R, Garrett ES, et al. Oral human papillomavirus infection in adults is associated with sexual behavior and HIV serostatus. *J Infect Dis* 2004;189(4):686-98.
34. Marur S, D'Souza G, Westra WH, Forastiere AA. HPV-associated head and neck cancer: a virus-related cancer epidemic. *The lancet oncology* 2010;11(8):781-9.
35. Brana I, Siu LL. Locally advanced head and neck squamous cell cancer: treatment choice based on risk factors and optimizing drug prescription. *Annals of Oncology* 2012;23(suppl 10):x178-x85.

36. Gillison ML, Chaturvedi AK, Lowy DR. HPV prophylactic vaccines and the potential prevention of noncervical cancers in both men andwomen. *Cancer* 2008;113(S10):3036-46.
37. de Villiers E-M, Fauquet C, Broker TR, Bernard H-U, zur Hausen H. Classification of papillomaviruses. *Virology* 2004;324(1):17-27.
38. Karagas MR, Waterboer T, Li Z, Nelson HH, Michael KM, Bavinck JNB, et al. Genus β human papillomaviruses and incidence of basal cell and squamous cell carcinomas of skin: population based case-control study. *BMJ*, 2010:c2986.
39. Hughes DS, Powell N, Fiander AN. Will vaccination against human papillomavirus prevent eye disease? A review of the evidence. *The British journal of ophthalmology* 2008;92(4):460-5.
40. Brabin BJ. Epidemiology of Infection in Pregnancy. *Review of Infectious Diseases* 1985;7(5):579-603.
41. Silverberg MJ, Thorsen P, Lindeberg H, Grant LA, Shah KV. Condyloma in pregnancy is strongly predictive of juvenile-onset recurrent respiratory papillomatosis. *Obstetrics and gynecology* 2003;101(4):645-52.
42. NHS Cervical Screening Programmes. Colposcopy and programme management: guidelines for the NHS cervical screening programme. Second ed: NHSCSP, 2010.
43. The NHS Information Centre. Cervical Screening Programme: England 2010-11: The Health and Social Care Information Centre, Screening and Immunisations team, 2011.
44. Joura EA, Leodolter S, Hernandez-Avila M, Wheeler CM, Perez G, Koutsky LA, et al. Efficacy of a quadrivalent prophylactic human papillomavirus (types 6, 11, 16, and 18) L1 virus-like-particle vaccine against high-grade vulval and vaginal lesions: a combined analysis of three randomised clinical trials. *Lancet* 2007;369(9574):1693-702.
45. Dillner J, Kjaer SK, Wheeler CM, Sigurdsson K, Iversen OE, Hernandez-Avila M, et al. Four year efficacy of prophylactic human papillomavirus quadrivalent vaccine against low-grade cervical, vulvar, and vaginal intraepithelial neoplasia and anogenital warts: randomised controlled trial. *BMJ* 2010;341:c3493.
46. Kjaer SK, Sigurdsson K, Iversen O-E, Hernandez-Avila M, Wheeler CM, Perez G, et al. A Pooled Analysis of Continued Prophylactic Efficacy of Quadrivalent Human Papillomavirus (Types 6/11/16/18) Vaccine against High-grade Cervical and External Genital Lesions. *Cancer Prevention Research* 2009;2(10):868-78.
47. The Burden of HPV-Associated Cancers. Eurogin; 2010 February 2010; Monte Carlo, Monaco.
48. Ferlay J SH, Bray F, Forman D, Mathers C and Parkin DM. GLOBOCAN 2008, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10 [Internet]. Lyon, France: International Agency for Research on Cancer, 2010.
49. Powell N, Boyde A, Tristram A, Hibbitts S, Fiander A. The potential impact of human papillomavirus vaccination in contemporary cytologically screened populations may be underestimated: an observational retrospective analysis of invasive cervical cancers. *Int J Cancer* 2009;125(10):2425-7.
50. Hibbitts S. Should boys receive the human papillomavirus vaccine? Yes. *BMJ* 2009;339:b4928.
51. Cuschieri K. Should boys receive the human papillomavirus vaccine? No. *BMJ* 2009;339:b4921.
52. Koch R. An Address on Bacteriological Research. *Br Med J* 1890;2(1546):380-3.
53. Burns J, Maitland N. Human papillomaviruses and cancer. *Microbiology Today*, August 2005:116-20.

APPENDIX II – PATIENT INFORMATION, CONSENT FORM, CLINIC STUDY

FLOW CHART

GIG CYMRU NHS WALES
Bwrdd Iechyd Prifysgol Caerdydd a'r Fro
Cardiff and Vale University Health Board

Ysbyty Athrofaol Llandochau
University Hospital Llandough

Penlan Road, Llandough
Penarth, Vale of Glamorgan
CF64 2XX

Heol Penlan, Llandochau
Penarth, Bro Morgannwg
CF64 2XX

SurePath Plus ™ in **Per**sistent **L**ow-grade Cytolog**y** (SuPerLy)

A Study to Evaluate BD SurePath Plus™ as a Predictor of High-grade Cervical Disease in Women with Persistent Low-grade Squamous Cytological Abnormalities

(A study of a new way of looking at cervical smear tests in the laboratory)

We would like to invite you to take part in a research study. Before you decide you need to understand why the research is being done and what it would involve for you. Please take time to read the following information carefully. Talk to others about the study if you wish.

What is the purpose of the study?
We are trying to find out if a new way of looking at cervical smears can reduce the number of women who are referred to hospital, without missing women who need to be seen. Essentially, we are asking to use your sample to test a new test.

Cervical Screening has greatly reduced the number of cervical cancers in the UK, but a lot of women are seen in the hospital who do not need to have treatment. The purpose of the screening programme is to pick out women who would most benefit from having treatment to prevent cervical cancer developing in years to come. The cervical smear taken at your doctors can identify whether you have an increased chance of having abnormal cells in the cervix (the neck of the womb). You can then be referred to hospital, to see if you need to have treatment to prevent cancer developing in years to come. Not all women picked up by the current cervical smear test and referred to hospital will actually have abnormal cells. Even fewer women will need to have treatment.

We will compare the results of the new test with what is found when you are seen in the hospital.

Why Have I Been Chosen?
You have had an abnormal result and are being referred to hospital to see if you need treatment to prevent you getting cancer in the years to come. Most women in your situation will not need to have treatment. We are trying to find out if a new test can help identify which women most need to be seen and which women do not.

What are the tests being used?
Current cervical smear testing involves looking at the shapes of the cells in your cervical smear to see if they look abnormal. The new test is called BD SurePath Plus™ and picks out abnormal cells in a cervical smear sample by testing for raised levels of some proteins in the cell.

We would also like to do a test to see if the virus Human Papillomavirus (HPV) that can cause abnormal smears and cervical cancer is present. HPV is very common and most women will be infected at some point in their lives. HPV usually clears from the body in about a year, without any treatment. Having an HPV infection only very rarely leads to cancer. As HPV is so common, just testing for the virus would mean too many women would have to go to hospital to be checked. We are therefore looking to see if using the new test (BD SurePath Plus™) as well will reduce the number of women who have to be referred.

We can do both these tests on the cervical smear sample that you have already had taken. We will then compare these results with the results of all the tests you have done on your cervix routinely, either at the hospital or your GP for the next two years.

The tests will be done on the sample you have already had taken. The people testing your sample will not know who you are. Because the purpose of this study is to see if the new tests actually work we cannot use the results to manage your current care or treatment. The results of these tests will therefore not be known to the staff in the colposcopy clinic and will not be given to you.

Will any genetic tests be done?
Because HPV infects the DNA within cells a HPV test involves some DNA being extracted from cells contained in your sample. DNA is extracted purely to see if the virus is present and to check the quality of the sample.

What Would I Have To Do?
You would have to sign a consent form to say that we can do these two extra tests on the sample that has already been taken. You would also have to agree to the study having access to the results of all the tests you have done on your cervix routinely, either at the hospital or your GP for the next two years. You would **not** have to have any extra samples taken. The results of the new test will not change what happens to you as a result of the abnormal smear you have had.

What are the Possible Disadvantages of Taking Part?
Taking part in the study will not change the way you are treated.

What Extra Tests will I need if I Take Part?
You will not need any extra samples taken from you if you agree to take part in the study.

What Are The Possible Benefits of Taking Part?
Taking part in this study will not change the way you are treated. The study will not help you now, but the information we get from this study will help improve the way cervical screening is carried out in the future.

Do I Have To Take Part?
It is up to you to decide whether you want to take part in the study or not. If, after reading this information sheet, you decide you would like to take part we will ask you to sign a consent form when you attend the colposcopy clinic. If you decide not to take part or you decide to withdraw at any time, it will not affect the standard of care you receive, now or in the future.

What Would Happen To Me If I Decide To Take Part?
The way you are treated by the hospital will not change. The study team will test your most recent smear test and will be told the results of any tests related to your cervix done routinely at the hospital or your GP over the next two years.

What If Something Goes Wrong?
If you are unhappy with any aspect of the study we would like you to tell us. Regardless of this, if you wished to complain about any aspect of the way in which you have been approached or treated, the normal NHS complaints mechanism is available to you.

Would My Taking Part In This Study Be Kept Confidential?
All information which is collected about you during the course of the research will be kept strictly confidential. The fact you are included in the study will be documented in your hospital notes. You would be assigned a study number which will be used on information or samples that leave the hospital. Your medical records might be inspected by regulatory authorities to check the study is properly carried out. Signing the consent form to take part in the study means you agree to this access.

What Happens to the Results of the Study?
The results may be published in a medical journal and/or presented at a scientific meeting. It would not be possible to identify you from any of the information published or presented.

Who Has Reviewed the Study?
All research in the NHS is looked at by independent group of people, called a Research Ethics Committee to protect your safety, rights, wellbeing and dignity. This study has been reviewed and given favourable opinion by South East Wales Research Ethics Committee.

Who Is Organising and Funding the Research?
The research is being coordinated by the Academic Department of Obstetrics and Gynaecology at Wales College of Medicine, Cardiff University.

It is being supported by Cervical Screening Wales, Becton Dickinson (the company that makes the test) and Source BioScience (the company that supply and carry out the test).

Further Information

Further information about cervical smear tests, results and treatment is available on the Cervical Screening Wales website: www.screeningservices.org.uk/csw

If you have any questions at any time during the study please contact:

Dr Amanda Tristram 02920 742337
Research Nurse Sue Ashman 02920 745365

Thank you for taking the time to read this information and considering whether or not you would like to take part in this study of a new test for cervical screening.

GIG CYMRU NHS WALES
Bwrdd Iechyd Prifysgol Caerdydd a'r Fro
Cardiff and Vale University Health Board

Ysbyty Athrofaol Llandochau
University Hospital Llandough

Penlan Road, Llandough
Penarth, Vale of Glamorgan
CF64 2XX

Heol Penlan, Llandochau
Penarth, Bro Morgannwg
CF64 2XX

Consent Form

SurePath Plus™ in Persistent Low-grade Cytology (SuPerLy)

(A study of a new test for cervical screening)

Centre Number: ☐

Study Number: ☐

Referral Smear Laboratory Number: ☐

Name of Researcher: Dr Amanda Tristram

Please initial each box

1. I confirm that I have read and understand the information sheet dated 13.10.10 version 1.1 for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected. ☐
3. I understand that relevant sections of my medical notes and data collected during the study, may be looked at by individuals from Cardiff University, from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. ☐
4. I agree to take part in the above study. ☐

Name of Patient	Date	Signature

Name of Person taking consent	Date	Signature

When completed, 1 for patient; 1 to Cytology Laboratory; 1 (original) to be kept in medical notes

☐

SurePath Plus™ in Persistent Low Grade Cytology

SuPerLy (Overall study flow chart)

Cytology samples from women fitting study criteria are already being stored by cytology lab

CSAD will notify women of smear result as normal

CSAD will produce 3 study consent forms per woman which will accompany referral list to colposcopy admin

Colposcopy admin to send patient information leaflet for the study with appointment letter for colposcopy

Notes to be flagged with 'SuPerLy' label

3 consent forms plus one additional patient information leaflet to be placed in colposcopy notes

Women will attend colposcopy appointment

Attending colposcopist to enquire if woman received patient information for the study and if she is happy to participate

All 3 consent forms to be signed by woman and colposcopist

One copy of consent to woman

One copy to be filled in notes with copy of patient information leaflet

3rd consent form to be forwarded to cytology laboratory

Addressed envelopes will be in clinic
Consent forms to be placed in envelope and sent with the normal cytology samples for that day

When consent received by cytology laboratory sample will be identified and released for testing

If women do not wish to participate then unsigned consent to be returned in envelope and sample will then be discarded by the lab

APPENDIX III – HYBRID CAPTURE 2 PROCEDURE

Before starting the denaturation step 5 drops of indicator dye was added to the denaturation buffer. Within the kit was a negative calibrator, a positive calibrator, a low-risk positive control and a high-risk positive control. Into each of these vials denaturation buffer was added in a 2:1 ratio. The samples previously prepared in 250µl aliquots (see 7.3.1) were thawed and centrifuged for 15 minutes at 3000 rpm. The supernatant was removed and 100 µl of Specimen Transport Medium (STM) was added. In addition, 50 µl of denaturation buffer was added to each sample before vortexing. All the samples, controls and calibrators were incubated in a water bath preheated to 65 °C for 45 minutes. Following that the samples were vortexed again.

A 96-well hybridisation microplate was prepared with 75 µl of each calibrator, control, and sample as follows: A1–A3 negative calibrator, A4–A6 positive calibrator, A7 high-risk control, A8 low-risk control, A9 onwards up to 88 samples. A lid was placed on the plate and left at room temperature for 10 minutes to allow further denaturation. During the denaturation the high and low-risk probe was briefly centrifuged. The final probe mix was made up by 1:25 dilution with probe diluent into polypropylene tubes and vortexed gently. A reservoir was used in order to multi-dispense pipette 25 µl of probe mix into each well in the microplate. With the lid replaced the microplate was put in the plate shaker for 3 minutes at 1100 rpm. The colour of the samples should change at this point from purple to yellow because of neutralization. If this reaction did not occur then an additional 25 µl of the probe mix was added and the microplate was shaken for a further 1 minute. Following this the microplate was placed in the heated block at 65 °C with the lid off for 1 hour.

After one hour the entire contents of each well were transferred into the ELISA coated wells in the pre-labeled Capture Microplate. The plate was then shaken at 1100 rpm for 1 hour at room temperature. Excess liquid was removed by emptying it into a sink and banging the plate on paper towels. Into each well 75 µl on the conjugate was added, the lid was replaced and left to stand for 30 minutes at room temperature. During this time the wash buffer was prepared by mixing 100 ml of wash buffer concentrate with 2900 ml of sterile water. After the well contents were discarded the wash buffer was applied and then discarded again. This process was repeated five times. To each well 75 µl of the detection reagent 2 was added and the plate was left in the dark for 15 minutes. The plate was then read in the luminometer using the provided software.