# A.M. Matthews

# May 2015

# Doctor of Medicine

# The use of salivary biomarkers in the detection of oral squamous cell carcinoma

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor of Medicine by

**April Marie Matthews**

**May 2015**

## Contents

## List of Tables

## List of Tables

## List of Figures

## List of abbreviations

| | |
|---|---|
| Cm | Centimetre |
| ml | Millilitre |
| mm | Millimetre |
| ng | nanogram |
| nM | nanomole |
| $^{o}$C | Degrees celcius |
| rpm | Revolutions per minute |
| µg | Microgram |
| µl | Microlitre |
| +ve | Positive |
| -ve | Negative |
| 2DE | 2D gel electrophoresis |
| 5-Aza-dc | 5-aza-2-deoxycytidine |
| *ACTB* | Beta actin gene |
| *ACVR1* | Activin A receptor, type 1 gene |
| *ADAMTS9* | ADAM Metallopeptidase with Thrombospondin Type 1 Motif, 9 gene |
| *AIM1* | Absent in melanoma 1 gene |
| *APC* | Adenomatous polyposis coli gene |
| *ATM* | ATM serine/threonine kinase gene |
| *BRCA1* | Breast cancer 1 early onset gene |
| *BRCA2* | Breast cancer 2 early onset gene |
| *CALCA* | Calcitonin-related polypeptide alpha gene |
| *CASP8* | Caspase 8, apoptosis-related cysteine peptidase gene |
| *CCNA1* | Cyclin A1 gene |
| *CCND1* | Cyclin D1 gene |
| *CCND2* | Cyclin D2 gene |

| | |
|---|---|
| *CD44* | CD44 molecule (Indian blood group) gene |
| *CDH1* | Cadherin 1 type 1 gene |
| *CDH4* | Cadherin 4, type 1, R-cadherin gene |
| *CDH13* | Cadherin 13 gene |
| CDK4 | Cyclin-dependent kinase 4 enzyme |
| CDK6 | Cyclin-dependent kinase 6 enzyme |
| *CDKN2B* | Cyclin - dependent kinase inhibitor 2B gene |
| C-erbB-2 | Receptor tyrosine kinase erbB2 protein |
| CGH | Comparative genomic hybridisation |
| *CHL1* | Cell adhesion molecule L1-like |
| CI | Confidence interval |
| *COL1A2* | Collagen type I alpha 2 gene |
| *COL4A1* | Collagen type IV alpha 1 gene |
| *COL5A2* | Collagen type V alpha 2 gene |
| CRUK | Cancer research UK |
| *CSMD1* | Cub and Sushi multiple domains 1 gene |
| CT | Cycle threshold |
| *CYGB* | Cytoglobin gene |
| *DAPK* | Death associated protein kinase gene |
| *DBC1* | Deleted in bladder cancer 1 gene |
| *Dcc* | Deleted in colorectal cancer gene |
| DNA | Deoxyribonucleic acid |
| DOD | Died of disease |
| DOC | Died of other causes |
| DOH | Department of Health |
| *DPM1* | Dolichyl-phosphate mannosyltransferase polypeptide 1, catalytic subunit gene |
| *DUSP1* | Dual specificity phosphatase 1 gene |

| | |
|---|---|
| *ECAD* | Cadherin 1 type 1 gene |
| ECS | Extra capsular spread |
| ED | Emergency department |
| *EDNRB* | Endothelin receptor type B gene |
| *EGFR* | Epidermal growth factor receptor gene |
| ELISA | Enzyme-linked immunosorbent assay |
| *ESR* | Oestrogen receptor 1 gene |
| FAM | Fluorescein fluorescent dye |
| *FANCD2* | Fanconi anaemia complementation group D2 gene |
| *FANCG* | Fanconi anaemia complementation group G gene |
| *FAT1* | FAT atypical cadherin 1 gene |
| *FHIT* | Fragile histidine triad gene |
| *FKBP4* | FK506 binding protein 4, 59KDa gene |
| FOM | Floor of mouth |
| FU | Follow up |
| *GABRB3* | Gamma-aminobutyric (GABA) A receptor, beta 3 gene |
| *GATA4* | GATA binding protein 4 gene |
| *GSTP1* | Glutathione S-transferase pi 1 gene |
| *H3F3A* | H3 histone family 3A gene |
| HIF | Hypoxia inducible factor proteins |
| HNSCC | Head and neck squamous cell carcinoma |
| *hmlH1* | mutL homolog 1 gene |
| *HOXA9* | Homeobox A9 gene |
| *H.pylori* | Helicobacter pylori bacteria |
| HPV | Human Papilloma virus |
| HRE | Hypoxic related element |
| *INSR* | Insulin receptor gene |

| | |
|---|---|
| Kb | Kilobase |
| *KIF1A* | Kinesin family member 1A gene |
| IL-1Beta | Interleukin 1 Beta |
| IL-6 | Interleukin 6 |
| IL-8 | Interleukin 8 |
| *IL-11* | Interleukin 11 gene |
| *KRAS* | Kirsten rat sarcoma viral oncogene homolog gene |
| LOH | Loss of heterozygosity |
| LPREC | Liverpool paediatric research ethics committee |
| LREC | Liverpool research ethics committee |
| *MAL* | Mal, T-cell differentiation protein gene |
| MBD | Methyl-CpG binding protein |
| *MBD3L2* | Methyl-CpG binding domain protein 3-like 2 gene |
| *MCAM* | Melanoma cell adhesion molecule gene |
| MDM2 | Mouse double minute 2 homolog protein |
| Me-DIP | Methyl-DNA immunoprecipitation |
| MEP | Methylation enriched pyrosequencing |
| *MET* | MET proto-oncogene, receptor tyrosine kinase gene |
| *MGMT* | O-6-methylguanine-DNA methyltransferase |
| *MINT1* | Amyloid beta (A4) precursor protein binding, family A, member 1 gene |
| Mi-RNA | Micro Ribonucleic acid |
| MISCC | Micro-invasive squamous cell carcinoma |
| *MLH-1* | MutL homolog 1 gene |
| *MMP-1* | Matrix metalloproteinase-1 gene |
| *MMP-9* | Matrix metalloproteinase-9 gene |
| MRNA | Messenger Ribonucleic acid |
| MS-AP-PCR | Methylation sensitive arbitrarily primed PCR |

| | |
|---|---|
| MSI | Microsatellite instability |
| MSRE | Methylation sensitive restriction enzymes |
| MSP | Methylation specific PCR |
| MTI | Methylation Index |
| MTR | Malignant transformation rate |
| NA | Not applicable |
| NED | Fluorescent dye |
| *NISCH* | Nischarin gene |
| *NEUROD1* | Neurogenic differentiation 1 gene |
| *NID2* | Nidogen 2 (osteonidogen) gene |
| N0 | Clinically node negative neck |
| NPC | Nasopharyngeal carcinoma |
| NSC | National screening committee |
| NSCLC | Non-small cell lung carcinoma |
| NTC | No template control |
| *NTRK3* | Neurotrophic tyrosine kinase, receptor, type 3 gene |
| *OAZ1* | Ornithine decarboxylase antizyme 1 gene |
| OED | Oral epithelial dysplasia |
| OSCC | Oral squamous cell carcinoma |
| $P14^{ARF}$ | Cyclin-dependent kinase inhibitor 2A gene |
| *P15* | Cyclin-dependent kinase inhibitor 2B gene |
| *P16* | Cyclin-dependent kinase inhibitor 2A gene |
| *PAK3* | p21 protein (Cdc42/Rac)-activated kinase 3 gene |
| PCR | Polymerase chain reaction |
| PET | Positron emission tomography |
| *PGP9.5* | Ubiquitin carboxy-terminal esterase L1 gene |
| PIK3CA alpha | Phosphatidylinositol-4,5-bisphosphonate 3-kinase, catalytic subunit gene |

| | |
|---|---|
| pN | Pathologic nodal stage |
| PORT | Post-operative radiotherapy |
| pT | Pathologic tumour stage |
| *PTEN* | Phosphatase and tensin homolog gene |
| *PTPRD* | Protein tyrosine phosphatise receptor type D gene |
| *PVA* | Pemphigus vulgaris antigen gene |
| PVL | Proliferative verrucous leukoplakia |
| *PXN* | Paxillin gene |
| qMSP | Quantitative methylation specific polymerase chain reaction |
| qPCR | Quantitative polymerase chain reaction |
| *RARB* | Retinoic acid receptor, beta gene |
| *RASSF1A* | Ras association domain family member 1 gene |
| RCT | Randomised control trial |
| *RIZ1* | PR domain containing 2, with 2F domain gene |
| RLGS | Restriction landmark genome scanning |
| ROC | Receiver operating curve |
| RR | Relative risk |
| *S100P* | S100 calcium binding protein gene |
| *SAT* | Spermidine/spermine N-acetyl transferase 1 gene |
| SCC | Squamous cell carcinoma |
| SERPINE1 | Serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1 protein |
| *SERPINB* | Serpin peptidase inhibitor, clade B (ovalbumin) gene |
| *SMAD4* | SMAD family member gene |
| SNLB | Sentinel lymph node biopsy |
| SNP | Single-nucleotide polymorphism |
| *SPP1* | Secreted phosphoprotein 1 gene |
| SPSS | Statistical package for the social sciences |

| | |
|---|---|
| SPT | Second primary tumour |
| STAT | Signal transducer and activation transcription factor |
| *TACSTD1* | Tumour associated signal transducer 1 gene |
| *TERT* | Telomerase reverse transcriptase gene |
| *TIMP* | TIMP metallopeptidase inhibitor 3 gene |
| *TMEFF2* | Transmembrane Protein with EGF-Like and Two Follistatin-Like Domains 2 gene |
| TNM | Classification of malignant tumours |
| *TP53* | Tumour protein 53 gene |
| *TP73* | Tumour protein 73 gene |
| TSG | Tumour suppressor gene |
| *TTK* | TTK protein kinase gene |
| UK | United Kingdom |
| UNG | Uracil-N-glycosylase |
| UV | Ultra violet |
| WGA | Whole genome amplified |
| WP | Weighted prevalence |
| *WT1* | Wilms tumour 1 gene |
| *VGF* | VGF nerve growth factor inducible gene |
| VIC | Fluorescent dye |

## Abstract

### Background

Oral squamous cell carcinoma (OSCC) is the 15th most common cancer worldwide but has poor five year survival (50%). Late stage presentation and limitations of early diagnostic techniques are persistent clinical problems. Sixty percent of patients present with advanced stage disease and with the attendant increase in mortality, morbidity and risk of recurrent disease it is particularly burdensome for both patients and health economies. Early diagnosis and treatment of OSCC improves prognosis. There is an opportunity to diagnose OSCC early in patients with oral epithelial dysplasia however currently there is no way of accurately predicting which lesions will undergo malignant transformation. Aberrant methylation of tumour suppressor genes plays a significant role in the biology of early cancer and is detectable in both tumour and saliva. Saliva is a non-invasive method of longitudinal sampling and has potential as a tumour surrogate in disease surveillance programmes. This study aims to compare rates of methylation of a panel of genes in OSCC patients and a normal cohort to establish a threshold by which we could determine future disease testing in a dysplastic population.

### Methods

Saliva samples were collected from 219 individuals from three diagnostic groups: Normal (defined as no oral malignant or premalignant disease) n=97, OSCC n=62 and dysplasia n=60. For statistical analysis the dysplasia cohort was sub-divided into lesions of low and high risk of malignant transformation based on the histological diagnosis of the index lesion. DNA was extracted and bisulphite treated from 258 saliva samples before duplex quantitative methylation specific PCR (qMSP) assays were performed on all samples to detect the frequency of methylation in saliva of a panel of genes. The five target genes (*ADAMTS9, CCNA1, CYGB, P16, TMEFF2*) were selected using a candidate approach on the basis of tumour specificity from studies on tumour/normal matched tissue pairs. Clinicopathological data was correlated with the qMSP data and analysed using SPSS v.21 statistical software to look for associations with tumour and survival characteristics.

**Results**

Only 3/97 individuals from the control normal cohort had saliva samples with detectable methylation above the analytical sensitivity of the *P16* assay. Methylation of the remaining target genes (*ADAMTS9, CCNA1, CYGB, TMEFF2*) was not detected in normal saliva at levels above the analytical sensitivity of the qMSP assays. The most significant finding in this study was that methylation of four of the target genes (*CCNA1, CYGB, P16, TMEFF2*) in saliva, individually and when considered as a panel, was significantly associated with OSCC and as such could aid discrimination between malignant disease and normal saliva samples. Methylation of at least one gene in the panel was discovered in 29/67 of the binned OSCC saliva samples but only 3/97 of normal samples (Fisher's exact p=0.001). Furthermore methylation of the gene panel is associated with high risk lesions when detected in saliva of patients with premalignant lesions (Fisher's exact p=0.03).

**Conclusions**

This exploratory data supports the utility of duplex qMSP as a detection method for methylation markers in saliva. The detection of methylation of this gene panel in saliva is significantly more associated with oral malignancy and high risk premalignant lesions than normal and low risk disease. This implies saliva may have merit as a surrogate tissue in an adjunctive role to clinical assessment and biopsy. The assays are specific but have limited sensitivity. However with further work, inclusive of additional genes, this methodology may identify predictive biomarkers that can be introduced into a trial surveillance of premalignant lesions.

## Declaration

I declare that the work contained in this thesis is my own. It was carried out in the Department of Molecular and Clinical Cancer Studies in the Institute of Translational Medicine. Professor Richard Shaw and Dr Janet Risk provided guidance for the study design and commentary on the written thesis. Dr Lakis Liloglou provided the qMSP primer/probe assay designs and support with statistical analysis of the qMSP data. Dr Ratna Veeramachaneni provided the pyrosequencing data for the target gene *ADAMTS9.*

## Acknowledgements

Thank you to my supervisors Dr Janet Risk and Professor Richard Shaw for providing me with the opportunity to do this MD. I am extremely grateful for their patience, encouragement and support to complete it. I would like to express my gratitude to the British Association of Oral and Maxillofacial Surgeons and the British Association of Head and Neck Oncologists for funding this work. Thank you to Dr Lakis Lilogou who has patiently taught me about qPCR and provided me with reassurance and encouragement when I got disheartened with misbehaving qPCR reactions. Dr George Nikolaidis and Mr Ben Brown were also very helpful in providing technical advice in the laboratory. Thank you to all of the patients who provided samples for this study and the clerical and nursing staff who helped me with the logistics of collection.

Massive thanks to Tristan and my mum and dad who have cheered me on to the finish line, providing endless hours of childcare and cups of tea. Thank you to Tallulah who has patiently tolerated my sitting at a desk instead of playing princesses and knights.

# 1. Introduction

## 1.1 Oral Squamous Cell Carcinoma and Oral Premalignant Disease

Oral cancer is malignancy within the oral cavity which is defined as the structures occupying the space between the vermillion border of the lips to the junction between the hard and soft palate. The most common anatomical site for a lesion to arise is oral tongue and floor of mouth representing approximately sixty percent of presentations worldwide (De Camargo Cancela, Voti et al. 2010). The floor of mouth and associated sulci are considered areas of increased susceptibility due to the theory that carcinogens are likely to pool in this area (Barnes 2005) exposing cells to longer periods of tumourigenic substances. Squamous cell carcinoma represents ninety percent of oral cancer. It is an invasive epithelial neoplasm with varying degrees of squamous differentiation with a propensity to metastasize to regional lymph nodes (Barnes 2005) and is the disease of interest in this thesis.

Oral cancer can develop de novo from ostensibly normal mucosa or from premalignant lesions. The most common precursor clinical lesions of oral squamous cell carcinoma (OSCC) are leukoplakia (white oral mucosa lesion) and erythroplakia (red oral mucosal lesion) (Brennan, Migliorati et al. 2007). The global prevalence of leukoplakia is estimated between 1.7-2.7% with an annual malignant transformation rate of 1.36% (Petti 2003). Leukoplakia most commonly occurs on the buccal mucosa but floor of mouth and lateral border of the tongue are associated with high risk lesions (Napier, Speight 2008). Other risk factors of statistical significance for the transformation to OSCC are female gender, non-smokers, non-homogenous appearance, lesion size greater than 200$mm^3$, DNA aneuploidy and duration of lesion (van der Waal 2010). Dysplasia is defined as architectural change associated with cytological atypia within the epithelial cells and presents as a spectrum of change (Warnakulasuriya S, Reibel J et al. 2008). However, it is histologically graded as mild, moderate, severe and carcinoma-in-situ (Barnes 2005). The presence and severity of dysplasia within leukoplakic lesions as determined by histopathological assessment is currently the best predictor of malignant transformation (Smith, Rattay et al. 2009) but the presence of dysplasia does not guarantee progression to cancer as some lesions without dysplasia or low grade dysplasia will transform whilst a proportion of high grade dysplasia will not (Reibel 2003).

### 1.1.1 Epidemiology of OSCC

Globocan reports the global burden of lip and oral cancer as 300373 new cases and 145353 deaths of the disease per annum in 2012 (Ferlay, Soerjomataram et al. 2012). It is the 15th most common cancer globally. In the UK there were 4986 new cases of lip and oral cancer and 1296 deaths (Ferlay, Soerjomataram et al. 2012). In the UK it more commonly affects men at a ratio of 20:10 and is more prevalent in deprived areas which tend to have heavier alcohol and smoking habits (CRUK 2014b). It presents a formidable health problem not least because there has been marginal improvement in survival despite advancement in surgical and adjunctive therapy. This is in part because late stage disease presentation is common with up to 60% of patients presenting with stage III/IV disease (McGurk, Chan et al. 2005, Rogers, Brown et al. 2009). The clinical sequelae of late stage disease and its treatment can also have a negative impact on quality of life which can require long term post-treatment clinical support for swallowing, nutrition and speech function.

Cancer Research UK (CRUK) reports the incidence of OSCC has steadily increased in the UK (CRUK 2014b). An expanding aging population is likely to contribute to a rise in cancer in future years. A population, surviving longer with chronic disease, provides a challenge for the current, primary therapeutic modality of surgery, as people at the limits of physiological reserve are unlikely to withstand the aggressive surgical or chemo-radiation therapy required for advanced oral cancer therefore early disease detection may be particularly important in this group. Although primarily a disease of older age, with a peak incidence in UK males at age 60-64 years, there has recently been an observed increase of OSCC in subgroups of young people who have had little or no exposure to the dominant risk factors of tobacco and alcohol consumption (Schantz, Yu 2002, Llewellyn, Linklater et al. 2003). An increased incidence of non-HPV related oral tongue SCC was noted in a US population of young white women; of particular note they are non-smokers and non-drinkers (Patel, Carpenter et al. 2011). It may be that intrinsic molecular factors are responsible for the onset of OSCC in these patients. Promoter methylation of *P16* has been associated with non-HPV anterior tongue lesions in female patients under 40 (O'Regan, Toner et al. 2008).

The socioeconomic impact of an increasing incidence of OSCC, particularly late stage disease in a young population would be significant and highlights the need for risk stratification and early disease detection. Risk stratification becomes more challenging in individuals developing OSCC who are not exposed to accepted risk factors. As knowledge of

the molecular profiles of these lesions becomes more advanced molecular biomarkers may increase the sensitivity of clinical assessment for the detection of OSCC in these atypical groups.

### 1.1.2 Aetiological Factors

There is global variation in the aetiological factors of OSCC however worldwide tobacco and alcohol consumption remain the dominant risk factors for OSCC. Other factors such as oral hygiene, diet and viruses have also been implicated in the aetiology of OSCC but these remain speculative rather than established factors (Warnakulasuriya 2009).

#### 1.1.2.1 Tobacco and Alcohol

The overall risk of oral cancer among smokers is 7–10 times higher than for those who never smoked (Warnakulasuriya, Sutherland et al. 2005). A large case-control study from the U.S.A showed that tobacco and alcohol consumption in isolation increased the risk of oral cancer with a dose-response relationship. This risk was further increased if substances were consumed together (35 fold increase in those who smoke 40 + cigarettes/day and consume > 28 drinks/week) as they have a synergistic action (Blot, McLaughlin et al. 1988). Alcohol may have a locally permeabilizing effect on oral mucosa which facilitates the penetration of tobacco-specific and other carcinogens across oral mucosa (Du, Squier et al. 2000, Seitz, Stickel 2007). Chemical carcinogens are thought to induce molecular changes which can interfere with DNA repair (Seitz, Stickel 2007) and normal function of tumour suppressor genes (Brennan, Boyle et al. 1995). Consumers of tobacco and alcohol are therefore an 'at risk' population for OSCC and may benefit from screening programmes to detect early disease.

Tobacco use is the commonest pre-disposing factor for the development of leukoplakia (Napier, Speight 2008) and the relative risk of having a dysplastic lesion for smokers compared with non-smokers or ex-smokers was estimated at 7.0 (Kulasegaram, Downer et al. 1995). However, a recent study of patients with biopsy proven oral epithelial dysplasia (OED) observed that non-smokers were seven times more likely to develop malignant transformation than heavy smokers (Ho, Risk et al. 2012). Earlier longitudinal studies also support the finding that OSCC develops at higher frequencies in oral premalignant lesions from non-smoking patients (Silverman Jr., Gorsky et al. 1984) particularly female

(Schepman, Van Der Meij et al. 1998). This data suggests an aetiological mechanism other than tobacco carcinogens in this group of OED patients.

#### 1.1.2.2 HPV

In recent decades Europe has witnessed an upsurge in young patients presenting with head and neck squamous cell carcinoma (HNSCC) not entirely attributable to the risk factors commonly associated with the disease (Franceschi et al. 1994). A viral role in the aetiology of HNSCC was put forward after correlations were noted between cancers known to have a Human Papilloma Virus (HPV) driven pathogenesis and oral tumours (Bjørge, Hennig et al. 1995)

High risk strains such as HPV 16 are thought to be contributing to the increase in HNSCC diagnosis but this appears to be predominantly oropharyngeal and tonsillar SCC (Schwartz, Daling et al. 1998, Gillison ML, Koch WM et al. 2000, Kreimer, Clifford et al. 2005). The aetiological role in oral cavity cancer and clinical outcomes is less certain, with great variation in the published rates of HPV found in OSCC lesions. The wide variation in the reported HPV rates is likely due to methods of detection as PCR techniques are extremely sensitive and vulnerable to contamination. A multi-centre case control study found HPV in 3.9% of OSCC versus 18.3% of oropharynx and tonsil cancers (Herrero, Castellsagué et al. 2003). Isayeva (Isayeva, Li et al. 2012) performed a systematic review on sixty publications limited to PCR studies of HPV in non-oropharyngeal SCC. They found a 20.2% weighted prevalence (WP) of HPV (all types) in OSCC which compared to a 6.9% WP in normal controls suggesting HPV may contribute to oral carcinogenesis. Other reviews support this figure; Kreimer et al (Kreimer, Clifford et al. 2005) report a 23.3% cumulative pooled prevalence and Syrjanen 3.98 odds ratio of HPV (all types) in OSCC versus controls (Syrjänen, Lodi et al. 2011).

However the current evidence only supports an association between HPV and OSCC and not a causal role. Importantly, unlike in oropharyngeal and tonsil SCC, there is no strong body of evidence that the presence of HPV predicts for improved clinical outcome in OSCC (Isayeva, Li et al. 2012) therefore HPV as a biomarker to risk stratify OSCC patients is currently contentious.

#### 1.1.2.3 Oral Hygiene and Microbiotia

Poor oral hygiene and the resultant disruption in the normal oral microbiotia have been mooted as a potential risk factor for OSCC. A causal link between carcinogenesis and infection has been established in other cancers (Parkin 2006). The evidence for infection and OSCC is less clear; whether it is aetiological or a bystander event has yet to be answered (Meurman, Bascones-Martinez 2011). Tezal et al (Tezal, Sullivan et al. 2009) report the presence and severity of chronic periodontitis as an independent risk factor for HNSCC when compared with a control population (odds ratio 4.36). Additionally the presence of periodontitis was more likely to be associated with poorly differentiated OSCC. However, the data should be interpreted with caution as the measures of periodontitis were also significantly associated with smoking. The significance of the association between periodontitis and HNSCC was weakened when corrected for current smoking habit (odds ratio 2.86). A 'modest' increase in the association of HNSCC with periodontitis was noted in a separate study but it did not persist in non-smokers (Divaris, Olshan et al. 2010). Similarly a recent large multi-centre case control study concluded that indicators of poor oral health and dental care were independent risk factors for upper aerodigestive cancers but this risk was greatly attenuated when correcting for the confounders of smoking and alcohol (Ahrens, Pohlabeln et al. 2014).

#### 1.1.2.4 Diet

The majority of the evidence regarding oral cancer and dietary habits is from retrospective case-control studies therefore attributing a direct causal role to diet is difficult. Petridou et al (Petridou, Zavras et al. 2002) performed a questionnaire based observational study on dietary intake of 106 individuals with OSCC and matched controls. The researchers report a statistically significant inverse relationship between riboflavin, magnesium and iron intake and oral carcinoma. However, the sample size is small and the data reliant on patient reports, which is vulnerable to information bias. It is unclear exactly what element of vegetables is beneficial. Vegetables are high in antioxidants and vitamins; this has led some to suggest a diet deficient of carotenoids, folates, vitamins and fibre may be linked with upper aerodigestive carcinogenesis (Franceschi, Favero et al. 1999, Kane 2005, Rossi, Garavello et al. 2007). Others speculate that those individuals who indulge in high risk lifestyles such as heavy smoking and alcohol intake are likely to consume less fruit and vegetables and rather than a causal role for diet in cancer, it is merely associative (Serdula, Byers et al. 1996, Boffetta, Couto et al. 2010). Currently, there is only speculative evidence that dietary factors, other than alcohol intake is causative in OSCC.

### 1.1.3 Diagnosis and Treatment of OSCC

The current gold standard for OSCC diagnosis is clinical examination and tissue biopsy followed by staging investigations. A typical presentation of OSCC is a discrete lump or ulcer with raised margins present for a number of weeks (Shaw, Pace-Balzan et al. 2011) which may be noticed by a vigilant patient but many go unnoticed. Symptomology is often lacking until advanced stages and is very likely a contributing factor to the persistent trend of late presentation disease (McGurk, Chan et al. 2005). A large retrospective study observed pain was only reported at initial presentation in 20% of 1412 oral cancer cases and predominantly in advanced disease (Cuffari, Tesseroli de Siqueira et al. 2006). A recent review of the literature revealed that the reasons for late presentation are complex and multi-factorial; health beliefs, deprivation, lack of education and alcohol consumption are commonly implicated (Noonan 2014). It is accepted that the incidence of OSCC is highest in lower socioeconomic groups (O'Hanlon, Forster et al. 1997, Thorne, Etherington et al. 1997, Conway, Brewster et al. 2007) and established risk factors for OSCC such as high alcohol and tobacco consumption are also associated with deprivation (O'Hanlon, Forster et al. 1997) which are very likely, but not exclusively, causative factors in this group. Conway et al (Conway, Brewster et al. 2007) consider those subject to deprivation alone, regardless of lifestyle choices, represent a high risk group for OSCC.

The initial risk stratification for potentially malignant oral lesions is currently reliant on clinical examination and history. Clinicians have to be vigilant in high risk groups but also have a high index of suspicion, when presented with persistent oral lesions in high risk anatomical sites, in patients who are not exposed to the common risk factors. Koo et al (Koo, Barrowman et al. 2013) report a subgroup of non-smoking, elderly female patients with OSCC of the oral tongue and maxillary alveolus who had worse disease-specific mortality than those exposed to alcohol and tobacco. It is in these atypical groups that better understanding of the molecular biology of tumours may provide prognostic biomarkers to tailor treatment choices. Two-week referral pathways are established in the UK to expediate access to specialist opinion for suspicious oral lesions with an aim to reduce late stage presentations. An audit of 1079 two week referrals indicated 118/1079 (10.9%) referred as possible cancer had HNSCC whereas only 21.4% of HNSCC diagnoses over the same time-period were identified by the two-week referral system, the majority diagnosed from routine referrals (McKie, Ahmad et al. 2008). These findings could reflect the observation that some non-specialists have limited awareness of oral cancer (Carter,

Ogden 2007) or that clinical discrimination between non-malignant and malignant lesions is difficult. This adds further support for adjunctive biomarker tests which could be used in the community setting.

Primary surgery remains the first-line treatment in the majority of OSCC cases. In a review of ten years of clinical practice in a large head and neck unit; four hundred and eighty nine of five hundred and forty one patients (90%) had primary surgery for OSCC, 40% of these had adjuvant radiotherapy (Rogers, Brown et al. 2009). Neck dissection is performed routinely if there is nodal involvement but there is still debate about the most appropriate management of the clinically node negative neck (N0). Diagnostic imaging techniques are limited in the clinically negative neck as a minimum size of disease is required for detection which means micrometastases can be missed (Takes 2004). Occult metastasis has been found in 34% of oral tongue carcinoma cases with clinically N0 necks (Greenberg, El Naggar et al. 2003). Generally, neck dissection is performed in the N0 neck if there is a 20% or more chance the nodes harbour occult metastasis (Shaw, Pace-Balzan et al. 2011) which will mean some patients receive a neck dissection unnecessarily. Conversely, not treating the neck and adopting a watchful waiting strategy risks progression of occult metastasis to incurable disease (Takes, Rinaldo et al. 2008). In recent years multiple studies have shown the ability of sentinel lymph node biopsy (SNLB) to accurately stage the neck when compared to elective neck dissection in early stage oral cavity cancer (T1/T2) with the exception of floor of mouth tumours. Negative predictive values between 91-100% are reported without a reduction in survival (Monroe, Lai 2014). The advantage of SNLB is it potentially has reduced morbidity. Molecular analysis of the lymph node intra-operatively may further improve SNLB accuracy. A recent study of HNSCC lymph nodes has shown rapid qPCR techniques using a panel of two markers pemphigus vulgaris antigen (*PVA*) and tumour associated signal transducer 1 (*TACSTD1*) genes can detect nodal spread with a negative predictive value of 96% (Ferris, Stefanika et al. 2012). The molecular biology of the primary tumour may provide further insight into the issue of detecting cervical metastasis in the clinical N0 neck. Gene expression profiles within the primary HNSCC tumour may predict the absence or presence of lymph node metastasis (Roepman, Wessels et al. 2005). In a microarray study of primary OSCC tissue it was noted that increased expression of extracellular matrix-degrading enzymes were associated with OSCC cases with lymph node metastasis. In particular *MMP-1* was highly expressed in these cases and

may be a prognostic marker of lymph node involvement (Nagata, Fujita et al. 2003). Expression of SERPINE1 and SMA proteins in the invading front of the primary OSCC tumour are associated with increased sensitivity (95% and 82% respectively) for the detection of extra-capsular spread (ECS) when compared to the standard clinical technique of magnetic resonance imaging (MRI) (56%) (Dhanda, Triantafyllou et al. 2014).

### 1.1.4 Prognosis of OSCC

The difference in disease specific five year survival for late stage cancer (stage 4) 57% and early disease (stage 1) 96% is considerable (Rogers, Brown et al. 2009). In patients with nodal involvement the single most important measurable prognostic factor is the presence of ECS in nodal metastasis. In one study the presence of ECS doubled the incidence of local and regional recurrence, tripled the incidence of distant metastasis and patients with macroscopic ECS had only a 23% overall survival (Shaw, Lowe et al. 2010). Other tumour and patient characteristics that impact on recurrence and survival are involved margins, advanced T-stage inclusive of depth of invasion and age of patient (Hicks Jr., North Jr. et al. 1998, Koo, Lim et al. 2006, Rogers, Brown et al. 2009 ). The sequelae of current therapeutic modalities for late stage disease can be functionally and aesthetically mutilating despite advances in reconstructive techniques. The compact nature of the anatomy also lends itself to involvement of other functionally important structures which unfortunately, can render late stage disease inoperable. Loco-regional recurrence for OSCC has been reported at rates of 21-28% (Hicks Jr., North Jr. et al. 1998, Koo Lim et al. 2006, Rogers, Brown et al. 2009) with the majority of recurrent disease occurring within the first year of primary treatment (Koo, Lim et al. 2006, Kissun, Magennis et al. 2006). Second primary tumours (SPT) have been observed in approximately 10-20% of OSCC patients (León, Quer et al. 1999, Lin, Patel et al. 2005, Rogers, Brown et al. 2009) with head and neck primary tumours being particularly susceptible to the development of second primary malignancy (Sturgis, Miller 1995).

Currently there is no method of predicting, with certainty, which patients will experience secondary disease; generally this has poor survival outcomes. Ten year survival for HNSCC patients who developed a secondary primary tumour (22%) was half that of those without a secondary neoplasm (55%) (León, Quer et al. 1999).

In a large study of long-term survivors of HNSCC (defined as > 3 years from primary treatment) the cause of death from other site second primary malignancy was 23% (3007/13120) (Baxi, Pinheiro et al. 2014). Overall survival for recurrent OSCC is approximately 30% (Mücke, Wagenpfeil et al. 2009, Yanamoto, Yamada et al. 2012) with early stage (I and II) recurrence (as measured by recurrent tumour size) having significantly improved overall survival (p=0.0001) than advanced stage (III and IV) recurrence (Sun, Tang et al. 2009). A prospective study of salvage surgery outcomes in a mixed head and neck cohort (n=109) showed that median 2 year disease free survival decreased with increasing stage of recurrent disease: recurrence stage I: 73%, stage II: 67%, stage III: 33% and stage IV: 22% (Goodwin Jr. 2000). The study contained 21 oral cavity cancers and although this data was not stratified by anatomical site, it does highlight the importance of early detection and the need for close disease surveillance. Current post-treatment surveillance regimes involve regular attendance to out-patient appointments for clinical examination to detect signs of recurrent/second primary disease for a period of five years. However, disease surveillance by clinical examination becomes problematic in the aftermath of treatment because normal anatomical architecture is altered. A subtle early malignant lesion may be camouflaged by surgically scarred or irradiated tissue or alternatively clinicians may biopsy post-treatment friable tissues unnecessarily.

### 1.1.5 Precancerous fields and disease surveillance

The propensity for recurrent and second primary tumours in HNSCC has been acknowledged since the 1950s when Slaughter et al introduced the concept of field cancerisation (Slaughter, Southwick et al. 1953). They reported, of the 783 oral tumour cases reviewed, all had histologically atypical epithelium surrounding the primary tumour and 11.2% of patients (88/783) had independent multiple tumours (Slaughter, Southwick et al. 1953); a figure which has not been improved upon today. More recently numerous studies of surgical resection margins and premalignant lesions have shown that along with histological atypia there is also a molecular basis for the process of field cancerisation involving *TP53* mutation and LOH at 3p,9p,17p (Brennan, Mao et al. 1995, Tabor, Brakenhoff et al. 2002, Tabor et al. 2002, van Houten, Tabor, Brakenhoff et al. 2004, van Houten, Leemans et al. 2004) which may predict for recurrence or second field tumours. An analysis of tissue from 28 primary HNSCC tumours and associated macroscopically normal mucosa biopsies was performed for the presence of LOH at chromosome loci 3p, 9p and 17p. Thirty six percent of patients (10/28) had detectable chromosomal aberrations in at

least one of the markers in the mucosa surrounding the tumour with 70% (7/10) of cases extending beyond the surgical margins (Tabor, Brakenhoff et al. 2001). In 25% (4/10) of the patients the field contained genetic losses not found in the tumour (Tabor, Brakenhoff et al. 2001) which is some evidence for a separate precursor field. In a follow-up study of clinically determined second primary tumours they found 60% (6/10) of patients had similar chromosomal aberrations in the index tumour, second tumour and intervening mucosa which suggests these tumours are better defined as a second field tumour, as they share the same clonal origin as the primary disease (Tabor, Brakenhoff et al. 2002). The authors also note the field can be as large as 3-6cm (Tabor, Brakenhoff et al. 2002) which has implications for the suitability of these areas for surgical treatment.

Oral precursor lesions represent a group which requires disease surveillance, as in some cases, these lesions are a clinical expression of field change. A recent retrospective longitudinal surveillance study of 91 patients with OED under specialist review showed that early detection of malignancy (23/91) resulted in all of the patients receiving treatment at stage one (clinical T1N0M0) disease. Twenty one of twenty three of the patients required minimal surgery with wide-local excision of the lesion. The authors report one death resulting in an overall survival of 96% and a disease-specific survival of 100% (Ho, Field et al. 2013). Their findings support the need for long-term disease surveillance for OED by specialists which can be costly. Twenty five percent of patients underwent malignant transformation in this cohort. However, 75% did not and it is the future hope that new techniques will develop to accurately predict lesions at risk of malignant transformation to allow limited resources to be focused on this subgroup of patients. Reported follow-up intervals of premalignant lesions are widely variable as they are dictated by subjective clinical assessment and are not evidence based; some studies report surveillance periods of up to 18 years (Holmstrup, Vedtofte et al. 2006).

Accurately determining which lesions will progress is an ongoing problem and many patients are reviewed for long periods of time without malignant transformation. Histopathological examination remains the most valid method of predicting malignant change within premalignant lesions, although it is vulnerable to subjectivity and lacks intra and inter-observer reproducibility due to an insufficiency of validated morphological criteria (Fleskens, Slootweg 2009). It is generally accepted that severe dyplasia is associated with increased malignant transformation (Schepman, Van Der Meij et al. 1998, Mehanna, Rattay et al. 2009, Warnakulasuriya, Kovacevic et al. 2011), however some studies have

found the grade of dysplasia does not influence this outcome (Holmstrup, Vedtofte et al. 2006). Rosin et al (Rosin, Lam et al. 2002) noted that 22/47 lesions with mild or no dysplasia at previous oral cancer sites progressed to malignancy. These contradictory findings may be explained by the limitations of surgical biopsy. Researchers analysing 101 lesions, which underwent surgical biopsy by experienced oral surgeons followed by total excision of the lesion, observed that 35% of the initial biopsy samples under diagnosed the severity of the lesion including 8% of oral cancer. In 17% of lesions the biopsy over diagnosed the severity of the entirety of the lesion (Holmstrup, Vedtofte et al. 2007). This study points to the cellular heterogeneity of the dysplastic lesion and also the limitations of current sampling techniques upon which the majority of treatment planning is based.

In the context of field cancerisation, uncertainty also surrounds the benefit of surgical excision of premalignant lesions for conferring protection against the development of OSCC. A published review of the literature comments on the lack of randomised controlled studies on this subject and identified only two single-centre retrospective studies which specifically addressed outcomes of treatment for OED. Both studies found no difference in the development of OSCC between OED patients treated with surgery and those without treatment (Balasundaram, Payne et al. 2014). In contradiction to this, a meta-analysis of the literature found that patients who did not have excision of lesions had significantly higher rates of malignant transformation compared to excised lesions (Mehanna, Rattay et al. 2009). Surgery remains the mainstay of treatment as there is currently no evidence for a more successful chemoprevention therapy (Brennan, Migliorati et al. 2007, Sheth, Johnson et al. 2014). There are still large gaps in the knowledge of the natural history of OED (Napier, Speight 2008) but with better understanding of the molecular biology of lesions these gaps may be filled. The presence of molecular markers such as LOH at 3p and/or 9p and matrix metalloproteinases in particular (MMP-9) in precursor lesions have shown promise in cancer risk stratification with the relative risk of transformation reported at 17-19 with the presence of these markers, but further validation in clinical studies is needed (Brennan, Migliorati et al. 2007, Smith, Rattay et al. 2009, Dionne, Warnakulasuriya et al. 2015).

### 1.1.6 Screening

Screening is a process of identifying apparently healthy people who may be at increased risk of a disease or a condition. They can be offered information, further tests and

appropriate treatment to reduce their risk and or complications arising from the disease or condition (www.screening.nhs.uk/screening). The current UK National Screening Committee (NSC) criteria to which a valid screening test must adhere is an expansion and modernisation of the Wilson and Jungner (Wilson, Jungner 1968) principles of screening for disease. In brief, the condition should be important, have a recognisable early symptomatic stage and the natural course of the disease from latent to declared disease should be adequately understood. A suitable test or examination which is acceptable to the population should exist and the case findings should be continuous. An effective treatment for patients with recognised disease, facilities for diagnosis and treatment and agreed policy of whom to treat should be available. The cost of case finding should be economically balanced in relation to expenditures on medical care as a whole (Wilson, Jungner 1968).

The aim of screening is to detect oral cancer early and improve health outcomes. Visual inspection is the most common method of oral screening as it is not surgically invasive, painless and is socially acceptable (Brocklehurst, Kujan et al. 2013). Authors of a meta-analysis selected eight clinical examination screening studies and reported high sensitivity and specificity for this technique at 85% (95% CI: 73-91.9%) and 96.5% (95% CI: 93-98.2%) (Downer, Moles et al. 2004). A more recent Cochrane review reported a widely varying sensitivity at 50-99% and specificity of 98% for clinical examination (Walsh, Liu et al. 2013). Several clinical adjuncts such as autofluorescence, chemiluscence, toluidine blue and brush biopsy are under investigation to enhance clinical examination and improve the screening process. However, none have proven to have increased sensitivity or specificity above clinical examination and there is lack of evidence for impact on survival and recurrence outcomes (Lingen, Kalmar et al. 2008, Walsh, Liu et al. 2013, Messadi 2013).

Oral cancer screening studies are limited by methodological heterogeneity, inadequate follow-up and lack of reported population mortality (Downer, Moles et al. 2004, Downer MC, Moles DR et al. 2006). A recent Cochrane review to assess the effectiveness of oral screening methods identified only one randomised control trial (RCT) from the global literature. The RCT was executed in Kerala India, with 191,873 participants divided into an intervention and a control arm studied over 15 years. The difference in mortality rate as reported by the Cochrane review was 15.4/100000 person years for the intervention group

and 17.1/100000 person years for the control group which did not reach statistical significance. In those who consumed alcohol, the mortality rate was 30/100000 person years in the intervention group and 39/100000 person years in the control group which was statistically significant RR=0.76 (CI: 0.60-0.97). Individuals diagnosed with oral cancer stage III or above numbered 147/279 (52.6%) in the intervention arm and 159/244 (65.2%) in the control arm which also reached statistical significance RR=0.81 (95% CI: 0.70-0.93). The sensitivity of visual inspection in detecting oral cancer was 67.4% (Brocklehurst, Kujan et al. 2013). This study showed evidence of cost-effectiveness, stage-shift effect and improved survival but this finding was limited to alcohol and tobacco users. The Cochrane group felt the study was limited by bias and further RCTs in different populations are required to establish the effectiveness of oral cancer screening by visual inspection.

Currently a screening programme for oral cancer in the UK does not exist. Following an external review of screening for oral cancer in 2010, a systematic population screening programme was not recommended but is currently under review by the UK NSC. In summary, the authors of the 2010 evaluation implied that oral cancer met some of the screening criteria, in that OSCC is an important health problem; being a debilitating disease; with poor prognosis and in the majority of cases is preceded by a preclinical phase manifest as a potentially malignant lesion. However, visual screening inspection studies have not yet yielded convincing evidence for improvement in survival or stage shift and there are currently no alternatives or adjunctive diagnostic techniques that have been evaluated as screening tests. Amongst the recommendations is the development of point-of-care tests using biomarkers which could identify which screen-detected lesions are most likely to progress to malignancy (Speight, Warnakulasuriya 2010).

### 1.2 Biomarkers

The National Institute of Health Biomarkers Definitions Working group defined a biomarker as "*a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes or a pharmacologic process response to a therapeutic intervention*" (Atkinson, Colburn et al. 2001). An ideal biomarker for early detection of cancer would be one secreted by tumour but not normal tissue therefore highly specific and be easily and cheaply detected in a body fluid so it can be detected non-invasively (Pepe, Etzioni et al. 2001). The issue of specificity is important when considering biomarkers for screening programmes as even a small false positive rate when multiplied

on a population scale will be a large financial burden and potentially cause psychological stress to individuals (Pepe, Etzioni et al. 2001).

The improvement of morbidity and mortality is a primary aim of clinical research and as such are clinical endpoints (Strimbu, Tavel 2010). Biomarkers become surrogate endpoints when they have been proved to have well evaluated clinical relevance (Atkinson, Colburn et al. 2001); there must be robust scientific evidence that clinical outcome is consistently and accurately predicted by a biomarker (Strimbu, Tavel 2010). The process of developing a clinical biomarker is therefore necessarily rigorous and must undergo several phases (Pepe, Etzioni et al. 2001) from discovery to clinical trials before they become part of routine patient management. At present there are no molecular biomarkers routinely used for the diagnosis or prognostication of oral cancer.

### 1.2.1 Diagnostic biomarkers

In a disease that often presents late with associated poor prognosis, has identifiable high risk groups and limitations on detection from clinical examination, a biomarker with satisfactory accuracy for identifying OSCC early would be valuable. Clinical examination is the mainstay of oral cancer diagnosis at present but disease can be occult and evade the naked eye. The pressing need for improved diagnostics is because of the marked differential in survival rates between early and advanced stage cancer. There is a trend for late presentation which is likely multifactorial (Noonan 2014) but could be due to the fact that OSCC is largely an asymptomatic condition, especially in the early stages. Another reason could be that aggressive tumours progress rapidly. The sensitivity of visual inspection for the detection of oral cancer in primary care settings can be variable (Walsh, Liu et al. 2013) with limitations of non-specialists to recognise cancerous oral lesions (Carter, Ogden 2007). To date studies of opportunistic oral screening studies based on visual inspection in unstratified populations have not made significant improvement in mortality through early detection (Brocklehurst, Kujan et al. 2013). Added to this, risk stratification based on age and lifestyle factors associated with OSCC may not be sufficient; as a subgroup of young patients are developing OSCC with little or no exposure to the common risk factors (Llewellyn, Linklater et al. 2003).

### 1.2.2 Prognostic biomarkers

Prognostic biomarkers predict how the malignancy will progress and the patients' overall outcome without therapy (Oldenhuis, Oosting et al. 2008). HPV in oropharyngeal SCC is an example of a prognostic molecular marker that successfully stratifies patients into predicted outcomes and informs treatment strategy (Rios Velazquez, Hoebers et al. 2014). Currently, OSCC prognostication is informed by the American Joint Committee on Cancer and the Union for International Cancer Control TNM staging criteria which is an anatomical classification. In a comprehensive review Takes and colleagues (Takes, Rinaldo et al. 2010) discuss the limitations of this system in the context of prognosis and patient management. A key criticism is that the TNM stages are too broad and as a result there is a wide range of survival and treatment response within any given stage. They propose this is contributory to continued poor survival rates and morbidity as a result of over or under treatment. They also raise the lack of consideration for the biological nature of the tumour and highlight the fact that a T1N2C tumour would receive the same 'stage' as a T4N0 when they are biologically, very different. Clinical markers of tumour biology which are known prognostic indicators, such as depth of invasion and nodal extracapsular spread, are notably missing from the staging method (Takes, Rinaldo et al. 2010).

Another need for prognostic biomarkers in OSCC is frequent recurrence and second primary tumours (Koo, Lim et al. 2006, Rogers, Brown et al. 2009) which have poor survival. The main theory for this propensity is field change in which tumours develop within areas of histologically atypia (Slaughter, Southwick et al. 1953) and molecular aberrations (Tabor, Brakenhoff et al. 2002, Braakhuis, Tabor et al. 2002). Molecular biomarkers may be able to identify those patients at risk of second field tumours at diagnosis of the primary disease and filter these patients into more intense post-operative disease surveillance follow-up, new chemotherapy treatments or future gene therapies (Braakhuis, Tabor et al. 2003).

### 1.2.3 Predictive biomarkers

Predictive biomarkers suggest how a cancer will respond to a therapeutic intervention (Oldenhuis, Oosting et al. 2008). Some cases of OSCC are preceded by a visible oral lesion (Napier, Speight 2008) and histopathological diagnosis is used to predict the risk of malignant change (Warnakulasuriya S, Reibel J et al. 2008) but severity of dysplasia does not always predict transformation (Holmstrup, Vedtofte et al. 2006). Surgical incisional biopsy can under diagnose cancerous oral lesions (Holmstrup, Vedtofte et al. 2007) and the

literature is divided on the value of preventative surgical excision of precursor lesions in the context of preventing malignant change (Holmstrup, Vedtofte et al. 2006, Mehanna, Rattay et al. 2009, Balasundaram, Payne et al. 2014). Alternative predictive markers are required to accurately assess which oral precursor lesions will respond to excision, require close surveillance or do not require therapy. *P16* methylation and LOH at 3p and/or 9P in oral precursor lesions have been observed to predict for malignant change (Rosin, Cheng et al. 2000, Rosin, Lam et al. 2002, Hall, Shaw et al. 2008, Cao, Zhou et al. 2009, Zhang, Poh et al. 2012) and may have potential as predictive markers. Loss of heterozygosity in 3p and 9p in oral cancers are currently part of a phase III trial of an epidermal growth factor (EGFR) inhibitor, Ertinolab (www.clinicaltrials.gov/ct2/show/NCT00402779).

### 1.3 Molecular biology of OSCC

Oral carcinogenesis is a multistep process (Califano, Van Der Riet et al. 1996) inclusive of both genetic and epigenetic changes and evidenced by the myriad of associated molecular markers in the literature. Many dysfunctional, interrelated pathways may be involved in the aetiology of OSCC (Glazer, Chang et al. 2009).

#### 1.3.1 Copy number

One of the hallmarks of malignancy is chromosomal aberrations detected as deletions, amplifications or structural re-arrangement (Silva, Ferlito et al. 2011). Loss of genetic material is commonly found in oral cancer and is thought to be in chromosomal areas that harbour tumour suppressor genes. The study of loss of heterozygosity (LOH) was one of the most widely used molecular approaches to detect allelic imbalances associated with malignancy (Campo-Trapero, Cano-Sánchez et al. 2008) and is detected using polymorphic markers such as microsatellites or single nucleotide polymorphisms (Kasamatsu, Uzawa et al. 2011). Allelic imbalance at various loci has been observed on most autosomes in oral cancer studies (Kasamatsu, Uzawa et al. 2011) but most frequently on chromosomes 3, 9, 11, 13 and 17 (Massano, Regateiro et al. 2006). Early studies identified frequent deletions on chromosome 3 associated with oral premalignant lesions that underwent malignant transformation ( Mao, Lee et al. 1996, Partridge, Emilion et al. 1998, Rosin, Cheng et al. 2000), early stage OSCC (Partridge, Kiguwa et al. 1994) and development of secondary oral carcinoma at previously treated sites (Rosin, Lam et al. 2002). These findings suggest chromosome 3 contains genes important in progression of oral cancer and could have a role in disease surveillance. Authors have noted LOH at 13q present in 67% of OSCC cases

which have lymph node metastasis (Uzawa, Yoshida et al. 1998) which may also indicate a marker for progression of the disease.

LOH at the *P16* locus of 9p21 has been reported in two thirds of HNSCC (Reed, Califano et al. 1996) and OSCC tumours (Ohta, Uemura et al. 2009). Few tumours contained *P16* mutations (9%) as opposed to methylated *P16* in 64% of the tumours (Ohta, Uemura et al. 2009) which supports the theory of methylation as the second hit in OSCC. Early studies for detection of LOH are limited in cohort size and number of markers because standard techniques rely on lots of good quality tumour tissue (Reed, Califano et al. 1996). However, the value of these early findings is not diminished as similar chromosomal aberrations have been found using more advanced techniques. Comparative genomic hybridisation (CGH) and single nucleotide polymorphism (SNP) arrays can evaluate entire genome gains and losses thereby identifying potential oncogenes represented by gains and tumour suppressor genes (TSG) by losses (Patmore, Cawkwell et al. 2005). Three p loss and 3q gains are amongst the most consistent findings in OSCC in the genome wide studies of OSCC (Garnis, Coe et al. 2004, Sparano, Quesnelle et al. 2006, Martin, Reshmi et al. 2008) which implies they contain genes important in oral carcinogenesis. The genome wide approach provides the opportunity to discover novel candidate genes in areas of the genome previously un-investigated. Genes known to be involved in hereditary cancers; *FANCD2, FANCG, BRCA1* and *BRCA2* were discovered at high frequency in 21 primary OSCC using array CGH and may be involved in the initiation of sporadic OSCC (Sparano, Quesnelle et al. 2006). Recently loss at 3p.26.3 at the *CHL1* locus was identified as an independent prognostic factor in disease specific survival of OSCC (Uchida, Oga et al. 2011).

### 1.3.2 Whole genome mutations

Genetic mutation in DNA repair genes is known to be a driver mechanism in hereditary cancers but this does not appear to be the case in sporadic cancers. A review of high throughput techniques performed on non-hereditary cancers found that mutation in repair genes was infrequent; the authors suggest that the first event in sporadic carcinogenesis may be mutations in oncogenes or anti-oncogenes which trigger activated growth signalling which then leads to mutations in tumour suppressor genes (Negrini S, Gorgoulis VG et al. 2010). This could mean several mutations contribute to oral carcinogenesis and may display chronological variation according to the stage/risk profile of the disease. Whole genome next generation sequencing techniques accelerate the process of studying DNA and RNA

(Jessri, Farah et al. 2014) with the ability to identify all genetic variants (Zhang, Chiodini et al. 2011) and therefore may identify markers to help risk stratify patients.

Two recent studies have used exome sequencing to detect mutations in HNSCC tumours. Some of the data confirmed existing knowledge that *TP53* is the commonest mutation in HNSCC; being absent in HPV driven tumours whilst present in the majority (78%) of HPV negative tumours. Smokers have more mutations, thus providing further evidence that tumours that are morphologically similar are distinct diseases at the molecular level (Agrawal, Frederick et al. 2011) which could influence future prognostication. *NOTCH 1* was identified as a novel tumour suppressor gene, after *TP53* it was the commonest mutation but was present in only 15% of tumours. One of the most interesting results was that of 28 non HPV tumours, only 18% contained mutations in true oncogenes compared to 89% of tumours displaying mutated true TSGs (Agrawal, Frederick et al. 2011). Stransky et al (Stransky, Egloff et al. 2011) also identified mutations associated with loss of function at *NOTCH1*. Similarly they found twice as many mutations in HPV negative tumours than positive ones. Although these high-throughput genome wide techniques are extremely attractive to decipher the genes involved in oral carcinogenesis from a clinical application perspective, high cost and complicated analysis currently limit the routine application of next generation sequencing as many studies still verify data using Sanger techniques (Zhang, Chiodini et al. 2011, Jessri, Farah et al. 2014) .

### 1.3.3 Transcriptomics and Proteomics

The appeal of protein markers is that they are the functional element of the cancer pathway and may provide key information how genetic aberration is translated into cellular overgrowth, tissue invasion, metastatic spread and information about the tumour microenvironment. In the literature 14-3-3 sigma and 14-3-3 zeta/delta are potential diagnostic protein markers in HNSCC as they are consistently upregulated and are involved in many signalling pathways including cell cycle regulation (Schaaij-Visser, Brakenhoff et al. 2010). Keratin 4 and 13 are consistently downregulated in HNSCC and may be potential prognostic biomarkers but as yet have not been clinically validated (Schaaij-Visser, Brakenhoff et al. 2010). The presence of low expression levels of keratin 4 and cornulin proteins in surgical margin tissue from 46 HNSCC patients was associated with disease relapse; keratin 4 hazard ratio 3.8 (95% CI: 1.6-9.5) and cornulin hazard ratio 2.7 (95% CI:

1.1-6.5). In combination the markers had a hazard ratio of 8.8 (95% CI: 2-37.6) which was significantly associated with local disease-free survival ($p<0.0005$). Histopathological grading of the margin tissue was also performed, which did not show any association with disease free survival, even when considered as a binary classification (low/high risk of malignant transformation) and therefore was outperformed by protein markers as it had no prognostic relevance in this cohort of patients (Schaaij-Visser, Graveland et al. 2009). In support of these findings another study performed immunohistochemical analysis of 65 OSCC surgical margins which found the frequency of a second malignant event was 90.9% (10/11) in patients with keratin 4 negative margins and 51.8% (28/54) in patients with keratin 4 positive margins (Fisher's exact $p=0.0197$) (Polachini, Sobral et al. 2012). Garbis et al (Garbis, Lubec et al. 2005) summarise some of the limitations of proteomic analysis; proteins are vulnerable to environmental changes and therefore sampling protocols are difficult to standardise; small proteins can be difficult to detect and high abundance proteins frequently mask low abundant proteins which are often the target of interest. The proteome is very complex, analysis can be time consuming and the bioinformatic analysis remains relatively involved.

Studies of the transcriptome have provided insight into the mechanistic effects of genetic aberrations in HNSCC and aid understanding of interactive cancer pathways. Twelve regulatory pathways commonly dysfunctional in OSCC were identified from a meta-analysis of four public gene expression micro-array datasets which had all analysed greater than eight OSCC tissue samples and normal tissue (Liu, Niu et al. 2012). The authors categorised the common dysregulated pathways into three broad groups based on function: 1) Blood coagulation and prothrombin: Platelet amyloid precursor protein pathway 2) Cell surface and cell communication: Extracellular matrix ECM_receptor interaction, HS_matrix_metalloproteinases, cell communication, focal adhesion, intracellular proteasome, proteasome, proteasome complex and HS_Proteasome degradation 3)Metabolism related pathways: Bile acid synthesis, glycolipid metabolism and arginine and proline metabolism pathways. In the analysis, the extracellular matrix ECM_receptor interaction pathway was the most dysregulated in OSCC. This pathway has an effect on cellular processes of adhesion, differentiation proliferation and apoptosis. Four genes (*SPP1, COL5A2, COL4A1* and *COL1A2*) were identified in this pathway with significant dysregulation in all four of the microarray datasets. The authors validated the

downregulation of these genes in a small cohort of OSCC and normal samples using quantitative polymerase chain reaction (qPCR) (Liu, Niu et al. 2012). In a separate study, *SPP1* was found to be expressed at significantly higher levels in OSCC tissue (n=43) than normal oral mucosa (n=17) but had no relationship with tumour pathological characteristics or overall survival (Huang, Yu et al. 2014). Reduced expression of *COL1A2* was associated with promoter methylation of the gene in head and neck SCC cell lines. In addition 46/98 (47%) primary HNSCC tissues had detectable *COL1A2* promoter methylation and furthermore it was significantly related to disease free survival (p=0.005) (Misawa, Kanazawa et al. 2011) . Further exploratory or validation data for *COL5A2* and *COL4A1* was unavailable in the current HNSCC literature.

### 1.3.4 Epigenetics

Epigenetics is defined as inheritable changes in gene expression that are not accompanied by changes in DNA sequence (Jones, Baylin 2007). Three epigenetic modifications are described in the literature: DNA methylation, histone modifications and altered expression of micro-RNAs (MiRNA). Histone modifications alter the tertiary structure of DNA which inhibits the binding of transcription factors. Modifications to histones include methylation, acetylation, ADP_ribosylation, phosphorylation, ubiquination and sumolyation of specific residues within histone tails and occur in each of the four histone complexes (H3,H4,H2A,H2B) (Gasche, Goel 2012). MiRNA are short, non-coding RNAs which function in post-transcription gene expression (Jansson, Lund 2012). To date DNA methylation is the most studied of the three epigenetic modifications in HNSCC cancer biology. DNA methylation occurs in cytosines that precede guanine bases known as dinucleotide CpGs. They are found in CpG rich regions known as CpG islands (Esteller 2008) which are approximately 500bp in length, characterised by a GC content > 55% with a CpG ratio (observed CpG)/expected CpG) of 65% (Takai, Jones 2002) and are found in gene promoter regions. DNA methylation represses genes expression by impeding binding of transcriptional factors (Cedar 1988) therefore methylation in the promoter regions of tumour suppressor genes prevents their expression and has a role in carcinogenesis (Esteller 2008). Promoter methylation of several genes has been implicated in OSCC; *P16, CDH1, MGMT, DAPK, DBC1, P14$^{ARF}$, DCKN2B, RARB, RASSF1A, MLH1, P73, DCC, FHIT, SERPINB* (Ha, Califano 2006, Gasche, Goel 2012) which reflects the complex molecular nature of this disease. There is often a wide variation in reported methylation rates of these genes which could reflect biological variation but is also dependent on sample selection and

methodology (Gasche, Goel 2012) and some of these changes may represent passenger rather than driver events. Further functional studies would be required to establish their role in cancer biology.

### 1.3.5 Potential biomarkers in OSCC

There is a wealth of literature on potential candidate genes in HNSCC. Authors of a recent review of the common aberrant genes in head and neck squamous cell carcinoma (HNSCC) literature designated them *established* or *candidate* based on the strength of evidence for a mechanistic role in HNSCC (Leemans CR, Braakhuis BJ et al. 2011). Eight genes in total were considered as *established* cancer genes, of them four are tumour suppressor genes: *P16, PTEN, TP53, SMAD4* and four oncogenes: *PIK3CA, EGFR, MET* and *CCND1* (Leemans CR, Braakhuis BJ et al. 2011). Amongst them *TP53*, *P16* and *EGFR* are perhaps the most well described in the broader HNSCC literature. A lot of the early and informative studies applied discovery techniques to mixed cohorts of HNSCC. As OSCC is the most common subgroup of this disorder it is often well represented in these studies but seldom are the results stratified by anatomical subsite.

*TP53* is a well established tumour suppressor gene and is commonly mutated in all types of cancer (Negrini S, Gorgoulis VG et al. 2010). It is one of the most commonly mutated genes in HPV negative HNSCC (Agrawal, Frederick et al. 2011) and has been described in pre-invasive HNSCC and invasive HNSCC (Somers, Merrick et al. 1992, Boyle, Mao et al. 1994, van Houten, Tabor et al. 2002). It interrelates with other genes in cancer pathways known to have roles in HNSCC carcinogenesis such as *P16* and *Cyclin D* (Zhang, Xiong et al. 1998, Opitz, Suliman et al. 2001). Inactivated p53 confers expanded lifespan on oral keratinocytes *in vitro* (Opitz, Suliman et al. 2001) and in knock-out murine models (*TP53*+/-) the loss of *TP53* was associated with earlier onset of tumorigenesis and a more aggressive, metastatic HNSCC phenotype (Ku, Nguyen et al. 2007). In terms of potential as a biomarker it may have a role in prognostication. The presence of *TP53* mutations in tumours has been associated with reduced survival (Poeta, Manola et al. 2007) and in surgical margins, is associated with increased risk of locoregional recurrence even in those considered histopathologically free of tumour (Brennan, Mao et al. 1995, van Houten, Leemans et al. 2004) .

The *P16* gene is a cell cycle regulator located at the 9p21 locus which is known to have a high incidence of genetic loss (Van Der Riet, Nawroz et al. 1994). The *P16* (*INK4a-ARF*) gene locus encodes two unrelated proteins which regulate the retinoblastoma and P53 tumour suppressor pathways respectively (Zhang, Xiong et al. 1998). *P16 INK4a* binds to and inhibits the activity of cyclin dependent kinase 4 (CDK4 and CDK6). ARF binds to and promotes the degradation of the MDM2 protein which is responsible for the degradation of the P53 tumour protein (Zhang, Xiong et al. 1998). Fifty three percent (17/32) of HNSCC tumours had LOH at 9p21 locus (Miracca, Kowalski et al. 1999) which supports other authors' observations that it is frequently inactivated in HNSCC, likely through chromosome loss and methylation (Reed, Califano et al. 1996) as mutation is relatively rare (Miracca, Kowalski et al. 1999, Ohta, Uemura et al. 2009). Methylation is generally considered an inactivation mechanism of the tumour suppressor function of *P16* in HNSCC (Merlo, Herman et al. 1995, El-Naggar, Lai et al. 1997). *P16* methylation may have a role as a predictive biomarker being present in precursor lesions (Kresty, Mallery et al. 2002) and is associated with, and predictive of malignant progression in these lesions (Hall, Shaw et al. 2008, Cao, Zhou et al. 2009). Other authors cite promoter methylation of *P16* as a potential prognostic indicator in HNSCC tumours based on an association with disease recurrence (Sinha, Bahadur et al. 2009) and poor clinical prognostic indicators such as advanced stage disease and nodal metastasis (Sailasree, Abhilash et al. 2008, Huang, Yu et al. 2014 ).

Epidermal growth factor receptor (*EGFR*) is considered an oncogene in HNSCC on the basis it is overexpressed in tumours (Grandis, Tweardy 1993, Hama, Yuza et al. 2009) and has shown ability to confer malignant traits on oral keratinocytes *in vitro* (Goessel, Quante et al. 2005). It is present on chromosome 7p; a locus shown to have chromosomal gains in OSCC (Baldwin, Garnis et al. 2005, Martin, Reshmi et al. 2008). Using SNP array Sheu et al (Sheu, Hua et al. 2009) observed the most frequent amplification in OSCC patients (9/29) was at 7p11.2. Genomic mapping identified *EGFR* at that locus and fluorescence in situ hybridisation assay confirmed upregulation of *EGFR* associated with those tumours. Functionally, *EGFR* is a member of the c-erb family of transmembrane proteins which are involved in transcriptional regulation of proteins and cytokines implicated in tumour invasion and angiogenesis (Rogers, Harrington et al. 2005). It has been associated with recurrence and poor survival (Grandis, Melhem et al. 1998) and HNSCC patients treated with *EGFR* antibodies showed improved response to radiotherapy (Bonner, Harari et al. 2006). Inactivation of *EGFR,* with the chemical inhibitor AG1748, has been shown to reduce the progression from precursor lesions to oral cancer in a mouse oral cancer model;

suggesting *EGFR* has a role in early carcinogenesis and secondly may provide a therapeutic target in oral cancer (Sheu, Hua et al. 2009).

### 1.3.6 Gene panel selected for this study –a candidate gene approach

The gene panel investigated in this study was selected using a candidate gene approach. Candidate gene studies focus upon a selection of genes identified in previous studies to be, or likely to be, related to the disease of interest (Patnala, Clements et al. 2013). The candidate gene approach has been defined as the study of the genetic influences on a complex trait by generating hypotheses about and identifying candidate genes that might have a role in the aetiology of the disease (Tabor, Risch et al. 2002). In contrast to a genome wide approach which identifies markers throughout the genome without regard to their function or context in a specific gene, candidate gene studies focus on genes that are selected because of *a priori* hypotheses about their aetiological role in disease (Tabor, Risch et al. 2002). The gene variant is commonly verified for disease association in case-control studies to establish its association with diagnosis and prognosis and future potential as a biomarker (Patnala, Clements et al. 2013). Four of the genes *CCNA1*, *CYGB, P16* and *TMEFF2* had previously been identified from the literature as genes of interest in aerodigestive carcinoma by our research group and investigated on a series of primary HNSCC tissue. *ADAMTS9* was identified as a novel gene in OSCC as part of a collaborative work with Dr Ratna Veeramachenini, involving cell lines from premalignant and malignant oral tumours and primary HNSCC tissue, which will be discussed in depth in the results section.

Cyclin A1 (*CCNA1*) maps to chromosome 13q.12.3-q13 and is part of a family of genes that control the progression of the cell cycle through activation of cyclin dependent kinases. Cyclins activate cyclin dependent kinases (CDKs) in an ordered sequence which is both determined by events in the cell cycle and important for its progression (Yang, Morosetti et al. 1997). In the physiological state it appears to be expressed in significant concentrations in a tissue specific manner and was first discovered in testis germ cell cycles (Yang, Morosetti et al. 1997) and later in normal and leukaemic hameopoietic cells (Yang, Nakamaki et al. 1999) in which it appears to have a role in somatic G1 to S cell cycle progression (Ji, Agrawal et al. 2005). Cyclin A1 is a complex protein with roles in multiple molecular pathways including phosphorylating Histone 1, RB proteins and E2F-1 transcription factor (Yang, Müller et al. 1999). It is also a P53 induced gene (Maxwell, Davis 2000, Müller-Tidow, Ji et al. 2004) and through a P53 mediated pathway is associated with G2 cell cycle arrest, apoptosis and mitotic catastrophe in renal, ovarian and non-small cell

lung cancer cell lines; the mechanism is thought to be unscheduled activation of CDK1 (Rivera, Mavila et al. 2006). Earlier studies on animal and HeLa cell lines implied a role for Cyclin A proteins in apoptosis via activation of the cyclin-A-dependent kinases (Meikrantz, Gisselbrecht et al. 1994). Furthermore murine studies revealed that dominant negative mutants for CDK2 had suppressed apoptosis which suggests cyclin A proteins may act as cell cycle dependent facilitators of apoptosis (Meikrantz, Schlegel 1996). Murine studies revealed *CCNA1* expression is induced by gamma radiation via a p53 pathway. The cyclin A1-CDK-2 complex binds to Ku70 a known DNA repair protein. DNA double-strand break (DSB) repair was deficient in cells from cyclin A1 -/- mice (Müller-Tidow, Ji et al. 2004) suggesting *CCNA1* may have a role in DNA repair. It would appear that the chronology of *CCNA1* protein expression is important for its normal function. It is possible to see how ectopic or altered expression of the protein due to gene modification, perhaps by an epigenetic event, may prevent cell death or appropriate DNA double strand repair and result in uncontrolled cellular growth and proliferation.

Studies on primary HNSCC tissue have shown an inverse relationship between *CCNA1* expression and P53 mutation (Tokumaru, Yamashita et al. 2004, Farhadieh, Smee et al. 2009). One of the first papers to publish a role for *CCNA1* in HNSCC was by Tokumaru et al (Tokumaru, Yamashita et al. 2004) they identified *CCNA1* as one of 6 genes from an original 278 that were epigenetically silenced in a tumour specific pattern in HNSCC tissue. Using bisulphite-sequence analysis they determined 9/20 (45%) HNSCC tumours displayed *CCNA1* promoter methylation. This study influenced our research group's early work on developing *CCNA1* as a potential marker in OSCC. However, more recently it has been discovered that higher methylation of *CCNA1* and lower expression of *CCNA1* is found in HPV positive HNSCC tumours (Sartor, Dolinoy et al. 2011) which are more likely to have wild type *TP53* (Agrawal, Frederick et al. 2011). OSCC is associated with *TP53* mutations and may explain the relatively low levels of observed methylation for this gene in our study of saliva from OSCC patients. *CCNA1* methylation has also been reported as a potential biomarker in high grade cervical epithelial neoplasia and cervical cancer which is another HPV mediated disease (Yang, Eijsink et al. 2009).

Discovery of over expression combined with enhanced S phase entry in leukaemia cell lines has led some to propose a tumourigenic role for *CCNA1* (Ji, Agrawal et al. 2005). Over expression of the protein is associated with more aggressive prostate cancer (Wegiel,

Bjartell et al. 2005). However in clinical outcome studies over expression of *CCNA1* in leukaemia patients was associated with a survival advantage (Nakamaki, Hamano et al. 2003) and in laryngeal SCC with a reduction in recurrence (Weiss, Koopmann et al. 2012) which could suggest expression is a response rather than causative factor. There is also reported association between increased risk of second primary tumours and *CCNA1* methylation in primary HNSCC cases (Rettori, de Carvalho et al. 2013) which would further support a tumour suppressor role. It appears that the function of *CCNA1,* in different cell types and malignant disease states, is not fully characterised and that there is potential for a tumour suppressor or oncogenic role for *CCNA1* which some authors propose could depend on tissue type and context.

The Cytoglobin(*CYGB*) gene is located on chromosome 17q25 and encodes an intracellular globin found in vertebrates which has a putative role in oxygen metabolism at the cellular level (Burmester, Haberkamp et al. 2004). The functional role is uncertain but in the *in vitro* setting *CYGB* has been shown to be upregulated in hypoxic conditions (Schmidt, Gerlach et al. 2004) and may afford cytoprotection from oxidative DNA damage (Hodges, Innocent et al. 2008). In cell line studies, hypoxic related elements (HREs) have been identified at positions -141, -144 and -448 on the *CYGB* gene which bind to HIF-1 proteins and are related to *CYGB* expression under hypoxic stress (Guo, Philipsen et al. 2007). Latterly an anti-fibrotic role has also been discovered for CYGB whereby it was observed to be upregulated in response to renal ischaemic reperfusion injury (Mimura, Nangaku et al. 2010). The mechanism by which it contributes to carcinogenesis is not completely understood but appears to be complex. In a comprehensive review of the literature Oleksiewicz and colleagues (Oleksiewicz, Liloglou et al. 2011) speculate that down regulation of *CYGB* may result in a higher burden of oxidative or nitrosative stress in tissues which contribute to DNA damage and mutations. They provide a second hypothesis that impaired healing leads to fibrosis induced inflammation which in turn stimulates proliferative signals and causes aberrant cellular overgrowth (Oleksiewicz, Liloglou et al. 2011, Oleksiewicz, Liloglou et al. 2013).

A functional tumour suppressor role has been observed in a study on lung and breast cancer cell lines which showed that loss of *CYGB* expression in *CYGB* positive lung cell lines resulted in colony formation. Conversely, *CYGB* transfection in *CYGB* negative lung and breast cancer cell lines resulted in reduced colony formation. Furthermore, cell lines with promoter methylation of the gene showed significantly reduced expression of *CYGB*

(Shivapurkar, Stastny et al. 2008). Promoter methylation of *CYGB* has also been associated with HNSCC, oesophageal, lung and ovarian malignancy (Shaw, Liloglou et al. 2006, McRonald, Liloglou et al. 2006, Xinarianos, McRonald et al. 2006, Wojnarowicz, Provencher et al. 2012, Oleksiewicz, Liloglou et al. 2013). Our research group reported that promoter methylation of *CYGB* is significantly elevated in HNSCC, with 44% of tumours (35/80) having significantly higher levels than adjacent normal tissue (Shaw, Liloglou et al. 2006). Additionally, a correlation between HNSCC tumour hypoxia and *CYGB* mRNA expression and a negative correlation between C*YGB* expression and gene promoter methylation was observed; implying that the *CYGB* gene is regulated by both tissue hypoxia and methylation (Shaw, Omar et al. 2009).

*TMEFF2*, transmembrane protein with EGF-like and two follistatin-like domains 2 is located on chromosome 2q33 (Liang, Robertson et al. 2000). The physiological role of *TMEFF2* and its involvement in carcinogenesis remains unclear. *TMEFF2* has been shown to have androgen dependent expression and anti-proliferative activity in prostate cancer cell lines (Gery, Sawyers et al. 2002). *TMEFF2* induced apoptosis has also been noted in colon cancer cell line studies and found to induce a signal transducer and activation transcription factor (STAT1) which is known to stimulate other interferon inducible genes (Elahi, Zhang et al. 2008). Further involvement with STAT pathways has been seen in infection driven *H.pylori* gastric carcinoma cell lines and tissues. *H.pylori* activated STAT3 which was observed to bind to the *TMEFF2* promoter and inhibited its expression. In normal mucosa the expression of *TMEFF2* suppressed phosphorylation of STAT3 (Sun, Tang et al. 2015). The combined findings of these studies may suggest *TMEFF2,* in a response to chemical mediation from the tumour microenvironment, exerts a tumour suppressor role through an interferon mediated pathway. *TMEFF2* is epigenetically silenced in several cancers. Early studies observed methylation of *TMEFF2* in cell lines of bladder and colon cancers (Liang, Robertson et al. 2000) and it is differentially methylated in tissue (Young, Biden et al. 2001, Sato, Shibata et al. 2002) and blood (Lofton-Day, Model et al. 2008, Elliott, Johnson et al. 2013) from colorectal cancer patients. Allelic imbalance at sites on the 2q chromosome has been associated with HNSCC (Ransom, Barnett et al. 1998). However, epigenetic aberrations of *TMEFF2* have rarely been reported in HNSCC. It was found unmethylated in a small UK series of carcinoma expleomorphic adenomas (Schache, Hall et al. 2010). High methylation rates were reported in saliva samples from a Japanese cohort of OSCC patients. Twenty nine of thirty four OSCC samples were methylated versus 3/34 controls as measured by MSP. Promising sensitivity 85.3% (95% CI: 73.4-97.2%), positive predictive

value of 87.5% (95% CI: 74.3-100%) and negative predictive value of 90.6% (95% CI: 80.5-100%) data are reported as derived by a ROC analyses. However, the sensitivity and predictive values have not been validated on a separate cohort of patients (Nagata, Hamada et al. 2012). In contrast modest informativity for *TMEFF2* is reported in a recent surgical margins study of OSCC tissue; 13/48 tumours (26%) had detectable *TMEFF2* methylation using qMSP and a 5% methylation cut-off, but presence in margin tissue as part of a panel, inclusive of *CYGB* and *P16*, was not associated with tumour recurrence or survival outcome (Shaw, Hobkirk et al. 2013).

*P16* (cyclin dependent kinase inhibitor 2A) is a tumour suppressor gene located on chromosome 9p21.3 which is involved in cell cycle control (Serrano, Hannon et al. 1993, Rayess, Wang et al. 2012). *P16* is a cyclin dependent kinase inhibitor that regulates the passage from G1 to S phase in the cell cycle by binding to and inhibiting the cyclin D-cyclin dependent kinase 4 or 6 complex which inactivates the retinoblastoma protein and in normal function would initiate cell cycle arrest (Rocco, Sidransky 2001). Inactivation of *P16* by promoter methylation has as a role in carcinogenesis. Epigenetic silencing of *P16* in HNSCC is well documented in the literature (Rocco, Sidransky 2001, Maruya, Issa et al. 2004, Kulkarni, Saranath 2004, Kato, Hara et al. 2006, Shaw, Liloglou et al. 2006) . It is thought to be an early event in carcinogenesis (Rocco, Sidransky 2001) and methylated *P16* has been detected in oral epithelial dysplasia that progresses to OSCC (Hall, Shaw et al. 2008, Cao, Zhou et al. 2009) . The specificity of *P16* methylation as a tumour marker is under debate with some authors claiming little or no methylation in normal samples (Shaw, Liloglou et al. 2006, Righini, De Fraipont et al. 2007, Carvalho, Jeronimo et al. 2008) and others asserting it is associated with tobacco damaged tissue (Belinsky, Palmisano et al. 2002, von Zeidler, Miracca et al. 2004). This is an issue which will be discussed in a later section.

*ADAMTS9* is an anti-angiogenic metalloprotease (Koo, Coe et al. 2010) which maps to 3p14.2-3p14.3 (Clark, Kelner et al. 2000). It may have a role in tumour angiogenesis. In studies using oesophageal and nasopharyngeal carcinoma cell lines Lo et al (Lo, Lung et al. 2010) describe *ADAMTS9* as having a critical role in the 'angiogenic switch' when a tumour progresses from non-angiogenic to an angiogenic phenotype which is pivotal to cancer progression. In keeping with this phenomenon, methylation of this gene has been demonstrated in locally metastatic nasopharyngeal carcinoma (Hong, Lo et al. 2008). The findings in both of these studies may infer a role for *ADAMTS9* in the later stages of

carcinogenesis in these tumour types. Dr Veeramachenini (Veeramachaneni 2010) found it was epigenetically silenced in both premalignant and malignant HNSCC cell lines suggesting an earlier role for *ADAMTS9* in HNSCC carcinogenesis. It is however known that the methylation status of cell lines and primary tumour tissue may differ. Further support for *ADAMTS9* having an early role in HNSCC is the finding of allelic imbalance at 3p14.12 in oral premalignant lesions which undergo malignant transformation (Rosin, Cheng et al. 2000, Tsui, Rosin, Lam et al. 2002, Rosin et al. 2008). In view of this, *ADAMTS9* was considered a gene of interest in our study; investigating the presence of methylation of these genes in saliva from patients with premalignant lesions and those with TNM stage I-IV OSCC.

## 1.4 Body fluids as tumour surrogates

The relative difficulty in obtaining tissue samples makes tissue surrogates an attractive proposition for early disease detection and disease monitoring. This is particularly true with bodily fluids which are in intimate contact with potential tumour sites. The feasibility of detecting potential methylation biomarkers in several body fluids has been explored in the literature.

### 1.4.1 Advantages of surrogates

Patient acceptance is pivotal for the successful development of a screening or disease surveillance biomarker. From a clinical and research perspective surrogates are appealing because their non-invasive acquisition generally meets this criterion. Exclusive of blood derivatives, surrogates are largely collected by non-invasive means, which also renders them suitable for repeat sampling. Tumour tissue is finite and this restricts its use for disease follow-up whereas saliva is unlimited, being continually produced and secreted into the easily accessible oral cavity. Disease surveillance may necessitate chronological tissue sampling but sequential biopsy is not realistic or always representative of the extent of disease. Tumour surrogates may provide a solution to these limitations. Bodily fluids have intimate contact with the target tumour but importantly beyond this they bathe the whole tissue field, which could provide a more objective perspective of malignant disease in the context of molecular field changes. There is evidence across many tumour types that surrogate tissues share concordant molecular profiles with tumour tissue (Liloglou, Field 2010). It is envisaged that biomarkers detectable in surrogate tissues may eventually

provide accessible means of dose and response-to-therapy monitoring for cancer treatments. Along with quantitative measures of the biomarker, repeat samples are likely to be required to accurately assess therapeutic responses. Aside from blood products most other body fluids lend themselves to self collection which could lessen the cost of repeat sampling. Saliva DNA has proven sufficiently stable after self collection and postal return of samples to be used in downstream applications such as PCR (García-Closas, Egan et al. 2001) and self-collected faecal occult blood sampling is routinely used in the UK for colorectal carcinoma screening.

There is increasing interest for surrogate biomarker research in malignant pathology where screening and disease surveillance have proven difficult or unsatisfactory and early detection of disease would significantly improve prognostics, quality of life or expand treatment options. Accessibility and inexhaustible supply of surrogate tissues such as blood, saliva, faeces and urine hold obvious attractions to researchers but inherent in these attributes is also the scope for use in clinical outpatient and community settings. In many assays, the definition of a negative sample is informed by the level of the target found in control tissue. The definition of 'normal' tissue can differ markedly across research fields and often includes control tissue from the same site as the tumour, albeit geographically remote. However, the potential for field change calls into question the reliability of these assumed normal controls to provide 'cut-offs' to distinguish a positive cancer result from negative 'background methylation'. The level of methylation in normal tissue remains a relative unknown. To date ethical limitations prevent the harvesting of like for like normal tissue in enough breadth and quantity to find the definitive answer. Surrogate tissue may provide a suitable compromise as it is relatively easily obtained from a disease free population.

### 1.4.2 Disadvantages of surrogates

The prevailing theory for the origin of body fluid DNA is exfoliated epithelium. In oral cavity work on *TP53* mutation supports the assertion that the detectable abnormal DNA in surrogate tissue is from the tumour of origin because the uniqueness of the target aberration makes it unlikely to be from any other source (Boyle, Mao et al. 1994). Larger tumours are likely to shed greater quantities of epithelia as expanding size outstrips nutrient supply. Extending this principle, the concern might be that premalignant, field and

early stage lesions may go undetected by body fluid assays due to a relative lack of exfoliation. However, epigenetic biomarkers have been successfully detected in saliva from precursor oral lesions (Pattani, Zhang et al. 2010, Liu, Zhou et al. 2012). The origin of DNA in surrogate samples remains unproven and as a result sampling methods remain non-standardised. Even in a relatively small field such as saliva cancer diagnostics, the range of sampling methods, storage and processing are different enough to prevent reliable direct comparisons of the results.

DNA extracted from surrogate tissue can lack the integrity and quantity of that derived from tumour. The chemical and physical stresses it endures in its journey from source to excretion could cause fragmentation. However, the ready supply of surrogate material would permit multiple samples and pooling of DNA. The pre-analytical preparation of DNA can further compromise it. Bisulphite modification imposes a chemical assault on the structural integrity of DNA which can result in degradation and loss of DNA during the clean-up process (Munson, Clark et al. 2007). This can present difficulties for downstream applications when only a small amount of target DNA is available (Liloglou, Field 2010). A problem with DNA of uncertain origin is how to know exactly what process it is heralding. It is entirely feasible that the surrogate derived DNA marker is not from the tumour of interest and may counsel of a tumour elsewhere. This is a particular concern in a non-tissue specific surrogate such as blood. Further compounding this issue, is the non-tissue specific nature of certain genes for example *RASSF1A* found in numerous tissues which may be an unlikely candidate as a suitable singular screening marker for a cancer of unknown site (Hesson, Cooper et al. 2007).

Surrogate biomarker research remains an evolving field because the epigenetic and genetic profile of cancer tissue is yet to be fully described, in part due to the complex molecular heterogeneity of tumours. As a result of this heterogeneity, it is thought that a panel of markers infers a greater sensitivity than singular markers (Esteller, Corn et al. 2001) but the inclusion of multiple markers can negatively affect the specificity (Carvalho, Jeronimo et al. 2008) there is currently a need to improve this balance in the surrogate literature. Perhaps the biggest technical challenge presented by surrogate tissues is dilution of the target material by thousands of other types of cell. A difficulty with body fluids from non-sterile cavities, like faecal DNA derivatives, is the relatively small amount of tumour DNA. It is estimated that only 0.1 -0.01% of the extracted DNA is human and only 1% of that

proportion is tumour cell; the rest is contributed to by bacteria (Whitney, Skoletsky et al. 2004) additionally, bile pigments and salts may also act as PCR inhibitors. Saliva is also not a sterile tissue and samples may contain 7-49% bacterial DNA content (García-Closas, Egan et al. 2001, Iwasiow, Tayeb et al. 2011). This published broad range is likely a result of different storage conditions. Once the sample is properly prepared the problem of dilution of the target DNA by thousands of normal cells still exists and this requires a high degree of sensitivity from detection assays.

### 1.4.3 Biomarkers in body fluids

#### 1.4.3.1 Faeces

Faecal matter is easily collected and faecal occult blood testing is currently used as a screening tool in the UK which has limited sensitivity in part due to a reliance on the tumour being a bleeding phenotype (Young, Bosch 2011). Potential diagnostic molecular markers have been detected in faeces. In a case control study of 94 cancers and 198 healthy controls methylation of the *vimentin* gene had a sensitivity of 46% and a specificity of 90% for the detection of colorectal cancer (Chen, Han et al. 2005). Another group found sensitivity improved to 75% when *vimentin* was considered as part of a methylation panel inclusive of *MGMT* and *hmlH1* specificity of 87% was comparable to the single marker study. This panel also detected adenoma with a sensitivity of 60% and may be useful in detection of premalignant disease (Baek, Chang et al. 2009). A panel of genes containing hypermethylated *vimentin* and mutation of *KRAS* and *APC* was found to be more sensitive for the detection of colorectal cancer than faecal occult blood testing (Ahlquist, Sargent et al. 2008). In future testing molecular markers may increase the sensitivity of the existing screening tools for colorectal carcinoma.

#### 1.4.3.2 Urine

Urine is sterile, accessible and already has a place in clinical diagnostics being used for hormonal assays and screening for glucosuria and proteinuria which are early indicators of chronic disease. Promising molecular markers for the detection of urogenital cancers have now emerged in the literature. Promoter methylation of *GSTP1* is tumour specific in prostate cancer and has been detected in the urine of patients with biopsy confirmed prostate cancer with a sensitivity of 75% and specificity of 98% (Woodson, O'Reilly et al.

2008). In a case control study of 52 patients with prostate cancer and 91 age-matched controls the presence of at least one positive marker from a panel of four genes *GSTP1, ARF, MGMT* and *P16* had a sensitivity of 87% and a specificity of 100% for the detection of biopsy proven prostate cancer (Hoque, Topaloglu et al. 2005). The same group also note that methylation of *GSTP1* detected in tumours at primary surgery is a significant factor in the time to progression of the disease (Rosenbaum, Hoque et al. 2005) and may have a role as a prognostic marker.

#### 1.4.3.3 Sputum

Sputum is primarily used in lung cancer research and methylation at the promoter of *P16* shows promise for the detection of lung cancer from sputum samples (Palmisano, Divine et al. 2000, Belinsky, Liechty et al. 2006, Belinsky, Grimes et al. 2007). A sensitivity of 66.7% for the detection of non-small-cell lung carcinoma (NSCLC) is reported using *P16* methylation detected by MSP as a marker in the sputum of 50 patients. Chronic heavy smokers were used as a control group (n=100). Four percent had detectable *P16* promoter methylation in sputum samples (Destro, Bianchi et al. 2004). No follow-up data is provided for this high risk group so it is unclear if they represent undiagnosed early cancer or a high risk group as *P16* could be a marker of field cancerisation. More recently a large case control study of promoter methylation of *P16, TERT, WT1 and RASSF1* in bronchial washings established this panel of biomarkers is more sensitive for the detection of lung cancer than cytological assessment which is currently part of the lung cancer diagnostic pathway in the UK (Nikolaidis, Raji et al. 2012).

#### 1.4.3.4 Blood

The actual origin of free DNA in blood is unknown but there is a theory of phagocytic ingestion of solid tumour cells or circulating tumour cells that undergo necrosis and release DNA (Wong, Dennis Lo et al. 2001). Blood is non tissue specific and both serum and plasma have been used to detect molecular markers associated with solid tumours in various organs such as detection of ovarian cancer (De Caceres, Battagli et al. 2004), prostate cancer (Bryzgunova, Morozkin et al. 2008, Payne, Serth et al. 2009), colorectal cancer (Tänzer, Balluff et al. 2010) and HNSCC (Carvalho, Jeronimo et al. 2008). In addition to diagnostics, methylation markers in blood have also been used as predictive biomarkers in chemotherapy trials for solid tumours.

Fiegl et al (Fiegl, Jones et al. 2008) investigated serum *NEUROD 1* methylated DNA as a predictor of chemosensitivity in a cohort of 107 breast cancer patients using quantitative MSP. High levels of *NEUROD 1* methylation in oestrogen receptor negative breast cancer tissue is associated with a 10.8 fold increase in response following neoadjuvant chemotherapy (sensitivity: 80% specificity: 72%). Oestrogen receptor negative patients with *NEUROD 1* methylation present in pre and post treatment sera had a significantly worse relapse-free and overall survival compared with those who had become *NEUROD 1* free in post treatment sera p=0.01. Pre-treatment and post-treatment plasma samples from multistage ovarian cancer patients receiving carboplatin taxoid chemotherapy as part of a stage III trial were collected. One hundred and thirty eight patients with relapse provided a matched pre-treatment and an at relapse plasma sample. 16/138 (12%) of patients were positive pre-treatment and 45/138 (33 %) at relapse which represents a significant increase ($p<0.001$) in *hMLH1* methylation at relapse what is more the post-treatment acquisition of hypermethylated *hMLH1* in plasma samples is associated with worsened survival (p=0.007) (Gifford, Paul et al. 2004).

In terms of translation into clinical use, surrogate derived epigenetic biomarker research is still evolving. This is in part because tissue epigenetic research continues to answer questions about tissue heterogeneity and normal levels of methylation. Currently in the HNSCC surrogate literature there are differences in sample collection, assay types and conditions and statistical analysis that make direct inter-study comparisons difficult and any definitive conclusions about specific genes difficult to draw.

### 1.4.4 Saliva as a source of biomarkers

Saliva is a complex fluid with contribution from major and minor salivary glands, gingival crevicular fluid, nasal and bronchial secretions, blood and desquamated epithelial linings (Loo, Yan et al. 2010). It is a rich source of electrolytes, immunoglobulins, proteins, enzymes, mucins and nitrogenous products with an average daily flow of up to 1.5 L (Humphrey, Williamson 2001). It is therefore a regularly replenished source of a myriad of potential disease markers which can be sampled repeatedly. Saliva has been used as a surrogate for OSCC across research disciplines such as genomics, proteomics, transcriptomics and epigenomics. It is also a source of potential biomarkers to detect

tumours remote from the oral cavity. Soluble C-erbB-2 protein, a prognostic marker for breast cancer, was detected at higher levels in saliva samples from breast cancer patients than healthy controls or patients with benign breast disease (Streckfus, Bigler et al. 2000). Messenger RNA markers (*KRAS, MBD3L2, ACRV1* and *DPM1*) detectable in saliva supernatant are differentially raised between early stage resectable pancreatic cancer and healthy controls and may have a role in non-invasive detection of pancreatic cancer (Zhang, Farrell et al. 2010).

#### 1.4.4.1 Genomics

### Gene mutations

One of the first exploratory studies of saliva as a tumour surrogate was performed on a small cohort of HNSCC patients to detect *TP53* tumour specific mutations present in pre-operative saliva. Five out of seven saliva samples were positive for *TP53* mutation. The data suggests the DNA extracted from saliva must have been from shed tumours cells to display the same *TP53* point mutation as the tumour (Boyle, Mao et al. 1994). Another feasibility study compared the presence of *TP53* mutations at exon 4 and intron 6 in pre-treatment saliva from a cohort of 10 OSCC patients and an unmatched cohort of young healthy controls (n=27). Five out of eight (63%) OSCC and 5/27 (19%) normal saliva samples revealed a *TP53* exon 4 codon 63 mutation; a difference which reached statistical significance ($p<0.05$) (Liao, Chang et al. 2000). However, the data should be treated with caution as this study is limited by small numbers and an unmatched control group. If genetic mutations are acquired over time with exposure to environmental risk factors then a young control group are unlikely to have had time to acquire genetic aberrations associated with OSCC. Furthermore, in terms of utility as a biomarker point mutations can occur anywhere along the gene and vary between tumours therefore it may be difficult and time consuming to detect these mutations and predict which of them are clinically relevant.

#### Microsatellite analysis

Authors of an early proof of principle study used microsatellite analysis to interrogate 23 microsatellite loci to detect tumour specific genetic alterations in mouth rinse from HNSCC patients (n=44) and healthy controls (n=43: 20 smokers 23 non-smokers). Forty four percent of the matched tumour and salivary rinse pairs had concordance of at least one of

the 23 markers. Some of the positive markers were only positive in one tumour e.g D20585 this variability in markers may reflect the heterogeneity of the tissue and cohort (Spafford, Koch et al. 2001) and highlights the need for a panel rather than singular markers for detection of disease. Another group analysed 37 matched OSCC tumour and saliva samples for the presence of LOH. Ten saliva samples from 5 smokers and 5 non-smoker healthy controls were included for comparison. The highest incidence of LOH in both tumour and saliva occurred on chromosomes 9p, 3p and 17p. LOH of at least 1 of the 25 markers was detected in 18/37 (49%) of saliva samples and 32/37 (86%) of the matched tumours. LOH was not found in any of the control saliva samples (El-Naggar, Mao et al. 2001) but again the sample size is very small. If the limitations of the studies are accepted, both illustrate that cancer-related genetic aberrations are detectable in saliva.

### 1.4.4.2 Transcriptomics and Proteomics

It could be argued that tumorigenesis and cancer pathways are ultimately controlled by proteins. As proteins undergo post translational modification and degradation then cancer biology is potentially more complex than genetic modification alone. Proteomics may be more reflective of cellular behaviour and better placed to generate useful biomarkers for cancer. Several studies have focused on the detection of HNSCC related protein and transcript markers in saliva.

Using ELISA assays, Katukuro et al (Katakura, Kamiyama et al. 2007) compared OSCC saliva (n=19) against control saliva (n=20) performed for the presence of four potential cytokine protein markers: IL-6, IL-8, IL1Beta and Osteopontin. All four markers were raised in the OSCC group but only IL-6 (86.5pg/ml) compared to zero in the control group was statistically significant ($p<0.05$). The lack of patients with benign inflammatory disease in the control group may mask potential confounding variables as interleukin proteins are also raised in non-malignant oral inflammatory conditions (Giannopoulou, Kamma et al. 2003). However, this rise in fluid phase salivary interleukins correlates with another study in which raised levels of IL-8 in the saliva and IL-6 in the sera of OSCC patients (n=19) was observed when compared with a matched healthy cohort (n=32) . The concentration of IL-8 in OSCC saliva was 720pg/ml and in normal saliva 250pg/ml. There was also a difference in mRNA expression-$1.1x10^8$ copies of IL8 in saliva from OSCC compared to $2.6 x10^6$ present in the control group (St. John, Li et al. 2004). The lack of statistical concordance for the IL-8 marker in saliva between these two studies may be a result of the small cohort size but may also reflect the large inter-individual variation of saliva proteins which is a potential limitation in biomarker development.

The presence of CD44 protein was investigated in the saliva of a large cohort of HNSCC (n=102) and matched controls with benign aerodigestive inflammatory conditions (n=69). Differential CD44 concentrations of 24.4ng/ml in HNSCC and 9.9ng/ml in benign disease are reported. Dependant on the cut off value for 'elevated' sol CD44 sensitivity ranged from 62-70% and specificity between 75-88% this was also dependent on HNSCC site. In those patients who had low salivary sol CD44 expression the authors hypothesized that transcription was switched off by epigenetic alteration of the *CD44* promoter. MSP assays were performed on a limited cohort of saliva samples from HNSCC patients (n=11) and benign controls (n=10). Nine out of eleven OSCC patients showed methylation whereas 0/10 controls exhibited methylation of the *CD44* promoter. They did not extend this assay to the remaining 91 HNSCC saliva samples therefore the utility of methylation of *CD44* as a potential marker for OSCC was not fully explored (Franzmann, Reategui et al. 2007) .

Hu et al (Hu, Arellano et al. 2008) identified five candidate biomarkers (M2BP, MRP14, CD59, Calatase and profilin ) from the pooled saliva of a group of 16 OSCC patients and 16 matched controls using mass spectrometry. This panel of biomarkers had a 90% sensitivity and 83% specificity when validated on an OSCC cohort (n= 48) and matched controls (n=48). They utilised mass spectrometry-LC-MS/MS and 2D gel electrophrosis (2DE) to generate this panel of biomarkers from 461 OSCC and 438 normal non redundant proteins. Overlap between the two groups was noted for 409 proteins. The authors acknowledged difficulty in producing decent 2DE gel patterns because of the presence of high abundance proteins such as amylase and immunoglobulins in saliva. They also found a large number of proteins to be differentially expressed and techniques used to deplete high abundance proteins can deplete other proteins. This study highlights some of the issues with the logistics of developing protein biomarkers.

ELISA and immunohistochemistry techniques are already used in the clinical setting for proteomic analysis but tend to be used for singular or a few proteins. Antibody arrays, 2DE and mass spectrometry techniques are utilised for screening large number of proteins or samples but in the main are limited to the research sphere because of the massive amount of data that can be generated by one specimen. This requires complex bioinformatics software analysis, which has to be performed in specialist laboratories as do some of the mass-spectrometry techniques.

**Messenger RNA (mRNA)**

As a proof of principle of salivary transcriptome oral cancer diagnostics, nine candidate transcripts were selected from a comparison of early stage OSCC (n=10) and healthy controls (n=10) and subsequently tested on a validation cohort using qPCR . Thirty two patients with T1 and T2 oral squamous cell carcinoma and matched normal controls (n=32) comprised the validation cohort. Seven mRNAs (*IL8, ILIB, OAZ1, DUSP1, H3F3A, S100P, SAT* ) were significantly raised in OSCC. At least a 3.52 fold increase was observed in OSCC when compared with control saliva and can be considered cancer related transcripts. A synchronised elevation in the concentration of four of these mRNAs; *IL8, IL1B, OAZ1* and *SAT* achieve a 91% sensitivity and specificity in predicting OSCC samples (Li, St. John et al. 2004).

### 1.4.4.3 DNA Methylation

Mouth and throat rinsing fluid from patients with primary nasopharyngeal carcinoma (NPC) (n=30) and unmatched healthy controls (n=43) was analysed for the presence of *RASSF1A, DAPK, P16, P15* and *E-Cadherin* using MSP. Six normal tissue biopsies were also analysed. All six of the normal tissue samples were free of methylation and only 1/43 normal mouth rinsing fluids displayed promoter methylation in *DAPK* and *P15* but further details of this individual are not given. Methylation of *P15:* 80% *(*24/30), *DAPK:* 73% *(*22/30), *RASSF1A:* 67% *(*20/30), *E-Cadherin:* 53% (16/30) and *P16:* 33% (10/30) were found in the NPC tumour tissue assays. Twenty nine (97%) of the tumours had at least one gene methylated. Of thirty NPC tumour patients, mouth rinsing fluid was positive in *DAPK:* 50%, *E-Cadherin:* 43%, in *RASSF1A:* 37%, in *P15:* 40% and *P16:* 17%. A sensitivity of 90% and specificity of 98% is reported for the gene panel in this cohort of NPC patients (Chang, Chan et al. 2003b). A small normal, unmatched cohort (age 11-84 years) may have favourably affected the specificity results in this study. A second study by the same group examined a single marker *RIZ1* promoter methylation using MSP in mouth-rinsing fluid and tissue from a cohort of NPC patients (n=30). They included 5-Aza-dC treatment of NPC cell lines to verify the re-expression of *RIZ-1*. Oral rinses from healthy volunteers (n=20) were included in this study and 0/20 were positive for promoter methylation of *RIZ1* whereas 18/30 of the mixed stage primary NPC tumour group were positive for *RIZ1* methylation. Of the 18 positive tumour samples 7/18 of the matched mouth rinses were also positive for *RIZ1* methylation. The 12 negative tumours also had negative matching mouth rinse samples. The expression

of *RIZ1* was reactivated by the treatment of 5-Aza-dC in cell lines suggesting it is silenced by methylation (Chang, Chan et al. 2003a).

### 1.4.5 Saliva collection methods

In the process of developing a diagnostic tool, consideration must be paid to patient acceptability, ease of execution and the acquisition of appropriate information. Bisulphite treatment is a requirement of some methylation assays but it has demands on the integrity of DNA therefore, it is a prerequisite of any methylation assay that the harvested DNA be of good quantity and quality. To ensure reproducibility of results, duplicate or triplicate assays are performed therefore the sample DNA yield must be large enough to accommodate multiple runs. If saliva is to be useful as a surrogate source of DNA it is imperative that the collection method will satisfy these conditions.

Saliva collection is well tolerated by patients, requires minimal preparation and equipment, does not require trained personnel and is therefore easily carried out. Saliva also has other practical advantages over the collection of the common surrogate, blood. It is likely to be better tolerated by needle-phobics or groups with compliance issues such as children, mentally impaired or very elderly. Saliva collection is likely to be less time-consuming and easier in those who have poor vasculature secondary to chemotherapy or intravenous drug use and it does not present the same risk of needle stick injury to personnel that venepuncture does. Saliva may have its disadvantages, for example those patients who have been rendered xerostomic by radiotherapy or whose post-surgical, residual oral function and competence makes saliva donation a challenge.

Garcia-Closas et al (García-Closas, Egan et al. 2001) compared the DNA yield and quality from self-collected cytobrush and a self-collected 10ml mouthwash from two separate cohorts. In a cohort of female breast cancer patients the median DNA yield for mouthwash (n=35) measured 57.3μg and cytobrush (n=40) 13.6μg. In healthy males the results were similar mouthwash (n=25) yielded 38.7 μg and the cytobrush (n=28) 13.5μg of DNA. High molecular weight DNA (>23Kb) was reported from 93% mouthwash samples and 60% of cytobrushes. This study used a postal system for distribution and return of samples. Despite not immediately processing the sample and freezing it, high molecular weight DNA, which withstood PCR was obtained from greater than 93% of the mouthwash samples. They found no significant reduction of DNA yield following 1 year of storage at -80 $^{o}$C. This

is encouraging for the use of saliva derived DNA in longitudinal studies. However, a study on the effect of time from collection to DNA extraction on DNA yield from a 10ml mouthwash in 35 healthy people revealed DNA extracted at 10 and 30 days post collection showed a significantly lower yield than day one (P=0.01) (Feigelson, Rodriguez et al. 2001). Both studies show that saliva is suitable for self-collection and postal studies which are attractive features for a cost-effective screening or surveillance programme but DNA from mouthwash stored at room temperature degrades in days. Oragene™ whole saliva collection device was compared with cytobrush, foam brush and oral rinse in 17 healthy volunteers. The authors assessed DNA yield and quality using gel electrophoresis, spectrophotometry and PCR. The Oragene ™ kit had a significantly higher DNA yield (median 181.88μg) when compared to the next highest of 36.56μg from mouthwash (Rogers, Cole et al. 2007). The swab (10.72μg) and cytobrush yield (13.22 μg) was comparable to other studies using similar techniques (García-Closas, Egan et al. 2001).

The saliva collection methods utilised in OSCC studies at the time of our study design were mainly oral rinse samples (Rosas, Koch et al. 2001, Righini, De Fraipont et al. 2007, Carvalho, Jeronimo et al. 2008,) and the majority of published DNA yield data was from healthy individuals (Table 10 in results section). Studies of DNA yield and saliva collection methods in OSCC patients had not been considered in the literature and the majority of data on the DNA yield from collection methods was based on normal individuals.

## 1.5 The role of DNA methylation in saliva in the management of OSCC

### 1.5.1 OSCC diagnosis and prognosis

One of the first proof of principle studies of salivary methylation markers was performed by Rosas and colleagues (Rosas, Koch et al. 2001). Fresh tumour tissue and matched salivary rinses from 30 HNSCC patients with primary disease were interrogated for a panel of markers using MSP for the detection of *P16*, *MGMT* and *DAPK*. Thirty normal controls provided salivary rinses; 15 smokers and 15 non-smokers. Seventeen of thirty (57 %) of primary HNSCC tumours showed methylation of at least one of the three genes individually they were detected at rates of: *P16* in 14/30 (47%), *DAPK* in 10/30 (33%) and *MGMT* in 7/30 (23%). Strong correlation between tumour and saliva methylation status was reported with

11 patients displaying the same methylation profile in both saliva and tumour. The authors report a higher rate of abnormal methylation from the saliva of patients with oral cavity cancer 8/10 compared to other sites 3/20 $p<0.001$. Six of thirty tumours were positive when the matching saliva was negative. One individual from the control group displayed positive *P16* and *MGMT* promoter methylation. This individual was from the smoking cohort, the significance of which is uncertain as no follow-up data is provided. This is a common criticism of control data but is often imposed by medical ethics. This paper has a limited cohort size and it is not clear if the normal controls are matched to the cancer cohort. However, a strong correlation between tumour promoter methylation and saliva methylation supports the use of saliva as a suitable surrogate for head and neck cancer.

Similar findings were seen in methylation markers profiled in a small study of saliva and tumour tissue from a mixed cohort of oral epithelial dysplasia and OSCC (Viet, Jordan et al. 2007) using quantitative methylation specific PCR (qMSP) to detect: *APC, E-Cadherin, MGMT, P15* and *P16*. Concordance between tumour tissue and matched saliva was reported at; *P16:* 87.5%, *E-Cadherin:* 87.5%, *P15:* 62.5%, *MGMT:* 62.5% and *APC:* 12.5%. Considered as a panel the methylation rate is reported as 71% for detection of OSCC using saliva. Higher overall rates of methylation are found in *P16* and *MGMT* than reported by other authors but this could be a result of a more sensitive detection assay or that samples in this study are all from the oral cavity. It must be noted that this was in a mixed cohort of premalignant and OSCC disease and the results are not stratified by tumour type.

A larger, longitudinal study further bolstered saliva as a tumour surrogate for HNSCC and suggested saliva derived methylation markers may be useful in prognostication of HNSCC. A panel of genes (*P14, P15, P16, DAPK, RASSF1A, FHIT, MGMT, hMLH1, ECAD, APC,TIMP3, RARB, DCC, ATM, THSB1, CASP8*) was evaluated for the detection of early head and neck recurrence (Righini, De Fraipont et al. 2007). A cohort of 90 HNSCC patients provided tumour tissue and matched salivary rinses. Thirty healthy controls also provided a salivary rinse. Sixty nine of the ninety tumours were positive for promoter methylation of at least one gene as detected by MSP. *TIMP3* (46%), *ECAD* (36%), *MGMT* (29%), *P16* (29%), *DAPK* (27%) and *RASSF1A* (20%) were the most frequently methylated genes in disease but not in control samples and were used for post-operative surveillance. Using this panel of markers 24/47 (51%) individuals had the same methylation profile in paired saliva and tumour samples (Righini, De Fraipont et al. 2007). Post-operative samples positive for at least one gene predicted recurrence before clinical examination. The results are supportive of saliva

samples as a clinical adjunct beyond initial disease detection and this paper will be further explored in the discussion. It may be that saliva biomarker tests prove useful in targeted populations at high risk of recurrence rather than blanket screening.

An Illumina GoldenGate methylation array analysis of 807 genes was performed on tumour tissue and pre and post-operative whole saliva samples from 13 OSCC patients. Saliva from ten normal controls was included as part of the analysis. Thirty four of the eight hundred and seven genes were methylated in tumour tissue and pre-operative saliva but not in post-operative or normal saliva. The authors report the highest sensitivity of 77% and specificity of 87% was generated by a panel of 6 genes: *GABRB3_E42_F, IL11_P11_R, INSR_P1063, NOTCH3_E403_F, NTRK3_E131_F and PXN_P308_f* (Viet, Schmidt 2008). This study demonstrates the feasibility of using array technology on saliva samples for the discovery of novel genes that may be missed using a candidate approach.

A criticism of surrogate literature is the absence of controls which are matched for risk factor exposure. There is a tendency to publish attractive results on the total absence of methylation in control samples but often the numbers are small or the control cohort bears little resemblance to the disease cohort, which makes true comparison of results difficult and questions the validity of the proposed markers for clinical use (Carvalho, Jeronimo et al. 2008). This was addressed in a study containing a large cohort of age and risk factor matched controls. Using qMSP they compared the methylation profile of saliva and serum from a large HNSCC cohort (n=211) and a normal control group (n=527). In a multistep study they developed a set of potential gene panels that can distinguish cancer cases from normal controls with good specificity but limited sensitivity. One of the better panels contains *MINT 31, CCNA1, DCC, DAPK* and *P16* and offers a sensitivity of 34.1% and a specificity of 91.8%. In comparison to the other surrogate studies presented, the sensitivity is poor but may be a result of using a larger mixed HNSCC cohort. As a second arm to the study the authors looked at sera of the same cohorts, they note that the same biomarkers are not predictive for HNSCC in serum as they are in saliva, suggesting the phenomenon of *compartment specificity*. The authors suggest it may be that the genes found in blood derivatives may represent more aggressive, metastatic disease than the luminal fluids more specific to the site of the tumour (Carvalho, Jeronimo et al. 2008).

### 1.5.2 Premalignant monitoring

There is a need to diagnose oral cancer early to improve prognosis and survival. Earlier in the history of molecular biomarker research it was felt that more good quality studies are required on the molecular characteristics of premalignant lesions before an informative biomarker for OSCC can be found (Westra, Califano 2004) and this still remains the case today. Promising biomarkers are starting to emerge but require validation on larger cohorts, preferably in multi-centre trials. Hall et al (Hall, Shaw et al. 2008) studied promoter methylation of *P16*, *MGMT*, *CyclinA1* and *Cytoglobin* in biopsy proven oral epithelial dysplasia (OED) using MEP. Longitudinal data was available for up to 3 years on 38 patients with OED. Fourteen of thirty eight cases underwent malignant transformation of the lesion to OSCC. Eight of the fourteen cases that transformed (57%) had *P16* promoter methylation in at least one of the oral scrape samples taken prior to development of OSCC, leading to the conclusion that promoter methylation of *P16* has potential as a predictor for malignant transformation in OED. These findings are supported in a later study. In a cohort of mild and moderate oral epithelial dysplasia, 22/78 OED lesions transformed into OSCC. Forty four percent of patients with *P16* methylation transformed versus 17.4% of unmethylated cases. *P16* methylation was also associated with non malignant progression of dysplasia grade. In those patients who transformed, cancer free survival was shorter than those without *P16* methylation but statistical significance was not achieved for this outcome (Cao, Zhou et al. 2009). Nineteen patients with homogeneous oral leukoplakia (group L) and 15 patients with leukoplakia and a history of OSCC (group LCP) provided salivary rinse for detection of methylation of *P16, P14* and *MGMT* genes using MSP. The methylation rates for the gene panel are reported as 15 /19 (79%) patients in the L group and 13/15 (87%) in the LCP group. *P14* was infrequently methylated 4/34 cases. *P16* and *MGMT* were more highly methylated 15/34 and 19/34 cases respectively but were equally distributed amongst both groups (López, Aguirre et al. 2003). It is not made clear if these groups of patients were matched for risk factors such as smoking or alcohol intake. This data set shows that promoter methylation is detectable in the oral rinses of patients with oral leukoplakia but the small cohort, lack of normal control and follow-up data means we are unable to draw conclusions about whether methylation of these genes is predictive of malignant transformation. This data may suggest that these markers could have a role in early disease or field change.

The literature evidences that methylation of *P16* is associated with malignant transformation of OED and that it is detectable in saliva but studies on saliva are few, lack

longitudinal clinical follow up and include a small number of markers. Considering the existing evidence it is feasible that the detection of methylation markers in saliva could have a role in the monitoring of premalignant lesions but further studies are required and should include a panel of markers to increase detection and should include clinical follow-up data.

### 1.6 DNA methylation detection techniques

To date the relative detectability of methylation has ensured its prevalence over other epigenetic changes such as histone acetylation and miRNA in cancer research. The methylation pattern and DNA strand are inherently more stable than RNA and RNA molecules (Cottrell 2004). This is important because samples destined for DNA extraction require less specialist storage and handling than perhaps those whose fate is proteomic or RNA based assays.

Fan (Fan 2004) summarises the advantages of methylated DNA as a potential biomarker for detection of cancer. It has a) *High informity*- promoter methylation is more common than genetic alterations in cancer b) *Simplicity*- it occurs in the same well-defined region of any given gene whereas a wide range of mutation variations occur within a given gene among different cancers. c) *High sensitivity*- promoter methylation constitutes a 'positive signal' that can be detected in a background of normal cells where loss of heterozygosity (LOH), homozygous deletion and microsatellite Instability (MSI) are negative signals which present detection difficulties in the presence of normal cells a situation likely to be seen in surrogate samples. This positive signal means that the presence of hypermethylated DNA is a dichotomous 'yes' or 'no' result. This is favourable when compared with many of the proteomic markers whereby significance relies on a difference in levels and can require involved bio-informatic analysis.

DNA methylation is a chemical change to the DNA which occurs with the addition of a methyl group to the 5' carbon of cytosines within CpGs under the influence of DNA methyltransferase enzymes (Jones, Gonzalgo 1997, Okano, Xie et al. 1998). Broadly there are three techniques to discriminate methylated from unmethylated cytosine: digestion by methylation sensitive/insensitive enzymes, enrichment by DNA affinity or chemical conversion by bisulphite modification (Laird 2010, Hsu, Weng et al. 2014). The choice of analysis will depend on the need for quantitation, sensitivity and output (Cottrell 2004).

There are several techniques based on methylation sensitive restriction enzymes (MSRE): downstream gel electrophoresis and Southern blotting (De Bustros, Nelkin et al. 1988), restriction landmark genome scanning (RLGS)(Smiraglia, Plass 2003) and methylation sensitive arbitrarily primed PCR techniques (MS-AP-PCR)(Gonzalgo, Liang et al. 1997) but they generally require large amounts of high quality DNA and can have limited throughput (Laird 2010). The various affinity enrichment techniques based on methyl-DNA immunoprecipitation (MeDIP) and methyl-CpG binding proteins (MBD) (Nair, Coolen et al. 2011, Hsu, Weng et al. 2014) have similar demands of the DNA and are therefore not ideal for use with body fluids. The main quantitative methods of detection of DNA methylation at specific loci are based on modification by sodium bisulphite (Shen, Waterland 2007).

### 1.6.1 Bisulphite conversion applications

Bisulphite treatment induces a cytosine to uracil base change in unmethylated DNA but the methlyated cytosines remain intact thereby generating a sequence change (Herman, Graff et al. 1996) which can be detected by various methods.

#### 1.6.1.1 Methylation specific PCR

Endpoint methylation specific PCR was first described by Herman et al (Herman, Graff et al. 1996) and is a very sensitive technique which relies on primer sets designed to bind and amplify only the methylated or unmethylated target DNA. The PCR products are resolved by gel electrophoresis; the presence and intensity of a band determines the presence of methylated DNA therefore it is a subjective qualitative analysis of methylated DNA. At high cycle numbers it can become vulnerable to mis-priming which has been reported at 10% of samples (Shaw, Akufo-Tetteh et al. 2006). The specificity of the technique is limited due to the risk of failure of the bisulphite conversion of unmethylated cytosines to uracil. A false assumption that all cytosines are methylated can result in positive bias. If the amount of target DNA is small as in the context of surrogate samples this could significantly bias the results. However, as it is highly sensitive it was commonly used in earlier studies for detection of methylation in saliva (Rosas, Koch et al. 2001, Righini, De Fraipont et al. 2007).

#### 1.6.1.2 Pyrosequencing and Methylation Enrichment Pyrosequencing

Pyrosequencing is a sequencing-by-synthesis method which involves enzymatically synthesizing the complimentary strand of the target DNA. The use of chemiluscent enzymes release detectable light when a base is incorporated to the strand (Ronaghi, Karamohamed et al. 1996). As it directly sequences the target DNA it is capable of quantitative measurement of methylation and is useful in assessing tumour specificity of potential candidate genes (Shaw, Liloglou et al. 2006, Nikolaidis, Raji et al. 2012) but can be limited in

body fluids as the analytical sensitivity is approximately 5% (Markopoulou, Nikolaidis et al. 2012). Methylation enrichment pyrosequencing (MEP) involves interrogating the MSP amplicon with a pyrosequencing assay to assess the completeness of bisulphite conversion and therefore addresses the specificity issues of MSP. MEP has been successfully used to detect methylation in saliva (Shaw, Akufo-Tetteh et al. 2006) and from oral scrapes (Hall, Shaw et al. 2008). Although accurate, MEP is another multi-stage assay which is time and resource consuming.

#### 1.6.1.3 Quantitative methylation specific PCR

Real-time methylation specific PCR or quantitative MSP (qMSP) is based on similar design principles to MSP but it incorporates the use of a fluorescent probe which improves sequence specificity (Liloglou, Field 2010). The sequence specific probe binds to the target DNA and as the methylation specific primers extend the fluorescent chemistry of the probe is activated and can be detected in real time (Eads, Danenberg et al. 2000). It is an automated assay and the use of fluorescence renders it a highly sensitive technique suitable for samples with small amounts of target DNA. It has been discussed earlier that bisulphite treatment can degrade DNA. One approach to overcome this limitation is to increase the amount of DNA template. Genome wide amplification of bisulphite treated DNA from plasma samples followed by pyrosequencing to interrogate a number of markers has been described (Vaissière, Cuenin et al. 2009). Another approach is to conserve the available DNA by multiplexing the qMSP reaction with several target genes tested simultaneously (Fackler, Malone et al. 2006) but the optimisation process can be complex (Liloglou, Field 2010) due to the different abundance of targets and the differential consumption of reagents.

qMSP is the popular choice for the detection of methylation in body fluids. This is due to the sensitivity of the technique as the signal can be amplified therefore samples with low concentrations can be used as sources of methylated target DNA (Cottrell 2004) a common situation when using surrogate samples which are predominantly normal cells. It has proven reproducible and allows relatively high throughput processing. It is a one stage automated assay and can be used to quantitate methylation by relative comparison to a standard curve incorporated on each plate. Quantitation is an attractive and necessary function if methylation analysis is to be used to monitor clinical therapeutics or if methylation is present in the CpG promoter regions of normal tissues. The amount of methylation may then determine the clinical value of a marker.

After consideration of the literature there is a pressing need for earlier diagnosis of OSCC and recurrent disease to improve survival outcome. Improvement in the predictive and prognostic tests available for premalignant lesions at risk of transformation and post-operative OSCC is also required and molecular analysis may provide this. Body fluid derived molecules show promise for non-invasively, detecting and monitoring malignancy. DNA is a stable molecule and is readily detected in saliva which seems to be the most obvious choice for detection of malignancy in the oral cavity. Promoter methylation is a common event in cancer and is chemically stable lending itself to detection. There is evidence that it is detectable in saliva and is associated with OSCC and premalignant lesions at risk of transformation. To progress with developing a biomarker, the methylation level in DNA derived from healthy individuals is required. Many existing studies have used small control groups and have not specified a cut-off to define a positive result which this study hopes to address.

## 2. Research Strategy

### 2.1 Aims

The overall aim is to assess saliva as a surrogate tissue for the detection and surveillance of oral squamous cell carcinoma (OSCC).

### 2.1 Specific aims of the study

- To determine the rates of methylation of a panel of epigenetic markers in the saliva of a cohort of normal individuals.
- To compare rates of methylation in known OSCC patients with the normal cohort to establish a threshold by which we could determine disease status in future testing.
- To apply these thresholds in a series of patients with premalignant disease and assess if there are any differences between methylation in saliva from dysplastic lesions considered as low and high risk of malignant transformation.
- To assess methods of saliva collection in patients with OSCC with specific reference to patients who received post-operative radiotherapy with the aim of determining the validity of saliva collection methods for patients with xerostomia.

### 2.2 Research plan

### Target populations

The inclusion of a control cohort of patients was to establish the level of methylation in normal saliva to set a cut-off to discriminate normal from disease. At the time of this study design, the majority of similar studies used small control groups ranging from n=5 (Viet, Schmidt 2008) to n=30 (Rosas, Koch et al. 2001, Righini, De Fraipont et al. 2007). The reported specificity of potential candidate genes may be falsely high as a result of using young volunteers who have not been exposed to risk factors for the disease (Carvalho, Jeronimo et al. 2008). To address this limitation in the current literature our intention was to collect samples for the control cohort from individuals matched to the demographic and risk factor profile of typical OSCC patients. The control cohort is composed of normal saliva and oral scrape samples collected from patients with no clinical evidence of oral squamous cell carcinoma or pre-malignant oral lesions at the time of collection. Measures were taken to ensure that the demographics of the control population matched the disease population.

The inclusion of an OSCC cohort was to establish disease-status methylation profiles using saliva as a tumour surrogate. Power calculations were based on a pilot of a comparable, clinically acceptable, functioning screening programme into faecal occult blood for detecting bowel cancer which operated with a sensitivity of 60% (DOH report 2006, Pilot of Bowel Screening). The aim of our work was to achieve an 80% (+/- 10 %) sensitivity and specificity with a panel of methylation biomarkers able to discriminate normal from OSCC saliva samples. To achieve this, 65 patients were calculated to be required for the normal and OSCC group and were recruited (Table 1).

In a separate part of the study, post-treatment samples were also collected to assess the utility of this panel for the detection of residual disease. Published data from a Liverpool cohort of OSCC patients reported 21% of OSCC patients developed loco-regional recurrence (Rogers, Brown et al. 2009) and an earlier study found that the median time to recurrence was 8 months with 90% of recurrences occurring within 2 years (Kissun, Magennis et al. 2006). In view of the reported frequency of recurrence the intention was to collect pilot data from longitudinal post-operative saliva samples from 65 OSCC patients which would approximate 14 cases of recurrent disease.

A dysplasia cohort was included in this study to provide preliminary data on the utility of this panel of genes in saliva as a marker of malignant transformation before the clinical/histological transformation of the lesion. At the time of this study design there was a lack of longitudinal studies evaluating methylation in saliva from OED patients. There are varying reports of OED malignant transformation rate (MTR) in the literature; 12% MTR with a mean time to transformation of 4.3 years was reported for histologically confirmed OED in a meta-analysis of the literature (Mehanna, Rattay et al. 2009). In the context of disease surveillance, the aim was to collect between 2 to 5 years of longitudinal samples from a cohort of 50 patients to provide pilot data on the utility of this panel of methylation markers to predict malignant transformation.

<table>
<tr><th></th><th>Dysplasia cohort</th><th colspan="2">Normal cohort</th><th colspan="2">OSCC cohort</th></tr>
<tr><td>Number patients</td><td>63</td><td colspan="2">97</td><td colspan="2">104</td></tr>
<tr><td rowspan="3">Number of samples</td><td rowspan="3">87</td><td rowspan="3">Saliva</td><td rowspan="3">97</td><td>Pre-operative saliva</td><td>69</td></tr>
<tr><td rowspan="2">Scrapes</td><td rowspan="2">90</td><td>Post-operative saliva</td><td>47</td></tr>
<tr><td>Matched tumour tissue</td><td>34</td></tr>
</table>

**Table 1. Total number and type of samples collected and included in analysis.**

### Sample collection methods

The purpose of this arm of the study is to evaluate the best method of saliva collection to maximise DNA yield in OSCC patients. To ensure the utility of salivary biomarkers in post-operative disease surveillance, it is important to establish which saliva collection techniques provide sufficient DNA in both irradiated and non-irradiated OSCC patients. As patients at high risk of recurrence will usually have received post-operative radiotherapy and may be xerostomic as a result.

### Rationale for the chosen method of detection

A challenge presented by the investigation of saliva rather than tumour, is the presence of relatively large proportions of DNA from normal oral mucosa within which the occasional methylated DNA has to be detected. qMSP is a high throughput, semi-quantitative technique. Specificity is provided by methylation specific primers that will only amplify the methylated DNA and the sequence specific probe will only bind to the methylated target sequence. Our research group had considerable experience with qMSP assays in body fluid samples prior to this proposed study. It had proven reliable and reproducible with a technical sensitivity of 0. 1% established with non-clinical samples. In view of these attractive features it was selected as the best method of methylation detection in saliva for this study.

## Rationale for the methylation panel

Ideally a biomarker would display high informativity, in that it is present in most of the disease cases and high specificity being present in the cancer but not many of the normal controls. Such a discriminatory biomarker does not yet exist for HNSCC; the highest methylation rates for well described singular methylation markers such as *P16, MGMT, DAPK* and *CDH1* are consistently observed between 30-60% (Ha, Califano 2006, Demokan, Dalay 2011) however in normal tissues, methylation rates of between 0-41 % are reported (Viswanathan, Tsuchida et al. 2003, Maruya, Issa et al. 2004, Shaw, Liloglou et al. 2006, Kato, Hara et al. 2006, Righini, De Fraipont et al. 2007, Steinmann, Sandner et al. 2009). HNSCC tissue is molecularly and histologically heterogeneous. This feature may limit the use of singular markers; some studies observed increased informativity (75-77%) when using a panel inclusive of *P16, MGMT and DAPK* (Viswanathan, Tsuchida et al. 2003, Righini, De Fraipont et al. 2007). A combination of multiple methylation markers increased sensitivity to 85% in a large cohort of HNSCC salivary samples but impacted on specificity (30%) (Carvalho, Jeronimo et al. 2008). In this study, we have employed a panel of five genes which have previously proven discriminatory between normal and tumour HNSCC tissue using pyrosequencing methylation assays. Five genes were chosen with the aim to improve sensitivity of the methodology above that achievable with one or two markers.

# 3. Materials and methods

## 3.1 Collection of tissue samples

### 3.1.1 Normal cohort

Patients attending Liverpool University Dental Emergency Department (ED) were initially recruited over a period of 12 months. The experimental protocol was approved by Liverpool Paediatrics Research ethics committee (LPREC 08/H1002/42)(Appendix 1 A1.1). The surrogate samples in the form of saliva and oral scrapes which form the normal control cohort have been collected under LREC 08/H1002/42. I wrote the ethics application, attended the ethics meeting and was granted approval for this arm of the study in August 2008. Patients were provided with an information leaflet (Appendix 1 A1.2) half an hour prior to sample collection and informed written consent (Appendix 1 A1.3) was obtained. Patients were assessed for suitability to enrol in the study by clinical examination and a questionnaire (Appendix 1 A1.4). Patients aged 16 and over, who were able to provide a saliva and oral scrape sample and complete the questionnaire, were entered into the study. Patients were excluded from the study if they had a clinically detectable premalignant or malignant oral lesion or they had a history of premalignant or malignant pathology in the aerodigestive tract. Patients were not excluded if they had a history of malignancy remote from the aerodigestive tract. A saliva sample and a paired oral scrape were obtained from each patient together. Each patient completed an anonymous questionnaire detailing relevant demographic and disease risk factor information at the time of sample collection. Consent was not linked to the research number allocated to the samples and data from that individual. An interim assessment of the demographics of the first 72 patients, which were collected consecutively in the ED, determined that the age range was skewed towards younger patients (<40 years). To address this, a further 25 older patients were also recruited from the prosthetics and restorative dental clinics at Liverpool University Dental Hospital over a subsequent eight months which makes a total of 20 months.

Whole saliva was collected using a commercially available Oragene$^{TM}$ vial (Figure 1). Patients were asked to deposit 2ml of whole saliva into the vial as per manufacturer's guidelines (DNA Genotek Inc 2012). Oral scrapes were obtained after the deposition of the saliva sample by scraping the buccal mucosa using a plastic spatula housed in a collection tube (Figure 2). The spatula was pressed firmly on bilateral sites of the buccal mucosa and cells were harvested using five strokes per side. Cellular debris was visible on the spatula. A

similar technique has previously been used by our research group to obtain DNA for biomarker methylation assays (Hall, Shaw et al. 2008). Samples were stored at -80$^{o}$C prior to DNA extraction.

**Figure 1. Oragene$^{TM}$ vial used to collect whole saliva.**

**Figure 2. Plastic spatula used to collect oral scrapes.**

### 3.1.2 Squamous Cell Carcinoma Cohort

Pre-operative saliva samples (n=69) had previously been collected from patients with histologically confirmed HNSCC along with tumour tissue at the time of primary surgery by other researchers within our group. Patients were examined and diagnosed in the Department of Oral and Maxillofacial Surgery in the University Hospitals Aintree NHS Foundation Trust, Liverpool UK and enrolled into a study of molecular biomarkers of oral cavity cancer. The experimental protocol has ethical approval (Sefton REC ref.no. EC 47.01 and Liverpool (adult) LREC ref no.07/Q1501/15). Patients were selected for entry into the study if they could provide informed consent and had a histological diagnosis of HNSCC, with a primary tumour large enough to be divided for pathological diagnosis and research samples, therefore small T1 tumours were excluded. Patients who had provided saliva samples were identified from our research database. Twelve of sixty nine pre-operative patients provided a post-operative sample at four weeks after primary surgery prior to any radiotherapy or chemotherapy using the same collection method. A further eleven patients provided only post-operative saliva samples as part of the collection methods study. Saliva samples used in this arm of the study were collected prospectively between the years 2006-2008 and were stored at -80$^{o}$C at the University of Liverpool. Whole saliva was collected using the Oragene$^{TM}$ vial as described earlier. Clinical follow-up data was accessed from clinical records and the departmental research database.

### 3.1.3 Post-operative Radiotherapy (PORT) Cohort

Post-operative saliva samples were collected from an additional twenty four patients who had received surgical and post operative radiotherapy treatment for OSCC. Whole saliva samples were collected using Oragene™ sponges using the manufacturer's collection protocol (DNA Genotek Inc 2011). In brief, five sponge tips were placed individually in areas of pooled saliva in the oral cavity of each patient. Once loaded with saliva, each of the five tips were placed into a standard Oragene™ vial containing preservatives and stored at -80°C. A paired mouthwash sample was collected immediately after Oragene™ sponge harvesting of saliva by asking patients to swill 25ml 0.9% normal saline (Sterets Normasol, Medlock Medical Ltd, Oldham, UK) from a sterile 50ml tube around the mouth for 30 seconds before depositing the mouthwash back into the tube. Mouthwash samples were immediately stored at 4°C in clinical refrigeration for a maximum of three hours before being subjected to centrifugation at 1,200g in a bench top centrifuge for five minutes. The supernatant was removed and the cell pellet stored at -80°C.

### 3.1.4 Dysplasia Cohort

Patients who had provided saliva samples as part of an ongoing clinical and molecular biomarker study of oral epithelial dysplasia (OED) were identified from our research database. Patients had been enrolled into that study if they had histological confirmed OED, were available for clinical follow-up and could provide informed consent (Sefton REC ref.no. EC 47.01). Patients with a synchronous OSCC were excluded, but those with a history of oral cancer were not if they had received surgical curative treatment. Previous OSCC was not an exclusion factor because these patients present a high risk group and present an area of interest to molecular biomarker research; particularly as previous surgery and radiotherapy can make clinical and histopathological assessment of these lesions difficult. Eighty seven saliva samples were available from 63 patients attending the tertiary oral dysplasia clinic at the University of Liverpool dental school. Saliva samples used in this arm of the study were collected between the years 2006-2011 and were stored at -80°C at the University of Liverpool. Whole saliva was collected using two methods. The first method required patients to deposit saliva directly into a plastic vial which was immediately frozen using dry ice and transported to the onsite -80°C freezer (13/87 samples were collected with this method). As of October 2007 onwards whole saliva samples were collected using the Oragene™ vial (74/87) as a more convenient method to collect saliva in the clinical environment.

## 3.2 DNA Extraction

### 3.2.1 Oragene™ vial whole saliva

DNA extraction from the Oragene™ vials was performed using the manufacturer's protocol for manual purification from 0.5ml of sample (DNA Genotek Inc 2015a) with the exception that Oragene purifier™ was used instead of PT-L2P purifier. In brief, the total sample was incubated at 50°C for two hours in a shaking incubator to lyse cells and digest nuclear proteins. Twenty microlitres of Oragene™ purifier was added to 0.5ml of sample before precipitation with 100% ethanol. The sample was washed with 70% ethanol before final re-suspension in 100µl 1 x TE buffer. The prepared DNA was stored at -20°C until required.

### 3.2.2 Oragene™ sponge whole saliva

The vials containing the saliva soaked sponge tips were incubated in a 50 °C water bath for 1 hour. As much free liquid as possible was removed from the vial with a sterile pipette, along with excess liquid from the sponges, and transferred to a 15ml centrifuge tube. The saliva sponges were placed into the barrel of a 5ml plastic syringe and this was inserted into the 15ml centrifuge tube. Samples were centrifuged at 22°C, at 1,200g for five minutes. This step was repeated in an attempt to maximise the expulsion of liquid from the sponges into the 15ml collection tube. The 5ml barrel containing dried sponges was removed and discarded. The manufacturer's protocol for manual purification of DNA from the whole sample (DNA Genotek Inc 2015b) was followed for the entire liquid volume of the sample with the exception that 1/25th volume Oragene™ purifier was used in place of PT-L2P purifier. DNA was re-suspended in 200µl 1 x TE buffer and stored at -20°C until required.

### 3.2.3 Mouthwash

DNA was extracted from the processed mouthwash samples according to a cell spin column extraction protocol (DNeasy Blood and Tissue Kit, 2006 Qiagen, Maryland, USA). Briefly, phosphate buffered saline was added to each cell pellet to make a final volume of 500µl. Twenty five microlitres of proteinase K (20mg/ml) was added to an 180µl aliquot of the original sample, 200µl of buffer AL was added to the sample prior to vortexing and incubating the sample at 56°C for 10 minutes. The DNA was separated from cellular debris using Qiagen DNeasy spin columns with elution in 200µl of AE buffer (Qiagen). The centrifugation steps were carried out at 19°C. The prepared DNA was stored at -20°C until required.

### 3.2.4 Whole saliva with no preservative samples

Whole saliva samples in sterile vials were defrosted and the volume made up to 500µl with 1x TE buffer solution. An equal volume of 2x PK buffer (Appendix 2) was added followed by 20µl of Proteinase K (20mg/ml). The Oragene™ whole saliva 0.5ml protocol was then followed for each sample. The prepared DNA was stored at -20°C until required.

### 3.2.5 Oral scrapes

Oral scrape vials were defrosted at room temperature; the spatula used to collect the cells was snapped off and left in the vial. Two hundred microlitres of phosphate buffered saline was added to the vial followed by 25µl of proteinase K (20mg/ml). Two hundred microlitres of buffer AL was added to the sample prior to vortexing and incubating the sample at 56°C for 10 minutes. A cell spin column extraction protocol (Qiagen DNeasy Blood and Tissue Kit, 2006) was followed. Centrifugation was carried out at 19°C. The DNA was resuspended in 100µl of AE buffer and stored at -20°C.

### 3.2.6 Tumour tissue preparation

DNA was extracted from matched tumour and normal tissue (adjacent to tumour approx 10mm distance from macroscopic edge of tumour (Shaw, Liloglou et al. 2006) from 24 patients (48 samples) with histologically confirmed OSCC which were recruited as part of the epigenetic biomarkers of oral cavity cancer study (Sefton REC ref.no. EC 47.01). The tissue had been snap frozen at collection and stored in -80°C freezer until use. The tissue was micro-dissected in an extraction hood using a size 15 surgical blade to provide a sample of tumour tissue approximately $2mm^3$ in size which was cut into smaller pieces prior to lysis. A cell spin column extraction protocol (Qiagen DNeasy Blood and Tissue Kit, 2006) was followed. 100µl of AE buffer preheated to 70°C and centrifugation at 8000rpm for 1 minute was used to elute the DNA. A second elution step was added to maximise the yield. The DNA was stored at -20°C prior to use.

## 3.3 DNA Quantification

### 3.3.1 Spectrophotometry

Corrected UV absorbance spectrophotometry (260:280nm and 260:230nm) using a NanoDrop ND1000 spectrophotometer (Thermo Scientific, Waltham, MA) was used to

measure the DNA yield and purity in 5μl of each sample. This data was used to calculate the volume required to standardise each sample to 500ng of DNA for bisulphite treatment.

### 3.3.2 Absolute qPCR

In a subset of 90 saliva samples provided by 21 patients from the OSCC cohort and 24 patients from the PORT cohort; the amount of human DNA was quantified using a human ribonuclease P (*RNAse P*) gene quantitative PCR (qPCR) assay using the standard curve programme of an AB7500 Fast system according to the manufacturer's guidelines (Applied Biosystems, Foster City, CA). A standard curve of known DNA concentration was prepared from a stock of human lymphocyte DNA (500ng/μl). Serial dilutions were made to include seven standards of 100ng/μl, 50ng/μl, 25ng/μl, 12ng/μl, 6ng/μl, 3ng/μl, 1.6ng/μl and a 'no template' control (NTC). DNA was quantified using Picogreen (QuantIT-kit-Life Technologies, Paisley, UK). All of the experimental samples and the controls were run in triplicate. Premixed stock mastermix consisted of; Taqman Universal Master Mix II (Applied Biosystems) 10μl, RNAse P probe 1μl (Applied Biosystems #4403326), double distilled $H_2O$ 7.5μl and target DNA 1.5μl. Reactions were heated at 95$^{o}$C for 10 minutes and subject to 50 amplification cycles of 15 seconds at 95$^{o}$C and 60 seconds at 60$^{o}$C. The cycle threshold ($C_T$) was plotted against a $\log_2$ scale of the standard dilutions to create a linear model of their relationship. This linear model was used to estimate DNA concentration of the samples for any given $C_T$ value.

## 3.4 Bisulphite treatment

Five hundred nanograms of each DNA sample was treated with sodium bisulphite as per the manufacturer's protocol using Zymo EZ-96 DNA Methylation-Gold kit (Zymo Research Corporation, Orange, CA, USA), eluted in 30μl of M-elution buffer and stored at -80$^{o}$C until use.

## 3.5 Quantitative PCR Quality Control

### 3.5.1 Unmethylated Technical control

Whole genome amplified DNA was included as an absolute unmethylated DNA technical control for each assay. It was made using RepliG Screening Kit (Qiagen) as follows; up to 600ng of human lymphocyte DNA in 5μl 1 xTE added to 85μl of SB1. Heated to 65$^{o}$C for 5 minutes and cooled to room temperature before 85μl of SB2 and 5μl of polymerase was added. Finally the sample was then heated to 37$^{o}$C overnight and stored at -20$^{o}$C until use.

### 3.5.2 Methylated controls

Human lymphocyte DNA was methylated *in vitro* (SssI methylase) using a standard protocol. Up to 3μg of human lymphocyte DNA in 25 μl of 1 xTE was added to 3ml of buffer (10x), 1μl SAM and 1μl of SssI methylase (2000u/ml) and incubated at 37°C for 2 hours. A further 1μl SAM was added to saturate the reaction and left at 37°C overnight. As a final step the total sample was heated to 65°C for 20 minutes and stored at -20°C until use.

### 3.5.3 Quantitative measurement of methylation in control DNA

Pyrosequencing quantitation of the methylation levels of control DNA was undertaken so that serial dilutions could be modified to produce accurate methylated standards. End point PCR amplification of the *in vitro* methylated (undiluted, diluted 1:1, 'unmethylated' lymphocyte and WGA) DNA was conducted in triplicate using TTK protein Kinase(*TTK*) as the target gene which is known to be reliable and unmethylated in lymphocyte DNA (Dr T.Liloglou, personal communication); 1μl *TTK* primer mix (10μl 400nmol forward primer and 5μl of 200nmol biotinylated reverse primer), 1μl of 5mM deoxynucleoside triphosphates(Qiagen), 2μl template DNA 50ng/μl, 0.1μl Hot-starTaq polymerase plus (Qiagen), 2.5μl of 10x Coraload PCR buffer (Qiagen) and 18.4μl double distilled water in a total volume of 25 μl. PCR conditions were as follows: 95°C for 5 minutes, 94°C for 30 seconds, 57°C for 30 seconds, 72°C for 30 seconds 40 cycles, 72°C for 10 minutes and a hold step of 6°C. The PCR products were resolved on a 1% agarose gel using 50bp DNA ladder (Hyperladder 5 Bioline) to ensure there was DNA prior to pyrosequencing. PCR products were then sequenced according to a standard protocol using the PyroGold kit (Qiagen) and pyrosequencing system PSQ96MA (Qiagen).

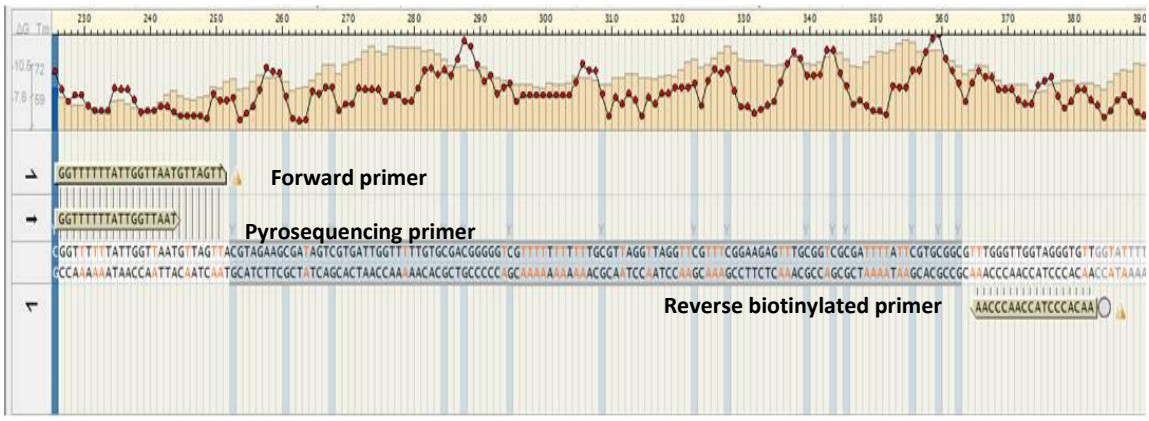



**Figure 3. *TTK* forward and reverse PCR primer and pyrosequencing primer sequence. The CpGs evaluated by the pyrosequencing are shaded in blue and denoted by a Y.**

The average percentage methylation of the *in vitro* methylated lymphocyte DNA was calculated from the pyrosequencing data (Figure 4); this informs the ratio of methylated DNA required to create accurate serial dilutions of 10%, 1%, 0.5%, and 0.25%. These serial dilutions, together with the 'unmethylated' lymphocyte DNA and WGA DNA, underwent bisulphite treatment simultaneously.

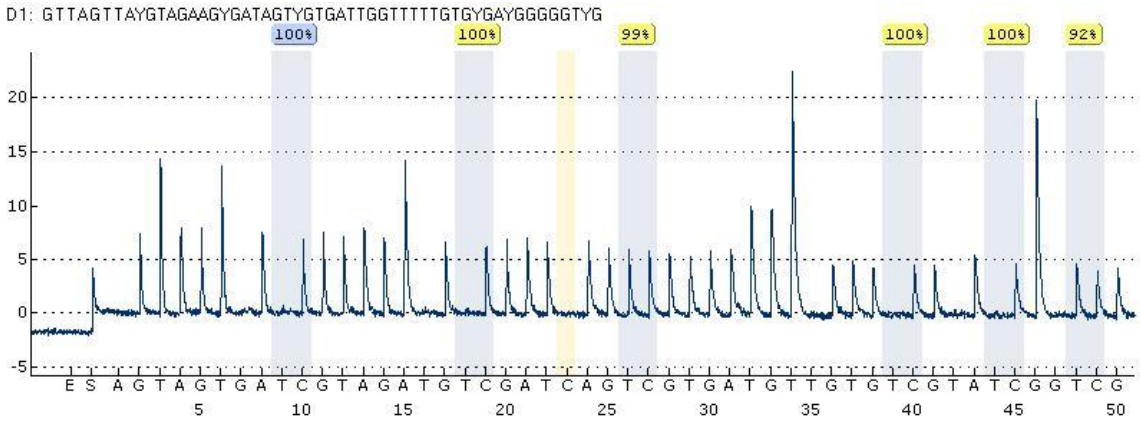



**Figure 4. An example TTK pyrogram of the *invitro* methylated control DNA showing average methylation of 99%. This value is used to inform the input DNA ratios for the standard DNA serial dilutions. The sequence to be interrogated is displayed in the top left corner of this diagram**

## 3.6 qMSP

### 3.6.1 Target gene validation for duplex qMSP assays

Previous work in our research group had identified primer-probe concentrations and thermal profiles for *p16, TMEFF2, CYGB, CCNA1* and *ADAMTS9* in duplex qPCR with beta actin. Validation of the qPCR conditions for the five candidate genes in the current study was based on this optimisation process to ensure reproducible specificity and sensitivity. In brief, 25μl qMSP reactions each contained: 12.5μl Taqman Universal Master Mix II No UNG (Applied Biosystems), 2.5 μl primer/probe mix (concentration range 300-900nM for target gene primers, 150nM target gene probe, 125nM endogenous control (*beta actin)* primers, 125nM *beta actin* probe and 2μl bisulphite treated DNA from serial dilutions of methylated DNA (1%, 0.5%, 0.25%, 'unmethylated' lymphocyte, WGA). The reactions were carried out on a 7500FAST StepOne real-time PCR machine (Applied Biosystems). Assays for each gene used PCR conditions: 95$^{o}$C 10 minutes, 95$^{o}$C 15 seconds and annealing temperatures

ranging from 57-60°C for 50 cycles and an extension step of 72°C. The lengthy optimisation process for the five target genes involved standard protocol primer-limiting assays (Applied Biosystems http://www3.appliedbiosystems.com) which had previously been performed by Mr Brown and Dr Liloglou, other researchers in our group. In our study the endogenous control (beta actin) gene is more abundant than the genes of interest. The beta actin assay amplifies better and consumes the reagents before the other genes can amplify. To prevent this, the beta actin primer concentration is sequentially limited to levels much lower than a standard assay e.g beta actin is 900nm in the singleplex methodology and 125nm in the duplex methodology. The aim is to establish the primer concentration of the more abundant target that provides the earliest cycle threshold value for beta actin, without distorting the cycle threshold value of the target genes (Applied Biosystems). The concentration of the probe was decided by a similar trial and error process starting with probe concentrations that had worked well in the singleplex assay. Once the concentrations of the primer/probe mix for both genes had been established, the most efficient annealing temperature for each gene was determined. Reactions were set up using DNA of known concentration as described above and run across three different temperatures (57-60°C) on the 7500FAST StepOne real-time PCR machine (Applied Biosystems). In the perfect assay, the cycle difference between the 1:2 serial dilutions should equate to one cycle difference between the $\delta C_T$ values. The thermal profile which gives the best cycle differentiation at the lower concentrations of the standard DNA was selected. Another factor which determines the accuracy of the chemistry of the reaction is the signal from the biological control lymphocyte DNA and the technical control WGA. It is known that background methylation is present in human lymphocyte DNA for certain genes therefore a technical control is included. The technical control is unmethylated and therefore should not produce an amplification signal. It was necessary to include the biological control because any assay to be used in a clinical setting must be able to significantly differentiate between background methylation and that associated with disease. It therefore informs the 'cut-off' level considered a positive result for clinical samples.

### 3.6.2 Establishing technical exclusion criteria

Research samples were excluded from analysis if a) beta actin assay failed to generate an amplification curve so there is no DNA in the sample b) the amplification curve of the target

gene was aberrant suggesting degraded DNA. Runs were repeated if a signal was generated by the non-template control.

### 3.6.3 Singleplex qMSP

This qMSP Taqman methodology is based on extensive work on bronchial lavage samples by our research group (Nikolaidis, Raji et al. 2012). The epigenetic biomarkers selected have been identified by a candidate gene approach from work already carried out on tumour/normal paired tissue samples by the Merseyside Head and Neck Oncology Research Group. These include gene promoter methylation of *P16, TMEFF2*, *CyclinA1* (*CCNA1*) and *Cytoglobin* (*CYGB*). *ADAMTS9* was selected as a novel gene for this study as part of a collaborative work which is described in the results chapter.

Methylation assays of the five target genes and the beta actin control were carried out on all 247 saliva samples and 90 oral scrapes using the singleplex methodology. In brief, 25μl qMSP reactions contained: 12.5μl Taqman Universal Master Mix II No UNG (Applied Biosystems), 2.5 μl primer/probe mix (300-900nM target gene primers, 250nM target gene probe and 2μl bisulphite treated DNA (5%, 1%, 0.5%, 0.1% methylated lymphocyte or sample). Primer and probe sequences were designed using Oligo 6.0 software (Molecular Biology Insights, USA) and are provided in Appendix 3. The PCR conditions for each gene are displayed in Table 2.

| Gene | PCR Conditions | N° of cycles | Dye |
|---|---|---|---|
| ***P16*** | 95°C 10 mins | 50 | FAM |
| | 95°C 15 secs | | |
| | 60°C 60 secs | | |
| ***TMEFF2*** | 95°C 10 mins | 50 | FAM |
| | 95°C 15 secs | | |
| | 58°C 15 secs | | |
| | 60°C 45secs | | |
| ***CYGB*** | 95°C 10 mins | 50 | FAM |
| | 95°C 15 secs | | |
| | 64°C 5 secs | | |
| | 61°C 40 secs | | |
| ***ADAMTS9*** | 95°C 10 mins | 50 | FAM |
| | 95°C 15 secs | | |
| | 62°C 60 secs | | |
| ***CCNA1*** | 95°C 10 mins | 50 | FAM |
| | 95°C 15 secs | | |
| | 62°C 60 secs | | |
| ***ACTB*** | 95°C 10 mins | 50 | VIC |
| | 95°C 15 secs | | |
| | 58°C 15 secs | | |
| | 60°C 45 secs | | |

**Table 2. qMSP PCR cycling conditions of singleplex reactions for 5 target genes and *beta actin* (*ACTB*) internal reference gene.**

Beta actin was included to normalise for input DNA. Limited quantities of sample DNA dictated that duplicate runs would be performed and any equivocal results would be tested further in order to obtain a consensus (Nikolaidis, Raji et al. 2012). Duplicate plates for all 6 genes were prepared at the same sitting and stored at 4°C for no more than 12 hours. In

order to run a total of 6 assays (5 target genes and 1 internal control gene) in duplicate, two AB7500 systems were used simultaneously and duplicates were run consecutively on the same machine. The amount of control DNA required to complete all of the assays was calculated and prepared (methylated and bisulphite treated) as one batch to reduce pipetting variation. The Primer/Probe mixture was also made up as a 1000µl master stock prior to preparing the plates (Table 3).

| Target Gene | Forward Primer 100nM | Reverse Primer 100nM | Probe 100nM | Double distilled water |
|---|---|---|---|---|
| | µl | | | |
| *P16* | 70 | 70 | 25 | 835 |
| *TMEFF2* | 90 | 90 | 25 | 795 |
| *CYGB* | 30 | 30 | 25 | 915 |
| *ADAMTS9* | 70 | 70 | 25 | 835 |
| *CCNA1* | 50 | 50 | 25 | 875 |
| *ACTB* | 90 | 90 | 25 | 795 |

**Table 3. Primer probe dilutions used in singleplex qMSP for 5 target genes and beta actin gene (ACTB).**

In order to minimise the effect of pipetting variance when setting up assays using the same DNA, a 'master premix' was made for each plate containing 157µl of Taqman mastermix and 19µl of sample DNA per well and included reference samples of 5%, 1%, 0.5%, 0.1% methylated DNA, 'unmethylated' lymphocyte DNA and a no-template control. Of note, lymphocyte DNA alone was used as the 'unmethylated' control as WGA was not routinely used in the qMSP protocol at this time. From this master plate, 28µl of each Taqman/DNA mix was placed in the corresponding well of 6 'primary' plates. Twenty two microlitres of primer/probe mix was then added to each well of the appropriate primary plate and pipetted to mix the reaction. This produced a final volume of 50µl per well in the primary plate (Figure 5). An exact duplicate plate containing 25µl per well was then made from each of the six primary plates (corresponding to each of the six genes used in this study) and centrifuged at 13000rpm for 1 minute immediately prior to the run.

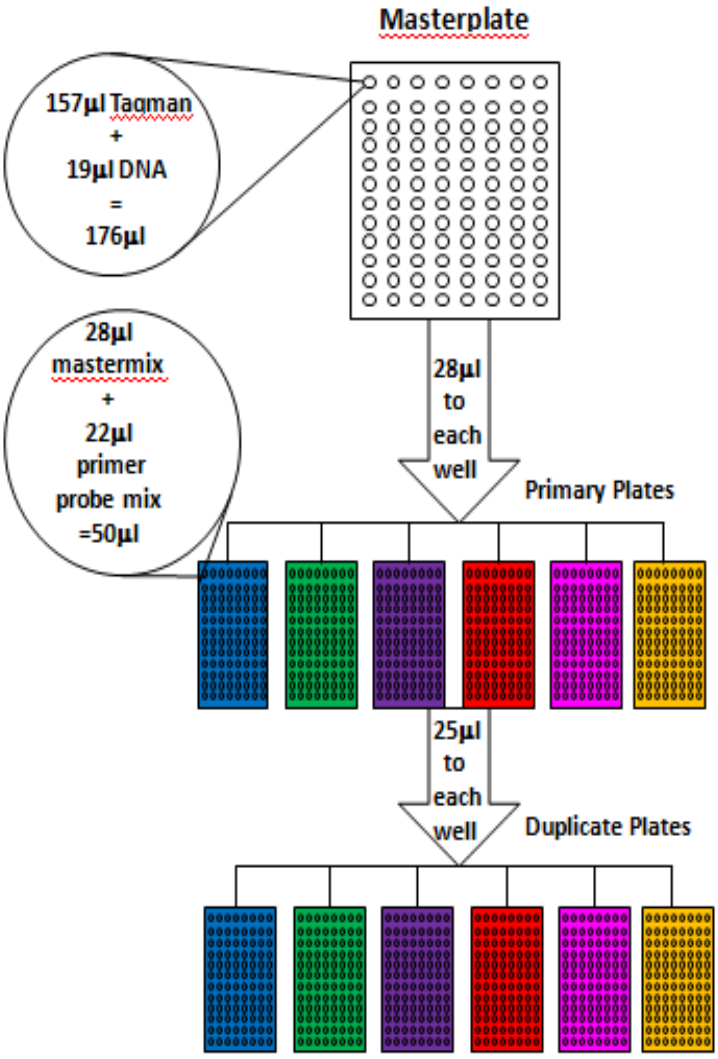



**Figure 5. Diagram of the duplicate plate set-up for singleplex methodology for 5 target genes and beta actin.**

### 3.6.4 Duplex qMSP

Five hundred nanograms of DNA from 258 saliva samples and DNA from 34 tumour samples (tumour prepared by Mr AJ Hobkirk) was bisulphite treated as previously described. The samples were randomised across four 96 well plates prior to qMSP. A stock of 10x probe mix inclusive of the requisite water for the total number of reactions was made up for each target gene in advance of running the reactions (Table 4-8). A mastermix containing 2500µl Taqman Universal Master Mix II No UNG (Applied Biosystems) and 1900µl probe mix was made up immediately prior to setting up duplicate plates for each gene. Forty four microlitres of this mastermix was pipetted into each well of the primary 96 well plate. Six microlitres of target DNA was pipetted into each well and mixed thoroughly before 25µl of the total mix was removed and placed in a duplicate plate. Duplicate plates were briefly centrifuged prior to each plate being run immediately after preparation and simultaneously

on separate, calibrated AB7500 qPCR machines (Applied Biosystems). The primer and probe sequences for each gene listed are the same as detailed in Table 1 above. Cycling conditions for each gene are shown in Tables 4-8 below.

| Gene | Reagents | Reagent final concentration (nM) | | Probe mix volume (µl) | | Cycling conditions | | |
|---|---|---|---|---|---|---|---|---|
| | | | | FAM dye | NED dye | Temp (°C) | Time | Cycles |
| | | *ADAMTS9* | *ACTB* | *ADAMTS9* | *ACTB* | | | |
| *ADAMTS9* | **F Primer** | 700 | 125 | 126 | 22.5 | 95 | 10 min | |
| | **R Primer** | 700 | 125 | 126 | 22.5 | 95 | 15 sec | 50 |
| | **Probe** | 150 | 125 | 27 | 22.5 | 60 | 90 sec | |
| | **$ddH_2O$ (µl)** | 6493.5 | | | | 72 | 15 sec | |
| | **Final vol (µl)** | **6840** | | | | | | |

**Table 4**

| Gene | Reagents | Reagent final concentration (nM) | | Probe mix volume (µl) | | Cycling conditions | | |
|---|---|---|---|---|---|---|---|---|
| | | | | FAM dye | NED dye | Temp (°C) | Time | Cycles |
| | | *CYGB* | *ACTB* | *CYGB* | *ACTB* | | | |
| *CYGB* | **F Primer** | 500 | 125 | 90 | 22.5 | 95 | 10 min | |
| | **R Primer** | 500 | 125 | 90 | 22.5 | 95 | 15 sec | 50 |
| | **Probe** | 150 | 125 | 27 | 22.5 | 58 | 90 sec | |
| | **$ddH_2O$ (µl)** | 6565.5 | | | | 72 | 15 sec | |
| | **Final vol (µl)** | **6840** | | | | | | |

**Table 5**

| Gene | Reagents | Reagent final concentration (nM) | | Probe mix volume (µl) | | Cycling conditions | | |
|---|---|---|---|---|---|---|---|---|
| | | | | FAM dye | NED dye | Temp (°C) | Time | Cycles |
| | | *CCNA1* | *ACTB* | *CCNA1* | *ACTB* | | | |
| *CCNA1* | F Primer | 700 | 125 | 126 | 22.5 | 95 | 10 min | |
| | R Primer | 700 | 125 | 126 | 22.5 | 95 | 15 sec | 50 |
| | Probe | 150 | 125 | 27 | 22.5 | 58 | 90 sec | |
| | $ddH_2O$ (µl) | 6493.5 | | | | 72 | 15 sec | |
| | Final vol (µl) | 6840 | | | | | | |

Table 6

| Gene | Reagents | Reagent final concentration (nM) | | Probe mix volume (µl) | | Cycling conditions | | |
|---|---|---|---|---|---|---|---|---|
| | | | | FAM dye | NED dye | Temp (°C) | Time | Cycles |
| | | *P16* | *ACTB* | *P16* | *ACTB* | | | |
| *P16* | F Primer | 700 | 125 | 126 | 22.5 | 95 | 10 min | |
| | R Primer | 700 | 125 | 126 | 22.5 | 95 | 15 sec | 50 |
| | Probe | 150 | 125 | 27 | 22.5 | 60 | 90 sec | |
| | $ddH_2O$ (µl) | 6493.5 | | | | 72 | 15 sec | |
| | Final vol (µl) | 6840 | | | | | | |

Table 7

| Gene | Reagents | Reagent final concentration (nM) | | Probe mix volume (µl) | | Cycling conditions | | |
|---|---|---|---|---|---|---|---|---|
| | | | | FAM dye | NED dye | Temp (°C) | Time | Cycles |
| | | *TMEFF2* | *ACTB* | *TMEFF2* | *ACTB* | | | |
| *TMEFF2* | F Primer | 900 | 125 | 162 | 22.5 | 95 | 10 min | |
| | R Primer | 900 | 125 | 162 | 22.5 | 95 | 15 sec | 50 |
| | Probe | 150 | 125 | 27 | 22.5 | 60 | 90 sec | |
| | $ddH_2O$ (µl) | 6421.5 | | | | 72 | 15 sec | |
| | Final vol (µl) | 6840 | | | | | | |

Table 8

**Tables 4-8: Primer/Probe concentrations and cycling conditions for duplex qMSP reactions of the five target genes.**

### 3.7 Statistical analysis

Statistical analysis was performed using SPSS (Version 21.0. Armonk, NY: IBM Corp.). Statistical comparisons between the methylation markers and clinicopathological data were made using the Chi-square and or Fisher's exact test (2-sided). Disease free survival curves were prepared by the Kaplan-Meier method. Statistical difference between curves was established using the Mantel-Cox log rank test. $P<0.05$ indicated statistical significance.

# 4. Results

## 4.1 Target gene panel selection

The five target genes in this panel were selected using a candidate gene approach. *CCNA1*, *CYGB, P16* and *TMEFF2* were genes of interest in our research group having shown tumour specific methylation in quantitative methylation studies on tumour/normal tissue from HNSCC patients (Shaw, Liloglou et al. 2006, Shaw, Hall et al. 2007, Shaw, Omar et al. 2009) and in lung cancer (Xinarianos, McRonald et al. 2006). A longitudinal study of OED lesions performed by our research group found *P16* methylation as a potential marker for early malignancy; 57% of patients that underwent malignant transformation had methylated *P16* in the lesions (Hall, Shaw et al. 2008). Preliminary unpublished pyrosequencing data from our research group revealed *TMEFF2* as an informative marker as it was significantly discriminatory between paired normal and tumour tissue (Wilcoxon signed ranks test p=0.002). Eighteen of twenty five (72%) tumour tissue samples contained methylated *TMEFF2* above the reference methylation index in a series of primary HNSCC samples (n=30) (Personal communication from Mr Mark Reid, student of the Department of Translational Medicine, University of Liverpool 2008).

### 4.1.1 Selecting a novel target gene: *ADAMTS9*

The selection of *ADAMTS9* (ADAM Metallopeptidase with Thrombospondin type 1 motif 9) as part of the panel of target genes in this saliva study was borne of a collaboration with Dr. Ratna Veeramachaneni, Department of Genetic Medicine, University of Manchester, Manchester, UK. Using SNP array analysis of dysplasia and primary cancer cell lines as part of her PhD thesis, Dr. Veeramachaneni selected seven genes that appeared to be differentially deleted as an early event in HNSCC (Veeramachaneni 2010), five of which were subsequently shown by pyrosequencing to be methylated in the gene promoter (Table 9).

| Gene | Sample type | Samples with MtI[a] greater than five % methylation | Number of cases when methylation of Tumour > Normal[b] |
|---|---|---|---|
| *ADAMTS9* | Tumour | 52.2 % | **7/23 (30 %)** |
| | Normal | 42.9 % | |
| *FAT1* | Tumour | 39.1 % | 7/23 (30 %) |
| | Normal | 21.7 % | |
| *PTPRD* | Tumour | 36.4 % | 2/23 (9 %) |
| | Normal | 26.4 % | |
| *CSMD1* | Tumour | 92 % | 2/24 (8 %) |
| | Normal | 79 % | |
| *CDH13* | Tumour | 50 % | 4/17 (24%) |
| | Normal | 25 % | |

**Table 9 Pyrosequencing data showing rates of gene promoter methylation in a Liverpool cohort of 24 paired fresh tumour OSCC tissue and adjacent normal tissue (from resection margins).**

**[a]The methylation index (MtI) for each sample was calculated as the mean percentage methylation over all CpG dinucleotides interrogated.**

**[b] Reference MtI is the methylation value below which 95 % of the 'normal' tissue sample data falls (Shaw, Liloglou et al. 2006).**

Column three of Table 9 shows the percentage of technical positive tumour and normal adjacent tissue samples for each gene. Any sample was considered a technical positive if the MtI was greater than 5%, as anything below this threshold could be considered 'biological noise'. (Shaw, Liloglou et al. 2006). It is accepted that the 'normal' tissue from resection margins is not a true normal; and that the level of methylation in 'true normal' tissue is also unknown; therefore the biological relevance of the methylation results in Table 9 (column four) was determined by establishing a cut-off methylation index (Reference Methylation Index) that is higher than 95% of the normal tissue values. This cut-off was required to determine if the target genes were sufficiently discriminatory between frank tumour and adjacent 'normal' tissue.

On analysis of the data, methylation of Cub and Sushi multiple domains 1 (*CSMD1),* Cadherin 13 *(CDH13)* and protein tyrosine phosphotase receptor type D *(PTPRD)* genes were not discriminatory enough to be considered as potential biomarkers. The results presented for Cadherin 13 *(CDH13)* were based on only 17 tumour/normal samples whereas the other four genes were based on 24 matched pairs. Moreover the mean methylation of *CDH13* in the tumour samples was low with only one sample with mean methylation > 10%.

Of the remaining three genes, *ADAMST9* was selected to develop into a qMSP assay with which to analyse the saliva samples, because it showed the highest tumour sensitivity (52.2%). Additionally 30% of tumours had a mean methylation greater than the reference MtI and therefore *ADAMTS9* was thought to be the most biologically relevant gene. Although 43% (9/23) of the normal samples had a recorded mean methylation >5% it must be noted that 4/9 of these normal samples were very close to the 5% cut-off (mean methylation 5.05-5.59).

Our hypothesis that *ADAMTS9* is the most biologically relevant of the genes as a marker of OSCC, was supported by Dr. Veeramachaneni's earlier work which showed that *ADAMTS9* promoter methylation was absent in mortal cell lines but present in 14% of the immortal premalignant oral lesion cell lines and 41% of the immortal OSCC cell lines (Veeramachaneni 2010). In addition, our research group has previously shown that methylation of FAT atypical cadherin 1 (*FAT1)* in 44 paired normal and tumour HNSCC, detected on an Illumina 'Golden Gate' methylation array, was observed to be low, with mean methylation in the tumour tissue of only 10% and 8% in paired tumour normal tissue samples (Jithesh, Risk et al. 2013).

Further support for the role of *ADAMTS9* in aerodigestive tumorigenesis has been described in Lo et al's (Lo, Leung et al. 2007) work on oesophageal squamous cell carcinoma, which showed loss or down regulation of *ADAMTS9* in oesophageal carcinoma cell lines. In that study, promoter methylation of *ADAMTS9* was detected in the cell lines which lacked gene expression; reinstatement of gene expression followed demethylation drug treatment. Additionally *ADAMTS9* maps to 3p14.2 (Clark, Kelner et al. 2000). The chromosomal region containing ADAMTS9 (3p14.2) commonly shows allelic imbalance in OSCC and may have a role in malignant transformation of dysplastic oral lesions (Rosin, Cheng et al. 2000, Rosin, Lam et al. 2002, Tsui, Rosin et al. 2008) and lung cancer (Field, Kiaris et al. 1995, Field, Neville et al. 1996, Nunn, Scholes et al. 1999).

## 4.2 Establishing thresholds to define a positive result

### 4.2.1 Defining a positive result as greater than 0.5 % methylation for all five target genes for the singleplex technique

The analytical sensitivity for the singleplex technique was defined by the lack of overlap between the positive standard picked for each assay and the biological control (lymphocyte DNA) as determined in the optimisation and validation process using standard curve analyses of serial dilutions of target DNA. A 'positive' result was defined as a $\delta$CT value equivalent to, or exceeding the analytical sensitivity for each gene. In the case of singleplex qPCR this was initially 0.1% however, it was noted in several of the clinical sample runs that the amplification signal for 0.1% was inconsistent, absent or $\delta\delta$CT<2 from the lymphocyte biological control. Therefore a level of 0.5% methylation was used as the cut-off threshold for a positive result for all 5 genes in singleplex qMSP.

### 4.2.2 Defining a positive result as greater than 0.25% methylation for *CYGB, P16* and *TMEFF2* and 0.5% methylation for *ADAMTS9* and *CCNA1* for the duplex technique

The analytical sensitivity was defined by the lack of overlap between the positive standard picked for each assay and the biological (lymphocyte DNA) and technical (WGA) controls. This value was different for each gene and is defined below. In order to determine if a reaction run was successful, a non template control, a technical control (WGA) and a biological control (blood lymphocyte DNA) were included, alongside the DNA standards of known methylation (0.25-1%). The biological control is a guide as to the amount of methylation that is present in white blood cells of normal, disease free individuals. The cut-off value of 0.25% was established for three genes in the panel (*P16, TMEFF2* and *CYGB)* and 0.5% for *ADAMTS9* and *CCNA1.* The biological control in the *ADAMTS9* and *CCNA1* assays had an amplification curve within 2 $\delta$CT of the 0.25% standard (in 5/35 and 2/35 runs respectively) therefore the 0.5% standard was used as a cut-off for these two genes.

## 4.3 DNA yield obtained from an OSCC and PORT cohort using different saliva collection methods

### 4.3.1 Patient cohorts

DNA yield was quantified using a human *RNAse P* quantitative PCR (qPCR) assay in a subset of patients from the OSCC cohort and the PORT cohort. This was performed to evaluate the

best method of saliva collection to maximise DNA yield in OSCC patients. Other published studies of DNA yield and saliva collection methods have focused on patients without oral malignancy (Table 10). However, certain sequelae of oral malignancy such as disrupted oral competency and xerostomia secondary to PORT can affect the patient's ability to produce a saliva sample. A total of 45 patients, representative of a typical Liverpool cohort (Rogers, Brown et al. 2009), participated in this study and provided 90 saliva samples collected by three methods. Nine of the forty five patients (all 9 were pre-operative samples) were also included in the qMSP biomarker detection element of the larger study.

Patients were divided into two groups; Group one (total n=21) contained pre-operative (n=10) and post-operative (n=11) OSCC patients and provided paired Oragene™ and saline mouthwash saliva samples. Four of the 11 post-operative patients in this group had received PORT. Group two contained 24 post-operative OSCC patients who had all completed PORT prior to the saliva collection. This group provided paired Oragene™ sponge and saline mouthwash saliva samples.

| Author | Participants | Sample | Average DNA Yield (μg) | Quantification method |
|---|---|---|---|---|
| (Philibert, Zadorozhnyaya et al. 2008) | Behavioural genetic study | Oragene whole saliva | 92 | Spectrophotometry |
| (Hansen, Simonsen et al. 2007) | Normal Cohort | Oragene whole saliva<br>Buccal swab<br>Foam Tipped Applicator | 10.8<br>64.4<br>0.36 | Spectrophotometry |
| (Rogers, Cole et al. 2007) | Normal Cohort | Oragene whole saliva<br>Mouthwash (10ml)<br>Cytobrush<br>Buccal swab | 182<br>36.5<br>13<br>11 | Spectrophotometry |
| (Quinque, Kittler et al. 2006) | Normal Cohort | Whole saliva<br>Buccal swab | 11.4<br>8.29 | PCR |
| (Rylander-Rudqvist, Håkansson et al. 2006) | Normal Cohort | Oragene whole saliva (2ml) | 135.9<br>29.4<br>19.2 | Spectrophotometry<br>Picogreen fluorescence<br>RTqPCR |
| (Ng, Koh et al. 2004) | Normal Cohort | Whole saliva | 2.9-6.8 | Spectrophotometry |
| (Cozier, Palmer et al. 2004) | Normal Cohort | Mouthwash<br>Buccal swab | 10.1<br>3.69 | Spectrophotometry |
| (García-Closas, Egan et al. 2001) | Breast Cancer Cohort | Mouthwash<br><br>Cytobrush | 57.3<br>27.5<br>13.6<br>1 | Spectrophotometry<br>PCR<br>Spectrophotometry<br>PCR |
| (Bauer, Rezaishiraz et al. 2004) | Smoking cessation trial | Mouthwash | 44.93 | Spectrophotometry |
| (Satia-Abouta, King et al. 2002) | Normal Cohort | Mouthwash<br>Cytobrush | 15.8<br>12 | Spectrophotometry |
| (Feigelson, Rodriguez et al. 2001) | Normal cohort | Mouthwash | 34 | Spectrophotometry |
| (Heath, Morken et al. 2001) | Normal Cohort | Mouthwash | 18.6 | PCR |
| (Le Marchand, Lum-Jones et al. 2001) | Normal cohort | Mouthwash | 35.1 | Spectrophotometry |
| (Lum, Le Marchand 1998) | Normal Cohort | Mouthwash | 49.7 | Spectrophotometry |
| (Freeman, Powell et al. 1997) | Behavioural genetic study | Mouthwash | 38 | Spectrophotometry |

**Table 10: Summary of published data of DNA yield from different saliva collection methods**

### 4.3.2 DNA yield from whole saliva Oragene™ collection method outperformed mouthwash collection in the OSCC cohort but not in the post-operative radiotherapy cohort

All 90 samples contained human DNA, as confirmed by *RNAse P* qPCR assay. In group one, the DNA concentration from the Oragene™ whole saliva samples (range 10-929ng/μl; median 306/μl; SD 291ng/μl) was significantly greater (Wilcoxon signed ranks test p=0.001) than that of the matched mouthwash samples (range 7-657ng/μl; median 62ng/μl; SD 169ng/μl) (Figure 6a).

In group two, a lower median DNA concentration was observed than for group one and there was no statistical significance (Wilcoxon signed ranks test p=0.5) in DNA concentrations obtained using Oragene™ sponge (range 0.3-306ng/μl; median 22ng/μl; SD 86 ng/μl) and matched mouthwash (range 1.6-267ng/μl; median 20ng/μl; SD 77 ng/μl) collection methods (Figure 6b). Both cohorts demonstrated a wide range of DNA concentrations which was irrespective of the method of collection.

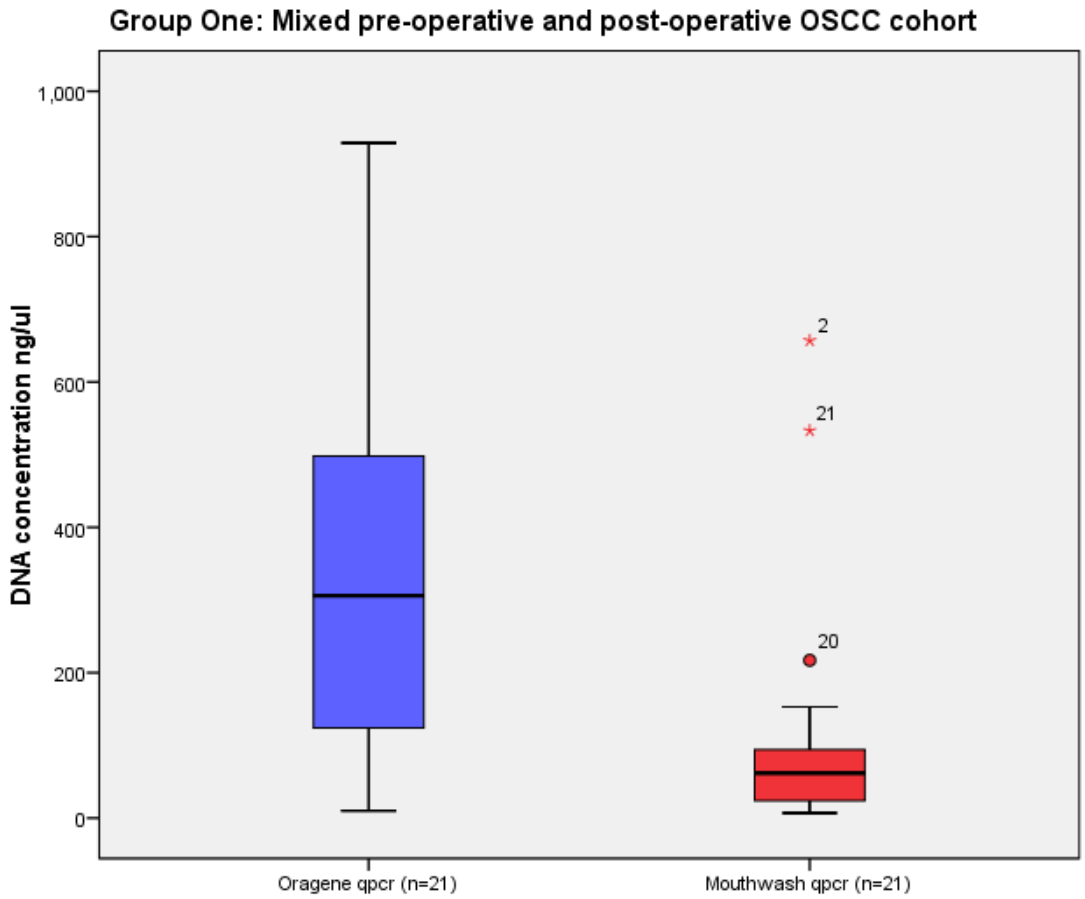



**Figure 6a: DNA concentration range (ng/μl) for the mixed pre-operative and post-operative OSCC cohort comparing Oragene™ to mouthwash saliva collection methods using qPCR ( * and ° signify samples that are outliers).**

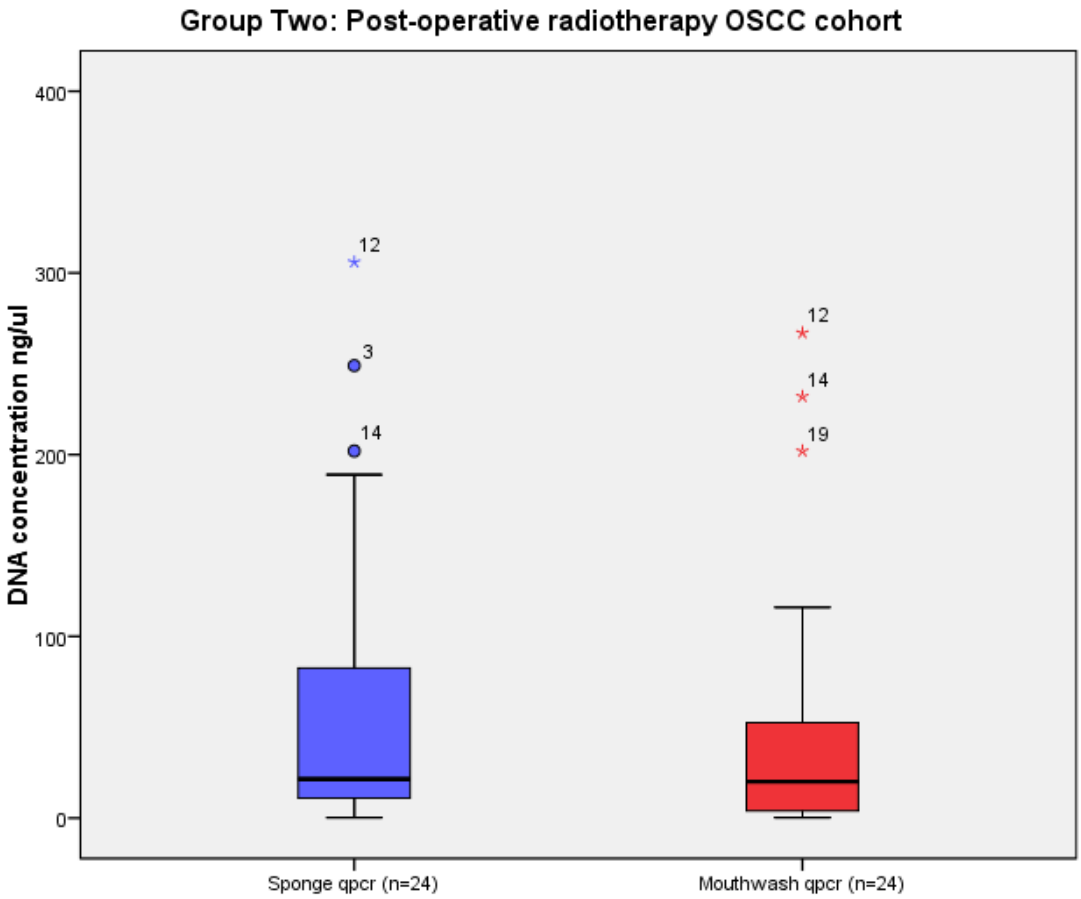



**Figure 6b : DNA concentration range (ng/µl) in post-radiotherapy cohort comparing Oragene™ sponge kits to mouthwash collection methods using qPCR ( * and ° signify samples that are outliers).**

The total DNA yield for each sample was calculated using the *RNAse P* qPCR concentration data and the total sample volume. This data mirrors the DNA concentration data, with the total amount of DNA in whole saliva collected using Oragene™ being significantly greater than mouthwash in group 1 cohort (p=0.0001 Wilcoxon signed ranks test) (Table 11). There was no significant difference between DNA yield from the Oragene™ sponge and mouthwash method in the PORT cohort (Table 12).

| DNA Yield μg | Oragene™ whole saliva | Mouthwash |
|---|---|---|
| Range | 4-372 | 2-184 |
| Median | 122 | 17 |
| Mean | 140 | 32 |
| SD | 116 | 47 |

**Table 11 Total DNA yield (μg) in paired Oragene™ whole saliva and mouthwash samples from the group 1 cohort (n=21) determined by *RNAse P* qPCR. p=0.0001 (Wilcoxon signed ranks test)**

| DNA Yield μg | Oragene™ Sponge | Mouthwash |
|---|---|---|
| Range | 0.1-61 | 0.1-75 |
| Median | 4 | 5.5 |
| Mean | 13 | 14 |
| SD | 17 | 21 |

**Table 12 Total DNA yield (μg) in paired Oragene™ sponge and mouthwash samples from the post-radiotherapy group 2 cohort (n=24) determined by *RNAse P* qPCR. p=0.9 (Wilcoxon signed ranks test)**

## 4.4 Concordance of methylation results between different sources of oral cellular material

### 4.4.1 Concordance of methylation between saliva and tumour in this series was limited

The concordance between methylation in head and neck tumours and paired saliva has previously been reported for other gene panels (Righini, De Fraipont et al. 2007, Rosas, Koch et al. 2001, Carvalho, Henrique et al. 2011). Righini et al (Righini, De Fraipont et al. 2007) detected *P16* methylation in tumour 20/60 (33%) and paired saliva 16/60 (27%) with a high concordance rate as represented by a K-coefficient of 0.8. Rosas et al (Rosas, Koch et al. 2001) report; of 14 tumours positive for *P16,* 11/14 paired saliva samples were also positive (79% concordance). Both authors used MSP, a qualitative method vulnerable to false positive results (Shaw, Akufo-Tetteh et al. 2006). Using qMSP, Viet (Viet, Jordan et al. 2007) reported an 87.5 % positive agreement in a small cohort (n=14) of matched saliva and oral tumours. Carvalho et al (Carvalho, Henrique et al. 2011) report methylation of at least 1 gene from a panel of 7 genes including *P16* and *CCNA1* in 59 tumours. At least one gene was positive in 33/59 (56%) paired saliva samples.

The frequency of methylation detected in pre-operative saliva and paired tumour samples for all target genes in our study is shown in Tables 13-17. The overall concordance levels, when all samples are included, for each gene in our study are comparable to genes used in

other studies (71-88%) however our data is dominated by negative agreement (i.e absence of methylation in both saliva and tumour)(Tables 13-17) and the relevance of these results is questionable. A positive saliva sample with a negative paired tumour sample was an infrequent event (six occurrences) and the significance of this is currently unknown but could be explained by the fact that saliva contains cells from the entire oral cavity and may contain cells from field change beyond the index tumour. The leading hypothesis is that methylation detected in saliva is from cellular material shed from the tumour mass. Working with this assumption a more useful view of our data would be that of tumour positive concordance (i.e the number of positive saliva samples when the paired tumour was methylated). *CYGB* and *P16* had the best tumour concordance at 70 % (7/10) and 64% (7/11) respectively which is in keeping with the published studies discussed above. *CCNA1* had a tumour concordance of 45% (5/11) and *TMEFF2* 33% (4/12) with *ADAMTS9* performing the worst with only 11% (1/9) concordance.

The assays appear to offer good specificity but variable sensitivity. The detection of methylation in saliva for *P16* in our data set 9/34 (26%) is comparable to that of Righini (Righini, De Fraipont et al. 2007) 16/60 (27%) and Viet (Viet, Jordan et al. 2007) 5/14 (35%) but much lower than Rosas (Rosas, Koch et al. 2001) 11/14 (79%) who reports on a smaller cohort using MSP. Concordance between positive results in tumour and saliva in our study was expected to be better and raises questions as to the suitability of saliva as a direct surrogate using this panel of genes, an issue which will be explored in the discussion.

| ***ADAMTS9*** | | **Tumour n=34** | |
|---|---|---|---|
| | | **Positive** | **Negative** |
| **Saliva n=34** | **Positive** | 1 | 0 |
| | **Negative** | 8 | 25 |

**Table 13**

| ***CCNA1*** | | **Tumour n=34** | |
|---|---|---|---|
| | | **Positive** | **Negative** |
| **Saliva n=34** | **Positive** | 5 | 1 |
| | **Negative** | 6 | 22 |

**Table 14**

| ***CYGB*** | | **Tumour n=34** | |
|---|---|---|---|
| | | **Positive** | **Negative** |
| **Saliva n=34** | **Positive** | 7 | 1 |
| | **Negative** | 3 | 23 |

**Table 15**

| *P16* | | Tumour n=34 | |
|---|---|---|---|
| | | Positive | Negative |
| Saliva n=34 | Positive | 7 | 2 |
| | Negative | 4 | 21 |

**Table 16**

| *TMEFF2* | | Tumour n=34 | |
|---|---|---|---|
| | | Positive | Negative |
| Saliva n=34 | Positive | 4 | 2 |
| | Negative | 8 | 20 |

**Table 17**

**Tables 13-17: Frequency of methylation in paired tumour and pre-operative saliva samples for the five target genes.**

### 4.4.2 Strong correlation between qMSP data was observed in matched saliva and buccal samples from a normal control cohort

Normal oral mucosa from disease free patients would be the ideal control to determine the level, if any, of methylation in normal oral tissue. Histologically normal tissue from non-diseased patients is difficult to obtain for practical and ethical reasons. As a compromise scrapes from the buccal mucosa of individuals in the normal cohort were obtained alongside a saliva sample to compare any difference between the methylation status of the scrapes versus the saliva. Oral scrapes have previously been shown to be an adequate surrogate for oral biopsy tissue with a reported 80% concordance rate for the detection of methylation in scrape samples taken directly from the lesion in OED patients (Hall, Shaw et al. 2008). It is assumed that the buccal scrapes have greater cellular density than the saliva samples. The purpose of the comparison between normal scrapes and saliva here was to see if the two methods yielded comparable information about methylation status or if the scrapes were a better surrogate tissue.

Ninety individuals from the normal cohort (n=97) provided a buccal mucosal scrape. Singleplex qMSP methodology was used and the data (Table 18) shows strong concordance between saliva and buccal scrapes for methylation status of each of the five target genes. However, as expected in a normal cohort both collection methods yielded predominantly negative results for methylation at all five target genes. Only five individuals with paired samples demonstrated a positive result (Table 18) and all five resulted in discordance between the saliva and scrape sample. Two individuals had a positive saliva and negative

scrape. Three individuals had a positive scrape and negative saliva. For 70 scrape samples and matched saliva only one run was available for analysis so the results are not in duplicate. It is also noteworthy that only 6/732 (1%) of saliva qMSP reactions failed but that 47/660 (7%) scrape reactions did so due to inadequate amount of input DNA (failure to generate amplification curve in beta actin) or poor quality of input DNA (amplification curve consistent with degraded DNA). Although there was a 7% failure rate in the scrapes; this data was retained for analysis as the sample size remains large enough to be informative. However, due to the factors stated above and the exhaustion of DNA from the scrape samples it was decided that any continued analysis would be performed on normal saliva alone. Mulki et al (Mulki, Shetty et al. 2013) report a significantly higher cell yield from oral rinses than oral scrapes in a normal cohort. Rogers et al (Rogers, Cole et al. 2007) also report higher DNA yield and quality from whole saliva samples compared to buccal cell collection methods (Table 10). The reason for the higher failure rate in scrapes in our series is likely to have been due to operator technique during the sample collection.

| Gene | Number of pairs available for analysis* | Concordant | Positive scrape | Positive saliva | Discordant |
|---|---|---|---|---|---|
| ADAMTS9 | 79 | 78 (99 %) | 0 | 1 | 1 |
| CCNA1 | 79 | 79 (100 %) | 0 | 0 | 0 |
| CYGB | 79 | 79 (100 %) | 0 | 0 | 0 |
| P16 | 78 | 77 (99 %) | 1 | 0 | 1 |
| TMEFF2 | 87 | 84 (98 %) | 2 | 1 | 3 |

***Missing data is due to technical failure of the reaction.**

**Table 18. Concordance of singleplex qMSP data between the presence of methylation in each of the 5 target genes between scrapes and paired saliva from the normal cohort (n=90). Concordant = either a +ve saliva +ve scrape or –ve saliva –ve scrape. Discordant =+ve saliva –ve scrape or–ve saliva +ve scrape**

### 4.4.3 Methylation results from the singleplex and duplex qMSP techniques were comparable

An interim analysis of the singleplex methodology results identified inconsistencies with the quantitative reliability of the control DNA at the lowest levels (0.1%). There was some discordance between 2% of the total number of samples on duplicate runs. These

inconsistencies might be expected due to relatively low levels of DNA in surrogate samples and an analytical sensitivity of 0.1% is at the stochastic part of these assays. In the high throughput clinical runs, small pipetting errors can be more profound which can result in an increase in variability between reactions. The standards that are close to the stochastic area of the PCR can become unstable. Additionally, the quantitative measurement of methylation in the standards had not yet been introduced at this stage of the protocol.

The singleplex methodology involved composing all of the plates simultaneously to standardise the component volumes of each, however this resulted in some plates being stored in the refrigerator for longer than others with a potential difference of several hours. Unpublished data from Dr. Lakis Liloglou (personal communication) revealed there was condition "slippage" of plates left refrigerated compared to those reactions which were run immediately after preparation which may have affected some of the discordant cases that fell very close to the analytical sensitivity of the assay.

A 'positive' sample was defined as that with at least two results above the assigned cut-off. If there was a discrepant result, between the duplicate runs for an individual sample, the protocol was to repeat the reaction if enough DNA remained in the sample. During the course of this study, the qPCR technology evolved to allow for multiplexing target genes in the same reaction (Fackler, Malone et al. 2006) which allows more efficient use of the small amounts of DNA in saliva. A quality control step, to quantitate the amount of methylation in the controls, was also incorporated into the methodology. Therefore, new assays were designed wherein beta actin, the internal control gene, was incorporated as a duplex reaction in the same tube as the target gene. In view of these factors, repeat runs were undertaken using duplex, rather than singleplex qMSP technology as this would satisfy the need for repeating samples and would maximise the use of minimal DNA. It also allowed eradication of the time difference factor as plates inclusive of the internal control gene could be composed and run immediately.

The duplicate runs of the duplex reactions were largely concordant, with only 13 of 1445 (1%) reactions across the five genes having a discordant result in relation to the chosen cut-off. The concordance between the singleplex and duplex data was measured using the McNemar $\chi 2$ test which tests for equality of correlated proportions and informs if the discord proportion between the two techniques is significant. All 5 target genes showed a low rate of discordance between the singleplex and duplex data which was not statistically significant in four out of the five genes ( *ADAMTS9*: discordance 2% p=0.125; *CCNA1*:

discordance 1.4% p = 0.250; *CYGB*: discordance 0.5% p=1; *P16*: discordance 3% p=0.65; *TMEFF2*: discordance 5% p=0.04 (Appendix 4 Tables A4.1-A4.5). The saliva DNA was limited. It was therefore decided that the singleplex data, where necessary, could be used in effect as a triplicate run to draw conclusions for any discrepant duplicate duplex results. This only applied to 13 samples. The duplex data is a more robust data set because of the additional quality control steps of the technical control WGA, quantitative measurement of methylation in the control DNA and every sample has at least a duplicate run. The duplex data was therefore used for further statistical analysis against demographic and clinicopathological data.

## 4.5 Observed patterns of methylation in saliva samples from the normal, dysplasia and OSCC cohorts

### 4.5.1 Patient populations

Two hundred and sixty four individuals were enrolled in this study. Prior to statistical analysis, six patients with oropharynx SCC were excluded as five of them were HPV positive and one had unknown HPV status. Three dysplasia patients were excluded as I was unable to confirm a histological diagnosis of OED from available pathology reports; one with unavailable histology, and two did not have dysplastic histology (chronic hyperplastic candidosis and basal keratinitic atypia ). Thirty five post-operative patients from the saliva collection methods study were not included in this arm of the study as they did not have matching pre-operative saliva samples and one pre-operative patient was excluded due to exhaustion of DNA in the saliva collection methods analysis. Two hundred and nineteen individuals were included in the methylation analysis of the target panel (Table 19). Statistical analysis of the demographic data using the Pearson chi-square test revealed a statistical difference between the normal and OSCC group and the normal and dysplasia group in all the demographic and risk factor data except gender (Appendix 5 Tables A5.1-A5.12). The OSCC cohort consumed more alcohol than the dysplasia cohort. (p=0.005 Appendix 5 Table A5.11).

#### 4.5.1.1 Clinical and pathological characteristics of the case and control cohorts

The pathology characteristics of the primary tumours in the OSCC cohort were similar to a typical Liverpool cohort (Rogers, Brown et al. 2009). Primary tumour site was available for 56 of 62 patients. The primary sites included in the cancer cohort were floor of mouth n=17 (30%) , tongue n=20 (35%), HPV negative oropharynx n=4 (7%), buccal mucosa n=3 (5%), other n=12 (21%) (mandible, maxilla, alveolus). Pathology characteristics were available on 55 tumours. There is a predominance of stage IV disease in this case series. Histopathological staging of the cases: stage I: 4 (7%), stage II: 12 (22%), stage III: 8 (14%) stage IV: 32 (57%). Twenty eight of fifty five (51%) patients were pN positive and of these 17/28 (61%) had extracapsular spread (ECS).

In the dysplasia cohort the site of the primary lesion was most commonly on the floor of mouth 25/60 (42%) with tongue as the next most common site 18/60 (30%), buccal mucosa 12/60 (20%) and other 5/60 (8%) which is a similar distribution to the series published by Ho et al (Ho, Risk et al. 2012).

| Demographic | | OSCC ( n=62) | Dysplasia (n=60) | Normal (n=97) |
|---|---|---|---|---|
| Age | 0-40 | 1 (2%) | 4 (7%) | 44 (45%) |
| | 41-60 | 28 (45%) | 31 (52%) | 29 (30%) |
| | 61+ | 28 (45%) | 25 (42%) | 24 (25%) |
| | No data* | 5 | NA | NA |
| Gender | Male | 39 (63%) | 34 (57%) | 56 (58%) |
| | Female | 22 (35%) | 26 (43%) | 41 (42%) |
| | No data* | 1 | NA | NA |
| Smoking status | None | 8 (13%) | 14 (23%) | 38 (39%) |
| | <= 20 pack year | 11 (18%) | 17 (28%) | 42 (43%) |
| | >= 20 pack year | 28 (45%) | 29 (48%) | 17 (18%) |
| | No data* | 15 | NA | NA |
| Alcohol | None | 5 (8%) | 19 (32%) | 36 (37%) |
| | <=28u week | 25 (40%) | 32 (53%) | 59 (61%) |
| | >=28u week | 16 (26%) | 8 (13%) | 2 (2%) |
| | No data* | 16 | 1 | NA |

**Table 19: Demographic and risk factor data for individuals within each of the three separate cohorts. * Clinical data was collected retrospectively from casenotes and was not available for every patient**

### 4.5.1.2 Methylation detected in saliva from the normal cohort was a rare event

A primary aim of this study was to establish a level of methylation of this gene panel in a normal population to determine a cut-off value which could be used to discriminate disease from normal saliva samples. In this normal cohort (n=97) four of the five genes did not have promoter methylation at a level higher than the technical thresholds described above. Thus, any result above the threshold is considered positive for *ADAMTS9, TMEFF2, CCNA1* and *CYGB*.

Three percent of this cohort (3/97) had 'positive' promoter methylation in *P16*. Two of the samples had methylation greater than the 1% methylated standard and one sample greater than the 0.5% methylated standard as determined by the $\delta$CT. All three patients were male and smoked in excess of 20 pack years. They were aged 48, 50 and 74 and had varied

alcohol intake. Despite three samples from the normal cohort having detectable methylation above the technical threshold of 0.25%: it was decided to accept this threshold for *P16* for the purpose of this exploratory study.

### 4.5.2 Clinicopathological parameters

#### 4.5.2.1 Diagnosis

Saliva samples were initially placed into seven groups based on diagnostic classification: normal, mild dysplasia, moderate dysplasia, severe dysplasia, proliferative verrucous leukoplakia (PVL), microinvasive SCC (MISCC) and SCC. *ADAMTS9* was infrequently methylated being detected in only one of 218 saliva samples. The one positive result was a pre-operative saliva sample from an OSCC patient who also had strong pattern of methylation in *CYGB* and *P16* in saliva and tumour samples. Statistical evaluation was not performed for *ADAMTS9* due to the likelihood of invalidity from such small numbers. This gene was henceforth discarded from the panel for further analyses of clinicopathological data. The presence of promoter methylation in the remaining four genes was able to discriminate between malignant disease and other diagnoses: Fisher's exact *CCNA1*: p=0.003; *CYGB*: p=0.001; *P16*: p<0.001; *TMEFF2*: p<0.001(Appendix 6 Tables: A6.1-A6.4).

Promoter methylation, above the declared threshold in the pre-operative saliva samples from known OSCC patients, was most frequently detected in *P16*: 27% (17/62). Methylation rates for the other genes in the panel are shown below (Table 20).

| Sample | Gene | | | |
|---|---|---|---|---|
| | CCNA1 | CYGB | P16 | TMEFF2 |
| **Pre-operative OSCC Saliva (n=62)** | 9/62 (15%) | 10/62 (16%) | 17/62 (27%) | 10/62 (16%) |
| **Pre-operative OSCC tumour tissue (n=34)** | 11/34 (32%) | 10/34 (29%) | 11/34 (32%) | 12/34 (35%) |

**Table 20: Methylation rates for each of the four genes in the panel in pre-operative OSCC saliva and tumour samples.**

**Binning the data**

It was noted that the numbers of patients within each diagnostic classification in this dataset were small in terms of statistical analysis. In order to strengthen the statistical tests of gene promoter methylation versus diagnosis, the groups were amalgamated as follows. The diagnoses were 'binned' into normal (n=97); low risk = mild and moderate dysplasia (n= 38); high risk = severe dysplasia and PVL (n=16) (defined as increased likelihood of malignant transformation (Hansen, Olson et al. 1985, Woolgar, Triantafyllou 2009, Ho, Risk et al. 2012) and malignancy (SCC and MISCC) n=68. It has been shown that a panel of genes is more sensitive than a single marker (Righini, De Fraipont et al. 2007) therefore further analysis was conducted as a gene panel, but with the omission of *ADAMTS9* for the reasons stated above.

A higher prevalence of promoter methylation at any marker was observed in high risk and malignant disease compared with the other groups (Appendix 7 Tables A7.1-A7.5). The presence of methylation of the gene panel in high risk premalignant lesions (7/18) versus low risk lesions (4/36) was statistically significant (Fisher's exact P=0.03). Methylation was present in at least one gene in 42% (36/85) of the combined high risk and malignancy samples and only 5% (7/133) of the normal and low risk samples (Fisher's exact p=0.001). In this series there is an observed trend towards increased number of methylated genes in saliva samples from patients with malignancy (Table 21).

| | | Diagnostic classification | | | | Total |
|---|---|---|---|---|---|---|
| | | Normal | Low risk | High risk | Malignancy | |
| Number of genes methylated | 0 | 94 | 32 | 11 | 38 | 175 |
| | 1 | 3 (3%) | 4 (11%) | 6 (33%) | 15 (22%) | 28 |
| | 2 | 0 | 0 | 1 (6%) | 9 (13%) | 10 |
| | 3 | 0 | 0 | 0 | 5 (7%) | 5 |
| Total | | 97 | 36 | 18 | 67 | 218 |

**Table 21: A contingency table of the number of methylated target genes versus the 'binned' diagnosis classification groups**

Of the three individuals in the normal cohort with positive saliva samples all had *P16* methylation and are commented on above (4.5.1.2). Of the four patients in the low risk group with a positive result, two are discussed in section 4.5.3.2 in the longitudinal dysplasia case studies. The two remaining patients were both female, heavy smokers (>20 packyears) and both had mild dysplasia which extended up to the biopsy margin. They were both lost to follow-up. The first aged 58 with a buccal mucosa lesion failed to attend after 3 months of follow-up. The second aged 51 with a FOM lesion had 20 months of follow up after an initial biopsy (mild dysplasia) and the saliva sample (*P16* positive) taken 02/03/2010. She had a further biopsy 24/11/11 which showed moderate squamous epithelial dysplasia and chronic candidosis, but failed to attend for further follow-up. Ten of fifty five (18%) (5/60 patients had histology of MISCC and were removed from this analysis and 'binned' with OSCC) had transformation of the premalignant lesion. Overall, methylation was not associated with transformation in this series (Fisher's exact p=0.67). Three out of ten patients with transformation had methylation of at least one gene in the panel in a saliva sample. Eight out of forty five patients who did not transform during

follow-up had methylation of at least one gene in the panel. One patient was excluded as there was insufficient DNA in the sample.

#### 4.5.2.2 Confounding variables

One of the primary outcomes of this work was to determine the frequency of methylation in a normal cohort, with the objective of determining if the selected candidate genes could be used to discriminate normal from a disease state in future clinical testing. It is important to investigate whether putative risk factors for methylation are associated with increased methylation rates in disease free individuals. If so, this would reduce the discriminatory value of methylation between those with and without the disease. We have therefore paid particular attention to the variation in methylation in the normal cohort depending on age, smoking history and alcohol consumption as this removes disease as a confounding variable. For completeness we have also included statistical analysis on each of the disease groups.

#### 4.5.2.3 Methylation rates were not associated with advancing age with this target gene panel

Cancer is an age related disease being more common in those of advancing age (CRUK 2014a). Aberrant epigenetic signalling has a role in cellular ageing (Fraga, Agrelo et al. 2007) and DNA methylation is known to be associated with cancer (Baylin, Ohm 2006, Esteller 2008). To answer the question set in our study: Can this panel of biomarkers be used to discriminate normal from malignant disease? It is pertinent to know if age is itself a confounding variable. The age categories were designated arbitrarily to distribute the normal cohort into young, middle aged and old. The young adult range of 16-40 years old approximates with other studies in the literature (Harris, Kimple et al. 2010, Patel, Carpenter et al. 2011). CRUK reports that age specific incidence rates of OSCC rise around age 40-44 years and peak at 60-64 for males (CRUK 2014b, CRUK 2014a). Promoter methylation of saliva samples from normal individuals in our study was found only in *P16*. The complete absence of methylation above the threshold in the remaining three genes, regardless of the age of the individual, supports the assertion that in the normal cohort methylation does not increase with age in these genes.

There did not appear to be a statistically significant relationship between advancing age and the presence of methylation in individual genes or the gene panel in the three disease

groups (Appendix Table A8.1-A8.5). As noted earlier, there is a comparatively large number of individuals in the under 40 age category in the normal cohort when compared with the disease groups (Table 19). Oral cancer is relatively uncommon in individuals under 40 in the UK but does occur (CRUK 2014b). A subanalysis of age versus methylation for each of the diagnosis groups was performed with the under 40 group removed (Appendix Table A8.6 ) and there was no relationship between advancing age and methylation.

#### 4.5.2.4 No correlation was observed between methylation rates and smoking history with this target gene panel

There is ongoing debate about the role of tobacco exposure in *P16* methylation. Kulkarni et al (Kulkarni, Saranth 2004) reported *P16* methylation in 50% (30/60 cases) of normal adjacent mucosa in a cohort of 60 individuals with OSCC all of whom had tobacco exposure. In a similar study Kato et al (Kato et al. 2006) report 27% (6/22 cases) with *P16* methylation in normal adjacent mucosa but state there was no correlation with tobacco consumption. Previous work by our research group does not support smoking as a cause of *P16* methylation. Methylation of *P16* was found to be tumour specific being detected in only 4% of clinically/histologically normal mucosa from 70 patients. Fifty seven percent of these patients were heavy smokers. However, some authors quote the frequency of *P16* methylation in oral tissue of smokers without oral malignancy at approximately 10% (von Zeidler, Miracca et al. 2004, Ruesga, Acha-Sagredo et al. 2007).

In this current data series there was no overall association between smoking and *P16* methylation detected in saliva samples (Chi-square p=0.51) (Figure 7). The smoking history of individuals in this study was arbitrarily stratified into non-smoker, 20 pack years or less and greater than 20 pack years with the aim of assessing if increased exposure to smoking has a relationship with methylation. The distribution of *P16* methylation as related to smoking in pack years in this series is shown in figures 8-11 below. In the normal cohort (Figure 8) *P16* methylation is only noted in the heavy smoking group and therefore was statistically significant (Fisher's exact p=0.005). This significance would be considered with caution as it is based on only 3/97 positive results.

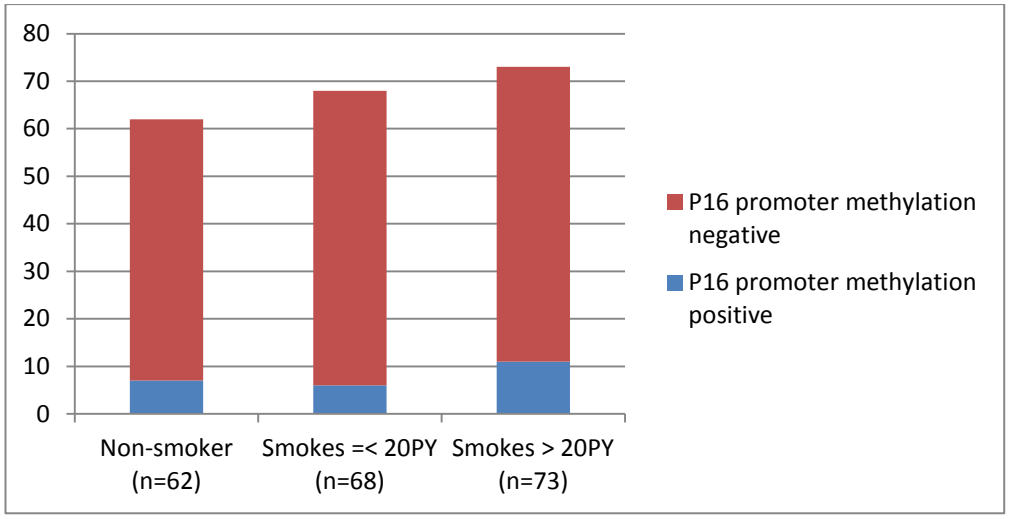



**Figure 7: Frequency (y-axis) of *P16* promoter methylation in saliva samples of smokers (PY =pack years) and non-smokers.**

In the disease groups (Figure 8-11) the general pattern is that methylation is not smoking related. In a longitudinal biomarker study of OED, which showed *P16* methylation as a potential predictive marker for malignant transformation, there was a higher proportion of non-smokers in those patients whose lesions underwent malignant transformation (Hall, Shaw et al. 2008). Ho et al (Ho, Risk et al. 2012), in a longitudinal outcome study of OED, noted that non-smokers were seven times more likely to undergo malignant transformation when compared to heavy smokers and suggest that it is other intrinsic factors that put these patients at higher risk of malignancy. The *P16* methylation smoking data from our data series is in support of this assertion. A weak statistical difference was noted between *P16* methylation and amount smoked in the 'high risk' group (Figure 10) with an observation that it is non-smokers with a tendency for *P16* methylation.

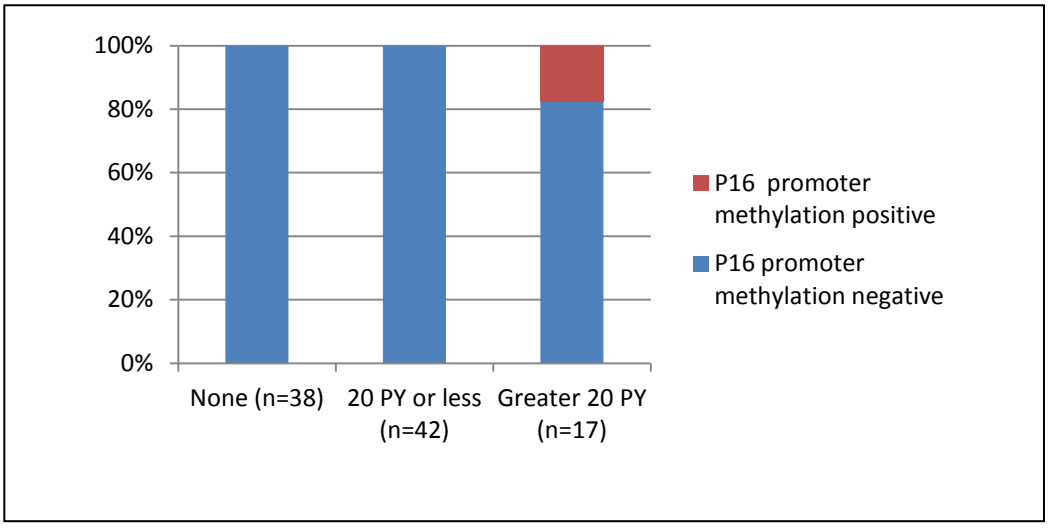



**Figure 8**. Normal cohort n=97

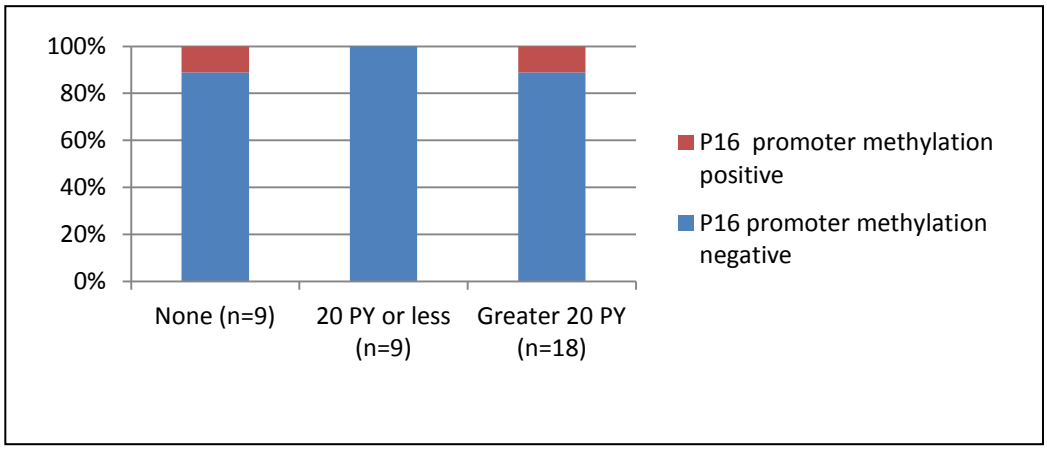



**Figure 9**. Low risk dysplasia cohort n=36

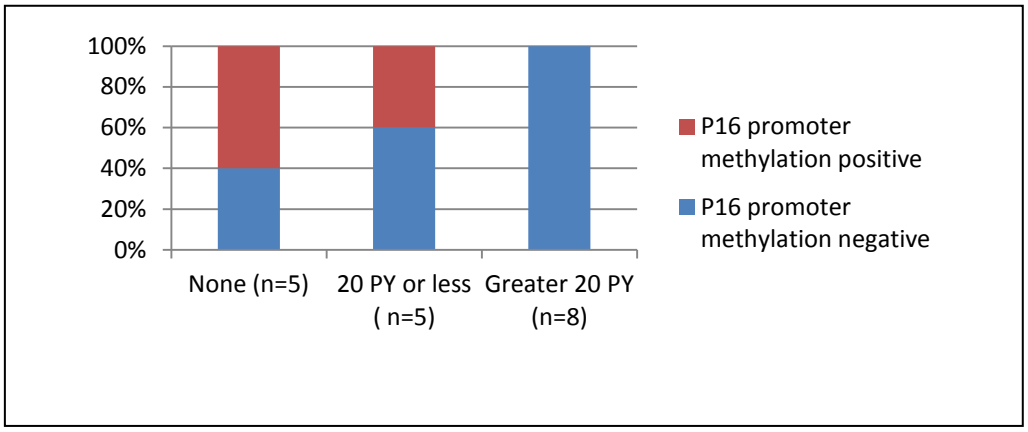



**Figure 10**. High risk dysplasia cohort n=18

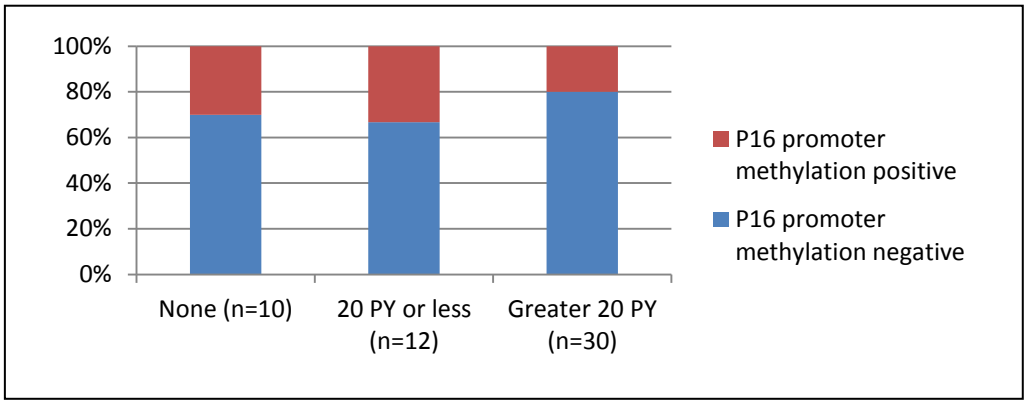



**Figure 11**. Malignancy n=52

**Figure 8-11: Frequency of *P16* methylation in saliva samples against smoking in pack years (PY) (horizontal axis) for each of the four diagnosis groups. n=individuals with smoking data**

Promoter methylation in the other three genes as related to smoking was not statistically significant across the four diagnosis groups (Appendix Table A9.1-A9.5). The combined gene panel showed statistical significance (Fisher's exact p=0.005) in the relationship between smoking and methylation in the normal cohort entirely on the strength of the three individuals with positive *P16* methylation.

#### 4.5.2.5 No significant association was observed between methylation rates and alcohol consumption for this target gene panel

The alcohol risk factor data was collected as three categories of intake: none, moderate (up to 28 units a week) or heavy (greater than 28 units a week). Twenty eight units a week is the upper limit of alcohol intake for an adult male as per the UK Department of Health. Heavy alcohol intake is associated with oral cancer (Blot, McLaughlin et al. 1988, Hindle, Downer et al. 2000, Bagnardi, Blangiardo et al. 2001) and it is therefore important to establish if it is independently associated with methylation. There was no statistical significance noted between alcohol and methylation in any of the individual target genes or the gene panel across all four diagnosis groups (Appendix Table A10.1-A10.5). As with the smoking data the numbers were small in some of the individual categories, notably the heavy drinkers in the non-malignancy diagnosis groups, therefore these statistics are treated with caution.

#### 4.5.2.6 No correlation was observed between methylation rates and stage of OSCC disease

A logical assumption is that a large tumour may exfoliate more cells into saliva and therefore be associated with increased amount of methylation. There was no observed pattern of increase in methylation as tumour size increased (Table 22). Although not statistically significant there did appear to be an increase in the number of positive genes as the TNM staging became more advanced (Table 23).

| Gene panel | pT* | | | | Total |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| 0 genes positive | 6 | 15 | 2 | 10 | 33 |
| 1 gene positive | 1 | 6 | 1 | 5 | 13 |
| 2 genes positive | 0 | 4 | 2 | 1 | 7 |
| 3 genes positive | 0 | 2 | 0 | 2 | 4 |
| At least one gene positive | 1/7 (14%) | 12/27 (44%) | 3/5 (60%) | 8/18 (44%) | NA |
| Total | 7 | 27 | 5 | 18 | 57 |

**Table 22. Tumour size versus the number of methylated target genes in the panel**
*** T=according to TNM classification**

| Gene panel | Stage | | | | Total |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| 0 genes positive | 3 | 7 | 4 | 19 | 33 |
| 1 gene positive | 1 | 4 | 1 | 7 | 13 |
| 2 genes positive | 0 | 1 | 3 | 3 | 7 |
| 3 genes positive | 0 | 0 | 0 | 4 | 4 |
| At least one gene positive | 1/4 (25%) | 5/12 (42%) | 3/8 (38%) | 14/33 (42%) | NA |
| Total | 4 | 12 | 8 | 33 | 57 |

**Table 23. TNM staging of OSCC versus number of genes positive in saliva samples**

Nodal status and in particular ECS are associated with disease aggression and survival outcomes (Rogers, Brown et al. 2009, Shaw, Lowe et al. 2010). In this saliva series *CYGB* was more commonly found in patients without ECS versus those with ECS (Table 24). There was weak statistical significance to this data (Fisher's exact p=0.05). The weight of this data is contentious due to the very small numbers involved and much larger numbers would be

required to draw any conclusions about association of methylation with ECS. There was no statistically significant relationship between methylation and the remaining three genes or the panel as a whole for either nodal status or ECS (Appendix 11 Table A11.1-A11.21).

| | | ECS | | Total |
|---|---|---|---|---|
| | | No | Yes | |
| CYGB | Neg | 8 | 17 | 25 |
| | Pos | 3 | 0 | 3 |
| Total | | 11 | 17 | 28 |

**Table 24. Frequency of *CYGB* methylation in pre-operative saliva samples in patients in comparison to nodal extracapsular spread (ECS) (Fisher's exact p=0.05)**

#### 4.5.2.7 No correlation was observed between methylation rates and survival or disease recurrence

Survival data was obtained for 53/62 patients in the OSCC cohort (Table 25). Follow-up (defined as date of primary surgery to the event: recurrence/death of disease/ died of other cause/ alive at last appointment) ranged from 1-65months with a median of 34 months. Sixteen of fifity three patients had follow-up less than 24 months, of these 14/16 patients died and 2/16 were lost to follow-up. Eleven of fourteen deaths from the disease (DOD) occurred within 24 months of the primary surgery date. The other three deaths occured at 25, 33 and 38 months.

| | | Fate | | | Total |
|---|---|---|---|---|---|
| | | Alive | DOD[a] | DOC[b] | |
| Gene Panel | Neg | 20 | 8 | 2 | 30 |
| | Pos | 14 | 6 | 3 | 23 |
| Total | | 34 | 14 | 5 | 53 |

**Table 25. Methylation status of pre-operative saliva samples versus patient survival status a=died of disease b=died of other causes.**

Recurrence data was available for 54 of 62 OSCC patients (Table 26). For the purpose of statistical analysis, 'recurrence' was defined as any further malignant event inclusive of local, regional and loco-regional lesions. It is accepted that this limits the ability to comment on whether markers predict for certain types of recurrent disease. However, it should be noted that the clinical data was acquired retrospectively from documentation of clinical notes spanning over 5 years. Time to recurrence was available for 19/22 patients and ranged from 3 to 35 months with a median of 7 months. Fourteen of twenty two patients with recurrence died of the disease. One of the twenty two died of other causes and 7 were alive at the last appointment with follow-up ranging from 34-55 months. Survival outcome data was not available for the remaining two patients. There was no statistical difference or observed patterns between methylation in this gene panel and recurrence (Table 26).

| | | recurrence | | Total |
|---|---|---|---|---|
| | | No recurrence | Recurrence | |
| Gene panel | Neg | 19 | 12 | 31 |
| | Pos | 13 | 10 | 23 |
| Total | | 32 | 22 | 54 |

**Table 26. Relationship between recurrence of disease (defined as local/loco-regional) and the presence of methylation in the target gene panel in pre-operative saliva samples (Fisher's exact p=0.724)**

Of 13/32 patients who did not have recurrence but did have a positive pre-operative saliva sample; the follow-up ranged from 11-57 months with a median of 43 months. It is noted that not every patient had a full 24 months of follow-up. Only 2 of these 13 patients had follow-up of less than 24 months. Of the 20/32 with no recurrence and a negative saliva result; only 2 of these patients had follow-up of less than 24 months both died of other causes within two months of surgery.

A Kaplan Meier survival analysis did not reveal a statistically significant difference (p=0.795 Mantel Cox) in disease related survival (defined as date of primary surgery to death from disease or last documented follow-up appointment) and the presence of methylation in pre-operative saliva for the gene panel (Figure 12) or individual genes (Appendix 14: F14.1-F14.4).

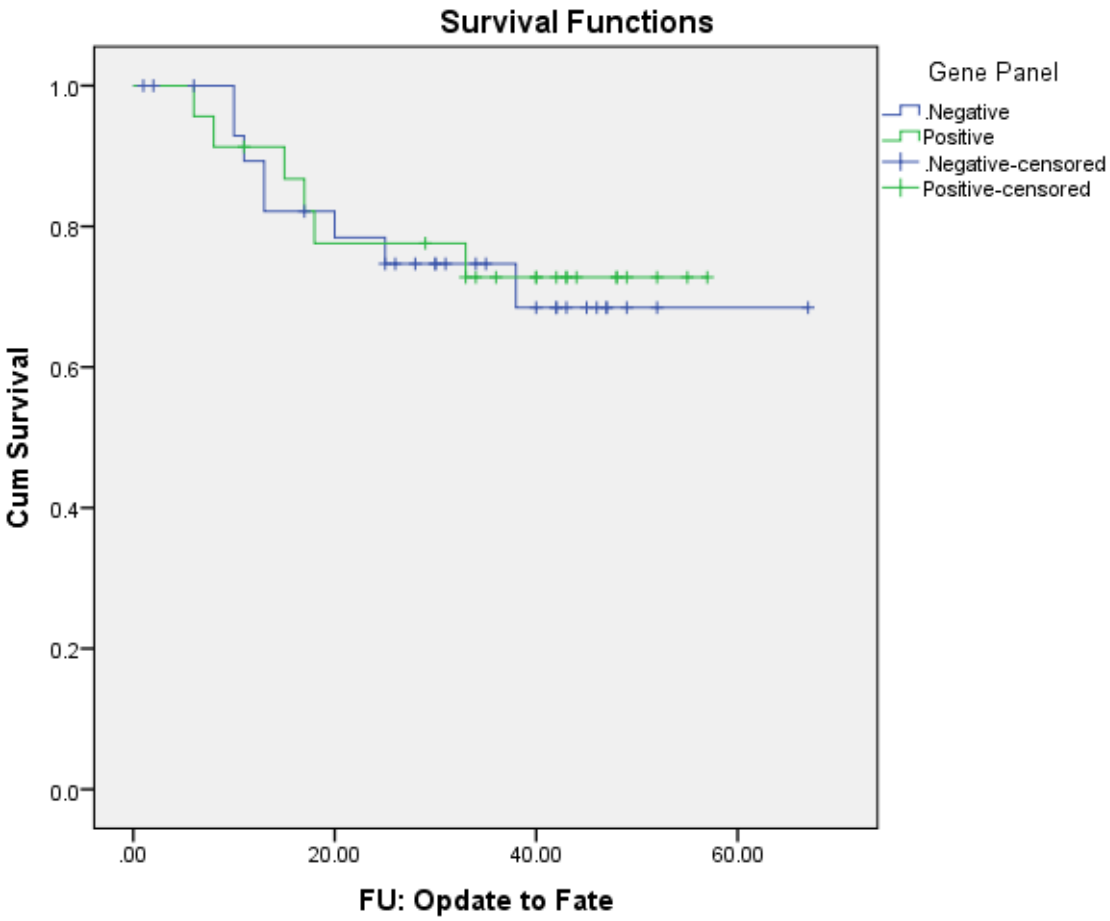



**Figure 12. Kaplan Meier curve of disease related survival as related to methylation in the target gene panel in pre-operative saliva samples.**

### 4.5.3 Clinical outcome

#### 4.5.3.1 No prognostic patterns of methylation were observed in a small OSCC cohort (n=12) of paired pre and post-operative OSCC patients

In a small subset (n=12) of the OSCC patients, matched pre and post-operative saliva samples (collected 4 weeks after primary surgery) were available along with clinical follow up data. The age range was 33-76 years with 8 (67%) male patients. The follow-up was 6-49 months. As the cohort is so small, statistical analysis is not appropriate. The results showed no strong patterns between methylation (either pre or post-operatively) and recurrence or death (Table 27).

| Pre-op saliva | Post-op saliva | Number of patients | Recurrence 0= No 1=Yes | Death 0=Alive 1=DOD[a] 2=DOC[b] | P Stage | Nodal Status 0=-ve 1=+ve | ECS 0=-ve 1=+ve | FU months | PORT 0= No 1= Yes |
|---|---|---|---|---|---|---|---|---|---|
| **Negative** | **Negative** | **3** | 0 | 0 | 2 | 0 | 0 | 47 | 0 |
| | | | 0 | 0 | 2 | 0 | 0 | 31 | 0 |
| | | | 1 | 1 | 4 | 1 | 0 | 13 | 1 |
| **Negative** | **Positive*** | **1** | 0 | 0 | 2 | 0 | 0 | 47 | 1 |
| **Positive** | **Positive** | **2** | 1 | 0 | 2 | 0 | 0 | 49 | 1 |
| | | | 0 | 0 | 2 | 0 | 0 | 40 | 0 |
| **Positive** | **Negative** | **6** | 0 | 0 | 4 | 0 | 0 | 36 | 1 |
| | | | - | 0 | 4 | 0 | 0 | 43 | 0 |
| | | | 1 | 1 | 4 | 1 | 1 | 33 | 1 |
| | | | 1 | 2 | 4 | - | - | 29 | 1 |
| | | | 0 | 0 | 3 | 0 | 0 | 48 | 0 |
| | | | 1 | 1 | 4 | 1 | 1 | 6 | 1 |
| **Total** | | **12** | **5** | **4** | | | | | |

*:positive means any of the four target genes in the panel are methylated. a=died of disease b=died of other causes

**Table 27. Methylation status of the matched pre and post-operative samples as related to disease recurrence and death.**

### 4.5.3.2 Methylation in saliva was not associated with malignant transformation of dysplastic lesions in this limited case series

Fourteen of the 60 patients in the dysplasia cohort donated more than one saliva sample during the course of their treatment. For the purpose of this study follow-up was defined as time elapsed from date of first saliva sample to date of last appointment and ranged from 17-73 months with a median of 47 months. Six of fourteen lesions transformed from a premalignant histological diagnosis to malignancy during this follow-up. Four out of six patients had negative saliva samples. Of these 2/4 had a history of invasive oral cancer and had previously undergone curative surgical treatment. One of the four had previous carcinoma-in-situ. The time from pre-transformation saliva sample to malignant change ranged from 16-21 months.

### Case studies of patients with methylation of target genes in saliva samples preceding malignant transformation of the index lesion

Two out of six patients (presented below) had gene promoter methylation in their saliva samples and in both cases it was present in a pre-transformation sample.

Patient 3447: a 72 year old male at first presentation with a lesion of moderate dysplasia on the ventral tongue. First sample was taken 18/11/2008 and was positive for *TMEFF2* methylation. The patient was kept under review. At a review appointment 10/02/2009 a

second saliva sample was taken and remained positive for *TMEFF2*. At this appointment an incisional biopsy of an extensive FOM lesion was taken which revealed PVL which progressed to proliferative verrucous carcinoma and was resected in May 2009. Patient remains in follow-up 21/07/2014 and is alive with subsequent incisional biopsies of the area confirming PVL but no suspicious clinical change. No further saliva samples were obtained after 10/02/2009.

Patient 3475 a 63 year old male first presented with moderate dysplasia ventral tongue 28/07/2008. A saliva sample was taken at this appointment and was negative for methylation in any of the target genes. A subsequent saliva sample was taken at review 27/01/2009 which remained negative and patient was discharged. In 2010 the patient was re-referred to the dysplasia clinic with a lesion on his ventral tongue. Histology from a biopsy 08/04/2010 described as moderate dysplasia which settled and the patient was kept under review. A saliva sample at this appointment was positive for *P16* methylation. No further saliva samples were collected. In 2012 the patient had an incisional biopsy of the area which revealed superficial invasive SCC. Complete excision of the lesion was performed November 2012 and histology of the complete lesion described mild dysplasia. The last documented follow-up for the patient was July 2014.

# 5.Discussion

## 5.1 Introduction

Oral cancer displays genetic and biological heterogeneity (Chung, Parker et al. 2004) which is not always evident at the histological level (Leemans CR, Braakhuis BJ et al. 2011). Clinical behaviour can be unpredictable within individual tumour types and implies a role for molecular profiling in diagnosis and prognostication. DNA biomarkers found in OSCC tumour tissue have been detected in saliva ( Rosas, Koch et al. 2001, Righini, De Fraipont et al. 2007, Carvalho, Jeronimo et al. 2008) and may have prognostic value (Righini, De Fraipont et al. 2007, Carvalho, Henrique et al. 2011).

Saliva is an accessible body fluid that can be collected non-invasively and has improved patient acceptability when compared with venepuncture (Koka, Beebe et al. 2008). DNA extracted from saliva is stable and suitable for room temperature storage without significant degradation (Feigelson, Rodriguez et al. 2001, Quinque, Kittler et al. 2006) allowing patients to self-sample at home which has the potential to reduce disease monitoring costs. Circulating DNA continues to attract attention as a source of biomarkers in several tumours and for a variety of purposes such as diagnostic, prognostic and predictive biomarkers. Saliva could potentially offer greater utility in OSCC and is even less invasive.

## 5.2 DNA yield from Oragene™ whole saliva and mouthwash collection methods from OSCC patients in the clinical setting is adequate for qMSP applications

In this arm of the study we aimed to compare three different methods of saliva collection in OSCC patients, in particular post-radiotherapy patients in whom xerostomia could potentially limit saliva collection. It is accepted that saliva contains adequate DNA for downstream applications. It was our intention to establish that adequate amounts of DNA can be obtained for downstream applications from saliva collected in a clinical, as opposed to a research setting and from the type of patients who would be the target of disease surveillance programmes. The 'ideal' conditions for sample collection are not always possible outside of the laboratory (Granger, Kivlighan et al. 2007) a problem compounded in the elderly population, where sampling is much more time-consuming, less well tolerated and often produces 'invalid' samples (Hodgson, Freedman et al. 2004).

Human saliva contains microbial content which can contribute to over-estimation of the DNA yield. UV spectrophotometry does not discriminate between human and non-human

DNA (García-Closas, Egan et al. 2001, Rylander-Rudqvist, Håkansson et al. 2006, Philibert, Zadorozhnyaya et al. 2008) therefore to improve accuracy of DNA yield calculations we used an *RNAse* P qPCR assay to estimate human DNA concentration. This technique allowed the differentiation of human from non-human DNA in contrast to spectrophotometry, which can be influenced by particulate matter or protein contamination.

The calculated median DNA yields (Table 11 and 12) are broadly comparable with those in the published literature for similar techniques but notably on predominantly healthy participants (Table 10). We found a large range of DNA concentrations across the samples in all three collection methods. A large variation in DNA concentrations has previously been observed with the Oragene™ method of collection (Rylander-Rudqvist, Håkansson et al. 2006, Viltrop, Krjutškov Kaarel et al. 2010). The sampling technique could have contributed to this variation as it is reliant on patients depositing the same amount of saliva in the collection pot which is challenging to standardise in the clinical setting. Equally It could represent the known individual variation of the number of oral epithelial cells in saliva (Dawes 2003).

Overall, the DNA yields (Table 10 and 11) from the PORT cohort (group 2) (mouthwash= 5.5μg and Oragene™ sponge= 4μg) are less than the OSCC cohort (group 1) (mouthwash=17μg and Oragene™ vial =122μg) but proved adequate for qPCR reactions. During the course of this study it was noted that the collection of saliva from patients with compromised oral competence in the clinical setting is logistically easier to undertake using sponges rather than mouthwash. The sponge method also requires less pre-storage processing. However, we did not collect any subjective patient opinion on their preferred method of collection.

The acknowledged limitations in this collection method study is a) the restricted time of ten minutes between sample collections and a swill time of 30 seconds for the mouthwash samples b) Oragene™ vial whole saliva collection technique was not compared in the radiotherapy cohort. The time restrictions were imposed by the clinical setting. Ideally an hour between sample collections would be preferable as a chance for saliva to replenish. Righini et al (Righini, De Fraipont et al. 2007) described swilling mouth wash for periods of up to 3 minutes however we found this was not tolerated by our patient group whose oral

competence was compromised by surgery. Ideally, to confidently conclude the best method of collection in the PORT cohort, whole saliva collection using the Oragene™ vials would also have been incorporated into the study design. It was considered unfeasible to test three methods of collection on the same radiotherapy cohort. Of note 4/11 post-operative patients in group 1 OSCC cohort (n=21: 10 pre-operative and 11 post-operative patients) which compared Oragene™ whole saliva collection to mouthwash had received PORT; the DNA yield was greater using the Oragene™ vials. This was an unexpected finding, the significance of which is uncertain with such a small population. It is possible in the context of reduced salivary flow that oral epithelial cells would be at greater concentration in saliva as they are not subject to frequent oral lavage. It was noted during saliva collection in group 1 that some PORT patients struggled to generate the recommended 2ml of saliva for the Oragene™ kits hence why whole saliva was collected using the sponge method in the follow-up study on the PORT group. To determine which whole saliva collection method yields the most DNA in the PORT group, Oragene™ vials would have to be compared to the sponge technique. However, in this study we have answered the clinically pertinent question of which collection method could be used to generate sufficient DNA for qMSP as part of a disease surveillance programme in OSCC patients. The population at risk from recurrence are those who have required PORT and they would potentially benefit most from surveillance biomarkers however they often cannot easily expectorate saliva nor have the oral competency to swish a volume of liquid in their mouth. In this study we show that the Oragene™ sponge method yields adequate DNA and could be used in this subgroup of patients.

### 5.3 The rarity of methylation in the normal cohort dictated that the technical sensitivity of the qMSP assays defines a positive result

A stated aim of this study was to determine the rates of methylation in saliva of a cohort of normal individuals with the objective of comparing a panel of epigenetic biomarkers in the saliva of known OSCC patients and a normal cohort using qMSP; to establish a threshold by which we could aid differentiation between normal and OSCC populations. The analytical sensitivity was defined by the lack of overlap between the positive standard picked for each assay and the biological (lymphocyte DNA) and technical (WGA) controls. In the case of *ADAMTS9* and *CCNA1* genes this was 0.5% for the remaining three genes: *CYGB, P16* and *TMEFF2*. Thereafter a positive result was defined as any sample with $\delta$CT equal to or less than the analytical sensitivity of each gene assay.

In this series there was no detectable methylation above the technical threshold for 4/5 genes: *ADAMTS9, CCNA1, CYGB, TMEFF2* in the control cohort (n=97). Therefore in these genes the analytical sensitivity of the gene assay could be used to define the threshold for disease-free samples in future research. Three individuals in the normal cohort had detectable methylation above the analytical sensitivity of the *P16* assay. Reports of detectable *P16* methylation in normal oral mucosa and saliva are mixed with some authors claiming a presence and suggesting *P16* has a role in field cancerisation and early malignancy (Maruya, Issa et al. 2004) and others claiming an absence or insignificant level in non-disease oral tissue/saliva (Shaw, Liloglou et al. 2006, Carvalho, Jeronimo et al. 2008). As the number of positive saliva samples in this normal cohort is so small (3/97) it is difficult to associate statistical relevance to them, additionally due to ethical constraints there is no follow-up of individuals in the normal cohort and therefore we are unable to comment on the prognostic relevance of positive samples in these individuals. At the time of sample collection a thorough intraoral examination was undertaken and no clinical lesion was present. However, *P16* methylation could potentially herald development of malignant/high risk disease in later life. In work by this and other research groups *P16* methylation, even in small amounts, has been identified in OED and predicted for oral cancer (Kresty, Mallery et al. 2002, Hall, Shaw et al. 2008, Cao, Zhou et al. 2009). This raises the ethical issue of whether these patients should have been followed up however as the samples were anonymised this was not possible in this study. Ideally, there would be clinical follow-up on a normal cohort of over five years but this requires heavy use of research resources and ethical considerations as it is not a screening programme. It is also likely that large numbers would be required as the risk of individuals being lost to follow-up is high if they are not part of a disease management programme. As the significance of the three individuals with *P16* methylation in saliva is unclear, more work on *P16* in a normal cohort would be required to set a threshold with confidence. However, for the purpose of this exploratory study the analytical sensitivity of 0.25% was accepted as a cut-off to define a positive sample in the disease groups.

## 5.4 The concordance of methylation of the selected gene panel was limited in matched saliva and tumour samples

The tumour positive concordance between tumour and saliva (n=34) in this series ranged from 11-70%. Two types of discordance exist: type A: positive saliva sample with negative tumour sample and type B: positive tumour sample with negative saliva sample. In this dataset, the former is unusual with only six recorded occurrences, of these three are from a single individual. A type A discordance may arise as a result of only a small part of the tumour being sampled. Tumour tissue is heterogeneous (Woolgar, Triantafyllou 2009) and the sample may not be molecularly representative of the whole spectrum of epigenetic change. A type A discordance may be a result of a representative area of the tumour having been omitted during the sampling stage but the saliva contains cells shed from the entirety of the tumour and may thus prove to be better than biopsy sampling as a means of detecting transformed cells.

Type B discordance is a more frequent occurrence in this data set which is perhaps more expected, as detection of methylation from tumour in saliva is reliant on the relevant cells being shed into the saliva in large enough quantities to be detectable amongst large numbers of normal cells. Another factor that may have contributed to the discordance in this dataset is the threshold by which a sample is considered positive. In the duplex qMSP assays used in this study the control DNA ranged from 0.25-1% methylated DNA because saliva requires a very sensitive assay to detect the small amounts of methylation present. However, in the case of tumour tissue some authors feel that very sensitive assays pick up biological “background noise” from the cell dense tumours and therefore a threshold for a positive result is set much higher at 5% methylation (Shaw, Liloglou et al. 2006, Shaw, Hobkirk et al. 2013). It could be argued that this is an arbitrary figure as an actual level of ‘normal’ methylation has not been established. The level of detection for tumours in the qMSP assays was set at 1% in this study and could have contributed to the rates of type B discordance seen here. Despite this the detection rate of methylation in tumours at a 1% threshold in this series is not markedly higher than that found in similar studies. The prevalence of *P16* in our tumour series as detected by duplex qMSP is 11/34 (32%). Kulkarni et al (Kulkarni, Saranath 2004) report much higher detection rates of *P16* methylation at 67% in OSCC tumour tissue. Righini et al (Righini, De Fraipont et al. 2007) publish a detection rate of 33% and Rosas et al (Rosas, Koch et al. 2001) 47% in *P16* using MSP. It is accepted that the detection technique used by these authors is different to our technique. MSP has been criticised for oversensitivity due to false positives as a result of mis-priming

(Shaw, Akufo-Tetteh et al. 2006) . However, Shaw and colleagues (Shaw, Liloglou et al. 2006) also report a comparable 28% prevalence of P16 in a tumour series using a more accurate quantitative pyrosequencing assay. In our dataset the presence of methylation in saliva was discriminatory to a statistically significant level for disease versus control (Fisher's exact p=0.001). It is recognised that we did not have matched tumour tissue for all 63 pre-operative OSCC saliva samples. Establishing the rate of tumour saliva concordance was not a primary aim of this study as it has already been explored in the literature (Rosas, Koch et al. 2001, Viet, Jordan et al. 2007, Righini, De Fraipont et al. 2007) . The tumour DNA available for this study had been used in other studies and was very limited. The qMSP assays were therefore not repeated with control DNA of higher methylation levels for the tumour tissue in this series.

It is an assumption that the methylated DNA detected in saliva is from exfoliated tumour cells. However, the DNA detected in saliva could be from a different type of cell such as immune cells that have responded to maliganancy induced inflammation. The actual source of methylated DNA in saliva is still an unknown and therefore the direct concordance of saliva and tumour is perhaps of questionable use. It is perhaps not the presence of the same methylated genes in saliva as tumour but the presence of methylation in saliva which heralds disease. This raises the question, for the purpose of discovering saliva methylation biomarkers for OSCC, whether saliva from OSCC should be interrogated by genome wide discovery assays to generate potential genes rather than OSCC tissue. As is discussed later in this chapter one study (Viet, Schmidt 2008) has used a methylation array to interrogate a small cohort of pre-operative and post-operative OSCC saliva samples to identify potential salivary methylation markers. Of note it identified several markers that have not previously been explored in saliva and did not detect markers which have commonly been studied in HNSCC such as *DAPK, MGMT* and *APC* at high enough methylation levels to be significant.

### 5.5 Saliva provided a greater quantity of good quality DNA than buccal scrapes

The inclusion of this data was to determine if buccal scrapes provide a superior surrogate tissue to saliva in a normal cohort. Normal mucosal biopsy would be the ideal tissue but cannot be justified in large enough numbers under UK ethics policy. Previous studies have shown this exact buccal scrape technique to be a satisfactory surrogate for dysplastic lesion tissue (Hall, Shaw et al. 2008). However, without a clinically identifiable lesion to scrape it is

possible that a scrape sample will omit potentially methylated cells which can be included in saliva samples with cellular contribution from the entire oral mucosa.

The poorer performance of the buccal scrapes in comparison to whole saliva in terms of DNA yield and failed assays, in this study, may have been a result of poor operator technique. However other authors have also shown that whole saliva out performs buccal cell collection methods in terms of DNA yield (Rogers, Cole et al. 2007, Mulki, Shetty et al. 2013). A comparison between detectable methylation of a three gene panel from oral scrape (27/53) samples versus saliva samples (33/53) in an HNSCC cohort (n=53) showed a trend toward increased detection in saliva but this was not statistically significant (Ovchinnikov, Cooper et al. 2012). Scrapes have the advantage of sampling a specific lesion but may miss other regions of the oral cavity which is why saliva may have had higher detection rates in this study. Saliva has coverage of and therefore cellular material from the whole oral cavity and is likely to be preferable to oral scrapes for the detection of occult lesions.

The methylation data is weakened by virtue of only one run having been performed in 70/90 scrape and saliva samples and without the benefit of duplicate or triplicate runs to confirm results, the single run data must be treated with a degree of caution. The concordance levels between buccal scrape and saliva are very high (98-100%) representing the extremely low rates of detectable methylation of the gene panel within the normal cohort. The five positive results (3 scrapes and 2 salivas) were all discordant so the seemingly high levels of concordance must be treated with caution. With this in mind our data does not support buccal scrapes outperforming whole saliva in terms of DNA yield nor assay failure. We are comfortable retaining saliva, in the absence of a detectable lesion, as having the greater potential as an oral surrogate tissue.

### 5.6 Patient populations: a young normal cohort and mixed histology in the dysplasia cohort

One of the aims of the study was to establish the presence of DNA methylation for this panel of genes in a normal cohort to set a cut-off by which to discriminate saliva samples from patients with disease. A directive at the beginning of the study was to collect saliva from patients who were matched to the typical OSCC profile for demographic and risk factor data. As noted in the demographic section the 44/97 individuals in this cohort were below 40 years old which is not comparable to the disease cohorts where only 1/62 patient

in the OSCC cohort and 5/60 in the dysplasia cohort were younger than 40 years. This was unforeseen in the initial planning when the clinical environment from which the saliva samples should be collected was decided. After an interim analysis of the demographics of the normal cohort revealed a bias towards the 40 and under age group, the clinical environment was changed to one where older patients routinely attend. In doing this similar numbers of individuals in the older categories were achieved: a) 41-60 years; OSCC n=29, Dysplasia n=31 Normal n=29 and b) 61 years and older: OSCC n=28, Dysplasia n=25, Normal n=24. Individuals under 40 years of age are less likely to accrue a smoking history of 20 pack years and a criticism levelled at this data might be that the absence of methylation in the normal cohort is due to the lack of exposure to risk factors. However, the methylation data of a young cohort may have value as there are reported increases of OSCC in younger individuals (18-45 years) without HPV or exposure to large amounts of alcohol and tobacco (Llewellyn, Linklater et al. 2003, Patel, Carpenter et al. 2011) which appear to have unique clinical profiles (Harris, Kimple et al. 2010). An understanding of the methylation profile of young, non-disease patients may therefore be useful for the development of biomarkers in groups not stratified by common risk factors.

The collection of the dysplasia saliva samples was performed at a clinical visit if the clinician clinically deemed the lesion to be dysplastic. Four of sixty individuals had a histological diagnosis of PVL but remained in the analysis as PVL is considered a high risk premalignant lesion and is of interest to this exploratory study on saliva as a surrogate for premalignant and malignant lesions. Similarly, samples from four patients with HPV negative oropharyngeal squamous cell carcinoma were retained as part of the analysis. Although not strictly oral cancer, all four were base of tongue or soft palate and therefore could feasibly contribute cellular material to saliva. As they have a non-HPV aetiology to the disease they could be considered both aetiologically and prognostically similar to the true OSCC group with nodal status being the most important prognostic factor (Rietbergen, Brakenhoff et al. 2013), in common with OSCC .

### 5.7 Detectable methylation in saliva is associated with OSCC and high risk dysplastic disease

The full clinicopathological data for each individual became available after the qMSP assays had been performed. As mentioned earlier, in some cases the dysplasia saliva samples were collected at the time of a clinical diagnosis and in 5/60 cases the clinically dysplastic lesion was histologically found to be MISCC. As the total number of samples in each diagnosis group was, in some cases, small and the total number of 'positive' samples was modest, the results were binned to facilitate more useful statistical and clinically relevant analysis of the data. The normal cohort remained as such, the mild and moderate dysplasia were classed as low risk on the basis that these lesions are relatively less likely to transform into malignancy. This classification is informed by a longitudinal study within our research group on the malignant transformation of lesions from 91 OED patients which showed that mild and moderate OED had similar transformation rates and were less likely than severe dysplasia to transform (Ho, Risk et al. 2012). The PVL and severe dysplasia were classed as high risk lesions. The MISCC and OSCC were collectively analysed as malignancy.

Eight patients in the dysplasia cohort had a history of OSCC but there was no association between methylation and previous disease in this data set. A history of OSCC was not an exclusion factor in the OED group because these patients present a high risk group and present an area of interest to molecular biomarker research; particularly as previous surgery and radiotherapy can make clinical and histopathological assessment of these lesions difficult. Rosin et al (Rosin, Lam et al. 2002) analysed 68 oral premalignant lesions for loss of heterozygosity from patients with previous oral cancer. Thirty six of sixty eight lesions progressed to a second oral malignancy (SOM). They found a 26.3 fold increase in the risk of developing an SOM if 3p +/9p loss was detected in the lesion which was a better risk prediction than histological diagnosis. Ho et al (Ho, Field et al. 2013) analysed the clinical outcomes of a cohort of patients who had developed OSCC in an area of OED. They comment on the high risk nature of this group of patients noting that 76% of patients had a further event (defined as SOM, OED or recurrence) in 5 years. In their series 5/23 patients developed an SOM and 5/23 developed further OED.

**5.8 Detection of DNA methylation in saliva appears to have limited sensitivity but good specificity in the detection of OSCC**

The potential advantages of saliva as a surrogate are clear from the work conducted in this study and the body of published literature; it is non-invasive therefore easy to collect with low risk of transmitting disease to healthcare professionals. The ease is such that patients can collect their own saliva which lends itself to community based disease surveillance programmes. It is a stable body fluid which yields sufficient DNA for downstream applications and requires minimal equipment to collect, store and process it.

Current opinion is that the suitability of saliva as an oral cancer tumour surrogate relies on the ability to detect markers from a tumour origin. In this study the methylation rates of the panel of markers in saliva samples from the OSCC cohort were low when considered as singular markers: *CCNA1* 9/67 (13%), *CYGB* 10/67 (15%), *P16* 17/67 (25%), *TMEFF2* 11/67 (16%) but other similar studies using qMSP on large normal cohorts have also reported similar methylation rates in saliva (Carvalho, Jeronimo et al. 2008, Rettori, De carvalho et al. 2013). The frequency of methylation in disease saliva samples was improved when considered as a panel 29/67 (43%) which is expected as OSCC is a molecularly heterogenous disease. However, the methylation rates in saliva under-represented the frequency of methylation in the tumour tissue for three of the genes: tumour positive concordance for *ADAMTS9*: 11%, *CCNA1*: 45% *TMEFF2*: 33% and was comparable to other published data for *P16: 70%* and *CYGB*: 60%. The possible reasons for which have been discussed earlier in section 5.4 of this chapter. Methylation of the gene panel in saliva appears to be specific as it was rarely detected (only 3/97 single marker *P16*) in the control cohort in this series. One of the strengths of this study is the relatively large control cohort n=97 as many comparable studies have used much smaller cohorts n= 5-30 (Rosas, Koch et al. 2001, Righini, De Fraipont et al. 2007, Nagata, Hamada et al. 2012, Ovchinnikov, Cooper et al. 2012) and the reported high levels of specificity in these studies may have been a result of using such small numbers. It should be noted that sensitivity and specificity have not been extracted from our pilot data as this is generated from a training set of clinical samples and a separate validation set would be required to calculate these parameters.

One of the strongest studies in the saliva methylation marker literature is presented by Carvalho and colleagues (Carvalho, Jeronimo et al. 2008) who analysed 21 genes using a candidate approach on salivary rinses and serum from a large cohort of mixed primary HNSCC patients (n=211) and matched controls (n=527) using qMSP assays. Targets that were methylated at high frequency in control samples were identified and eliminated by

comparison of primary HNSCC tissue and control salivary rinses and between HNSCC salivary samples and control salivary samples. A reduced cohort was used for this stage due to limited DNA and ranged from 11-136 tumour tissue cases and 30-500 salivary controls therefore not every gene was tested on every sample. Thirteen genes (*CCNA1, DAPK, DCC, MGMT, TIMP3, MINT31, P16, PGP9.5, AIM1, ESR, CCND2, MINT1 and CDH1*) were differentially methylated between HNSCC and control samples and used to analyse a limited cohort of salivary HNSCC cases and controls. Genes with an AUC >50 and a specificity >90% and sensitivity of at least 10% were selected for testing on the expanded cohort of case and control saliva. Of note two of the target genes in our data were included in their study; *P16* had a low sensitivity of 4.5% but was included in the final analysis as it was 100% specific and *CCNA1* had a sensitivity of 20% and specificity of 97%. In keeping with our findings the salivary gene panels displayed high specificity but limited sensitivity with the best performing panel (*MINT31,CCNA1,DAPK,DCC,P16*) having a sensitivity of 34% and a specificity of 92%. Different gene combinations achieved higher specificity but reduced sensitivity (Carvalho, Jeronimo et al. 2008). Another study achieved higher sensitivity in saliva samples using a panel of two markers. Ten genes (*KIF1A, EDNRB, CD44, TERT, CDH4, NISCH, PAK3, VGF, MAL, FKBP4*) initially identified via a candidate approach were evaluated for the presence of methylation in control saliva. *EDNRB* and *KIF1A* were selected for further analysis as they had the lowest levels of methylation in control saliva 1/47 (2%) and 3/45 (6%). *EDNRB* and *KIF1A* showed high tumour specificity despite methylation being present at low levels in greater than 80% of normal mucosal samples. Methylation of *KIF1A* was detected in 27/71 (38 %) of HNSCC saliva samples with a reported sensitivity of 37% and specificity of 98%. *EDNRB* methylation was present in 48/71 (68%) of the HNSCC saliva samples. Sensitivity was improved to 77% and specificity maintained at 93% when the genes were used in combination for the detection of HNSCC (Demokan, Chang et al. 2010). A large methylation array also identified *EDNRB* and *KIF1A* as genes with high sensitivity (68%, 77%) and specificity (100%, 92%) for the detection of OSCC using tissue samples but the genes were not further evaluated in saliva samples (Guerrero-Preston, Soudry et al. 2011).

Sensitivity and specificity of greater than 90% for the detection of OSCC in salivary rinses has been reported with a panel of three genes *ECAD, MGMT* and *TMEFF2* (Nagata, Hamada et al. 2012). Thirteen genes were identified by a candidate approach and analysed for differential methylation between OSCC and healthy saliva using MSP. ROC analysis was used to determine the most diagnostically useful genes. They report a much higher

frequency of *TMEFF2* 29/34 (85%) in OSCC saliva than was found in our study 11/67 (16%) but similarly few in the healthy cohort 3/24 saliva samples. However, this study has a limited OSCC cohort n=34 and an unmatched small control group n=24. It also calculates prediction efficiency, sensitivity and specificity data using a test set of samples. The reported sensitivity and specificity of this panel is not validated on a separate cohort of OSCC patients and may explain such optimistic results. Furthermore this study reports a gene sensitivity ranging from 24-94% for *ECAD, MGMT, DAPK, P16 and TIMP3* whereas Carvalho and colleagues (Carvalho, Jeronimo et al. 2008) report modest sensitivity of 5-30% for these genes. This variance may be a result of differences in the detection techniques, sampling techniques or case/control mix.

A concern raised from our data series is that detection of methylation in saliva lacks the requisite sensitivity for use as a clinical test because the presence of methylation in our study was associated with 43% (29/67) of the oral cancer group and 39% (7/18) of the high risk group which would potentially mean 57% of OSCC and 61% of high risk lesions would go undetected using this panel. However, this is exploratory work and the assays proved specific and reproducible and with an expanded gene panel the sensitivity may be improved. The observed modest levels of methylation in saliva could be panel dependent. The majority of salivary methylation biomarker studies in HNSCC have used the candidate approach for target gene selection (Rosas, Koch et al. 2001, Righini, De Fraipont et al. 2007, Carvalho, Jeronimo et al. 2008, Pattani, Zhang et al. 2010, Demokan, Chang et al. 2010, Nagata, Hamada et al. 2012). Similarly, the genes used in our data series were identified from the literature(Tokumaru, Yamashita et al. 2004, Maruya, Issa et al. 2004, Kulkarni, Saranath 2004, Kato, Hara et al. 2006) and previous work by our group using moderate numbers of tumour and adjacent normal tissue showing tumour specificity: *TMEFF2:* 72% (18/25) (personal communication Mark Reid), *CYGB:* 44% (35/80), *CCNA1:* 49% (38/78), *P16:* 28% (22/80) positive tumours (Shaw, Liloglou et al. 2006).

In our study at least one of the panel of candidate genes (*CCNA1, CYGB, P16, TMEFF2*) was positive in 24/34 (71%) tumours which is similar to other studies (Righini, De Fraipont et al. 2007) but is not high enough as a useful clinical biomarker. Shaw et al (Shaw, Hobkirk et al. 2013) have recently published on the modest informativity (defined as in all or high percentage of cases) of three genes also used in our saliva study. In the series presented by Shaw 30/48 (63%) of tumours were positive for at least one gene using qMSP to detect a panel containing *CDH1, P16, CCNA1* and *TMEFF2* in HNSCC tissue (Shaw, Hobkirk et al.

2013). If we are assuming the positive cells in saliva are directly from tumour, until the problem of tissue heterogeneity of primary cancer tissue (Esteller 2007, Esteller 2008, Woolgar, Triantafyllou 2009) is satisfactorily resolved in epigenetic studies (Ku, Naidoo et al. 2011) it remains a potential barrier to progression in surrogate tissue research. Loh et al (Loh, Liem et al. 2010) suggests better quality control is required for the purity of tumour tissue used in biomarker studies as the assessment of methylation is sensitive to cell purity. Using methylation array data on 98 gastric tumours Loh determined a tumour content of at least 70 % is required to achieve reliable sensitivity (Loh, Liem et al. 2010) thus studies that do not adhere to this level of quality control may identify markers of limited use at the discovery phase.

Genome wide technologies are becoming more accessible and may provide the answer to identifying relevant tumour markers from heterogenous disease. Two studies have used genome wide approaches to identify novel methylation markers for use in saliva. A high density promoter methylation platform which interrogates 27,578 CpG sites from 14,495 protein coding gene promoters and 110 mRNA gene was used to identify hypermethylated genes on small cohorts of tissue (HNSCC n=4, leukoplakia n=4 and normal n=4). In a complex analysis the data was combined with publicly available methylation and gene expression data to identify 140 hypermethylated and downregulated genes from this discovery screen. Eight of one hundred and forty genes (*EDNRB, HOXA9, GATA4, NID2, MCAM, KIF1A, DCC and CALCA*) were chosen using multiple criteria which is not made explicit in the paper, to establish methylation levels in OSCC (n=24) and normal mucosal tissue (n=12)using qMSP assays. Receiver operating characteristic (ROC) analyses were performed to establish those genes that were differentially methylated between these tissues; 4 genes were selected with an area under the curve (AUC) greater than 0.75 (*KIF1A, HOXA9, NID2, EDNRB*) and of these *HOXA9* and *NID2* had 100% specificity and sensitivity >70%. These two genes were tested in a separate cohort using HNSCC tissue (n=55) and non cancer tissue (n=37). Both genes had >80% sensitivity and >90% specificity individually and in combination greater than 90% However, when tested in saliva from OSCC (n=16) and non-cancer patients (n=19) the sensitivity and specificity of both genes decreased: *HOXA9* sensitivity was 75 % with a specificity of 53% and *NID2* sensitivity was 87% with a specificity of 50%. In combination the sensitivity decreased to 50% but specificity improved to 90% (Guerrero-Preston, Soudry et al. 2011). The saliva data is limited by very small numbers of cases and controls. A reduction in sensitivity is expected in saliva as there are a small number of methylated cells amongst vast numbers of normal cells. However, the reduction

in specificity for the individual genes in saliva is unexpected in the context of a high specificity in the tissue data. This may suggest that cells other than mucosal, present in saliva and not tissue, are contributing a methylation signal. This may be explained by the observed phenomenon of compartment specific methylation whereby some markers exhibit significant methylation in normal control subjects but in only one compartment (Carvalho, Jeronimo et al. 2008). This implies that the discovery phase for potential salivary biomarkers using a genome wide approach should interrogate saliva samples in addition to tissue.

Viet et al (Viet, Schmidt 2008) performed a proof of principle study using the Illumina Goldengate methylation array to interrogate 807 cancer associated genes in pre-operative OSCC and post-operative saliva samples with the aim of identifying highly methylated loci to produce a composite biomarker with diagnostic value. A limited cohort of OSCC patients (n=13) provided primary tumour tissue, pre-operative and post-operative whole saliva and ten healthy whole saliva samples provided a control on the array. Genes that were methylated in both pre-operative saliva and oral cancer tissue were selected and genes methylated in control or post-operative saliva were excluded. Forty one gene loci from 34 genes were identified using this method with sensitivity and specificity calculated for each locus. Nine gene panels of 4-10 genes were constructed and the sensitivity and specificity of each was calculated; this ranged from 62-77% and 83-100% respectively. The highest sensitivity was associated with the 6 gene panel: *GABRB3_E42_F, IL11_P11_R, INSR_P1063, NOTCH3_E403_F, NTRK3_E131_F and PXN_P308_F* with a sensitivity of 77% and specificity 87%. This sensitivity and specificity was not subsequently validated on a separate OSCC cohort so should be treated as preliminary data. However, these figures are an improvement on other published gene panels in saliva which achieved high specificity at the expense of sensitivity or vice versa using genes identified by a candidate approach in tumour tissue (Carvalho, Jeronimo et al. 2008) which is further support for the use of arrays in discovery of informative biomarkers in saliva. Genes previously methylated in HNSCC were included on this array such as *DAPK, MGMT, APC, P16, ESR* and *NOTCH3* but only the latter three were significantly methylated in this array. The author's explanation is that the probe design on the array requires higher levels of methylation in target genes to hybridise to the probes and therefore selects more biologically relevant markers (Viet, Schmidt 2008).

The selection of *ADAMTS9* as a potential marker was generated from collaborative work (Veeramachaneni 2010) as discussed in section 4.1.1. Forty three % of the normal adjacent tissue was noted to be methylated above the 5% cut-off using quantitative pyrosequencing methylation assays however it was not frequently methylated in normal saliva or OSCC saliva as detected by qMSP. The observed methylation in normal adjacent tissue could be due to the fact this tissue is not a true normal and could reflect molecular changes associated with field cancerisation. Despite being present in 26% (9/34) of tumours in this series only 1/218 saliva samples (pre-operative sample) was positive for methylation of *ADAMTS9* as detected by qMSP assays above the cut-off (0.5%) for this gene. This sample was positive for multiple genes in the panel. Primary cancer tissue is heterogenous and it could be that the methylation (52%) detected in the un-microdissected tumour samples was from cells other than tumour tissue such as stromal, immune or vascular cells. Myopepithelial and stromal fibroblasts in intimate proximation to malignant tumours, which are likely to be included in resections, have been shown to have distinct epigenetic changes of their own (Hu, Yao et al. 2005). A lack of presence in pre-operative saliva samples may be a result of the fact that these cells are not exfoliated in the numbers required to detect them in saliva. Allelic imbalance at 3p14.2 has been associated with malignant transformation of premalignant lesions (Rosin, Cheng et al. 2000, Rosin, Lam et al. 2002, Tsui, Rosin et al. 2008) which may suggest *ADAMTS9* has an earlier role in tumorigenesis, a possible reason why it was not detected in OSCC saliva and at relatively low numbers in the tumours (26%). However, *ADAMTS9* methylation was not detected in any of the dysplasia saliva samples in our dataset. As such it was considered an uninformative marker and removed from the panel for further analysis of saliva.

One of the difficulties of accurately comparing results from body fluid biomarker studies is the methodological differences, namely detection techniques and the use of mixed HNSCC disease cohorts. Earlier studies which have reported high rates of salivary markers such as *P16* methylation (Rosas, Koch et al. 2001, Righini, De Fraipont et al. 2007) which are not supported in other studies (Carvalho, Jeronimo et al. 2008) or our data may be a result of using MSP which is a qualitative technique prone to false positives especially if high cycle numbers are used. Often these studies have not clearly stated a threshold to define a positive result and the presence of methylation at any level defines a positive. We have maintained stringent thresholds to define a positive sample for the qMSP assays used in this study and this may be a reason why the detection of methylation in saliva in our dataset is lower than some other published works. It may be that we have achieved

specificity at the expense of sensitivity. However, this technique continues to improve and evolve and is likely to develop more sensitive assays.

The modest levels of detected methylation in saliva compared to tumour tissue in our study could also be a result of the limitations of saliva as a surrogate. It is a biological fluid with predominantly normal cells. There is an estimated 10,000 oral epithelial cells per ml of saliva with an estimated turnover of 2.7 hours (Dawes 2003) with potentially only 2-3 exfoliated from the tumour for every 100 cells found in saliva (Boyle, Mao et al. 1994). This requires very sensitive methods of detection which qMSP does provide. However, 'false-negatives' are a risk with this degree of dilution from normal cells.

## 5.9 Common risk factors for OSCC were not identified as confounding variables for methylation of this target gene panel

### Age

Cancer is associated with both methylation and ageing (Fraga, Agrelo et al. 2007). There is some evidence that methylation in normal tissues increases with age. *RARB2, RASSF1A* and *GSTP1* have been shown to be hypermethylated in an age related manner in normal prostate tissue (Kwabi-Addo, Chung et al. 2007). The authors of a methylation array study performed on mixed normal tissues (n=139) also observed an age-related increase in methylation in certain CpG loci within CpG islands of normal tissues. *ESR1, GSTP1, RARB* and *RASSF1* were amongst the genes observed to have increased methylation associated with aging (Christensen, Houseman et al. 2009). It is worthy of note that these changes were not observed in the normal head and neck tissues (n=11) included in the study.

It would be important, in selecting a gene panel for clinical use, that the methylation of the genes is not simply associated with ageing. The skewed demographic data (44/97 in the normal cohort under 40) in this study limits the usefulness of a statistical comparison of simply age versus methylated samples. In analysing this data for age as a cofounder, it is also important to acknowledge that the frequency of cancer increases with age and the age data has therefore been stratified according to diagnosis. If age alone were associated with methylation of this panel of genes one would expect there to be significantly higher number of positive samples in the older individuals in the normal cohort. This is not the case which has been shown with only 3/97 individuals having a positive sample in the *P16* gene assay. When stratified into the diagnostic categories of low-risk, high risk and

malignancy there is no statistical difference between the presence of methylation and increasing age for the disease groups. It is therefore reasonable to conclude that age does not appear to be a confounding factor for the detection of methylation for this panel of genes . This is supported by other authors who also did not find an age-related effect with a gene panel inclusive of *P16* and *CCNA1* (Carvalho, Jeronimo et al. 2008).

**Smoking and methylation**

There is conflicting data in the literature regarding the association between *P16* methylation and smoking. Reports of *P16* methylation in up to 50% of normal adjacent tissue in HNSCC studies have led authors to hypothesise a smoking aetiology for this aberrant methylation (Kulkarni, Saranath 2004). Promoter methylation of *P16* has also been observed in up to 10% of normal oral mucosa samples from smokers (von Zeidler, Miracca et al. 2004, Ruesga, Acha-Sagredo et al. 2007). However, a study from our research group using a more accurate and quantitative technique (pyrosequencing methylation assays) on a cohort of HNSCC tumours detected methylated *P16* in 1/26 (4%) (Shaw, Liloglou et al. 2006) in normal adjacent mucosa samples. In this cohort of patients 57% were heavy smokers. The reported high *P16* promoter methylation rates in normal tissue are from studies which utilised MSP which is vulnerable to false positive results from mis-priming (Shaw, Akufo-Tetteh et al. 2006). A recent large study of promoter methylation in 368 head and neck cancer cases using qMSP observed that *P16* methylation was negatively associated with tobacco use when subject to a multivariate analysis (Roh, Wang et al. 2013). In saliva samples an association with *P16* methylation and smoking has also been observed using MSP (Ovchinnikov, Cooper et al. 2012) but a large well designed study analysing saliva from up to 527 controls and 211 HNSCC cases with qMSP was unable to establish smoking as a confounding variable for *P16* methylation (Carvalho, Jeronimo et al. 2008). In the current study there was no association between *P16* promoter methylation detected in saliva and smoking (Chi-square p=0.51)**.** Three of ninety seven individuals in the control cohort were positive for *P16* methylation all of whom were smokers but do not to our knowledge have malignancy which is a potential confounder. It is difficult to draw any firm conclusions whether smoking is a contributory factor in *P16* methylation from such a small number of positives in a training set of samples without longitudinal follow-up, as the aberrant methylation could be heralding a malignancy. There was no observed association between smoking and methylation of any of the other target genes in this study (Appendix 9). To answer the question of whether *P16* methylation is smoking related, further

investigation of *P16* methylation in healthy individuals with a history of large numbers of pack years and a non-smoking cohort would be required.

### 5.10 Detection of methylated DNA in saliva samples was not associated with staging or prognosis of OSCC

#### Stage

In this study there was no observed association between methylation in pre-operative saliva and any of the staging pathology criteria. One might expect larger tumours to shed more cells and as such as the pT increased then an increase in methylation. Such a pattern was not observed in this dataset but could be due to the small numbers of positive samples. Although not statistically significant, a trend of increasing number of methylated genes with increase in disease stage was noted. This may be due to more pathways being involved as the tumour matures. It has been observed that the number of chromosomal aberrations increases with T2, T3 and T4 stage OSCC tumours displaying more aberrations than T1 tumours. Advanced stage tumours had additional deletions and gains of chromosomal regions not seen in earlier stage tumours (Okafuji, Ita et al. 2000). The numbers of patients in this study are too small to comment on particular methylation patterns at each stage.

One of the limitations of this and other published work on saliva and HNSCC tumour tissue is the use of mixed stage tumours at the discovery phase of markers. In this study of 62 tumours a predominance of stage 2 and 4 tumours was noted. However, this is often unavoidable because not many T1 tumours get tissue banked and T3 tumours are relatively rare because tumours of >4cm often invade adjacent structures. In searching for genetic and epigenetic aberration in HNSCC disease early studies have often used small, mixed cohorts of HNSCC tissue at various stages of the disease for the discovery phase of markers. Also a lot of early studies have been limited to evaluating targeted areas of genes rather than the entire genome and more relevant markers may have been missed (Patmore, Cawkwell et al. 2005). If one of the roles of salivary biomarkers would be to detect early stage disease it is wasted effort investigating genes that are only methylated in advanced stage disease therefore better characterisation of the methylome of tumours at different stages of the disease is required. Array studies have already shown that different DNA methylation profiles are associated with tobacco and alcohol exposure, early and advanced

stage disease (Marsit, Christensen et al. 2009) and tumour aggression as measured by ECS and recurrence (Jithesh, Risk et al. 2013). As genome wide techniques become cheaper and more widely available this problem could be more easily addressed.

**Prognosis**

In the OSCC cohort studied in this thesis, the time from primary surgery to disease recurrence ranged from 3 months to 34 months with a median of 7 months. Seventy eight percent recurred within 12 months and 89% within 24 months. This recurrence data is comparable to an earlier Liverpool cohort (Kissun, Magennis et al. 2006) and other studies (Hicks Jr., North Jr. et al. 1998, Koo, Lim et al. 2006). For the purpose of statistical analysis, local and loco-regional recurrence, were binned together as recurrence. Authors of a previous study have suggested the detection of methylation in pre-operative saliva is associated with more aggressive disease as measured by disease recurrence and poorer survival (Carvalho, Henrique et al. 2011). In our data series, 22 patients had a documented recurrence and 14 died of disease related causes. We were unable to establish an association between prognosis and methylation in pre-operative saliva; a finding which is supported by another similar study (Righini, De Fraipont et al. 2007).

The post-operative data in this study was limited to 12 patients with a single sample at 4 weeks post surgery with the aim of detecting residual disease. There was no observed pattern of methylation for the studied panel of genes associated with recurrence or disease specific survival in our data set. Considering the frequency of methylation of this panel of genes in saliva the post-operative cohort is too small with too few post-operative samples to identify potential methylation markers for post-operative OSCC disease surveillance. Samples would need to be collected over at least 24 months as most recurrences occur within this time frame and samples would need to be collected at regular intervals to demonstrate any disease relevant changes in the methylation patterns of saliva. This was originally planned as part of the experimental design but was not possible within the confines of the clinical setting as saliva collection was not a priority at the clinic and achieving reliable collection at fixed intervals for individuals in the study did not occur. Implementing the collection of regular saliva samples over large numbers of patients over extended periods of time to yield relevant information can be difficult. In the two published studies detailing post-operative longitudinal saliva sampling it is noted that 16-30% of the patients are unable for follow-up sampling for various reasons (Righini, De Fraipont et al. 2007, Rettori, De carvalho et al. 2013) which may impact on the utility of the data.

There are very few studies of prognostic salivary methylation markers for OSCC and the existing data is limited by lack of longitudinal post-operative sampling. Carvalho et al (Carvalho, Henrique et al. 2011) analysed pre-operative saliva samples from 61 patients with primary HNSCC for the presence of methylation in a panel of 7 genes (*DAPK, DCC, MINT31, TIMP3, P16, MGMT, CCNA1*) using qMSP. After correlation with prospectively collected clinical data and multivariate statistical analysis they established that methylation of this panel of genes was an independent prognostic factor for local recurrence (Hazard ratio 12.2, 95% CI: 1.8-80) and overall survival (Hazard ratio 2.8, 95% CI: 1.2-6.5). Local disease control rate at 5 years varied from 61% in cases with pre-operative saliva methylation to 92% for patients without methylation detectable in saliva. Survival at 5 years varied from 37% with methylated pre-operative saliva to 70% in cases without methylation in saliva. In this pre-operative cohort 38/61 (62%) of patients had received PORT and 22/61 patients experienced recurrent disease. Eighty two percent of recurrent disease was diagnosed before 24 months after initial treatment with a median of 15.7 months. The authors hypothesise that aggressive tumours may have increased rates of cellular shedding and those tumours with higher epigenetic burdens would be more detectable in saliva and may be more aggressive. They also suggest clonal expansion of premalignant cells extends beyond the primary tumour and are shed into saliva which both increases detectability in saliva and also predisposes to recurrent disease (Carvalho, Henrique et al. 2011). However, this prognostic association was not found in a follow-up validation study of this panel of genes performed by the same group using a similar experimental design on a larger cohort (n=197). In this study population 60/197 patients died with 34/60 reported as disease related deaths. Thirty six of one hundred and ninety seven (18%) of patients had recurrent disease which is a lower frequency than the previous study. Methylation of *TIMP3* in pre-operative saliva had an independent association with local recurrence free survival (Hazard ratio 2.5, 95% CI: 1.10-5.68) but none of the other markers were significantly associated with survival outcomes (Sun, Zaboli et al. 2012).

Two studies have utilised methylation detection in post-operative saliva for prognostication in HNSCC. The earliest study was performed by Righini et al (Righini, De Fraipont et al. 2007) who evaluated the prognostic significance of methylation in post-operative saliva in a French cohort of HNSCC patients. Initially, they identified a panel of 6 tumour specific genes (*TIMP3, ECAD, P16, MGMT, DAPK, RASSF1*) on a pre-operative tumour (n=90) and saliva cohort (n=60). Twenty two of sixty patients had post-operative saliva sampling performed at the first staging appointment which ranged from 8-20 months after diagnosis.

Saliva was obtained at 2-6 monthly intervals from this point. 6 patients had recurrent disease and 5/6 had methylation of the gene panel detected in saliva prior to recurrence detected by clinical examination and PET scanning. The median follow-up time was 25 months (19-31 months) which according to other published data is sufficient time for 80 % of recurrences to have occurred (Kissun, Magennis et al. 2006, Carvalho, Henrique et al. 2011). The recurrence cohort is limited in numbers but the data suggests salivary methylation markers may have a role in prognostication in HNSCC. The limitations of this study are that 85/90 tumours were verrucous carcinoma and the findings may not translate to squamous cell pathology. Additionally the first post-operative sample was taken at a median time of 14 months after treatment and 14 patients had recurred and 3 died before this time so were excluded from the study and there was no early post-operative saliva sampling to detect residual disease. Eleven patients were excluded because there was no methylation detected in the tumour or pre-operative saliva. Lack of methylation data earlier in the post-operative period would be useful to determine the role of methylation in saliva in prognostication.

Using qMSP, Rettori et al (Rettori, De carvalho et al. 2013) identified 5 genes (*CCNA1, DAPK, DCC, MGMT, TIMP3*) from a larger panel (24 genes) that were HNSCC tumour specific using pre-operative HNSCC saliva (n=146) and control saliva (n=60). Eighty of one hundred and forty six (55%) pre-treatment saliva samples had detectable methylation of at least one of these five genes. This panel was used to analyse post-treatment samples which were obtained at 7-15 days after treatment (n=142) and later at 6/12 (n=105) after treatment from HNSCC patients who had received various modalities of treatment (primary surgery 17% cases, surgery and PORT 39%, chemoradiotherapy 24%, radiotherapy 10% and other 10%). As would be expected the number of post treatment cases positive for methylation decreased in the saliva obtained immediately after treatment for all the genes except *DCC*. There was no association found between methylation in pre-treatment or post-treatment saliva of any of the five genes and overall survival or local-recurrence free survival. However, detection of *TIMP3* methylation in saliva obtained at 6/12 post treatment was an independent prognostic factor for local disease control rate at 3 years; 82% unmethylated versus 52% methylated p=0.008. After multivariate analyses *TIMP3* methylation remained an independent prognostic factor for local recurrence (Hazard ratio =2.66, 95% CI: 1.1-6.5 p=0.025). *TIMP3* methylation detected in saliva has been associated with prognosis in HNSCC across three studies however, each study has limited numbers and it would need to be validated on a separate larger cohort with more frequent post-operative sampling.

### 5.11 Lack of longitudinal data precluded comment on the usefulness of saliva for the surveillance of premalignant lesions

An objective of this study was to establish if methylation of this panel of genes in the saliva of patients with premalignant disease predicted for malignant change. This study demonstrated a significant difference (Fisher's exact p=0.03) in the observed methylation frequency of the gene panel in saliva from mild and moderate dysplasia patients 4/36 (11%) and saliva from high risk premalignant lesions 7/18 (39%). This may imply a role for saliva methylation markers in disease surveillance but longitudinal data would be required to confirm this. If we are to assume that methylation has an aetiologically linked association to malignancy and not merely a resulting factor then we would expect to see methylation become present as a lesion progressed. The two case studies presented in the results section 4.6.3.2 demonstrate this theory. Of course it is as yet an unknown at what stage methylation would appear in relation to the progression of the disease. The clinical samples were not available in large enough numbers in the entire dysplasia cohort to establish this pattern. The collection of saliva was not a priority in the dysplasia clinical setting and samples were collected at variable intervals in each patient's clinical journey. Ideally the collections would be at regular intervals over a time period from first diagnosis until at least the mean time period to malignant transformation of 4.3 (range 0.5-16years) (Mehanna, Rattay et al. 2009) had elapsed. This would require a research infrastructure within the clinic to acquire sufficient data which has cost implications and is beyond the scope of an MD project. It was therefore reasonable in the first instance to perform a pilot study on the available samples. In our study the saliva samples were not always collected at the first clinical appointment which would present a limitation only if the lesion were excised prior to saliva collection. In view of this the diagnostic group they are assigned to in this study is according to the histological diagnosis they had prior to the first saliva sample and not initial diagnosis. Due to the variable nature of the saliva collection some of the saliva samples were obtained 16-21months before the lesion progressed with no further sampling. Of course the temporal relationship between onset of methylation and disease is yet to be answered. This limits the interpretation of the methylation result in the context of malignant transformation as the saliva sample may be negative because the epigenetic change had not taken place at the point of collection. Authors of a review of dysplasia literature proposed that lack of good quality longitudinal data in large numbers was one of the factors hindering the progression and clinical application of biomarker research in OED

and was likely a combination of small numbers treated in individual centres and difficulty accessing complete follow-up data (Smith, Rattay et al. 2009). Meaningful data will likely be generated by large multi-centre studies.

## 6. Limitations of this study

Accepting that this is a pilot study, several limitations were noted during the lifetime of the study. The demographic profile of the normal control cohort was not as closely matched to the OSCC cohort as was anticipated in the design of the study. This was largely a result of the type of clinic initially selected for recruitment of the normal cohort. The large number of younger individuals (under 40 year old) and those with lesser or no smoking history in this cohort weakened the statistical comparison of confounding variables between it and the disease group.

A second limitation was a lack of availability of serial samples on post-operative and dysplasia patients. There was an insufficiently robust plan for sequential saliva collections and it became clear that saliva collection was not a priority for clinicians in a busy clinical environment. However, a lack of a sufficient number of chronological samples precluded any useful conclusions to be drawn on the role of salivary markers in predicting recurrence or prognosis in disease surveillance. In addition to this the follow-up period for the dysplasia cohort was insufficient in all samples to reliably detect adequate numbers of transforming lesions. As discussed in the literature a period of at least 5 years would be required and a larger number of patients would need to be recruited to the study.

At the onset of this study a power calculation was performed for the normal and OSCC cohort but did not include the dysplasia cohort, this omission presents a limitation and would need to be addressed in future work. A further limitation of this study may be a result of selecting the panel of target markers via a candidate approach. Following analysis of the results it was felt that the chosen panel of markers was insufficiently informative. The candidate approach has been criticised for non-replication of results and limitations on identifying all possible causative genes with the 'hypothesis-driven' approach less likely to yield results than the genome wide approaches (Tabor, Risch et al. 2002). At the time of this study design resource limitations restricted the use of genome wide applications but this is something that could be considered for future studies.

## 7. Conclusion

After consideration of the evidence produced in this study the author does not think the detection of this panel of biomarkers in saliva is a good enough surrogate to replace the gold standard of biopsy for diagnosis. However, this is pilot data and the fact that the detection of methylation in saliva was able to discriminate between normal/low risk disease and high risk/malignant disease (Chi square p=0.001) and low risk from high risk premalignant lesions (Fisher's exact p=0.03) suggests saliva is not entirely without merit as a surrogate tissue. With further work, inclusive of new biomarkers, it may represent a possible adjunct to clinical examination and biopsy for disease surveillance in well-defined disease cohorts.

## 8. Further studies

Appropriately resourced research could establish if the observed lack of sensitivity in saliva in this pilot data is a result of the limitations of the current gene panel or inherent limitations of saliva as a surrogate. The long term aim would be to identify predictive methylation biomarkers that can be introduced into a trial surveillance of premalignant lesions.

In the current study the panel of studied methylation markers appears to be specific for the detection of oral cancer but has limited sensitivity which was also the observation from a large case-control study of methylation markers in saliva from HNSCC patients (Carvalho, Jeronimo et al. 2008). The majority of saliva methylation marker literature is based on the candidate gene approach for panel selection which carries the risk of missing more relevant markers. A methylation array study of saliva identified a gene panel with improved sensitivity and specificity using saliva but this panel is unvalidated (Viet, Schmidt 2008). Another array study identified a panel of genes with high sensitivity and specificity in OSCC tumour tissue but selected only two genes to analyse in saliva; both had limited specificity (Guerrero-Preston, Soudry et al. 2011).

In future discovery studies there is a need to focus on markers with high specificity to avoid additional markers degrading the specificity of the test (Carvalho, Jeronimo et al. 2008). The framework for this proposed project is based on a study on methylation biomarkers in bronchial washings for the detection of lung cancer which had a sensitivity of 82% and specificity of 91% (Nikolaidis, Raji et al. 2012). I propose a genome wide approach for the

discovery phase using next generation bisulphite sequencing to identify markers of interest in early stage OSCC, OED tissue from lesions known to undergo malignant transformation and normal tissue. Bioinformatic analysis would assist the identification of 20-30 genes with the highest differential methylation between the disease groups and normal tissue. These markers would be validated on an independent cohort of early stage OSCC, OED and normal tissue using pyrosequencing methylation assays to determine those markers with high methylation rates in the OSCC and/or OED tissue and not in the normal tissue. This data would also be used to power the saliva sample study. qMSP assays would be designed for the chosen panel of genes. These genes would be analysed in a training set of saliva samples from a retrospectively collected cohort of demographic and risk factor matched healthy controls with 5 year follow-up data; an OED cohort with 5 year follow-up data and OSCC cohort with similar follow-up. This stage allows us to select the gene panel and remove markers with low diagnostic potential. The selected targets will be analysed in a validation set of saliva samples from a different cohort of controls, OED and OSCC patients with known outcomes and 5 year follow-up. Statistical analysis would identify the panel with adequate sensitivity and specificity for a clinical study. If satisfactory markers were identified these would be tested in a prospective clinical study. This would be designed to develop predictive biomarkers in OED as there is evidence that methylation in OED lesions is predictive of malignancy (Hall, Shaw et al. 2008, Cao, Zhou et al. 2009). The current clinical and histopathological methods of predicting the lesions that will transform have limitations (Warnakulasuriya S, Reibel J et al. 2008). A recent study has shown that patients under long-term specialist review in a tertiary referral dysplasia clinic who developed malignancy were all detected at stage 1 disease and had very good disease-specific survival as a result (Ho, Field et al. 2013). However, in that study the transformation rate was 25% which is relatively high compared to other published data (Mehanna, Rattay et al. 2009) but still means that 75% of patients may not need long-term, expensively delivered clinical surveillance. If a biomarker test detectable in saliva could be developed that could reliably predict those at risk of transformation, these patients could be stratified before they reach a tertiary level clinic. It is feasible that disease monitoring could be provided in the community.

In this current study we have already established that saliva collected from the proposed target populations using the Oragene$^{TM}$ vial outperforms mouthwash and yields sufficient DNA for use in multiple qMSP assays. The Oragene$^{TM}$ vial is a one stage system that allows storage at room temperature without significant degradation of DNA (Nunes, Oliveira et al.

2012) which makes it an attractive sampling method in the clinical environment. Data from the preclinical phase of the gene panel selection would be used to power the prospective clinical study of predictive biomarkers in histologically confirmed OED patients. Longitudinal saliva sampling would be required to determine if the methylation markers were predictive of a malignant event. Based on a meta-analysis of the OED literature there is a 12% (Mehanna, Rattay et al. 2009) methylation transformation rate and mean time to transformation is 4.3 years therefore OED saliva samples would need to be acquired from multiple centre oral medicine/maxillofacial surgery departments in order to generate enough data. This means a robust infrastructure would have to be in place to identify appropriate patients, standardise sample collection and storage, the acquisition of prospective clinical data and transport of the samples to the designated laboratory to perform the qMSP assays. Furthermore there would be large amounts of data that would have to be catalogued and analysed. All of which have considerable resource implications and to feasibly complete such a study, NIHR portfolio adoption would be required.

## 9. References


AGRAWAL, N., FREDERICK, M.J., PICKERING, C.R., BETTEGOWDA, C., CHANG, K., LI, R.J., FAKHRY, C., XIE, T.-., ZHANG, J., WANG, J., ZHANG, N., EL-NAGGAR, A.K., JASSER, S.A., WEINSTEIN, J.N., TREVIÑO, L., DRUMMOND, J.A., MUZNY, D.M., WU, Y., WOOD, L.D., HRUBAN, R.H., WESTRA, W.H., KOCH, W.M., CALIFANO, J.A., GIBBS, R.A., SIDRANSKY, D., VOGELSTEIN, B., VELCULESCU, V.E., PAPADOPOULOS, N., WHEELER, D.A., KINZLER, K.W. and MYERS, J.N., 2011. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. *Science,* **333**(6046), pp. 1154-1157.

AHLQUIST, D.A., SARGENT, D.J., LOPRINZI, C.L., LEVIN, T.R., REX, D.K., AHNEN, D.J., KNIGGE, K., LANCE, M.P., BURGART, L.J., HAMILTON, S.R., ALLISON, J.E., LAWSON, M.J., DEVENS, M.E., HARRINGTON, J.J. and HILLMAN, S.L., 2008. Stool DNA and occult blood testing for screen detection of colorectal neoplasia. *Annals of Internal Medicine,* **149**(7), pp. 441-450.

AHRENS, W., POHLABELN, H., FORAITA, R., NELIS, M., LAGIOU, P., LAGIOU, A., BOUCHARDY, C., SLAMOVA, A., SCHEJBALOVA, M., MERLETTI, F., RICHIARDI, L., KJAERHEIM, K., AGUDO, A., CASTELLSAGUE, X., MACFARLANE, T.V., MACFARLANE, G.J., LEE, Y.-.A., TALAMINI, R., BARZAN, L., CANOVA, C., SIMONATO, L., THOMSON, P., MCKINNEY, P.A., MCMAHON, A.D., ZNAOR, A., HEALY, C.M., MCCARTAN, B.E., METSPALU, A., MARRON, M., HASHIBE, M., CONWAY, D.I. and BRENNAN, P., 2014. Oral health, dental care and mouthwash associated with upper aerodigestive tract cancer risk in Europe: The ARCAGE study. *Oral oncology,* **50**(6), pp. 616-625.

ATKINSON, A.J., COLBURN, W.A., DEGRUTTOLA, V.G., DEMETS, D.L., DOWNING, G.J., HOTH, D.F., OATES, J.A., PECK, C.C., SCHOOLEY, R.T., SPILKER, B.A., WOODCOCK, J. and ZEGER, S.L., 2001. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clinical pharmacology and therapeutics,* **69**(3), pp. 89-95.

BAEK, Y.H., CHANG, E., KIM, Y.J., KIM, B.K., SOHN, J.H. and PARK, D.I., 2009. Stool methylation-specific polymerase chain reaction assay for the detection of colorectal neoplasia in Korean patients. *Diseases of the colon and rectum,* **52**(8), pp. 1452-1459.

BAGNARDI, V., BLANGIARDO, M., VECCHIA, C.L. and CORRAO, G., 2001. A meta-analysis of alcohol drinking and cancer risk. *British journal of cancer,* **85**(11), pp. 1700-1705.

BALASUNDARAM, I., PAYNE, K.F.B., AL-HADAD, I., ALIBHAI, M., THOMAS, S. and BHANDARI, R., 2014. Is there any benefit in surgery for potentially malignant disorders of the oral cavity? *Journal of Oral Pathology and Medicine,* **43**(4), pp. 239-244.

BALDWIN, C., GARNIS, C., ZHANG, L., ROSIN, M.P. and LAM, W.L., 2005. Multiple microalterations detected at high frequency in oral cancer. *Cancer research,* **65**(17), pp. 7561-7567.

BARNES, L., 2005. *World Health Organization Classification of tumours : pathology and genetics of head and neck tumours.* Lyon : IARC Press, 2005.

BAUER, J.E., REZAISHIRAZ, H., HEAD, K., COWELL, J., BEPLER, G., AIKEN, M., CUMMINGS, K.M. and HYLAND, A., 2004. Obtaining DNA from a geographically dispersed cohort of current and former smokers: Use of mail-based mouthwash collection and monetary incentives. *Nicotine and Tobacco Research,* **6**(3), pp. 439-446.

BAXI, S.S., PINHEIRO, L.C., PATIL, S.M., PFISTER, D.G., OEFFINGER, K.C. and ELKIN, E.B., 2014. Causes of death in long-term survivors of head and neck cancer. *Cancer,* **120**(10), pp. 1507-1513.

BAYLIN, S.B. and OHM, J.E., 2006. Epigenetic gene silencing in cancer - A mechanism for early oncogenic pathway addiction? *Nature Reviews Cancer,* **6**(2), pp. 107-116.

BELINSKY, S.A., GRIMES, M.J., CASAS, E., STIDLEY, C.A., FRANKLIN, W.A., BOCKLAGE, T.J., JOHNSON, D.H. and SCHILLER, J.H., 2007. Predicting gene promoter methylation in non-small-cell lung cancer by evaluating sputum and serum. *British journal of cancer,* **96**(8), pp. 1278-1283.

BELINSKY, S.A., LIECHTY, K.C., GENTRY, F.D., WOLF, H.J., ROGERS, J., VU, K., HANEY, J., KENNEDY, T.C., HIRSCH, F.R., MILLER, Y., FRANKLIN, W.A., HERMAN, J.G., BAYLIN, S.B., BUNN, P.A. and BYERS, T., 2006. Promoter hypermethylation of multiple genes in sputum precedes lung cancer incidence in a high-risk cohort. *Cancer research,* **66**(6), pp. 3338-3344.

BELINSKY, S.A., PALMISANO, W.A., GILLILAND, F.D., CROOKS, L.A., DIVINE, K.K., WINTERS, S.A., GRIMES, M.J., HARMS, H.J., TELLEZ, C.S., SMITH, T.M., MOOTS, P.P., LECHNER, J.F., STIDLEY, C.A. and CROWELL, R.E., 2002. Aberrant promoter methylation in bronchial epithelium and sputum from current and former smokers. *Cancer research,* **62**(8), pp. 2370-2377.

BJØRGE, T., HENNIG, E.M., SKARE, G.B., SØREIDE, O. and THORESEN, S.Ø., 1995. Second primary cancers in patients with carcinomain situ of the uterine cervix. The norwegian experience 1970–1992. *International Journal of Cancer,* **62**(1), pp. 29-33.

BLOT, W.J., MCLAUGHLIN, J.K., WINN, D.M., AUSTIN, D.F., GREENBERG, R.S., PRESTON-MARTIN, S., BERNSTEIN, L., SCHOENBERG, J.B., STEMHAGEN, A. and FRAUMENI JR., J.F., 1988. Smoking and drinking in relation to oral and pharyngeal cancer. *Cancer research,* **48**(11), pp. 3282-3287.

BOFFETTA, P., COUTO, E., WICHMANN, J., FERRARI, P., TRICHOPOULOS, D., BUENO-DE-MESQUITA, H.B., VAN DUIJNHOVEN, F.J.B., BÜCHNER, F.L., KEY, T., BOEING, H., NÖTHLINGS, U., LINSEISEN, J., GONZALEZ, C.A., OVERVAD, K., NIELSEN, M.R.S., TJØNNELAND, A., OLSEN, A., CLAVEL-CHAPELON, F., BOUTRON-RUAULT, M.-., MOROIS, S., LAGIOU, P., NASKA, A., BENETOU, V., KAAKS, R., ROHRMANN, S., PANICO, S., SIERI, S., VINEIS, P., PALLI, D., VAN GILS, C.H., PEETERS, P.H., LUND, E., BRUSTAD, M., ENGESET, D., HUERTA, J.M., RODRÍGUEZ, L., SÁNCHEZ, M.-., DORRONSORO, M., BARRICARTE, A., HALLMANS, G., JOHANSSON, I., MANJER, J., SONESTEDT, E., ALLEN, N.E., BINGHAM, S., KHAW, K.-., SLIMANI, N., JENAB, M., MOUW, T., NORAT, T., RIBOLI, E. and TRICHOPOULOU, A., 2010. Fruit and vegetable intake and overall cancer risk in the european prospective investigation into cancer and nutrition (EPIC). *Journal of the National Cancer Institute,* **102**(8), pp. 529-537.

BONNER, J.A., HARARI, P.M., GIRALT, J., AZARNIA, N., SHIN, D.M., COHEN, R.B., JONES, C.U., SUR, R., RABEN, D., JASSEM, J., OVE, R., KIES, M.S., BASELGA, J., YOUSSOUFIAN, H., AMELLAL, N., ROWINSKY, E.K. and ANG, K.K., 2006. Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck. *New England Journal of Medicine,* **354**(6), pp. 567-578.

BOYLE, J.O., MAO, L., BRENNAN, J.A., KOCH, W.M., EISELE, D.W., SAUNDERS, J.R. and SIDRANSKY, D., 1994. Gene mutations in saliva as molecular markers for head and neck squamous cell carcinomas. *American Journal of Surgery,* **168**(5), pp. 429-432.

BRAAKHUIS, B.J.M., TABOR, M.P., KUMMER, J.A., LEEMANS, C.R. and BRAKENHOFF, R.H., 2003. A genetic explanation of slaughter's concept of field cancerization: Evidence and clinical implications. *Cancer research,* **63**(8), pp. 1727-1730.

BRAAKHUIS, B.J.M., TABOR, M.P., LEEMANS, C.R., VAN DER WAAL, I., SNOW, G.B. and BRAKENHOFF, R.H., 2002. Second primary tumors and field cancerization in oral and oropharyngeal cancer: Molecular techniques provide new insights and definitions. *Head and Neck,* **24**(2), pp. 198-206.

BRENNAN, J.A., BOYLE, J.O., KOCH, W.M., GOODMAN, S.N., HRUBAN, R.H., EBY, Y.J., COUCH, M.J., FORASTIERE, A.A. and SIDRANSKY, D., 1995. Association between cigarette smoking and mutation of the p53 gene in squamous-cell carcinoma of the head and neck. *New England Journal of Medicine,* **332**(11), pp. 712-717.

BRENNAN, J.A., MAO, L., HRUBAN, R.H., BOYLE, J.O., EBY, Y.J., KOCH, W.M., GOODMAN, S.N. and SIDRANSKY, D., 1995. Molecular assessment of histopathological staging in squamous-cell carcinoma of the head and neck. *New England Journal of Medicine,* **332**(7), pp. 429-435.

BRENNAN, M., MIGLIORATI, C.A., LOCKHART, P.B., WRAY, D., AL-HASHIMI, I., AXÉLL, T., BRUCE, A.J., CARPENTER, W., EISENBERG, E., EPSTEIN, J.B., HOLMSTRUP, P., JONTELL, M., NAIR, R., SASSER, H., SCHIFTER, M., SILVERMAN, (.).B., THONGPRASOM, K., THORNHILL, M., WARNAKULASURIYA, S. and VAN DER WAAL, I., 2007. Management of oral epithelial dysplasia: a review. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology,* **103**(SUPPL.), pp. S19.e1-S19.e12.

BROCKLEHURST, P., KUJAN, O., O'MALLEY, L.A., OGDEN, G., SHEPHERD, S. and GLENNY, A.M., 2013. Screening programmes for the early detection and prevention of oral cancer. *The Cochrane database of systematic reviews,* **11**.

BRYZGUNOVA, O.E., MOROZKIN, E.S., YARMOSCHUK, S.V., VLASSOV, V.V. and LAKTIONOV, P.P., 2008. Methylation-specific sequencing of GSTP1 gene promoter in circulating/extracellular DNA from blood and urine of healthy donors and prostate cancer patients. *Annals of the New York Academy of Sciences*, 1137, pp. 222-225

BURMESTER, T., HABERKAMP, M., MITZ, S., ROESNER, A., SCHMIDT, M., EBNER, B., GERLACH, F., FUCHS, C. and HANKELN, T., 2004. Neuroglobin and cytoglobin: Genes, proteins and evolution. *IUBMB life,* **56**(11-12), pp. 703-707.

CALIFANO, J., VAN DER RIET, P., WESTRA, W., NAWROZ, H., CLAYMAN, G., PIANTADOSI, S., CORIO, R., LEE, D., GREENBERG, B., KOCH, W. and SIDRANSKY, D., 1996. Genetic progression model for head and neck cancer: Implications for field cancerization. *Cancer research,* **56**(11), pp. 2488-2492.

CAMPO-TRAPERO, J., CANO-SÁNCHEZ, J., PALACIOS-SÁNCHEZ, B., SÁNCHEZ-GUTIERREZ, J., GONZÁLEZ-MOLES, M.A. and BASCONES-MARTÍNEZ, A., 2008. Update on molecular pathology in oral cancer and precancer. *Anticancer Research,* **28**(2 B), pp. 1197-1205.

CAO, J., ZHOU, J., GAO, Y., GU, L., MENG, H., LIU, H. and DENG, D., 2009. Methylation of p16 CpG island associated with malignant progression of oral epithelial dysplasia: A prospective cohort study. *Clinical Cancer Research,* **15**(16), pp. 5178-5183.

CARTER, L.M. and OGDEN, G.R., 2007. Oral cancer awareness of general medical and general dental practitioners. *British dental journal,* **203**(5),pp. 248-249.

CARVALHO, A.L., HENRIQUE, R., JERONIMO, C., NAYAK, C.S., REDDY, A.N., HOQUE, M.O., CHANG, S., BRAIT, M., JIANG, W.-., KIM, M.M., CLAYBOURNE, Q., GOLDENBERG, D., KHAN, Z., KHAN, T., WESTRA, W.H., SIDRANSKY, D., KOCH, W. and CALIFANO, J.A., 2011. Detection of promoter hypermethylation in salivary rinses as a biomarker for head and neck squamous cell carcinoma surveillance. *Clinical Cancer Research,* **17**(14), pp. 4782-4789.

CARVALHO, A.L., JERONIMO, C., KIM, M.M., HENRIQUE, R., ZHANG, Z., HOQUE, M.O., CHANG, S., BRAIT, M., NAYAK, C.S., JIANG, W.-., CLAYBOURNE, Q., TOKUMARU, Y., LEE, J., GOLDENBERG, D., GARRETT-MAYER, E., GOODMAN, S., MOON, C.-., KOCH, W., WESTRA, W.H., SIDRANSKY, D. and CALIFANO, J.A., 2008. Evaluation of promoter hypermethylation detection in body fluids as a screening/diagnosis tool for head and neck squamous cell carcinoma. *Clinical Cancer Research,* **14**(1), pp. 97-107.

CEDAR, H., 1988. DNA methylation and gene activity. *Cell,* **53**(1), pp. 3-4.

CHANG, H.W., CHAN, A., KWONG, D.L.W., WEI, W.I., SHAM, J.S.T. and YUEN, A.P.W., 2003a. Detection of hypermethylated RIZ1 gene in primary tumor, mouth, and throat rinsing fluid, nasopharyngeal swab, and peripheral blood of nasopharyngeal carcinoma patient. *Clinical Cancer Research,* **9**(3), pp. 1033-1038.

CHANG, H.W., CHAN, A., KWONG, D.L.W., WEI, W.I., SUAM, J.S.T. and YUEN, A.P.W., 2003b. Evaluation of hypermethylated tumor suppressor genes as tumor markers in mouth and throat rinsing fluid, nasopharyngeal swab and peripheral blood of nasopharygeal carcinoma patient. *International Journal of Cancer,* **105**(6), pp. 851-855.

CHEN, W.-., HAN, Z.J., SKOLETSKY, J., OLSON, J., SAH, J., MYEROFF, L., PLATZER, P., LU, S., DAWSON, D., WILLIS, J., PRETLOW, T.P., LUTTERBAUGH, J., KASTURI, L., WILLSON, J.K.V., RAO, J.S., SHUBER, A. and MARKOWITZ, S.D., 2005. Detection in fecal DNA of colon cancer-specific methylation of the nonexpressed vimentin gene. *Journal of the National Cancer Institute,* **97**(15), pp. 1124-1132.

CHRISTENSEN, B.C., HOUSEMAN, E.A., MARSIT, C.J., ZHENG, S., WRENSCH, M.R., WIEMELS, J.L., NELSON, H.H., KARAGAS, M.R., PADBURY, J.F., BUENO, R., SUGARBAKER, D.J., YEH, R.-., WIENCKE, J.K. and KELSEY, K.T., 2009. Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CPG island context. *PLoS Genetics,* **5**(8)

CHUNG, C.H., PARKER, J.S., KARACA, G., WU, J., FUNKHOUSER, W.K., MOORE, D., BUTTERFOSS, D., XIANG, D., ZANATION, A., YIN, X., SHOCKLEY, W.W., WEISSLER, M.C., DRESSLER, L.G., SHORES, C.G., YARBROUGH, W.G. and PEROU, C.M., 2004. Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression. *Cancer Cell,* **5**(5), pp. 489-500.

CLARK, M.E., KELNER, G.S., TURBEVILLE, L.A., BOYER, A., ARDEN, K.C. and MAKI, R.A., 2000. ADAMTS9, a novel member of the ADAM-TS/metallospondin gene family. *Genomics,* **67**(3), pp. 343-350.

CONWAY, D.I., BREWSTER, D.H., MCKINNEY, P.A., STARK, J., MCMAHON, A.D. and MACPHERSON, L.M.D., 2007. Widening socio-economic inequalities in oral cancer incidence in Scotland, 1976-2002. *British journal of cancer,* **96**(5), pp. 818-820.

COTTRELL, S.E., 2004. Molecular diagnostic applications of DNA methylation technology. *Clinical biochemistry,* **37**(7), pp. 595-604.

COZIER, Y.C., PALMER, J.R. and ROSENBERG, L., 2004. Comparison of methods for collection of DNA samples by mail in the black women's health study. *Annals of Epidemiology,* **14**(2), pp. 117-122.

CRUK, 14 January 2014, 2014a-last update, Cancer incidence by age [Homepage of CRUK], [Online]. Available: http://www.cancerresearchuk.org/cancer-info/cancerstats/incidence/age/#Cancer [October,31, 2014].

CRUK, 21 May 2014, 2014b-last update, Oral cancer statistics [Homepage of Cancer Research UK], [Online]. Available: http://www.cancerresearchuk.org/cancer-info/cancerstats/types/oral/ [October 31, 2014].

CUFFARI, L., TESSEROLI DE SIQUEIRA, J.T., NEMR, K., RAPAPORT, A., CUFFARI, L., SIQUEIRA, JOSÉ TADEU TESSEROLI DE, NEMR, K. and RAPAPORT, A., 2006. Pain complaint as the first symptom of oral cancer: A descriptive study. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology,* **102**(1), pp. 56-61.

DAWES, C., 2003. Estimates, from salivary analyses, of the turnover time of the oral mucosal epithelium in humans and the number of bacteria in an edentulous mouth. *Archives of Oral Biology,* **48**(5), pp. 329-336.

DE BUSTROS, A., NELKIN, B.D., SILVERMAN, A., EHRLICH, G., POIESZ, B. and BAYLIN, S.B., 1988. The short arm of chromosome 11 is a 'hot spot' for hypermethylation in human neoplasia. *Proceedings of the National Academy of Sciences of the United States of America,* **85**(15), pp. 5693-5697.

DE CACERES, I.I., BATTAGLI, C., ESTELLER, M., HERMAN, J.G., DULAIMI, E., EDELSON, M.I., BERGMAN, C., EHYA, H., EISENBERG, B.L. and CAIRNS, P., 2004. Tumor cell-specific BRCA1 and RASSF1A hypermethylation in serum, plasma, and peritoneal fluid from ovarian cancer patients. *Cancer research,* **64**(18), pp. 6476-6481.

DE CAMARGO CANCELA, M., VOTI, L., GUERRA-YI, M., CHAPUIS, F., MAZUIR, M. and CURADO, M.P., 2010. Oral cavity cancer in developed and in developing countries: Population-based incidence. *Head and Neck,* **32**(3), pp. 357-367.

DEMOKAN, S., CHANG, X., CHUANG, A., MYDLARZ, W.K., KAUR, J., HUANG, P., KHAN, Z., KHAN, T., OSTROW, K.L., BRAIT, M., HOQUE, M.O., LIEGEOIS, N.J., SIDRANSKY, D., KOCH, W. and CALIFANO, J.A., 2010. KIF1A and EDNRB are differentially methylated in primary HNSCC and salivary rinses. *International Journal of Cancer,* **127**(10), pp. 2351-2359.

DEMOKAN, S. and DALAY, N., 2011. Role of DNA methylation in head and neck cancer. *Clinical Epigenetics,* **2**(2), pp. 123-150.

DESTRO, A., BIANCHI, P., ALLOISIO, M., LAGHI, L., DI GIOIA, S., MALESCI, A., CARIBONI, U., GRIBAUDI, G., BULFAMANTE, G., MARCHETTI, A., BOSARI, S., INFANTE, M., RAVASI, G. and RONCALLI, M., 2004. K-ras and p16INK4Aalterations in sputum of NSCLC patients and in heavy asymptomatic chronic smokers. *Lung Cancer,* **44**(1), pp. 23-32.

DHANDA, J., TRIANTAFYLLOU, A., LILOGLOU, T., KALIRAI, H., LLOYD, B., HANLON, R., SHAW, R.J., SIBSON, D.R. and RISK, J.M., 2014. SERPINE1 and SMA expression at the invasive front predict extracapsular spread and survival in oral squamous cell carcinoma. *British journal of cancer,* **111**(11), pp. 2114-2121.

DIONNE, K.R., WARNAKULASURIYA, S., ZAIN, R.B. and CHEONG, S.C., 2015. Potentially malignant disorders of the oral cavity: Current practice and future directions in the clinic and laboratory. *International Journal of Cancer,* **136**(3), pp. 503-515.

DIVARIS, K., OLSHAN, A., SMITH, J., BELL, M., WEISSLER, M., FUNKHOUSER, W. and BRADSHAW, P., 2010. Oral health and risk for head and neck squamous cell carcinoma: the Carolina Head and Neck Cancer Study. *Cancer Causes & Control,* **21**(4), pp. 567-575.

DNA GENOTEK INC, 2015a-last update,
Laboratory protocol for manual purification of DNA from 0.5 mL of sample [Homepage of DNA Genotek Inc, Ottawa], [Online]. Available: http://www.dnagenotek.com/us/pdf/PD-PR-006 [April/24, 2015].

DNA GENOTEK INC, 2015b-last update, Laboratory protocol for manual purification of DNA from whole sample [Homepage of DNA Genotek Inc, Ottawa], [Online]. Available: http://www.dnagenotek.com/US/pdf/PD-PR-015.pdf [April/24, 2015].

DNA GENOTEK INC, 2012-last update,
Oragene®•DNA Self-Collection Kit User Instructions (OG-250 Disc Format) [Homepage of DNA Genotek Inc, Ottawa], [Online]. Available: http://www.dnagenotek.com/ROW/pdf/PD-BR-014.pdf [April/24, 2015].

DNA GENOTEK INC, 2011-last update,
Using Saliva Sponges to Collect DNA Samples from Infants & Young Children - CS1with the OG-25 [Homepage of DNA Genotek Inc, Ottawa], [Online]. Available: http://www.dnagenotek.com/US/pdf/PD-PR-018.pdf [April/24, 2015].

DOWNER MC, MOLES DR, PALMER S and SPEIGHT PM, 2006. A systematic review of measures of effectiveness in screening for oral cancer and precancer. *Oral oncology; Oral oncology,* **42**(6), pp. 551-560.

DOWNER, M.C., MOLES, D.R., PALMER, S. and SPEIGHT, P.M., 2004. A systematic review of test performance in screening for oral cancer and precancer. *Oral oncology,* **40**(3), pp. 264-273.

DU, X., SQUIER, C.A., KREMER, M.J. and WERTZ, P.W., 2000. Penetration of N-nitrosonornicotine (NNN) across oral mucosa in the presence of ethanol and nicotine. *Journal of Oral Pathology and Medicine,* **29**(2), pp. 80-85.

EADS, C.A., DANENBERG, K.D., KAWAKAMI, K., SALTZ, L.B., BLAKE, C., SHIBATA, D., DANENBERG, P.V. and LAIRD, P.W., 2000. MethyLight: a high-throughput assay to measure DNA methylation. *Nucleic acids research,* **28**(8), e32.

ELAHI, A., ZHANG, L., YEATMAN, T.J., GERY, S., SEBTI, S. and SHIBATA, D., 2008. HPP1-mediated tumor suppression requires activation of STAT1 pathways. *International Journal of Cancer,* **122**(7), pp. 1567-1572.

ELLIOTT, G.O., JOHNSON, I.T., SCARLL, J., DAINTY, J., WILLIAMS, E.A., GARG, D., COUPE, A., BRADBURN, D.M., MATHERS, J.C. and BELSHAW, N.J., 2013. Quantitative profiling of CpG island methylation in human stool for colorectal cancer detection. *International journal of colorectal disease,* **28**(1), pp. 35-42.

EL-NAGGAR, A.K., LAI, S., CLAYMAN, G., LEE, J.-J., LUNA, M.A., GOEPFERT, H. and BATSAKIS, J.G., 1997. Methylation, a major mechanism of p16/CDKN2 gene inactivation in head and neck squamous carcinoma. *American Journal of Pathology,* **151**(6), pp. 1767-1774.

EL-NAGGAR, A.K., MAO, L., STAERKEL, G., COOMBES, M.M., TUCKER, S.L., LUNA, M.A., CLAYMAN, G.L., LIPPMAN, S. and GOEPFERT, H., 2001. Genetic heterogeneity in saliva from patients with oral squamous carcinomas. *Journal of Molecular Diagnostics,* **3**(4), pp. 164-170.

ESTELLER, M., 2008. Molecular origins of cancer: Epigenetics in cancer. *New England Journal of Medicine,* **358**(11), pp. 1148-1159.

ESTELLER, M., 2007. Cancer epigenomics: DNA methylomes and histone-modification maps. *Nature Reviews Genetics,* **8**(4), pp. 286-298.

ESTELLER, M., CORN, P.G., BAYLIN, S.B. and HERMAN, J.G., 2001. A gene hypermethylation profile of human cancer. *Cancer research,* **61**(8), pp. 3225-3229.

FACKLER, M.J., MALONE, K., ZHANG, Z., SCHILLING, E., GARRETT-MAYER, E., SWIFT-SCANLAN, T., LANGE, J., NAYAR, R., DAVIDSON, N.E., KHAN, S.A. and SUKUMAR, S., 2006. Quantitative multiplex methylation-specific PCR analysis doubles detection of tumor cells in breast ductal fluid. *Clinical Cancer Research,* **12**(11 I), pp. 3306-3310.

FAN, C.-., 2004. Epigenetic alterations in head and neck cancer: Prevalence, clinical significance, and implications. *Current oncology reports,* **6**(2), pp. 152-161.

FARHADIEH, R.D., SMEE, R., REES, C.G.G., SALARDINI, A., EGGLETON, S., YANG, J.-. and RUSSELL, P.J., 2009. Mutant p53 and cyclin A1 protein expression in primary laryngeal squamous cell carcinomas do not correlate to second primary tumours of the head and neck. *ANZ Journal of Surgery,* **79**(1-2), pp. 48-54.

FEIGELSON, H.S., RODRIGUEZ, C., ROBERTSON, A.S., JACOBS, E.J., CALLE, E.E., REID, Y.A. and THUN, M.J., 2001. Determinants of DNA yield and quality from buccal cell samples collected with mouthwash. *Cancer Epidemiology Biomarkers and Prevention,* **10**(9), pp. 1005-1008.

FERLAY, J., SOERJOMATARAM, I., ERVIK, M., DIKSHIT, R., ESER, S., MATHERS, C., REBELO, M., PARKIN, D., FORMAN, D. and BRAY, F., 2012-last update, GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet] [Homepage of Lyon, France: International Agency for Research on Cancer; 2013], [Online]. Available: http://globocan.iarc.fr [Februar/2, 2015].

FERRIS, R.L., STEFANIKA, P., XI, L., GOODING, W., SEETHALA, R.R. and GODFREY, T.E., 2012. Rapid molecular detection of metastatic head and neck squamous cell carcinoma as an intraoperative adjunct to sentinel lymph node biopsy. *Laryngoscope,* **122**(5), pp. 1020-1030.

FIEGL, H., JONES, A., HAUSER-KRONBERGER, C., HUTAREW, G., REITSAMER, R., JONES, R.L., DOWSETT, M., MUELLER-HOLZNER, E., WINDBICHLER, G., DAXENBICHLER, G., GOEBEL, G., ENSINGER, C., JACOBS, I. and WIDSCHWENDTER, M., 2008. Methylated NEUROD1 promoter is a marker for chemosensitivity in breast cancer. *Clinical Cancer Research,* **14**(11), pp. 3494-3502.

FIELD, J.K., KIARIS, H., RISK, J.M., TSIRIYOTIS, C., ADAMSON, R., ZOUMPOURLIS, V., ROWLEY, H., TAYLOR, K., WHITTAKER, J., HOWARDS, P., BEIRNE, J.C., GOSNEY, J.R., WOOLGAR, J., VAUGHAN, E.D., SPANDIDOS, D.A. and JONES, A., 1995. Allelotype of squamous cell carcinoma of the head and neck: Fractional allele loss correlates with survival. *British journal of cancer,* **72**(5), pp. 1180-1188.

FIELD, J.K., NEVILLE, E.M., STEWART, M.P., SWIFT, A., LILOGLOU, T., RISK, J.M., ROSS, H., GOSNEY, J.R. and DONNELLY, R.J., 1996. Fractional allele loss data indicate distinct genetic populations in the development of non-small-cell lung cancer. *British journal of cancer,* **74**(12), pp. 1968-1974.

FLESKENS, S. and SLOOTWEG, P., 2009. Grading systems in head and neck dysplasia: their prognostic value, weaknesses and utility. *Head & neck oncology,* **1**, pp. 11.

FRAGA, M.F., AGRELO, R. and ESTELLER, M., 2007. *Cross-talk between aging and cancer: The epigenetic language.*

FRANCESCHI, S., FRANCESCHI, S., LEVI, F., LUCCHINI, F., NEGRI, E., BOYLE, P. and LA VECCHIA, C., 1994. Trends in cancer mortality in young adults in Europe, 1955–1989. *European journal of cancer,* **30**(14), pp. 2096-2118.

FRANCESCHI, S., FAVERO, A., CONTI, E., TALAMINI, R., VOLPE, R., NEGRI, E., BARZAN, L. and LA VECCHIA, C., 1999. Food groups, oils and butter, and cancer of the oral cavity and pharynx. *British journal of cancer,* **80**(3-4), pp. 614-620.

FRANZMANN, E.J., REATEGUI, E.P., PEDROSO, F., PERNAS, F.G., KARAKULLUKCU, B.M., CARRAWAY, K.L., HAMILTON, K., SINGAL, R. and GOODWIN, W.J., 2007. Soluble CD44 is a potential marker for the early detection of head and neck cancer. *Cancer Epidemiology Biomarkers and Prevention,* **16**(7), pp. 1348-1355.

FREEMAN, B., POWELL, J., BALL, D., HILL, L., CRAIG, I. and PLOMIN, R., 1997. DNA by mail: An inexpensive and noninvasive method for collecting DNA samples from widely dispersed populations. *Behavior genetics,* **27**(3), pp. 251-258.

GARBIS, S., LUBEC, G. and FOUNTOULAKIS, M., 2005. Limitations of current proteomics technologies. *Journal of Chromatography A,* **1077**(1), pp. 1-8.

GARCÍA-CLOSAS, M., EGAN, K.M., ABRUZZO, J., NEWCOMB, P.A., TITUS-ERNSTOFF, L., FRANKLIN, T., BENDER, P.K., BECK, J.C., LE MARCHAND, L., LUM, A., ALAVANJA, M., HAYES, R.B., RUTTER, J., BUETOW, K., BRINTON, L.A. and ROTHMAN, N., 2001. Collection of genomic DNA from adults in epidemiological studies by buccal cytobrush and mouthwash. *Cancer Epidemiology Biomarkers and Prevention,* **10**(6), pp. 687-696.

GARNIS, C., COE, B.P., ZHANG, L., ROSIN, M.P. and LAM, W.L., 2004. Overexpression of LRP12, a gene contained within an 8q22 amplicon identified by high-resolution array CGH analysis of oral squamous cell carcinomas. *Oncogene,* **23**(14), pp. 2582-2586.

GASCHE, J.A. and GOEL, A., 2012. Epigenetic mechanisms in oral carcinogenesis. *Future Oncology,* **8**(11), pp. 1407-1425.

GERY, S., SAWYERS, C.L., AGUS, D.B., SAID, J.W. and KOEFFLER, H.P., 2002. TMEFF2 is an androgen-regulated gene exhibiting antiproliferative effects in prostate cancer cells. *Oncogene,* **21**(31), pp. 4739-4746.

GIANNOPOULOU, C., KAMMA, J.J. and MOMBELLI, A., 2003. Effect of inflammation, smoking and stress on gingival crevicular fluid cytokine level. *Journal of clinical periodontology,* **30**(2), pp. 145-153.

GIFFORD, G., PAUL, J., VASEY, P.A., KAYE, S.B. and BROWN, R., 2004. The acquisition of hMLH1 methylation in plasma DNA after chemotherapy predicts poor survival for ovarian cancer patients. *Clinical Cancer Research,* **10**(13), pp. 4420-4426.

GILLISON ML, KOCH WM, CAPONE RB, SPAFFORD M, WESTRA WH, WU L, ZAHURAK ML, DANIEL RW, VIGLIONE M, SYMER DE, SHAH KV, SIDRANSKY D, GILLISON, M.L., KOCH, W.M., GILLISON, M.L., KOCH, W.M., CAPONE, R.B., SPAFFORD, M., WESTRA, W.H., WU, L., ZAHURAK, M.L., DANIEL, R.W., VIGLIONE, M., SYMER, D.E., SHAH, K.V., SIDRANSKY, D. and GILLISON, M.L., 2000. Evidence for a Causal Association Between Human Papillomavirus and a Subset of Head and Neck Cancers. *Journal of the National Cancer Institute: JNCI,* **92**(9), pp. 709-720.

GLAZER, C.A., CHANG, S.S., HA, P.K. and CALIFANO, J.A., 2009. Applying the molecular biology and epigenetics of head and neck cancer in everyday clinical practice. *Oral oncology,* **45**(4-5), pp. 440-446.

GOESSEL, G., QUANTE, M., HAHN, W.C., KARADA, H., HEEG, S., SULIMAN, Y., DOEBELE, M., VON WERDER, A., FULDA, C., NAKAGAWA, H., RUSTGI, A.K., BLUM, H.E. and OPITZ, O.G., 2005. Creating oral squamous cancer cells: A cellular model of oral-esophageal carcinogenesis. *Proceedings of the National Academy of Sciences of the United States of America,* **102**(43), pp. 15599-15604.

GONZALGO, M.L., LIANG, G., SPRUCK III, C.H., ZINGG, J.-., RIDEOUT III, W.M. and JONES, P.A., 1997. Identification and characterization of differentially methylated regions of genomic DNA by methylation-sensitive arbitrarily primed PCR1. *Cancer research,* **57**(4), pp. 594-599.

GOODWIN JR., W.J., 2000. Salvage surgery for patients with recurrent squamous cell carcinoma of the upper aerodigestive tract: When do the ends justify the means? *Laryngoscope,* **110**(3 II), pp. 1-18.

GRANDIS, J.R., MELHEM, M.F., GOODING, W.E., DAY, R., HOLST, V.A., WAGENER, M.M., DRENNING, S.D. and TWEARDY, D.J., 1998. Levels of TGF-a and EGFR protein in head and neck squamous cell carcinoma and patient survival. *Journal of the National Cancer Institute,* **90**(11), pp. 824-832.

GRANDIS, J.R. and TWEARDY, D.J., 1993. Elevated levels of transforming growth factor a and epidermal growth factor receptor messenger RNA are early markers of carcinogenesis in head and neck cancer. *Cancer research,* **53**(15), pp. 3579-3584.

GRANGER, D.A., KIVLIGHAN, K.T., FORTUNATO, C., HARMON, A.G., HIBEL, L.C., SCHWARTZ, E.B. and WHEMBOLUA, G.-., 2007. Integration of salivary biomarkers into developmental and behaviorally-oriented research: Problems and solutions for collecting specimens. *Physiology and Behavior,* **92**(4), pp. 583-590.

GREENBERG, J.S., EL NAGGAR, A.K., MO, V., ROBERTS, D. and MYERS, J.N., 2003. Disparity in pathologic and clinical lymph node staging in oral tongue carcinoma: Implications for therapeutic decision making. *Cancer,* **98**(3), pp. 508-515.

GUERRERO-PRESTON, R., SOUDRY, E., ACERO, J., ORERA, M., MORENO-LÓPEZ, L., MACÍA-COLÓN, G., JAFFE, A., BERDASCO, M., ILI-GANGAS, C., BREBI-MIEVILLE, P., FU, Y., ENGSTROM, C., IRIZARRY, R.A., ESTELLER, M., WESTRA, W., KOCH, W., CALIFANO, J. and SIDRANSKY, D., 2011. NID2 and HOXA9 promoter hypermethylation as biomarkers for prevention and early detection in oral cavity squamous cell carcinoma tissues and saliva. *Cancer Prevention Research,* **4**(7), pp. 1061-1072.

GUO, X., PHILIPSEN, S. and TAN-UN, K.-., 2007. Study of the hypoxia-dependent regulation of human CYGB gene. *Biochemical and biophysical research communications,* **364**(1), pp. 145-150.

HA, P.K. and CALIFANO, J.A., 2006. Promoter methylation and inactivation of tumour-suppressor genes in oral squamous-cell carcinoma. *The Lancet Oncology,* **7**(1), pp. 77-82.

HALL, G.L., SHAW, R.J., FIELD, E.A., ROGERS, S.N., SUTTON, D.N., WOOLGAR, J.A., LOWE, D., LILOGLOU, T., FIELD, J.K. and RISK, J.M., 2008. p16 promoter methylation is a potential predictor of malignant transformation in oral epithelial dysplasia. *Cancer Epidemiology Biomarkers and Prevention,* **17**(8), pp. 2174-2179.

HAMA, T., YUZA, Y., SAITO, Y., O-UCHI, J., KONDO, S., OKABE, M., YAMADA, H., KATO, T., MORIYAMA, H., KURIHARA, S. and URASHIMA, M., 2009. Prognostic significance of epidermal growth factor receptor phosphorylation and mutation in head and neck squamous cell carcinoma. *Oncologist,* **14**(9), pp. 900-908.

HANSEN, L.S., OLSON, J.A. and SILVERMAN JR., S., 1985. Proliferative verrucous leukoplakia. A long-term study of thirty patients. *Oral Surgery, Oral Medicine, Oral Pathology,* **60**(3), pp. 285-298.

HANSEN, T.V.O., SIMONSEN, M.K., NIELSEN, F.C. and HUNDRUP, Y.A., 2007. Collection of blood, saliva, and buccal cell samples in a pilot study on the Danish nurse cohort: Comparison of the response rate and quality of genomic DNA. *Cancer Epidemiology Biomarkers and Prevention,* **16**(10), pp. 2072-2076.

HARRIS, S.L., KIMPLE, R.J., HAYES, D.N., COUCH, M.E. and ROSENMAN, J.G., 2010. Never-smokers, never-drinkers: Unique clinical subgroup of young patients with head and neck squamous cell cancers. *Head and Neck,* **32**(4), pp. 499-503.

HEATH, E.M., MORKEN, N.W., CAMPBELL, K.A., TKACH, D., BOYD, E.A. and STROM, D.A., 2001. Use of buccal cells collected in mouthwash as a source of DNA for clinical testing. *Archives of Pathology and Laboratory Medicine,* **125**(1), pp. 127-133.

HERMAN, J.G., GRAFF, J.R., MYÖHÄNEN, S., NELKIN, B.D. and BAYLIN, S.B., 1996. Methylation-specific PCR: A novel PCR assay for methylation status of CpG islands. *Proceedings of the National Academy of Sciences of the United States of America,* **93**(18), pp. 9821-9826.

HERRERO, R., CASTELLSAGUÉ, X., PAWLITA, M., LISSOWSKA, J., KEE, F., BALARAM, P., RAJKUMAR, T., SRIDHAR, H., ROSE, B., PINTOS, J., FERNÁNDEZ, L., IDRIS, A., SANCHEZ, M.J., NIETO, A., TALAMINI, R., TAVANI, A., BOSCH, F.X., REIDEL, U., SNIJDERS, P.J.F., MEIJER, C.J.L.M., VISCIDI, R., MUÑOZ, N. and FRANCESCHI, S., 2003. Human papillomavirus and oral cancer: The international agency for research on cancer multicenter study. *Journal of the National Cancer Institute,* **95**(23), pp. 1772-1783.

HESSON, L.B., COOPER, W.N. and LATIF, F., 2007. The role of RASSF1A methylation in cancer. *Disease markers,* **23**(1-2), pp. 73-87.

HICKS JR., W.L., NORTH JR., J.H., LOREE, T.R., MAAMOUN, S., MULLINS, A., ORNER, J.B., BAKAMJIAN, V.Y. and SHEDD, D.P., 1998. Surgery as a single modality therapy for squamous cell carcinoma of the oral tongue. *American Journal of Otolaryngology - Head and Neck Medicine and Surgery,* **19**(1), pp. 24-28.

HINDLE, I., DOWNER, M.C., MOLES, D.R. and SPEIGHT, P.M., 2000. Is alcohol responsible for more intra-oral cancer? *Oral oncology,* **36**(4), pp. 328-333.

HO, M.W., FIELD, E.A., FIELD, J.K., RISK, J.M., RAJLAWAT, B.P., ROGERS, S.N., STEELE, J.C., TRIANTAFYLLOU, A., WOOLGAR, J.A., LOWE, D. and SHAW, R.J., 2013. Outcomes of oral squamous cell carcinoma arising from oral epithelial dysplasia: Rationale for monitoring premalignant oral lesions in a multidisciplinary clinic. *British Journal of Oral and Maxillofacial Surgery,* **51**(7), pp. 594-599.

HO, M.W., RISK, J.M., WOOLGAR, J.A., FIELD, E.A., FIELD, J.K., STEELE, J.C., RAJLAWAT, B.P., TRIANTAFYLLOU, A., ROGERS, S.N., LOWE, D. and SHAW, R.J., 2012. The clinical determinants of malignant transformation in oral epithelial dysplasia. *Oral oncology,* **48**(10), pp. 969-976.

HODGES, N.J., INNOCENT, N., DHANDA, S. and GRAHAM, M., 2008. Cellular protection from oxidative DNA damage by over-expression of the novel globin cytoglobin in vitro. *Mutagenesis,* **23**(4), pp. 293-298.

HODGSON, N., FREEDMAN, V.A., GRANGER, D.A. and ERNO, A., 2004. Biobehavioral correlates of relocation in the frail elderly: Salivary cortisol, affect, and cognitive function. *Journal of the American Geriatrics Society,* **52**(11), pp. 1856-1862.

HOLMSTRUP, P., VEDTOFTE, P., REIBEL, J. and STOLTZE, K., 2007. Oral premalignant lesions: Is a biopsy reliable? *Journal of Oral Pathology and Medicine,* **36**(5), pp. 262-266.

HOLMSTRUP, P., VEDTOFTE, P., REIBEL, J. and STOLTZE, K., 2006. Long-term treatment outcome of oral premalignant lesions. *Oral oncology,* **42**(5), pp. 461-474.

HONG, L.L., LO, P.H.Y., XIE, D., APTE, S.S., CHEUNG, A.K.L., CHENG, Y., LAW, E.W.L., CHUA, D., ZENG, Y.-., SAI, W.T., STANBRIDGE, E.J. and LUNG, M.L., 2008. Characterization of a novel epigenetically-silenced, growth-suppressive gene, ADAMTS9, and its association with lymph node metastases in nasopharyngeal carcinoma. *International Journal of Cancer,* **123**(2), pp. 401-408.

HOQUE, M.O., TOPALOGLU, O., BEGUM, S., HENRIQUE, R., ROSENBAUM, E., VAN CRIEKINGE, W., WESTRA, W.H. and SIDRANSKY, D., 2005. Quantitative methylation-specific polymerase chain reaction gene patterns in urine sediment distinguish prostate cancer patients from control subjects. *Journal of Clinical Oncology,* **23**(27), pp. 6569-6575.

HSU, H.K., WENG, Y.I., HSU, P.Y., HUANG, T.H. and HUANG, Y.W., 2014. Detection of DNA methylation by MeDIP and MBDCap assays: an overview of techniques. *Methods in molecular biology (Clifton, N.J.),* **1105**, pp. 61-70.

HU, M., YAO, J., CAI, L., BACHMAN, K.E., VAN DEN BRÛLE, F., VELCULESCU, V. and POLYAK, K., 2005. Distinct epigenetic changes in the stromal cells of breast cancers. *Nature genetics,* **37**(8), pp. 899-905.

HU, S., ARELLANO, M., BOONTHEUNG, P., WANG, J., ZHOU, H., JIANG, J., ELASHOFF, D., WEI, R., LOO, J.A. and WONG, D.T., 2008. Salivary proteomics for oral cancer biomarker discovery. *Clinical Cancer Research,* **14**(19), pp. 6246-6252.

HUANG, C.-., YU, G.-., WANG, W.-., LIU, B. and SUN, Z.-., 2014. Prognostic and predictive values of SPP1, PAI and caveolin-1 in patients with oral squamous cell carcinoma. *International Journal of Clinical and Experimental Pathology,* **7**(9), pp. 6032-6039.

HUMPHREY, S.P. and WILLIAMSON, R.T., 2001. A review of saliva: Normal composition, flow, and function. *The Journal of prosthetic dentistry,* **85**(2), pp. 162-169.

ISAYEVA, T., LI, Y., MASWAHU, D. and BRANDWEIN-GENSLER, M., 2012. Human Papillomavirus in Non-Oropharyngeal Head and Neck Cancers: A Systematic Literature Review. *Head and Neck Pathology,* **6**(SUPPL. 1), pp. 104-120.

IWASIOW, R.M., TAYEB, M.A. and JAMES, C.M.P., 2011-last update, Quality, yield and bacterial content of DNA from human saliva collected and purified using Oragene®•ONE [Homepage of DNA Genotek, Ottawa], [Online]. Available: http://www.dnagenotek.com/ROW/pdf/PD-WP-013.pdf [April/24, 2015].

JANSSON, M.D. and LUND, A.H., 2012. MicroRNA and cancer. *Molecular Oncology,* **6**(6), pp. 590-610.

JESSRI, M., FARAH, C.S., JESSRI, M. and FARAH, C.S., 2014. Next generation sequencing and its application in deciphering head and neck cancer. *Oral oncology,* **50**(4), pp. 247-253.

JI, P., AGRAWAL, S., DIEDERICHS, S., BÄUMER, N., BECKER, A., CAUVET, T., KOWSKI, S., BEGER, C., WELTE, K., BERDEL, W.E., SERVE, H. and MÜLLER-TIDOW, C., 2005. Cyclin A1, the alternative A-type cyclin, contributes to G1/S cell cycle progression in somatic cells. *Oncogene,* **24**(16), pp. 2739-2744.

JITHESH, P.V., RISK, J.M., SCHACHE, A.G., DHANDA, J., LANE, B., LILOGLOU, T. and SHAW, R.J., 2013. The epigenetic landscape of oral squamous cell carcinoma. *British journal of cancer,* **108**(2), pp. 370-379.

JONES, P.A. and BAYLIN, S.B., 2007. The Epigenomics of Cancer. *Cell,* **128**(4), pp. 683-692.

JONES, P.A. and GONZALGO, M.L., 1997. Altered DNA methylation and genome instability: A new pathway to cancer? *Proceedings of the National Academy of Sciences of the United States of America,* **94**(6), pp. 2103-2105.

KANE, M.A., 2005. The role of folates in squamous cell carcinoma of the head and neck. *Cancer detection and prevention,* **29**(1), pp. 46-53.

KASAMATSU, A., UZAWA, K., USUKURA, K., KOIKE, K., NAKASHIMA, D., ISHIGAMI, T., FUSHIMI, K., OGAWARA, K., SHIIBA, M. and TANZAWA, H., 2011. Loss of heterozygosity in oral cancer. *Oral Science International,* **8**(2), pp. 37-43.

KATAKURA, A., KAMIYAMA, I., TAKANO, N., SHIBAHARA, T., MURAMATSU, T., ISHIHARA, K., TAKAGI, R. and SHOUNO, T., 2007. Comparison of salivary cytokine levels in oral cancer patients and healthy subjects. *The Bulletin of Tokyo Dental College,* **48**(4), pp. 199-203.

KATO, K., HARA, A., KUNO, T., MORI, H., YAMASHITA, T., TOIDA, M. and SHIBATA, T., 2006. Aberrant promoter hypermethylation of p16 and MGMT genes in oral squamous cell carcinomas and the surrounding normal mucosa. *Journal of cancer research and clinical oncology,* **132**(11), pp. 735-743.

KISSUN, D., MAGENNIS, P., LOWE, D., BROWN, J.S., VAUGHAN, E.D. and ROGERS, S.N., 2006. Timing and presentation of recurrent oral and oropharyngeal squamous cell carcinoma and awareness in the outpatient clinic. *British Journal of Oral and Maxillofacial Surgery,* **44**(5), pp. 371-376.

KOKA, S., BEEBE, T.J., MERRY, S.P., BEJESUS, R.S., BERLANGA, L.B., WEAVER, A.L., MONTORI, V.M. and WONG, D.T., 2008. The preferences of adult outpatients in medical or dental care settings for giving saliva, urine or blood for clinical testing. *Journal of the American Dental Association,* **139**(6), pp. 735-740.

KOO, B.-., COE, D.M., DIXON, L.J., SOMERVILLE, R.P.T., NELSON, C.M., WANG, L.W., YOUNG, M.E., LINDNER, D.J. and APTE, S.S., 2010. ADAMTS9 is a cell-autonomously acting, anti-angiogenic metalloprotease expressed by microvascular endothelial cells. *American Journal of Pathology,* **176**(3), pp. 1494-1504.

KOO, B.S., LIM, Y.C., LEE, J.S. and CHOI, E.C., 2006. Recurrence and salvage treatment of squamous cell carcinoma of the oral cavity. *Oral oncology,* **42**(8), pp. 789-794.

KOO, K., BARROWMAN, R., MCCULLOUGH, M., ISELI, T. and WIESENFELD, D., 2013. Non-smoking non-drinking elderly females: A clinically distinct subgroup of oral squamous cell carcinoma patients. *International journal of oral and maxillofacial surgery,* **42**(8), pp. 929-933.

KREIMER, A.R., CLIFFORD, G.M., BOYLE, P. and FRANCESCHI, S., 2005. Human papillomavirus types in head and neck squamous cell carcinomas worldwide: A systemic review. *Cancer Epidemiology Biomarkers and Prevention,* **14**(2), pp. 467-475.

KRESTY, L.A., MALLERY, S.R., KNOBLOCH, T.J., SONG, H., LLOYD, M., CASTO, B.C. and WEGHORST, C.M., 2002. Alterations of p16INK4a and p14ARF in patients with severe oral epithelial dysplasia. *Cancer research,* **62**(18), pp. 5295-5300.

KU, C.S., NAIDOO, N., WU, M. and SOONG, R., 2011. Studying the epigenome using next generation sequencing. *Journal of medical genetics,* **48**(11), pp. 721-730.

KU, T.K.S., NGUYEN, D.C., KARAMAN, M., GILL, P., HACIA, J.G. and CROWE, D.L., 2007. Loss of p53 expression correlates with metastatic phenotype and transcriptional profile in a new mouse model of head and neck cancer. *Molecular Cancer Research,* **5**(4), pp. 351-362.

KULASEGARAM, R., DOWNER, M.C., JULLIEN, J.A., ZAKRZEWSKA, J.M. and SPEIGHT, P.M., 1995. Case-control study of oral dysplasia and risk habits among patients of a dental hospital. *European Journal of Cancer.Part B: Oral Oncology,* **31**(4), pp. 227-231.

KULKARNI, V. and SARANATH, D., 2004. Concurrent hypermethylation of multiple regulatory genes in chewing tobacco associated oral squamous cell carcinomas and adjacent normal tissues. *Oral oncology,* **40**(2), pp. 145-153.

KWABI-ADDO, B., CHUNG, W., SHEN, L., ITTMANN, M., WHEELER, T., JELINEK, J. and ISSA, J.-.J., 2007. Age-related DNA methylation changes in normal human prostate tissues. *Clinical Cancer Research,* **13**(13), pp. 3796-3802.

LAIRD, P.W., 2010. Principles and challenges of genome-wide DNA methylation analysis. *Nature Reviews Genetics,* **11**(3), pp. 191-203.

LE MARCHAND, L., LUM-JONES, A., SALTZMAN, B., VISAYA, V., NOMURA, A.M.Y. and KOLONEL, L.N., 2001. Feasibility of collecting buccal cell DNA by mail in a cohort study. *Cancer Epidemiology Biomarkers and Prevention,* **10**(6), pp. 701-703.

LEEMANS CR, BRAAKHUIS BJ and BRAKENHOFF RH, 2011. The molecular biology of head and neck cancer. *Nature Reviews.Cancer,* **11**(1), pp. 9-22.

LEÓN, X., QUER, M., DIEZ, S., ORÚS, C., LÓPEZ-POUSA, A. and BURGUÉS, J., 1999. Second neoplasm in patients with head and neck cancer. *Head and Neck,* **21**(3), pp. 204-210.

LI, Y., ST. JOHN, M.A.R., ZHOU, X., KIM, Y., SINHA, U., JORDAN, R.C.K., EISELE, D., ABEMAYOR, E., ELASHOFF, D., PARK, N.-. and WONG, D.T., 2004. Salivary transcriptome diagnostics for oral cancer detection. *Clinical Cancer Research,* **10**(24), pp. 8442-8450.

LIANG, G., ROBERTSON, K.D., TALMADGE, C., SUMEGI, J. and JONES, P.A., 2000. The gene for a novel transmembrane protein containing epidermal growth factor and follistatin domains is frequently hypermethylated in human tumor cells. *Cancer research,* **60**(17), pp. 4907-4912.

LIAO, P., CHANG, Y., HUANG, M., TAI, K. and CHOU, M., 2000. Mutation of p53 gene codon 63 in saliva as a molecular marker for oral squamous cell carcinomas. *Oral oncology,* **36**(3), pp. 272-276.

LILOGLOU, T. and FIELD, J.K., 2010. 6 - Detection of DNA Methylation Changes in Body Fluids. *Advances in Genetics,* **71**(0), pp. 177-207.

LIN, K., PATEL, S.G., CHU, P.Y., MATSUO, J.M.S., SINGH, B., WONG, R.J., KRAUS, D.H., SHAHA, A.R., SHAH, J.P. and BOYLE, J.O., 2005. Second primary malignancy of the aerodigestive tract in patients treated for cancer of the oral cavity and larynx. *Head and Neck,* **27**(12), pp. 1042-1048.

LINGEN, M.W., KALMAR, J.R., KARRISON, T. and SPEIGHT, P.M., 2008. Critical evaluation of diagnostic aids for the detection of oral cancer. *Oral oncology,* **44**(1), pp. 10-22.

LIU, Y., ZHOU, Z.-., HE, Q.-. and JIANG, W.-., 2012. DAPK promoter hypermethylation in tissues and body fluids of oral precancer patients. *Medical Oncology,* **29**(2), pp. 729-733.

LIU, Z., NIU, Y., LI, C., YANG, Y. and GAO, C., 2012. Integrating multiple microarray datasets on oral squamous cell carcinoma to reveal dysregulated networks. *Head and Neck,* **34**(12), pp. 1789-1797.

LLEWELLYN, C.D., LINKLATER, K., BELL, J., JOHNSON, N.W. and WARNAKULASURIYA, K.A.A.S., 2003. Squamous cell carcinoma of the oral cavity in patients aged 45 years and under: A descriptive analysis of 116 cases diagnosed in the South East of England from 1990 to 1997. *Oral oncology,* **39**(2), pp. 106-114.

LO, P.H.Y., LEUNG, A.C.C., KWOK, C.Y.C., CHEUNG, W.S.Y., KO, J.M.Y., YANG, L.C., LAW, S., WANG, L.D., LI, J., STANBRIDGE, E.J., SRIVASTAVA, G., TANG, J.C.O., TSAO, S.W. and LUNG, M.L., 2007. Identification of a tumor suppressive critical region mapping to 3p14.2 in esophageal squamous cell carcinoma and studies of a candidate tumor suppressor gene, ADAMTS9. *Oncogene,* **26**(1), pp. 148-157.

LO, P.H.Y., LUNG, H.L., CHEUNG, A.K.L., APTE, S.S., CHAN, K.W., KWONG, F.M., KO, J.M.Y., CHENG, Y., LAW, S., SRIVASTAVA, G., ZABAROVSKY, E.R., TSAO, S.W., TANG, J.C.O., STANBRIDGE, E.J. and LUNG, M.L., 2010. Extracellular protease ADAMTS9 suppresses esophageal and nasopharyngeal carcinoma tumor formation by inhibiting angiogenesis. *Cancer research,* **70**(13), pp. 5567-5576.

LOFTON-DAY, C., MODEL, F., DEVOS, T., TETZNER, R., DISTLER, J., SCHUSTER, M., SONG, X., LESCHE, R., LIEBENBERG, V., EBERT, M., MOLNAR, B., GRÜTZMANN, R., PILARSKY, C. and SLEDZIEWSKI, A., 2008. DNA methylation biomarkers for blood-based colorectal cancer screening. *Clinical chemistry,* **54**(2), pp. 414-423.

LOH, M., LIEM, N., LIM, P.L., VAITHILINGAM, A., CHENG, C.L., SALTO-TELLEZ, M., YONG, W.P. and SOONG, R., 2010. Impact of sample heterogeneity on methylation analysis. *Diagnostic Molecular Pathology,* **19**(4), pp. 243-247.

LOO, J.A., YAN, W., RAMACHANDRAN, P. and WONG, D.T., 2010. Comparative human salivary and plasma proteomes. *Journal of dental research,* **89**(10), pp. 1016-1023.

LÓPEZ, M., AGUIRRE, J.M., CUEVAS, N., ANZOLA, M., VIDEGAIN, J., AGUIRREGAVIRIA, J. and MARTÍNEZ DE PANCORBO, M., 2003. Gene promoter hypermethylation in oral rinses of leukoplakia patients - A diagnostic and/or prognostic tool? *European journal of cancer,* **39**(16), pp. 2306-2309.

LUM, A. and LE MARCHAND, L., 1998. A simple mouthwash method for obtaining genomic DNA in molecular epidemiological studies. *Cancer Epidemiology Biomarkers and Prevention,* **7**(8), pp. 719-724.

MAO, L., LEE, J.S., FAN, Y.H., RO, J.Y., BATSAKIS, J.G., LIPPMAN, S., HITTELMAN, W. and HONG, W.K., 1996. Frequent microsatellite alterations at chromosomes 9p21 and 3p14 in oral premalignant lesions and their value in cancer risk assessment. *Nature medicine,* **2**(6), pp. 682-685.

MARKOPOULOU, S., NIKOLAIDIS, G. and LILOGLOU, T., 2012. DNA methylation biomarkers in biological fluids for early detection of respiratory tract cancer. *Clinical Chemistry and Laboratory Medicine,* **50**(10), pp. 1723-1731.

MARSIT, C.J., CHRISTENSEN, B.C., HOUSEMAN, E.A., KARAGAS, M.R., WRENSCH, M.R., YEH, R.-., NELSON, H.H., WIEMELS, J.I., ZHENG, S., POSNER, M.R., MCCLEAN, M.D., WIENCKE, J.K. and KELSEY, K.T., 2009. Epigenetic profiling reveals etiologically distinct patterns of DNA methylation in head and neck squamous cell carcinoma. *Carcinogenesis,* **30**(3), pp. 416-422.

MARTIN, C.L., RESHMI, S.C., RIED, T., GOTTBERG, W., WILSON, J.W., REDDY, J.K., KHANNA, P., JOHNSON, J.T., MYERS, E.N. and GOLLIN, S.M., 2008. Chromosomal imbalances in oral squamous cell carcinoma: Examination of 31 cell lines and review of the literature. *Oral oncology,* **44**(4), pp. 369-382.

MARUYA, S.-., ISSA, J.-J., WEBER, R.S., ROSENTHAL, D.I., HAVILAND, J.C., LOTAN, R. and EL-NAGGAR, A.K., 2004. Differential methylation status of tumor-associated genes in head and neck squamous carcinoma: Incidence and potential implications. *Clinical Cancer Research,* **10**(11), pp. 3825-3830.

MASSANO, J., REGATEIRO, F.S., JANUÁRIO, G. and FERREIRA, A., 2006. Oral squamous cell carcinoma: Review of prognostic and predictive factors. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology,* **102**(1), pp. 67-76.

MAXWELL, S.A. and DAVIS, G.E., 2000. Differential gene expression in p53-mediated apoptosis-resistant vs. apoptosis-sensitive tumor cell lines. *Proceedings of the National Academy of Sciences of the United States of America,* **97**(24), pp. 13009-13014.

MCGURK, M., CHAN, C., JONES, J., O'REGAN, E. and SHERRIFF, M., 2005. Delay in diagnosis and its effect on outcome in head and neck cancer. *British Journal of Oral and Maxillofacial Surgery,* **43**(4), pp. 281-284.

MCKIE, C., AHMAD, U.A., FELLOWS, S., MEIKLE, D., STAFFORD, F.W., THOMSON, P.J., WELCH, A.R. and PALERI, V., 2008. The 2-week rule for suspected head and neck cancer in the United Kingdom: Referral patterns, diagnostic efficacy of the guidelines and compliance. *Oral oncology,* **44**(9), pp. 851-856.

MCRONALD, F.E., LILOGLOU, T., XINARIANOS, G., HILL, L., ROWBOTTOM, L., LANGAN, J.E., ELLIS, A., SHAW, J.M., FIELD, J.K. and RISK, J.M., 2006. Down-regulation of the cytoglobin gene, located on 17q25, in tylosis with oesophageal cancer (TOC): Evidence for trans-allele repression. *Human molecular genetics,* **15**(8), pp. 1271-1277.

MEHANNA, H.M., RATTAY, T., SMITH, J. and MCCONKEY, C.C., 2009. Treatment and follow-up of oral dysplasia - A systematic review and meta-analysis. *Head and Neck,* **31**(12), pp. 1600-1609.

MEIKRANTZ, W., GISSELBRECHT, S., TAM, S.W. and SCHLEGEL, R., 1994. Activation of cyclin A-dependent protein kinases during apoptosis. *Proceedings of the National Academy of Sciences of the United States of America,* **91**(9), pp. 3754-3758.

MEIKRANTZ, W. and SCHLEGEL, R., 1996. Suppression of apoptosis by dominant negative mutants of cyclin-dependent protein kinases. *Journal of Biological Chemistry,* **271**(17), pp. 10205-10209.

MERLO, A., HERMAN, J.G., MAO, L., LEE, D.J., GABRIELSON, E., BURGER, P.C., BAYLIN, S.B. and SIDRANSKY, D., 1995. 5' CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers. *Nature medicine,* **1**(7), pp. 686-692.

MESSADI, D.V., 2013. Diagnostic aids for detection of oral precancerous conditions. *International Journal of Oral Science,* **5**(2), pp. 59-65.

MEURMAN, J. and BASCONES-MARTINEZ, A., 2011. Are oral and dental diseases linked to cancer? *Oral diseases,* **17**(8), pp. 779-784.

MIMURA, I., NANGAKU, M., NISHI, H., INAGI, R., TANAKA, T. and FUJITA, T., 2010. Cytoglobin, a novel globin, plays an antifibrotic role in the kidney. *American Journal of Physiology - Renal Physiology,* **299**(5), pp. F1120-F1133.

MIRACCA, E.C., KOWALSKI, L.P. and NAGAI, M.A., 1999. High prevalence of p16 genetic alterations in head and neck tumours. *British journal of cancer,* **81**(4), pp. 677-683.

MISAWA, K., KANAZAWA, T., MISAWA, Y., IMAI, A., ENDO, S., HAKAMADA, K. and MINETA, H., 2011. Hypermethylation of collagen a2 (I) gene (COL1A2) is an independent predictor of survival in head and neck cancer. *Cancer Biomarkers,* **10**(3-4), pp. 135-144.

MONROE, M.M. and LAI, S.Y., 2014. Sentinel lymph node biopsy for oral cancer: Supporting evidence and recent novel developments. *Current oncology reports,* **16**(5),.

MÜCKE, T., WAGENPFEIL, S., KESTING, M.R., HÖLZLE, F. and WOLFF, K.-., 2009. Recurrence interval affects survival after local relapse of oral cancer. *Oral oncology,* **45**(8), pp. 687-691.

MULKI, S., SHETTY, P. and PAI, P., 2013. Oral rinse as a simpler approach to exfoliative cytology: A comparative study. *Journal of Clinical and Diagnostic Research,* **7**(12), pp. 3036-3038.

MÜLLER-TIDOW, C., JI, P., DIEDERICHS, S., POTRATZ, J., BÄUMER, N., KÖHLER, G., CAUVET, T., CHOUDARY, C., VAN DER MEER, T., CHAN, W.-I., NIEDUSZYNSKI, C., COLLEDGE, W.H., CARRINGTON, M., KOEFFLER, H.P., RESTLE, A., WIESMÜLLER, L., SOBCZAK-THÉPOT, J., BERDEL, W.E. and SERVE, H., 2004. The cyclin A1-CDK2 complex regulates DNA double-strand break repair. *Molecular and cellular biology,* **24**(20), pp. 8917-8928.

MUNSON, K., CLARK, J., LAMPARSKA-KUPSIK, K. and SMITH, S.S., 2007. Recovery of bisulfite-converted genomic sequences in the methylation-sensitive QPCR. *Nucleic acids research,* **35**(9), pp. 2893-2903.

NAGATA, M., FUJITA, H., IDA, H., HOSHINA, H., INOUE, T., SEKI, Y., OHNISHI, M., OHYAMA, T., SHINGAKI, S., KAJI, M., SAKU, T. and TAKAGII, R., 2003. Identification of potential biomarkers of lymph node metastasis in oral squamous cell carcinoma by cDNA microarray analysis. *International Journal of Cancer,* **106**(5), pp. 683-689.

NAGATA, S., HAMADA, T., YAMADA, N., YOKOYAMA, S., KITAMOTO, S., KANMURA, Y., NOMURA, M., KAMIKAWA, Y., YONEZAWA, S. and SUGIHARA, K., 2012. Aberrant DNA methylation of tumor-related genes in oral rinse. *Cancer,* **118**(17), pp. 4298-4308.

NAIR, S.S., COOLEN, M.W., STIRZAKER, C., SONG, J.Z., STATHAM, A.L., STRBENAC, D., ROBINSON, M.D. and CLARK, S.J., 2011. Comparison of methyl-DNA immunoprecipitation (MeDIP) and methyl-CpG binding domain (MBD) protein capture for genome-wide DNA methylation analysis reveal CpG sequence coverage bias. *Epigenetics,* **6**(1), pp. 34-44.

NAKAMAKI, T., HAMANO, Y., HISATAKE, J.-., YOKOYAMA, A., KAWAKAMI, K.-., TOMOYASU, S., HONMA, Y. and KOEFFLER, P., 2003. Elevated levels of cyclin A1 and A (A2) mRNA in acute myeloid leukaemia are associated with increased survival. *British journal of haematology,* **123**(1), pp. 72-80.

NAPIER, S.S. and SPEIGHT, P.M., 2008. Natural history of potentially malignant oral lesions and conditions: An overview of the literature. *Journal of Oral Pathology and Medicine,* **37**(1), pp. 1-10.

NEGRINI S, GORGOULIS VG, HALAZONETIS TD, NEGRINI S, GORGOULIS VG, HALAZONETIS TD, NEGRINI, S., GORGOULIS, V.G., HALAZONETIS, T.D., NEGRINI, S., GORGOULIS, V.G., HALAZONETIS, T.D., NEGRINI, S., GORGOULIS, V.G. and HALAZONETIS, T.D., 2010. Genomic instability — an evolving hallmark of cancer. *Nature reviews.Molecular cell biology,* **11**(3), pp. 220-228.

NG, D.P.K., KOH, D., CHOO, S.G.L., NG, V. and FU, Q., 2004. Effect of storage conditions on the extraction of PCR-quality genomic DNA from saliva. *Clinica Chimica Acta,* **343**(1-2), pp. 191-194.

NIKOLAIDIS, G., RAJI, O.Y., MARKOPOULOU, S., GOSNEY, J.R., BRYAN, J., WARBURTON, C., WALSHAW, M., SHEARD, J., FIELD, J.K. and LILOGLOU, T., 2012. DNA methylation biomarkers offer improved diagnostic efficiency in lung cancer. *Cancer research,* **72**(22), pp. 5692-5701.

NOONAN, B. and NOONAN, B., 2014. Understanding the reasons why patients delay seeking treatment for oral cancer symptoms from a primary health care professional: An integrative literature review. *European Journal of Oncology Nursing,* **18**(1), pp. 118-124.

NUNES, A.P., OLIVEIRA, I.O., SANTOS, B.R., MILLECH, C., SILVA, L.P., GONZÁLEZ, D.A., HALLAL, P.C., MENEZES, A.M.B., ARAÚJO, C.L. and BARROS, F.C., 2012. Quality of DNA extracted from saliva samples collected with the Oragene DNA self-collection kit. *BMC Medical Research Methodology,* **12**.

NUNN, J., SCHOLES, A.G.M., LILOGLOU, T., NAGINI, S., JONES, A.S., VAUGHAN, E.D., GOSNEY, J.R., ROGERS, S., FEAR, S. and FIELD, J.K., 1999. Fractional allele loss indicates distinct genetic populations in the development of squamous cell carcinoma of the head and neck (SCCHN). *Carcinogenesis,* **20**(12), pp. 2219-2228.

O'HANLON, S., FORSTER, D.R., LOWRY, R.J., O'HANLON, S., FORSTER, D.P. and LOWRY, R.J., 1997. Oral cancer in the North-East of England: incidence, mortality trends and the link with material deprivation. *Community dentistry and oral epidemiology,* **25**(5), pp. 371-376.

OHTA, S., UEMURA, H., MATSUI, Y., ISHIGURO, H., FUJINAMI, K., KONDO, K., MIYAMOTO, H., YAZAWA, T., DANENBERG, K., DANENBERG, P.V., TOHNAI, I. and KUBOTA, Y., 2009. Alterations of p16 and p14ARF genes and their 9p21 locus in oral squamous cell carcinoma. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology,* **107**(1), pp. 81-91.

OKAFUJI, M., ITA, M., OGA, A., HAYATSU, Y., MATSUO, A., SHINZATO, Y., SHINOZAKI, F. and SASAKI, K., 2000. The relationship of genetic aberrations detected by comparative genomic hybridization to DNA ploidy and tumor size in human oral squamous cell carcinomas. *Journal of Oral Pathology and Medicine,* **29**(5), pp. 226-231.

OKANO, M., XIE, S. and LI, E., 1998. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases [1]. *Nature genetics,* **19**(3), pp. 219-220.

OLDENHUIS, C.N.A.M., OOSTING, S.F., GIETEMA, J.A. and DE VRIES, E.G.E., 2008. Prognostic versus predictive value of biomarkers in oncology. *European journal of cancer,* **44**(7), pp. 946-953.

OLEKSIEWICZ, U., LILOGLOU, T., FIELD, J.K. and XINARIANOS, G., 2011. Cytoglobin: Biochemical, functional and clinical perspective of the newest member of the globin family. *Cellular and Molecular Life Sciences,* **68**(23), pp. 3869-3883.

OLEKSIEWICZ, U., LILOGLOU, T., TASOPOULOU, K.-., DASKOULIDOU, N., BRYAN, J., GOSNEY, J.R., FIELD, J.K. and XINARIANOS, G., 2013. Cytoglobin has bimodal: Tumour suppressor and oncogene functions in lung cancer cell lines. *Human molecular genetics,* **22**(16), pp. 3207-3217.

OPITZ, O.G., SULIMAN, Y., HAHN, W.C., HARADA, H., BLUM, H.E. and RUSTGI, A.K., 2001. Cyclin D1 overexpression and p53 inactivation immortalize primary oral keratinocytes by a telomerase-independent mechanism. *Journal of Clinical Investigation,* **108**(5), pp. 725-732.

O'REGAN, E.M., TONER, M.E., FINN, S.P., FAN, C.Y., RING, M., HAGMAR, B., TIMON, C., SMYTH, P., CAHILL, S., FLAVIN, R., SHEILS, O.M. and O'LEARY, J.J., 2008. p16INK4A genetic and epigenetic profiles differ in relation to age and site in head and neck squamous cell carcinomas. *Human pathology,* **39**(3), pp. 452-458.

OVCHINNIKOV, D.A., COOPER, M.A., PANDIT, P., COMAN, W.B., COOPER-WHITE, J.J., KEITH, P., WOLVETANG, E.J., SLOWEY, P.D. and PUNYADEERA, C., 2012. Tumor-suppressor Gene Promoter Hypermethylation in Saliva of Head and Neck Cancer Patients. *Translational Oncology,* **5**(5), pp. 321-326.

PALMISANO, W.A., DIVINE, K.K., SACCOMANNO, G., GILLILAND, F.D., BAYLIN, S.B., HERMAN, J.G. and BELINSKY, S.A., 2000. Predicting lung cancer by detecting aberrant promoter methylation in sputum. *Cancer research,* **60**(21), pp. 5954-5958.

PARKIN, D.M., 2006. The global health burden of infection-associated cancers in the year 2002. *International Journal of Cancer,* **118**(12), pp. 3030-3044.

PARTRIDGE, M., EMILION, G., PATEROMICHELAKIS, S., A'HERN, R., PHILLIPS, E. and LANGDON, J., 1998. Allelic imbalance at chromosomal loci implicated in the pathogenesis of oral precancer, cumulative loss and its relationship with progression to cancer. *Oral oncology,* **34**(2), pp. 77-83.

PARTRIDGE, M., KIGUWA, S. and LANGDON, J.D., 1994. Frequent deletion of chromosome 3p in oral squamous cell carcinoma. *European Journal of Cancer.Part B: Oral Oncology,* **30**(4), pp. 248-251.

PATEL, S.C., CARPENTER, W.R., TYREE, S., COUCH, M.E., WEISSLER, M., HACKMAN, T., HAYES, D.N., SHORES, C. and CHERA, B.S., 2011. Increasing incidence of oral tongue squamous cell carcinoma in young white women, age 18 to 44 years. *Journal of Clinical Oncology,* **29**(11), pp. 1488-1494.

PATMORE, H., CAWKWELL, L., STAFFORD, N., GREENMAN, J., PATMORE, H.S., CAWKWELL, L., STAFFORD, N.D. and GREENMAN, J., 2005. Unraveling the Chromosomal Aberrations of Head and Neck Squamous Cell Carcinoma: A Review. *Annals of Surgical Oncology,* **12**(10), pp. 831-842.

PATNALA, R., CLEMENTS, J. and BATRA, J., 2013. Candidate gene association studies: A comprehensive guide to useful in silico tools. *BMC Genetics,* **14**.

PATTANI, K.M., ZHANG, Z., DEMOKAN, S., GLAZER, C., LOYO, M., GOODMAN, S., SIDRANSKY, D., BERMUDEZ, F., JEAN-CHARLES, G., MCCAFFREY, T., PADHYA, T., PHELAN, J., SPIVAKOVSKY, S., BOWNE, H.Y., GOLDBERG, J.D., ROLNITZKY, L., ROBBINS, M., KERR, A.R., SIROIS, D. and CALIFANO, J.A., 2010. Endothelin receptor type B gene promoter hypermethylation in salivary rinses is independently associated with risk of oral cavity cancer and premalignancy. *Cancer Prevention Research,* **3**(9), pp. 1093-1103.

PAYNE, S.R., SERTH, J., SCHOSTAK, M., KAMRADT, J., STRAUSS, A., THELEN, P., MODEL, F., DAY, J.K., LIEBENBERG, V., MOROTTI, A., YAMAMURA, S., LOGRASSO, J., SLEDZIEWSKI, A. and SEMJONOW, A., 2009. DNA methylation biomarkers of prostate cancer: Confirmation of candidates and evidence urine is the most sensitive body fluid for non-invasive detection. *Prostate,* **69**(12), pp. 1257-1269.

PEPE, M.S., ETZIONI, R., FENG, Z., POTTER, J.D., THOMPSON, M.L., THORNQUIST, M., WINGET, M. and YASUI, Y., 2001. Phases of biomarker development for early detection of cancer. *Journal of the National Cancer Institute,* **93**(14), pp. 1054-1061.

PETRIDOU, E., ZAVRAS, A.I., LEFATZIS, D., DESSYPRIS, N., LASKARIS, G., DOKIANAKIS, G., SEGAS, J., DOUGLAS, C.W., DIEHL, S.R. and TRICHOPOULOS, D., 2002. The role of diet and specific micronutrients in the etiology of oral carcinoma. *Cancer,* **94**(11), pp. 2981-2988.

PETTI, S., 2003. Pooled estimate of world leukoplakia prevalence: A systematic review. *Oral oncology,* **39**(8), pp. 770-780.

PHILIBERT, R.A., ZADOROZHNYAYA, O., BEACH, S.R.H. and BRODY, G.H., 2008. Comparison of the genotyping results using DNA obtained from blood and saliva. *Psychiatric genetics,* **18**(6), pp. 275-281.

POETA, M.L., MANOLA, J., GOLDWASSER, M.A., FORASTIERE, A., BENOIT, N., CALIFANO, J.A., RIDGE, J.A., GOODWIN, J., KENADY, D., SAUNDERS, J., WESTRA, W., SIDRANSKY, D. and KOCH, W.M., 2007. TP53 mutations and survival in squamous-cell carcinoma of the head and neck. *New England Journal of Medicine,* **357**(25), pp. 2552-2561.

POLACHINI, G.M., SOBRAL, L.M., MERCANTE, A.M.C., PAES-LEME, A.F., XAVIER, F.C.A., HENRIQUE, T., GUIMARÃES, D.M., VIDOTTO, A., FUKUYAMA, E.E., GÓIS-FILHO, J.F., CURY, P.M., CURIONI, O.A., MICHALUART JR., P., SILVA, A.M.A., WÜNSCH-FILHO, V., NUNES, F.D., LEOPOLDINO, A.M. and TAJARA, E.H., 2012. Proteomic Approaches Identify Members of Cofilin Pathway Involved in Oral Tumorigenesis. *PLoS ONE,* **7**(12),.

QUINQUE, D., KITTLER, R., KAYSER, M., STONEKING, M. and NASIDZE, I., 2006. Evaluation of saliva as a source of human DNA for population and association studies. *Analytical Biochemistry,* **353**(2), pp. 272-277.

RANSOM, D.T., BARNETT, T.C., BOT, J., DE BOER, B., METCALF, C., DAVIDSON, J.A. and TURBETT, G.R., 1998. Loss of heterozygosity on chromosome 2q: Possibly a poor prognostic factor in head and neck cancer. *Head and Neck,* **20**(5), pp. 404-410.

RAYESS, H., WANG, M.B. and SRIVATSAN, E.S., 2012. Cellular senescence and tumor suppressor gene p16. *International Journal of Cancer,* **130**(8), pp. 1715-1725.

REED, A.L., CALIFANO, J., CAIRNS, P., WESTRA, W.H., JONES, R.M., KOCH, W., AHRENDT, S., EBY, Y., SEWELL, D., NAWROZ, H., BARTEK, J. and SIDRANSKY, D., 1996. High frequency of p16 (CDKN2/MTS-1/INK4A) inactivation in head and neck squamous cell carcinoma. *Cancer research,* **56**(16), pp. 3630-3633.

REIBEL, J., 2003. Prognosis of oral pre-malignant lesions: Significance of clinical, histopathological, and molecular biological characteristics. *Critical Reviews in Oral Biology and Medicine,* **14**(1), pp. 47-62.

RETTORI, M.M., DE CARVALHO, A.C., BOMFIM LONGO, A.L., DE OLIVEIRA, C.Z., KOWALSKI, L.P., CARVALHO, A.L. and VETTORE, A.L., 2013. Prognostic significance of TIMP3 hypermethylation in post-treatment salivary rinse from head and neck squamous cell carcinoma patients. *Carcinogenesis,* **34**(1), pp. 20-27.

RETTORI, M.M., DE CARVALHO, A.C., LONGO, A.L.B., DE OLIVEIRA, C.Z., KOWALSKI, L.P., CARVALHO, A.L. and VETTORE, A.L., 2013. TIMP3 and CCNA1 hypermethylation in HNSCC is associated with an increased incidence of second primary tumors. *Journal of Translational Medicine,* **11**(1),.

RIETBERGEN, M.M., BRAKENHOFF, R.H., BLOEMENA, E., WITTE, B.I., SNIJDERS, P.J.F., HEIDEMAN, D.A.M., BOON, D., KOLJENOVIC, S., BAATENBURG-DE JONG, R.J. and LEEMANS, C.R., 2013. Human papillomavirus detection and comorbidity: Critical issues in selection of patients with oropharyngeal cancer for treatment De-escalation trials. *Annals of Oncology,* **24**(11), pp. 2740-2745.

RIGHINI, C.A., DE FRAIPONT, F., TIMSIT, J.-., FAURE, C., BRAMBILLA, E., REYT, E. and FAVROT, M.-., 2007. Tumor-specific methylation in saliva: A promising biomarker for early detection of head and neck cancer recurrence. *Clinical Cancer Research,* **13**(4), pp. 1179-1185.

RIOS VELAZQUEZ, E., HOEBERS, F., AERTS, H.J.W.L., RIETBERGEN, M.M., BRAKENHOFF, R.H., LEEMANS, R.C., SPEEL, E.-., STRAETMANS, J., KREMER, B. and LAMBIN, P., 2014. Externally validated HPV-based prognostic nomogram for oropharyngeal carcinoma patients yields more accurate predictions than TNM staging. *Radiotherapy and Oncology,* **113**(3), pp. 324-330.

RIVERA, A., MAVILA, A., BAYLESS, K.J., DAVIS, G.E. and MAXWELL, S.A., 2006. Cyclin A1 is a p53-induced gene that mediates apoptosis, G2/M arrest, and mitotic catastrophe in renal, ovarian, and lung carcinoma cells. *Cellular and Molecular Life Sciences,* **63**(12), pp. 1425-1439.

ROCCO, J.W. and SIDRANSKY, D., 2001. p16(MTS-1/CDKN2/INK4a) in cancer progression. *Experimental cell research,* **264**(1), pp. 42-55.

ROEPMAN, P., WESSELS, L.F.A., KETTELARIJ, N., KEMMEREN, P., MILES, A.J., LIJNZAAD, P., TILANUS, M.G.J., KOOLE, R., HORDIJK, G.-., VAN DER VLIET, P.C., REINDERS, M.J.T., SLOOTWEG, P.J. and HOLSTEGE, F.C.P., 2005. An expression profile for diagnosis of lymph node metastases from primary head and neck squamous cell carcinomas. *Nature genetics,* **37**(2), pp. 182-186.

ROGERS, N.L., COLE, S.A., LAN, H.-. and CROSSA, A., 2007. New saliva DNA collection method compared to buccal cell collection techniques for epidemiological studies. *American Journal of Human Biology,* **19**(3), pp. 319-326.

ROGERS, S.J., HARRINGTON, K.J., RHYS-EVANS, P., O-CHAROENRAT, P. and ECCLES, S.A., 2005. Biological significance of c-erbB family oncogenes in head and neck cancer. *Cancer and metastasis reviews,* **24**(1), pp. 47-69.

ROGERS, S.N., BROWN, J.S., WOOLGAR, J.A., LOWE, D., MAGENNIS, P., SHAW, R.J., SUTTON, D., ERRINGTON, D. and VAUGHAN, D., 2009. Survival following primary surgery for oral cancer. *Oral oncology,* **45**(3), pp. 201-211.

ROH, J., WANG, X.V., MANOLA, J., SIDRANSKY, D., FORASTIERE, A.A. and KOCH, W.M., 2013. Clinical correlates of promoter hypermethylation of four target genes in head and neck cancer: A cooperative group correlative study. *Clinical Cancer Research,* **19**(9), pp. 2528-2540.

RONAGHI, M., KARAMOHAMED, S., PETTERSSON, B., UHLÉN, M. and NYRÉN, P., 1996. Real-Time DNA Sequencing Using Detection of Pyrophosphate Release. *Analytical Biochemistry,* **242**(1), pp. 84-89.

ROSAS, S.L.B., KOCH, W., DA COSTA CARVALHO, M.D.G., WU, L., CALIFANO, J., WESTRA, W., JEN, J. and SIDRANSKY, D., 2001. Promoter hypermethylation patterns of p16, $O^6$-methylguanine-DNA-methyltransferase, and death-associated protein kinase in tumors and saliva of head and neck cancer patients. *Cancer research,* **61**(3), pp. 939-942.

ROSENBAUM, E., HOQUE, M.O., COHEN, Y., ZAHURAK, M., EISENBERGER, M.A., EPSTEIN, J.I., PARTIN, A.W. and SIDRANSKY, D., 2005. Promoter hypermethylation as an independent prognostic factor for relapse in patients with prostate cancer following radical prostatectomy. *Clinical Cancer Research,* **11**(23), pp. 8321-8325.

ROSIN, M.P., CHENG, X., POH, C., LAM, W.L., HUANG, Y., LOVAS, J., BEREAN, K., EPSTEIN, J.B., PRIDDY, R., LE, N.D. and ZHANG, L., 2000. Use of allelic loss to predict malignant risk for low-grade oral epithelial dysplasia. *Clinical Cancer Research,* **6**(2), pp. 357-362.

ROSIN, M.P., LAM, W.L., POH, C., LE, N.D., LI, R.J., ZENG, T., PRIDDY, R. and ZHANG, L., 2002. 3p14 and 9p21 loss is a simple tool for predicting second oral malignancy at previously treated oral cancer sites. *Cancer research,* **62**(22), pp. 6447-6450.

ROSSI, M., GARAVELLO, W., TALAMINI, R., NEGRI, E., BOSETTI, C., DAL MASO, L., LAGIOU, P., TAVANI, A., POLESEL, J., BARZAN, L., RAMAZZOTTI, V., FRANCESCHI, S. and LA VECCHIA, C., 2007. Flavonoids and the risk of oral and pharyngeal cancer: A case-control study from Italy. *Cancer Epidemiology Biomarkers and Prevention,* **16**(8), pp. 1621-1625.

RUESGA, M.T., ACHA-SAGREDO, A., RODRÍGUEZ, M.J., AGUIRREGAVIRIA, J.I., VIDEGAIN, J., RODRÍGUEZ, C., MARTÍNEZ DE PANCORBO, M.D.L.A. and AGUIRRE, J.M., 2007. p16INK4a promoter hypermethylation in oral scrapings of oral squamous cell carcinoma risk patients. *Cancer letters,* **250**(1), pp. 140-145.

RYLANDER-RUDQVIST, T., HÅKANSSON, N., TYBRING, G. and WOLK, A., 2006. Quality and quantity of saliva DNA obtained from the self-administrated oragene method - A pilot study on the cohort of Swedish men. *Cancer Epidemiology Biomarkers and Prevention,* **15**(9), pp. 1742-1745.

SAILASREE, R., ABHILASH, A., SATHYAN, K.M., NALINAKUMARI, K.R., THOMAS, S. and KANNAN, S., 2008. Differential roles of p16INK4A and p14ARF genes in prognosis of oral carcinoma. *Cancer Epidemiology Biomarkers and Prevention,* **17**(2), pp. 414-420.

SARTOR, M.A., DOLINOY, D.C., JONES, T.R., COLACINO, J.A., PRINCE, M.E.P., CAREY, T.E. and ROZEK, L.S., 2011. Genome-wide methylation and expression differences in HPV(+) and HPV(-) squamous cell carcinoma cell lines are consistent with divergent mechanisms of carcinogenesis. *Epigenetics,* **6**(6), pp. 777-787.

SATIA-ABOUTA, KING, I.B., ABOUTA, J.S., THORNQUIST, M.D., BIGLER, J., PATTERSON, R.E., KRISTAL, A.R., SHATTUCK, A.L., POTTER, J.D. and WHITE, E., 2002. Buccal cell DNA yield, quality, and collection costs: Comparison of methods for large-scale studies. *Cancer Epidemiology Biomarkers and Prevention,* **11**(10 I), pp. 1130-1133.

SATO, F., SHIBATA, D., HARPAZ, N., XU, Y., YIN, J., MORI, Y., WANG, S., OLARU, A., DEACU, E., SELARU, F.M., KIMOS, M.C., HYTIROGLOU, P., YOUNG, J., LEGGETT, B., GAZDAR, A.F., TOYOOKA, S., ABRAHAM, J.M. and MELTZER, S.J., 2002. Aberrant methylation of the HPP1 gene in ulcerative colitis-associated colorectal carcinoma. *Cancer research,* **62**(23), pp. 6820-6822.

SCHAAIJ-VISSER, T.B.M., GRAVELAND, A.P., GAUCI, S., BRAAKHUIS, B.J.M., BUIJZE, M., HECK, A.J.R., KUIK, D.J., BLOEMENA, E., LEEMANS, C.R., SLIJPER, M. and BRAKENHOFF, R.H., 2009. Differential proteomics identifies protein biomarkers that predict local relapse of head and neck squamous cell carcinomas. *Clinical Cancer Research,* **15**(24), pp. 7666-7675.

SCHAAIJ-VISSER, T.B.M., BRAKENHOFF, R.H., LEEMANS, C.R., HECK, A.J.R. and SLIJPER, M., 2010. Protein biomarker discovery for head and neck cancer. *Journal of Proteomics,* **73**(10), pp. 1790-1803.

SCHACHE, A.G., HALL, G., WOOLGAR, J.A., NIKOLAIDIS, G., TRIANTAFYLLOU, A., LOWE, D., RISK, J.M., SHAW, R.J. and LILOGLOU, T., 2010. Quantitative promoter methylation differentiates carcinoma ex pleomorphic adenoma from pleomorphic salivary adenoma. *British journal of cancer,* **103**(12), pp. 1846-1851.

SCHANTZ, S.P. and YU, G.-., 2002. Head and neck cancer incidence trends in young Americans, 1973-1997, with a special analysis for tongue cancer. *Archives of Otolaryngology - Head and Neck Surgery,* **128**(3), pp. 268-274.

SCHEPMAN, K.P., VAN DER MEIJ, E.H., SMEELE, L.E. and VAN DER WAAL, I., 1998. Malignant transformation of oral leukoplakia: A follow-up study of a hospital-based population of 166 patients with oral leukoplakia from The Netherlands. *Oral oncology,* **34**(4), pp. 270-275.

SCHMIDT, M., GERLACH, F., AVIVI, A., LAUFS, T., WYSTUB, S., SIMPSON, J.C., NEVO, E., SAALER-REINHARDT, S., REUSS, S., HANKELN, T. and BURMESTER, T., 2004. Cytoglobin Is a Respiratory Protein in Connective Tissue and Neurons, Which Is Up-regulated by Hypoxia. *Journal of Biological Chemistry,* **279**(9), pp. 8063-8069.

SCHWARTZ, S.M., DALING, J.R., DOODY, D.R., WIPF, G.C., CARTER, J.J., MADELEINE, M.M., MAO, E.-., FITZGIBBONS, E.D., HUANG, S., BECKMANN, A.M., MCDOUGALL, J.K. and GALLOWAY, D.A., 1998. Oral cancer risk in relation to sexual history and evidence of human papillomavirus infection. *Journal of the National Cancer Institute,* **90**(21), pp. 1626-1636.

SEITZ, H.K. and STICKEL, F., 2007. Molecular mechanisms of alcohol-mediated carcinogenesis. *Nature Reviews Cancer,* **7**(8), pp. 599-612.

SERDULA, M.K., BYERS, T., MOKDAD, A.H., SIMOES, E., MENDLEIN, J.M. and COATES, R.J., 1996. The Association between Fruit and Vegetable Intake and Chronic Disease Risk Factors : *Epidemiology,* **7**(2), pp. 161-165.

SERRANO, M., HANNON, G.J. and BEACH, D., 1993. A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. *Nature,* **366**(6456), pp. 704-707.

SHAW, R.J., AKUFO-TETTEH, E.K., RISK, J.M., FIELD, J.K. and LILOGLOU, T., 2006. Methylation enrichment pyrosequencing: combining the specificity of MSP with validation by pyrosequencing. *Nucleic acids research.,* **34**(11),.

SHAW, R.J., HALL, G.L., WOOLGAR, J.A., LOWE, D., ROGERS, S.N., FIELD, J.K., LILOGLOU, T. and RISK, J.M., 2007. Quantitative methylation analysis of resection margins and lymph nodes in oral squamous cell carcinoma. *British Journal of Oral and Maxillofacial Surgery,* **45**(8), pp. 617-622.

SHAW, R.J., HOBKIRK, A.J., NIKOLAIDIS, G., WOOLGAR, J.A., TRIANTAFYLLOU, A., BROWN, J.S., LILOGLOU, T. and RISK, J.M., 2013. Molecular staging of surgical margins in oral squamous cell carcinoma using promoter methylation of p16INK4A, cytoglobin, E-cadherin, and TMEFF2. *Annals of Surgical Oncology,* **20**(8), pp. 2796-2802.

SHAW, R.J., LILOGLOU, T., ROGERS, S.N., BROWN, J.S., VAUGHAN, E.D., LOWE, D., FIELD, J.K. and RISK, J.M., 2006. Promoter methylation of P16, RARβ, E-cadherin, cyclin A1 and cytoglobin in oral cancer: Quantitative evaluation using pyrosequencing. *British journal of cancer,* **94**(4), pp. 561-568.

SHAW, R.J., LOWE, D., WOOLGAR, J.A., BROWN, J.S., VAUGHAN, E.D., EVANS, C., LEWIS-JONES, H., HANLON, R., HALL, G.L. and ROGERS, S.N., 2010. Extracapsular spread in oral squamous cell carcinoma. *Head and Neck,* **32**(6), pp. 714-722.

SHAW, R.J., OMAR, M.M., ROKADIYA, S., KOGERA, F.A., LOWE, D., HALL, G.L., WOOLGAR, J.A., HOMER, J., LILOGLOU, T., FIELD, J.K. and RISK, J.M., 2009. Cytoglobin is upregulated by tumour hypoxia and silenced by promoter hypermethylation in head and neck cancer. *British journal of cancer,* **101**(1), pp. 139-144.

SHAW, R.J., PACE-BALZAN, A. and BUTTERWORTH, C., 2011. Contemporary clinical management of oral squamous cell carcinoma. *Periodontology 2000,* **57**(1), pp. 89-101.

SHEN, L. and WATERLAND, R.A., 2007. Methods of DNA methylation analysis. *Current opinion in clinical nutrition and metabolic care,* **10**(5), pp. 576-581.

SHETH, S.H., JOHNSON, D.E., KENSLER, T.W. and BAUMAN, J.E., 2014. Chemoprevention targets for tobacco-related head and neck cancer: Past lessons and future directions. *Oral oncology*, **50**, pp. 77-83.

SHEU, J.J.-., HUA, C.-., WAN, L., LIN, Y.-., LAI, M.-., TSENG, H.-., JINAWATH, N., TSAI, M.-., CHANG, N.-., LIN, C.-., LIN, C.-., HSIEH, L.-., WANG, T.-., SHIH, I.-. and TSAI, F.-., 2009. Functional genomic analysis identified epidermal growth factor receptor activation as the most common genetic event in oral squamous cell carcinoma. *Cancer research,* **69**(6), pp. 2568-2576.

SHIVAPURKAR, N., STASTNY, V., OKUMURA, N., GIRARD, L., XIE, Y., PRINSEN, C., THUNNISSEN, F.B., WISTUBA, I.I., CZERNIAK, B., FRENKEL, E., ROTH, J.A., LILOGLOU, T., XINARIANOS, G., FIELD, J.K., MINNA, J.D. and GAZDAR, A.F., 2008. Cytoglobin, the newest member of the globin family, functions as a tumor suppressor gene. *Cancer research,* **68**(18), pp. 7448-7456.

SILVA, S.D.D., FERLITO, A., TAKES, R.P., BRAKENHOFF, R.H., VALENTIN, M.D., WOOLGAR, J.A., BRADFORD, C.R., RODRIGO, J.P., RINALDO, A., HIER, M.P. and KOWALSKI, L.P., 2011. Advances and applications of oral cancer basic research. *Oral oncology,* **47**(9), pp. 783-791.

SILVERMAN JR., S., GORSKY, M. and LOZADA, F., 1984. Oral leukoplakia and malignant transformation. A follow-up study of 257 patients. *Cancer,* **53**(3), pp. 563-568.

SINHA, P., BAHADUR, S., THAKAR, A., MATTA, A., MACHA, M., RALHAN, R. and GUPTA, S.D., 2009. Significance of promoter hypermethylation of p16 gene for margin assessment in carcinoma tongue. *Head and Neck,* **31**(11), pp. 1423-1430.

SLAUGHTER, D.P., SOUTHWICK, H.W. and SMEJKAL, W., 1953. Field cancerization in oral stratified squamous epithelium; clinical. *Cancer,* **6**(5), pp. 963-968.

SMIRAGLIA, D.J. and PLASS, C., 2003. *The development of CpG island methylation biomarkers using restriction landmark genomic scanning.*

SMITH, J., RATTAY, T., MCCONKEY, C., HELLIWELL, T. and MEHANNA, H., 2009. Biomarkers in dysplasia of the oral cavity: A systematic review. *Oral oncology,* **45**(8), pp. 647-653.

SOMERS, K.D., MERRICK, M.A., LOPEZ, M.E., INCOGNITO, L.S., SCHECHTER, G.L. and CASEY, G., 1992. Frequent p53 mutations in head and neck cancer. *Cancer research,* **52**(21), pp. 5997-6000.

SPAFFORD, M.F., KOCH, W.M., REED, A.L., CALIFANO, J.A., XU, L.H., EISENBERGER, C.F., YIP, L., LEONG, P.L., WU, L., LIU, S.X., JERÓNIMO, C., WESTRA, W.H. and SIDRANSKY, D., 2001. Detection of head and neck squamous cell carcinoma among exfoliated oral mucosal cells by microsatellite analysis. *Clinical Cancer Research,* **7**(3), pp. 607-612.

SPARANO, A., QUESNELLE, K.M., KUMAR, M.S., WANG, Y., SYLVESTER, A.J., FELDMAN, M., SEWELL, D.A., WEINSTEIN, G.S. and BROSE, M.S., 2006. Genome-wide profiling of oral squamous cell carcinoma by array-based comparative genomic hybridization. *Laryngoscope,* **116**(5), pp. 735-741.

SPEIGHT, P. and WARNAKULASURIYA, S., May 2010, 2010-last update, Evaluation of screening for oral cancer against NSC Criteria [Homepage of UK National Screening Committee], [Online]. Available: http://www.screening.nhs.uk/oralcancer [April/14, 2015].

ST. JOHN, M.A.R., LI, Y., ZHOU, X., DENNY, P., HO, C.-., MONTEMAGNO, C., SHI, W., QI, F., WU, B., SINHA, U., JORDAN, R., WOLINSKY, L., PARK, N.-., LIU, H., ABEMAYOR, E. and WONG, D.T.W., 2004. Interleukin 6 and interleukin 8 as potential biomarkers for oral cavity and oropharyngeal squamous cell carcinoma. *Archives of Otolaryngology - Head and Neck Surgery,* **130**(8), pp. 929-935.

STEINMANN, K., SANDNER, A., SCHAGDARSURENGIN, U., DAMMANN, R.H. and DAMMANN, R.H., 2009. Frequent promoter hypermethylation of tumor-related genes in head and neck squamous cell carcinoma. *Oncology reports,* **22**(6), pp. 1519-1526.

STRANSKY, N., EGLOFF, A.M., TWARD, A.D., KOSTIC, A.D., CIBULSKIS, K., SIVACHENKO, A., KRYUKOV, G.V., LAWRENCE, M.S., SOUGNEZ, C., MCKENNA, A., SHEFLER, E., RAMOS, A.H., STOJANOV, P., CARTER, S.L., VOET, D., CORTÉS, M.L., AUCLAIR, D., BERGER, M.F., SAKSENA, G., GUIDUCCI, C., ONOFRIO, R.C., PARKIN, M., ROMKES, M., WEISSFELD, J.L., SEETHALA, R.R., WANG, L., RANGEL-ESCAREÑO, C., FERNANDEZ-LOPEZ, J.C., HIDALGO-MIRANDA, A., MELENDEZ-ZAJGLA, J., WINCKLER, W., ARDLIE, K., GABRIEL, S.B., MEYERSON, M., LANDER, E.S., GETZ, G., GOLUB, T.R., GARRAWAY, L.A. and GRANDIS, J.R., 2011. The mutational landscape of head and neck squamous cell carcinoma. *Science,* **333**(6046), pp. 1157-1160.

STRECKFUS, C., BIGLER, L., DELLINGER, T., DAI, X., KINGMAN, A. and THIGPEN, J.T., 2000. The presence of soluble c-erbB-2 in saliva and serum among women with breast carcinoma: A preliminary study. *Clinical Cancer Research,* **6**(6), pp. 2363-2370.

STRIMBU, K. and TAVEL, J.A., 2010. What are biomarkers? *Current Opinion in HIV and AIDS,* **5**(6), pp. 463-466.

STURGIS, E.M. and MILLER, R.H., 1995. Second primary malignancies in the head and neck cancer patient. *Annals of Otology, Rhinology and Laryngology,* **104**(12), pp. 946-954.

SUN, G.-., TANG, E.-., YANG, X.-. and HU, Q.-., 2009. Salvage treatment for recurrent oral squamous cell carcinoma. *Journal of Craniofacial Surgery,* **20**(4), pp. 1093-1096.

SUN, T.-., TANG, J.-., DU, W., ZHAO, H.-., ZHAO, G., YANG, S.-., CHEN, H.-., HONG, J. and FANG, J.-., 2015. Bidirectional regulation between TMEFF2 and STAT3 may contribute to Helicobacter pylori-associated gastric carcinogenesis. *International Journal of Cancer,* **136**(5), pp. 1053-1064.

SUN, W., ZABOLI, D., WANG, H., LIU, Y., ARNAOUTAKIS, D., KHAN, T., KHAN, Z., KOCH, W.M. and CALIFANO, J.A., 2012. Detection of TIMP3 promoter hypermethylation in salivary rinse as an independent predictor of local recurrence-free survival in head and neck cancer. *Clinical Cancer Research,* **18**(4), pp. 1082-1091.

SYRJÄNEN, S., LODI, G., VON BÜLTZINGSLÖWEN, I., ALIKO, A., ARDUINO, P., CAMPISI, G., CHALLACOMBE, S., FICARRA, G., FLAITZ, C., ZHOU, H., MAEDA, H., MILLER, C. and JONTELL, M., 2011. Human papillomaviruses in oral carcinoma and oral potentially malignant disorders: A systematic review. *Oral diseases,* **17**(SUPPL. 1), pp. 58-72.

TABOR, H.K., RISCH, N.J. and MYERS, R.M., 2002. Candidate-gene approaches for studying complex genetic traits: Practical considerations. *Nature Reviews Genetics,* **3**(5), pp. 391-397.

TABOR, M.P., BRAKENHOFF, R.H., RUIJTER-SCHIPPERS, H.J., KUMMER, J.A., LEEMANS, C.R. and BRAAKHUIS, B.J.M., 2004. Genetically altered fields as origin of locally recurrent head and neck cancer: A retrospective study. *Clinical Cancer Research,* **10**(11), pp. 3607-3613.

TABOR, M.P., BRAKENHOFF, R.H., RUIJTER-SCHIPPERS, H.J., VAN DER WAL, J.E., SNOW, G.B., LEEMANS, C.R. and BRAAKHUIS, B.J.M., 2002. Multiple head and neck tumors frequently originate from a single preneoplastic lesion. *American Journal of Pathology,* **161**(3), pp. 1051-1060.

TABOR, M.P., BRAKENHOFF, R.H., VAN HOUTEN, V.M.M., KUMMER, J.A., SNEL, M.H.J., SNIJDERS, P.J.F., SNOW, G.B., LEEMANS, C.R. and BRAAKHUIS, B.J.M., 2001. Persistence of genetically altered fields in head and neck cancer patients: Biological and clinical implications. *Clinical Cancer Research,* **7**(6), pp. 1523-1532.

TAKAI, D. and 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**(6), pp. 3740-3745.

TAKES, R.P., 2004. Staging of the neck in patients with head and neck squamous cell cancer: Imaging techniques and biomarkers. *Oral oncology,* **40**(7), pp. 656-667.

TAKES, R.P., RINALDO, A., RODRIGO, J.P., DEVANEY, K.O., FAGAN, J.J. and FERLITO, A., 2008. Can biomarkers play a role in the decision about treatment of the clinically negative neck in patients with head and neck cancer? *Head and Neck,* **30**(4), pp. 525-538.

TAKES, R.P., RINALDO, A., SILVER, C.E., PICCIRILLO, J.F., HAIGENTZ JR., M., SUÁREZ, C., VAN DER POORTEN, V., HERMANS, R., RODRIGO, J.P., DEVANEY, K.O. and FERLITO, A., 2010. Future of the TNM classification and staging system in head and neck cancer. *Head and Neck,* **32**(12), pp. 1693-1711.

TÄNZER, M., BALLUFF, B., DISTLER, J., HALE, K., LEODOLTER, A., RÖCKEN, C., MOLNAR, B., SCHMID, R., LOFTON-DAY, C., SCHUSTER, T. and EBERT, M.P.A., 2010. Performance of epigenetic markers SEPT9 and ALX4 in plasma for detection of colorectal precancerous lesions. *PLoS ONE,* **5**(2)

TEZAL, M., SULLIVAN, M.A., HYLAND, A., MARSHALL, J.R., STOLER, D., REID, M.E., LOREE, T.R., RIGUAL, N.R., MERZIANU, M., HAUCK, L., LILLIS, C., WACTAWSKI-WENDE, J. and SCANNAPIECO, F.A., 2009. Chronic periodontitis and the incidence of head and neck squamous cell carcinoma. *Cancer Epidemiology Biomarkers and Prevention,* **18**(9), pp. 2406-2412.

THORNE, P., ETHERINGTON, D. and BIRCHALL, M.A., 1997. Head and neck cancer in the South West of England: influence of socio-economic status on incidence and second primary tumours. *European Journal of Surgical Oncology,* **23**(6), pp. 503-508.

TOKUMARU, Y., YAMASHITA, K., OSADA, M., NOMOTO, S., SUN, D.-., XIAO, Y., HOQUE, M.O., WESTRA, W.H., CALIFANO, J.A. and SIDRANSKY, D., 2004. Inverse correlation between cyclin A1 hypermethylation and p53 mutation in head and neck cancer identified by reversal of epigenetic silencing. *Cancer research,* **64**(17), pp. 5982-5987.

TSUI, I.F.L., ROSIN, M.P., ZHANG, L., NG, R.T. and LAM, W.L., 2008. Multiple aberrations of chromosome 3p detected in oral premalignant lesions. *Cancer Prevention Research,* **1**(6), pp. 424-429.

UCHIDA, K., OGA, A., NAKAO, M., MANO, T., MIHARA, M., KAWAUCHI, S., FURUYA, T., UEYAMA, Y. and SASAKI, K., 2011. Loss of 3p26.3 is an independent prognostic factor in patients with oral squamous cell carcinoma. *Oncology reports,* **26**(2), pp. 463-469.

UZAWA, N., YOSHIDA, M.A., HOSOE, S., OSHIMURA, M., AMAGASA, T. and IKEUCHI, T., 1998. Functional evidence for involvement of multiple putative tumor suppressor genes on the short arm of chromosome 3 in human oral squamous cell carcinogenesis. *Cancer genetics and cytogenetics,* **107**(2), pp. 125-131.

VAISSIÈRE, T., CUENIN, C., PALIWAL, A., VINEIS, P., HAINAUT, P. and HERCEG, Z., 2009. Quantitative analysis of DNA methylation after whole bisulfitome amplification of a minute amount of DNA from body fluids. *Epigenetics,* **4**(4), pp. 221-230.

VAN DER RIET, P., NAWROZ, H., HRUBAN, R.H., CORIO, R., TOKINO, K., KOCH, W. and SIDRANSKY, D., 1994. Frequent loss of chromosome 9p21-22 early in head and neck cancer progression. *Cancer research,* **54**(5), pp. 1156-1158.

VAN DER WAAL, I., 2010. Potentially malignant disorders of the oral and oropharyngeal mucosa; present concepts of management. *Oral oncology,* **46**(6), pp. 423-425.

VAN HOUTEN, V.M.M., LEEMANS, C.R., KUMMER, J.A., DIJKSTRA, J., KUIK, D.J., VAN DEN BREKEL, M.W.M., SNOW, G.B. and BRAKENHOFF, R.H., 2004. Molecular diagnosis of surgical margins and local recurrence in head and neck cancer patients: A prospective study. *Clinical Cancer Research,* **10**(11), pp. 3614-3620.

VAN HOUTEN, V.M.M., TABOR, M.P., VAN DEN BREKEL, M.W.M., KUMER, J.A., DENKERS, F., DIJKSTRA, J., LEEMANS, R., VAN DER WAAL, I., SNOW, G.B. and BRAKENHOFF, R.H., 2002. Mutated p53 as a molecular marker for the diagnosis of head and neck cancer. *Journal of Pathology,* **198**(4), pp. 476-486.

VEERAMACHANENI, R., 2010. *High resolution genomic analyses of oral cancer.* PhD edn. United Kingdom: The University of Manchester.

VIET, C.T., JORDAN, R.C. and SCHMIDT, B.L., 2007. DNA promoter hypermethylation in saliva for the early diagnosis of oral cancer. *Journal of the California Dental Association,* **35**(12), pp. 844-849.

VIET, C.T. and SCHMIDT, B.L., 2008. Methylation array analysis of preoperative and postoperative saliva DNA in oral cancer patients. *Cancer Epidemiology Biomarkers and Prevention,* **17**(12), pp. 3603-3611.

VILTROP, T., KRJUTŠKOV KAAREL, K., PALTA, P. and METSPALU, A., 2010. Comparison of DNA extraction methods for multiplex polymerase chain reaction. *Analytical Biochemistry,* **398**(2), pp. 260-262.

VISWANATHAN, M., TSUCHIDA, N. and SHANMUGAM, G., 2003. Promoter hypermethylation profile of tumor-associated genes p16, p15, hMLH1, MGMT and E-cadherin in oral squamous cell carcinoma. *International Journal of Cancer,* **105**(1), pp. 41-46.

VON ZEIDLER, S.V., MIRACCA, E.C., NAGAI, M.A. and BIRMAN, E.G., 2004. Hypermethylation of the p16 gene in normal oral mucosa of smokers. *International journal of molecular medicine,* **14**(5), pp. 807-811.

WALSH, T., LIU, J.L., BROCKLEHURST, P., GLENNY, A.M., LINGEN, M., KERR, A.R., OGDEN, G., WARNAKULASURIYA, S. and SCULLY, C., 2013. Clinical assessment to screen for the detection of oral cavity cancer and potentially malignant disorders in apparently healthy adults. *The Cochrane database of systematic reviews,* **11**.

WARNAKULASURIYA S, REIBEL J, BOUQUOT J and DABELSTEEN E, 2008. Oral epithelial dysplasia classification systems: predictive value, utility, weaknesses and scope for improvement. *Journal of oral pathology & medicine : official publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology; Journal of Oral Pathology & Medicine,* **37**(3), pp. 127-133.

WARNAKULASURIYA, S., KOVACEVIC, T., MADDEN, P., COUPLAND, V.H., SPERANDIO, M., ODELL, E. and MØLLER, H., 2011. Factors predicting malignant transformation in oral potentially malignant disorders among patients accrued over a 10-year period in South East England. *Journal of Oral Pathology and Medicine,* **40**(9), pp. 677-683.

WARNAKULASURIYA, S., 2009. Global epidemiology of oral and oropharyngeal cancer. *Oral oncology,* **45**(4–5), pp. 309-316.

WARNAKULASURIYA, S., SUTHERLAND, G. and SCULLY, C., 2005. Tobacco, oral cancer, and treatment of dependence. *Oral oncology,* **41**(3), pp. 244-260.

WEGIEL, B., BJARTELL, A., EKBERG, J., GADALEANU, V., BRUNHOFF, C. and PERSSON, J.L., 2005. A role for cyclin A1 in mediating the autocrine expression of vascular endothelial growth factor in prostate cancer. *Oncogene,* **24**(42), pp. 6385-6393.

WEISS, D., KOOPMANN, M., BASEL, T. and RUDACK, C., 2012. Cyclin A1 shows age-related expression in benign tonsils, HPV16-dependent overexpression in HNSCC and predicts lower recurrence rate in HNSCC independently of HPV16. *BMC Cancer,* **12**:259.

WESTRA, W.H. and CALIFANO, J., 2004. Toward early oral cancer detection using gene expression profiling of saliva: A thoroughfare or dead end? *Clinical Cancer Research,* **10**(24), pp. 8130-8131.

WHITNEY, D., SKOLETSKY, J., MOORE, K., BOYNTON, K., KANN, L., BRAND, R., SYNGAL, S., LAWSON, M. and SHUBER, A., 2004. Enhanced retrieval of DNA from human fecal samples results in improved performance of colorectal cancer screening test. *Journal of Molecular Diagnostics,* **6**(4), pp. 386-395.

WILSON, J.M.G. and JUNGNER, G., 1968-last update, Principles and practice of screening for disease [Homepage of World Health Organisation], [Online]. Available: www.who.int/iris [April/14, 2015].

WOJNAROWICZ, P.M., PROVENCHER, D.M., MES-MASSON, A.-. and TONIN, P.N., 2012. Chromosome 17q25 genes, RHBDF2 and CYGB, inovarian cancer. *International journal of oncology,* **40**(6), pp. 1865-1880.

WONG, I.H.N., DENNIS LO, Y.M. and JOHNSON, P.J., 2001. *Epigenetic tumor markers in plasma and serum: Biology and applications to molecular diagnosis and disease monitoring.*

WOODSON, K., O'REILLY, K.J., HANSON, J.C., NELSON, D., WALK, E.L. and TANGREA, J.A., 2008. The Usefulness of the Detection of GSTP1 Methylation in Urine as a Biomarker in the Diagnosis of Prostate Cancer. *The Journal of urology,* **179**(2), pp. 508-512.

WOOLGAR, J.A. and TRIANTAFYLLOU, A., 2009. Pitfalls and procedures in the histopathological diagnosis of oral and oropharyngeal squamous cell carcinoma and a review of the role of pathology in prognosis. *Oral oncology,* **45**(4-5), pp. 361-385.

XINARIANOS, G., MCRONALD, F.E., RISK, J.M., BOWERS, N.L., NIKOLAIDIS, G., FIELD, J.K. and LILOGLOU, T., 2006. Frequent genetic and epigenetic abnormalities contribute to the deregulation of cytoglobin in non-small cell lung cancer. *Human molecular genetics,* **15**(13), pp. 2038-2044.

YANAMOTO, S., YAMADA, S., TAKAHASHI, H., YOSHITOMI, I., KAWASAKI, G., IKEDA, H., MINAMIZATO, T., SHIRAISHI, T., FUJITA, S., IKEDA, T., ASAHINA, I. and UMEDA, M., 2012. Clinicopathological risk factors for local recurrence in oral squamous cell carcinoma. *International journal of oral and maxillofacial surgery,* **41**(10), pp. 1195-1200.

YANG, N., EIJSINK, J.J.H., LENDVAI, Á., VOLDERS, H.H., KLIP, H., BUIKEMA, H.J., VAN HEMEL, B.M., SCHUURING, E., VAN DER ZEE, A.G.J. and WISMAN, G.B.A., 2009. Methylation markers for CCNA1 and C13ORF18 are strongly associated with high-grade cervical intraepithelial neoplasia and cervical cancer in cervical scrapings. *Cancer Epidemiology Biomarkers and Prevention,* **18**(11), pp. 3000-3007.

YANG, R., MOROSETTI, R. and KOEFFLER, H.P., 1997. Characterization of a second human cyclin A that is highly expressed in testis and in several leukemic cell lines. *Cancer research,* **57**(5), pp. 913-920.

YANG, R., MÜLLER, C., HUYNH, V., FUNG, Y.K., YEE, A.S. and KOEFFLER, H.P., 1999. Functions of cyclin A1 in the cell cycle and its interactions with transcription factor E2F-1 and the Rb family of proteins. *Molecular and cellular biology,* **19**(3), pp. 2400-2407.

YANG, R., NAKAMAKI, T., LÜBBERT, M., SAID, J., SAKASHITA, A., FREYALDENHOVEN, B.S., SPIRA, S., HUYNH, V., MÜLLER, C. and KOEFFLER, H.P., 1999. Cyclin A1 expression in leukemia and normal hematopoietic cells. *Blood,* **93**(6), pp. 2067-2074.

YOUNG, G.P. and BOSCH, L.J.W., 2011. Fecal tests: From blood to molecular markers. *Current Colorectal Cancer Reports,* **7**(1), pp. 62-70.

YOUNG, J., BIDEN, K.G., SIMMS, L.A., HUGGARD, P., KARAMATIC, R., EYRE, H.J., SUTHERLAND, G.R., HERATH, N., BARKER, M., ANDERSON, G.J., FITZPATRICK, D.R., RAMM, G.A., JASS, J.R. and LEGGETT, B.A., 2001. HPP1: A transmembrane protein-encoding gene commonly methylated in colorectal polyps and cancers. *Proceedings of the National Academy of Sciences of the United States of America,* **98**(1), pp. 265-270.

ZHANG, J., CHIODINI, R., BADR, A. and ZHANG, G., 2011. The impact of next-generation sequencing on genomics. *Journal of Genetics and Genomics,* **38**(3), pp. 95-109.

ZHANG, L., FARRELL, J.J., ZHOU, H., ELASHOFF, D., AKIN, D., PARK, N., CHIA, D. and WONG, D.T., 2010. Salivary Transcriptomic Biomarkers for Detection of Resectable Pancreatic Cancer. *Gastroenterology,* **138**(3), pp. 949-957.e7.

ZHANG, L., POH, C.F., WILLIAMS, M., LARONDE, D.M., BEREAN, K., GARDNER, P.J., JIANG, H., WU, L., LEE, J.J. and ROSIN, M.P., 2012. Loss of heterozygosity (LOH) profiles-validated risk predictors for progression to oral cancer. *Cancer Prevention Research,* **5**(9), pp. 1081-1089.

ZHANG, Y., XIONG, Y. and YARBROUGH, W.G., 1998. ARF promotes MDM2 degradation and stabilizes p53: ARF-INK4a locus deletion impairs both the Rb and p53 tumor suppression pathways. *Cell,* **92**(6), pp. 725-734.

## 10. Publication

Available online at www.sciencedirect.com

SciVerse ScienceDirect

British Journal of Oral and Maxillofacial Surgery 51 (2013) 394–398

BRITISH Journal of Oral and Maxillofacial Surgery

www.bjoms.com

# Saliva collection methods for DNA biomarker analysis in oral cancer patients

April M. Matthews [a], Harpal Kaur [b], Michael Dodd [b], Jacob D'Souza [c], Triantafillos Liloglou [d], Richard J. Shaw [c,d], Janet M. Risk [d,*]

[a] *School of Dentistry, University of Liverpool and Regional Maxillofacial, Unit, Aintree University Hospitals NHS Foundation Trust, Liverpool, UK*
[b] *School of Dentistry, University of Liverpool, UK*
[c] *Regional Maxillofacial Unit, Aintree University Hospitals NHS Foundation Trust, Liverpool, UK*
[d] *Department of Molecular & Clinical Cancer Medicine & School of Dentistry, University of Liverpool, UK*

Accepted 24 September 2012
Available online 12 October 2012

**Abstract**

Patients with head and neck cancers are predisposed to local recurrence and second primaries because of the phenomenon of field cancerisation, and clinical detection of recurrence remains challenging. DNA biomarkers in saliva may prove to be an adjunct to current diagnostic methods, but irradiation of the primary site often leads to xerostomia. We assessed 3 methods of collecting saliva for their ability to generate DNA of sufficient quantity and quality to use in biomarker assays. Paired saliva samples were collected from 2 groups of patients with oral squamous cell carcinoma (SCC). In the first group saliva was collected in Oragene® vials and as saline mouthwash from non-irradiated patients ($n = 21$) (4 had had radiotherapy before collection); in the second group it was collected using Oragene® sponge kits and as mouthwash from irradiated patients ($n = 24$). Quantitative polymerase chain reaction (qPCR) showed that Oragene® vials contained DNA in significantly greater amounts (median 122 μg, range 4–379) than mouthwash (median 17 μg, range 2–194) ($p = 0.0001$) in the non-irradiated patients, while Oragene® sponge kits (median 4 μg, range 0.1–61) and mouthwash (median 5.5 μg, range 0.1–75) generated comparable concentrations of DNA from the irradiated group. All 90 samples contained DNA of sufficient quantity and quality for p16 promoter quantitative methylation-specific PCR (qMSP). While Oragene® vials contained the most DNA, all 3 methods yielded enough to detect DNA biomarkers using qMSP. The method of collection should depend on the compliance of the patient and oral competency.


*Keywords:* Saliva; Collection method; Mouthwash; Biomarker; Oral cancer; DNA; PCR

## Introduction

Oral squamous cell carcinoma (SCC) is a molecularly heterogeneous disease[1] that often develops in areas of field change where there is a propensity for recurrence and second primary tumours.[2] Molecular biomarkers associated with oral SCC have been detected in saliva[3–5] despite evidence that more than 70% of the DNA contained in samples of saliva from normal people is from bacteria,[6] and that the human genomic DNA component is derived as much from immune as from epithelial cells.[7] Genomic DNA is a stable macromolecule which has been extracted by a variety of methods in normal people,[8–10] and has been successfully obtained from saliva returned by postal services and stored at room temperature without preservatives.[11,12] These robust features fit it for translation into the clinical setting as a source of biomarkers.

It is accepted that early diagnosis of oral SCC can significantly improve clinical outcome.[13] The main clinical value of saliva biomarkers in this disease would be in early detection and surveillance of disease when multiple,

* Corresponding author at: Department of Molecular & Clinical Cancer Medicine, Institute of Translational Medicine, University of Liverpool, Daulby Street, L69 3GN, UK. Tel.: +44 0151 706 5265; fax: +44 0151 706 5809.
*E-mail addresses:* j.m.risk@liverpool.ac.uk, j.m.risk@liv.ac.uk (J.M. Risk).

0266-4356/$ – see front matter 
http://dx.doi.org/10.1016/j.bjoms.2012.09.017

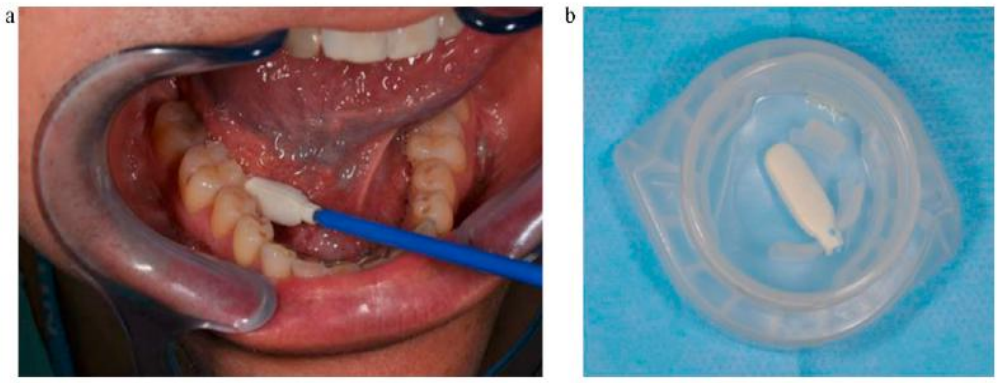


Fig. 1. (a) Collection of saliva using Oragene® sponge kit; (b) sponge tip loaded with saliva in the Oragene® collection pot.

chronological samples are required, and repeated biopsy examination is impractical. Collecting saliva is not invasive and is known to be more acceptable to patients than venipuncture.[14] Saliva bathes the whole oral cavity and is therefore more likely to be representative of the entire exposed field than invasive biopsy examination of local tissue, but concerns about sensitivity in relation to contamination by bacteria and immune cells remain unanswered.

Postoperative radiotherapy in the treatment of oral SCC identifies patients at risk of recurrence, and who need close follow-up,[15] but the clinical detection of early recurrence or second primaries in the surgically altered, irradiated field can be challenging. Xerostomia, a known sequela of radiotherapy, could make the collection of saliva for adjunctive molecular methods of detection difficult.

We aimed to establish the quality and quantity of DNA that could be extracted from the saliva of patients with oral cancer, including a group who had had radiotherapy, using 3 different methods of collection. One new technique uses a sponge that was developed for patients who find conventional methods difficult.

## Methods

Two separate groups of 21 and 24 patients (45 in total) were recruited from consecutive patients attending Aintree University Hospital Maxillofacial Unit for treatment of oral SCC. The study was given ethical approval (EC 47.01) and all patients agreed to participate.

The first group comprised 10 preoperative and 11 postoperative patients, 4 of whom had had radiotherapy before the saliva was collected. Two samples were collected from each patient as follows: 0.9% normal saline 25 ml was swilled in the mouth for 30 s and deposited into a tube.[4] Ten minutes later, whole saliva was deposited directly into an Oragene® vial (DNA Genotek Inc, Ontario, Canada) until the 2 ml fill line was reached. Samples were immediately stored at 4 °C for a maximum of 3 h before being placed in a centrifuge at 1200 × *g* for 5 min and the cell pellet stored at −80 °C. Oragene® vials do not require processing before storage.

The second group comprised 24 patients who had had operation and radiotherapy for oral SCC. Paired saliva samples were collected from each patient as follows: 5 individual sponges (DNA Genotek Inc, Ontario, Canada) were placed into pools of whole saliva in the patient's mouth (Fig. 1a), then removed and placed in a standard Oragene® vial (Fig. 1b). Mouthwash was collected and the samples stored as before.

DNA was extracted from the mouthwash in the cell pellets according to a spin column protocol (DNeasy Blood & Tissue Kit: Qiagen, Crawley, UK). Phosphate buffered saline was added to each pellet to make a final volume of 500 μl, then proteinase K (20 mg/ml) 25 μl was added to an aliquot (180 μl) of this, and separation continued using spin columns with elution in AE buffer (10 mM Tris–Cl, 0.5 mM EDTA; pH 9.0) 200 μl.

DNA was extracted from the Oragene® vials using the manufacturer's protocol for a sample of 0.5 ml. Vials were incubated at 50 °C for 1 h to lyse the cells and digest nuclear proteins, then Oragene® purifier 20 μl was added to 0.5 ml of sample before precipitation in ethanol and final resuspension in Tris–EDTA buffer (1 mM Tris, 0.1 mM EDTA, pH 8.0) 200 μl.

DNA was extracted from the Oragene® sponges using the manufacturer's protocol. The vials were incubated in a 50 °C water bath for 1 h as above, and the liquid was recovered from the sponges by centrifugation in a 5 ml plastic syringe placed in a 15 ml centrifuge tube. The manufacturer's Oragene®/saliva protocol for 4 ml samples was followed for the entire volume. DNA was resuspended in Tris–EDTA buffer (1 mM Tris, 0.1 mM EDTA, pH 8.0) 200 μl.

To quantify accurately the total concentration of human DNA, a quantitative polymerase chain reaction (qPCR) using the human ribonuclease P (RNAse P) gene[16] was done using the Applied Biosystems 7500 Fast Real-Time PCR system on the standard curve program according to the manufacturer's guidelines (Applied Biosystems, Foster City, CA, USA). Controls of known DNA concentration were diluted from

a stock of human lymphocyte DNA (500 ng/μl), and samples were run in triplicate. The cycle threshold was plotted against a $\log_2$ scale for the standard dilutions to create a linear model of their relation, and was used to calculate the concentration of DNA in the samples based on the cycle threshold value.

From each DNA sample 500 ng was treated with bisulphite according to the manufacturer's protocol using Zymo EZ-96 DNA Methylation-Gold™ kit (Zymo Research Corporation, Irvine, CA, USA), eluted in M-elution buffer 30 μl. Lymphocyte DNA 100 ng/μl, methylated in vitro with Sss-I methylase (Zymo Research Corporation), was diluted to 0.1%, 0.5%, 1.0%, and 5.0% methylation to act as reference samples, and was treated with bisulphite at the same time as the sample DNA.

To assess the quality of DNA for downstream application, TaqMan® p16 methylation qMSP assays[17] (Life Technologies Ltd, Paisley, UK) were done on the bisulphite-modified DNA with beta actin as an internal control gene (probe details on request). The delta delta cycle threshold method (ΔΔCT) was used to correct for amount of starting DNA, and samples were deemed to be positive for p16 promoter methylation if the mean delta CT was equal to or more than that obtained for control DNA methylated to a level of 0.5%. This value was calculated from data on levels of p16 methylation in saliva from a larger group of normal people with similar smoking histories (data not shown).

Results were analysed for statistical significance using the Wilcoxon signed rank test in SPSS v18 software package.

## Results

A total of 45 patients provided 90 saliva samples. Group 2 targeted patients who had had radiotherapy after problems that patients in group 1 had encountered when collecting saliva before and after operation.

All 90 samples contained human DNA as measured by the RNAse P qPCR assay, and both groups showed a large variation in the final concentrations. In group 1 the DNA concentration from the Oragene® whole saliva samples (range 10–929 ng/μl; median 306 ng/μl; SD 291 ng/μl) was significantly greater ($p = 0.001$) than that of the matched mouthwash samples (range 7–657 ng/μl; median 62 ng/μl; SD 169 ng/μl) (Fig. 2). Median DNA concentrations were lower in group 2 than in group 1, and there was no significant difference in the concentrations obtained using the Oragene™ sponge (range 0.3–306 ng/μl; median 22 ng/μl; SD 86 ng/μl) and matched mouthwash (range 1.6–267 ng/μl; median 20 ng/μl; SD 77 ng/μl) (Fig. 2).

The total DNA yield for each sample was calculated, and it reflected the DNA concentration data. In group 1 the total amount of DNA in whole saliva collected using Oragene™ (range 4–379 μg; median 122 μg; SD 116 μg) was significantly greater than the mouthwash (range 2–194 μg; median 17 μg; SD 47 μg) ($p = 0.0001$). Again, there was no significant difference between the DNA yield from the

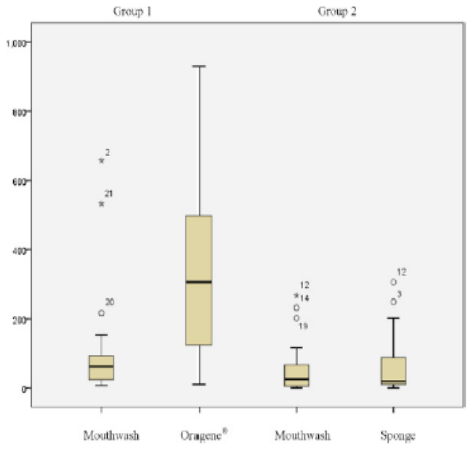



Fig. 2. DNA concentration range (ng/l) using qPCR. Group 1 compares mouthwash with Oragene® (*$p = 0.001$), group 2 compares mouthwash with Oragene® sponge kits (* and ° signify samples that are outliers).

Oragene® sponge and mouthwash methods in group 2 (range 0.1–61 μg; median 4 μg; SD 17 μg and range 0.1–75 μg; median 5.5 μg; SD 21 μg, respectively).

All 90 of the bisulphite-treated saliva DNA were amplified within 37 cycles using the beta actin probe (range 24–37; mean/median 27 to 2 SD). Four patients had p16 methylation in matched Oragene® and mouthwash samples; 3 of them had their samples collected preoperatively. Three other positive results were obtained from single samples from 3 patients who had had postoperative radiotherapy (two sponge samples and one mouthwash sample).

## Discussion

Oral cancer displays genetic and biological heterogeneity which is not always evident at the histological level.[1] Clinical behaviour can be unpredictable in some individual tumours, which implies a role for molecular profiling in diagnosis and prognostication in conjunction with clinical observation and screening. DNA biomarkers specific to oral SCC have been detected in saliva[2,3,18,19] and may have prognostic value.[4]

The main strength of our study is that it is the first to compare 3 different methods of collecting saliva in patients with oral SCC, particularly after radiotherapy when xerostomia could potentially limit the amount, and after operation when oral competence may be compromised and make collection difficult. To date, and to our knowledge, it is also the first to use Oragene® sponges in this application. To improve the accuracy of the estimation of DNA concentration, we used an RNAse P qPCR assay,[16] which allowed the differentiation of human from non-human DNA, while spectrophotometry can be influenced by particulate matter, contamination of protein, or the presence of non-human DNA.

The median DNA yields found in this study are broadly comparable with those published for predominantly healthy participants.[8,10,12] However, the median yields obtained from the radiotherapy group (group 2) were lower than those from group 1 even when data from the 4 patients who had had radiotherapy in that group were excluded (data not shown). The wide range of DNA concentrations in the samples across all 3 methods has previously been observed with the Oragene® method of collection.[10]

All 3 methods are subject to variation because of the sampling technique, individual variation in the number of oral epithelial cells in saliva,[7,20] or the availability of saliva. While the site of collection may also affect the sponge method, all our samples were taken from the floor of the mouth, and a recent study has shown that sampling at different locations in the mouth gives similar quantity and quality of DNA.[21] Objectively, it is easier to collect saliva from patients who are compromised orally using sponges rather than mouthwash in a clinical setting, and it requires less processing before storage. However, we did not ask patients which method they preferred.

One of the concerns about the use of surrogate samples for biomarker analysis is the relatively small amount of DNA for use in downstream applications. The DNA from every sample in our series retained enough integrity after modification by bisulphites to amplify in beta actin qPCR methylation assays, and produced enough DNA for more than 50 such assays in each case. The observation that 3/10 preoperative patients showed 16 promoter methylation in their paired saliva samples is comparable to our previous tumour data, where 28% of tumours show p16 methylation.[22] It should be noted that we did not collect matched tumour methylation data for this series of patients, but concordance between saliva and tumour tissue has previously been shown,[4,18] and was not a primary aim of the study. The detection of p16 promoter gene methylation in 3 specimens, but not in their paired samples, is probably indicative of levels of methylation at the limits of sensitivity of the test rather than a function of the amount or integrity of the DNA collected, or non-concordance between methods of collection, but the sample size is too small to comment. The clinical relevance of these results is part of a larger study.

A considerable body of evidence now exists to support saliva as an adequate source of DNA suitable for use in PCR[9,10] to detect potential biomarkers.[4,5,23] The advantage of this study is that it shows the feasibility of these collection techniques in a working clinical setting with patients who have problems with oral competence and saliva production. Ideal conditions for collecting samples are not always possible outside the laboratory where sampling is much more time-consuming, less well tolerated, and often produces "invalid" samples.[24] One acknowledged limitation of the study is the 10 min interval between collections imposed by the clinical setting, and a swill time of 30 s for the mouthwash samples. An hour between sample collections would be preferable to give the saliva time to replenish, and other authors have described swilling the mouthwash for periods of up to 3 min.[3] We found that this could not be tolerated by patients whose oral competence had been compromised by operation.

This paper shows that sufficient DNA can be isolated from saliva collected from patients with oral cancer at various stages of treatment, including those who have had radiotherapy, for use in testing a substantial panel of biomarkers. While collection using Oragene® vials provided the most DNA in this study, all 3 methods yielded enough to detect biomarkers using qMSP. Therefore, the method of collection for any given situation should be selected according to the compliance of the patient and oral competency. Saliva could have a role as a tumour surrogate in applications for longitudinal clinical biomarker testing. Further work is currently underway to establish a panel of DNA methylation biomarkers that can be used reliably to detect and monitor oral SCC.

## Conflict of interest

DNA Genotek provided 25 Oragene™ sponge collection kits free of charge.

## Acknowledgements

The British Association of Head and Oncologists and the British Association of Oral and Maxillofacial Surgeons funded this work. The study sponsors (University of Liverpool & Royal Liverpool and Broadgreen University Hospitals NHS Trust) had no involvement in any experimental aspects of this study nor in preparation of the manuscript.

## References

1. Leemans CR, Braakhuis BJ, Brakenhoff RH. The molecular biology of head and neck cancer. *Nat Rev Cancer* 2011;**11**:9–22.
2. Braakhuis BJ, Brakenhoff RH, Leemans CR. Second field tumors: a new opportunity for cancer prevention? *Oncologist* 2005;**10**:493–500.
3. El-Naggar AK, Mao L, Staerkel G, et al. Genetic heterogeneity in saliva from patients with oral squamous carcinomas: implications in molecular diagnosis and screening. *J Mol Diagn* 2001;**3**:164–70.
4. Righini CA, de Fraipont F, Timsit JF, et al. Tumor-specific methylation in saliva: a promising biomarker for early detection of head and neck cancer recurrence. *Clin Cancer Res* 2007;**13**:1179–85.
5. Carvalho AL, Jeronimo C, Kim MM, et al. Evaluation of promoter hypermethylation detection in body fluids as a screening/diagnosis tool for head and neck squamous cell carcinoma. *Clin Cancer Res* 2008;**14**:97–107.
6. Hu Y, Ehli EA, Nelson K, et al. Genotyping performance between saliva and blood-derived genomic DNAs on the DMET array: a comparison. *PLoS ONE* 2012;**7**:e33968.
7. Thiede C, Prange-Krex G, Freiberg-Richter J, Bornhäuser M, Ehninger G. Buccal swabs but not mouthwash samples can be used to obtain pretransplant DNA fingerprints from recipients of allogeneic bone marrow transplants. *Bone Marrow Transplant* 2000;**25**:575–7.

8. Feigelson HS, Rodriguez C, Robertson AS, et al. Determinants of DNA yield and quality from buccal cell samples collected with mouthwash. *Cancer Epidemiol Biomarkers Prev* 2001;**10**:1005–8.
9. García-Closas M, Egan KM, Abruzzo J, et al. Collection of genomic DNA from adults in epidemiological studies by buccal cytobrush and mouthwash. *Cancer Epidemiol Biomarkers Prev* 2001;**10**:687–96.
10. Rylander-Rudqvist T, Håkansson N, Tybring G, Wolk A. Quality and quantity of saliva DNA obtained from the self-administrated oragene method – a pilot study on the cohort of Swedish men. *Cancer Epidemiol Biomarkers Prev* 2006;**15**:1742–5.
11. Freeman B, Powell J, Ball D, Hill L, Craig I, Plomin R. DNA by mail: an inexpensive and noninvasive method for collecting DNA samples from widely dispersed populations. *Behav Genet* 1997;**27**:251–7.
12. Ng DP, Koh D, Choo SG, Ng V, Fu Q. Effect of storage conditions on the extraction of PCR-quality genomic DNA from saliva. *Clin Chim Acta* 2004;**343**:191–4.
13. Rogers SN, Brown JS, Woolgar JA, et al. Survival following primary surgery for oral cancer. *Oral Oncol* 2009;**45**:201–11.
14. Koka S, Beebe TJ, Merry SP, et al. The preferences of adult outpatients in medical or dental care settings for giving saliva, urine or blood for clinical testing. *J Am Dent Assoc* 2008;**139**:735–40.
15. Shaw RJ, Pace-Balzan A, Butterworth C. Contemporary clinical management of oral squamous cell carcinoma. *Periodontol* 2011;**2000**(57):89–101.
16. Philibert RA, Zadorozhnyaya O, Beach SR, Brody GH. Comparison of the genotyping results using DNA obtained from blood and saliva. *Psychiatr Genet* 2008;**18**:275–81.
17. Schache AG, Hall G, Woolgar JA, et al. Quantitative promoter methylation differentiates carcinoma ex pleomorphic adenoma from pleomorphic salivary adenoma. *Br J Cancer* 2010;**103**:1846–51.
18. Shaw RJ, Akufo-Tetteh EK, Risk JM, Field JK, Liloglou T. Methylation enrichment pyrosequencing: combining the specificity of MSP with validation by pyrosequencing. *Nucleic Acids Res* 2006;**34**:e78.
19. Viet CT, Schmidt BL. Methylation array analysis of preoperative and postoperative saliva DNA in oral cancer patients. *Cancer Epidemiol Biomarkers Prev* 2008;**17**:3603–11.
20. Dawes C, Estimates. from salivary analyses, of the turnover time of the oral mucosal epithelium in humans and the number of bacteria in an edentulous mouth. *Arch Oral Biol* 2003;**48**:329–36.
21. Nemoda Z, Horvat-Gordon M, Fortunato CK, Beltzer EK, Scholl JL, Granger DA. Assessing genetic polymorphisms using DNA extracted from cells present in saliva samples. *BMC Med Res Methodol* 2011;**11**:170.
22. Shaw RJ, Liloglou T, Rogers SN, et al. Promoter methylation of P16, RARbeta E-cadherin, cyclin A1 and cytoglobin in oral cancer: quantitative evaluation using pyrosequencing. *Br J Cancer* 2006;**94**:561–8.
23. Viet CT, Jordan RC, Schmidt BL. DNA promoter hypermethylation in saliva for the early diagnosis of oral cancer. *J Calif Dent Assoc* 2007;**35**:844–9.
24. Hodgson N, Freedman VA, Granger DA, Erno A. Biobehavioral correlates of relocation in the frail elderly: salivary cortisol, affect, and cognitive function. *J Am Geriatr Soc* 2004;**52**:1856–62.

# Appendices

# Appendix 1

## A1.1 Ethics confirmation

**Liverpool Paediatric Research Ethics Committee**
Bishop Goss Complex
Victoria Building
Rose Place
Liverpool
L3 3AN

Telephone: 0151 330 2071
Facsimile: 0151 330 2075

05 August 2008

Dr Anne Field
Honorary Senior Lecturer
University of Liverpool
School of Dental Sciences
Liverpool University Dental Hospital
Pembroke Place, Liverpool
L3 5PS

Dear Dr Field

**Full title of study:** **Cancer biomarkers in saliva: establishing a control group**

**REC reference number:** **08/H1002/42**

Thank you for your letter of 04 August 2008, responding to the Committee's request for further information on the above research and submitting revised documentation, subject to the conditions specified below.

The further information has been considered on behalf of the Committee by the Chair.

**Confirmation of ethical opinion**

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised.

**Ethical review of research sites**

The Committee has designated this study as exempt from site-specific assessment (SSA. There is no requirement for [other] Local Research Ethics Committees to be informed or for site-specific assessment to be carried out at each site.

**Approved documents**

The final list of documents reviewed and approved by the Committee is as follows:

| *Document* | *Version* | *Date* | |
|---|---|---|---|
| Application | | 24 April 2008 | |
| Application | 5.6 | 21 April 2008 | |
| Investigator CV | | 01 January 2008 | |
| Protocol | 2 | 20 July 2008 | |
| Protocol | 1 | 01 January 2008 | |
| Covering Letter | | | |
| Letter from Sponsor | | 29 April 2008 | |
| Questionnaire | 1 | 20 July 2008 | |
| Participant Information Sheet | 2 | 02 June 2008 | |
| Participant Information Sheet | 1 | 01 January 2008 | |
| Participant Consent Form | 1 | 01 January 2008 | |
| Participant Consent Form | 2 | 20 July 2008 | |
| Response to Request for Further Information | | 04 August 2008 | |
| C.V. for Supervisor | | 03 September 2007 | |
| Peer Review | | 30 June 2008 | |

**Statement of compliance**

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

**After ethical review**

Now that you have completed the application process please visit the National Research Ethics Website > After Review

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email referencegroup@nres.npsa.nhs.uk.

| **08/H1002/42** | **Please quote this number on all correspondence** |
|---|---|

With the Committee's best wishes for the success of this project

Yours sincerely

**Mrs Jean Harkin**

**Chair**

Email: adam.lewis@liverpoolpct.nhs.uk

*Enclosures:* "After ethical review – guidance for researchers"

*Copy to:* *Mr Tony Grayson*

*R andD Department*

*Royal Liverpool University Hospital*

**A1.2 Information sheet**

# Liverpool University Dental Hospital

## Cancer biomarkers in saliva: establishing a control group

**INFORMATION SHEET**

*Principal Investigator*
Miss April Matthews

*Chief Investigator*
Dr A.E. Field

*Educational Supervisors*
Janet M Risk
Richard J Shaw

Thank you for reading this information leaflet.

You are being invited to take part in a research study that is investigating saliva collection as a means of screening for oral disease.

**Background Information and purpose**

The UK has seen a recent increase of cases of oral cancer. It is generally accepted that smoking and excess drinking contribute to the cause of oral cancer, which occurs by the gradual build up of genetic changes in the cells that make up the lining of the mouth. It is known that late diagnosis can be a barrier to treatment and that early cancers are more likely to be cured and require less debilitating treatments. However, to date there is no accepted method of reliably detecting these changes or of identifying those people at risk of developing them before oral cancer has occurred.

We are researching how much genetic information we can obtain from the non-invasive methods of saliva and oral swab collection, with the future aim of testing these methods as potential screening tests for oral cancer in the population. In order to move towards this goal, we need ***normal*** saliva samples to work with (i.e. saliva from people who do not have oral cancer).

**It is important to understand that this work is experimental and does not aim to predict your risks of oral cancer.**

**Do I have to take part?**

No, it is entirely up to you to decide whether or not you wish to partake. If you do, you will be given this sheet to keep and a consent form to sign.

**What will happen if I take part?**

If you agree to participate, we would like you to give us a sample of your saliva, which you will deposit in a small plastic pot, and an oral swab (taken by the researcher) before your dental appointment. When we receive the samples we will use them to obtain genetic material. The genetic material from the saliva and the oral swab will be used for research. **We will only use these samples for this purpose.** You will **not** be required to attend for any additional appointments. We are not attempting to find a diagnosis for you, and your dental treatment will be no different if you agree to participate. We will also require you to answer a few questions about your smoking and drinking habits and your oral health.

**What are the benefits / risks?**

The study will not benefit you in any way, but may help with diagnosis and treatment of future patients. If you decide to give a sample you will be provided with privacy in a room when producing your sample into the pot.

**What if there is a problem?**

The collection of these samples are non invasive and unlikely to cause you harm. If you wish to complain about any aspect of the way you have been approached or treated during the course of this study, the National Health Service complaints mechanisms will be available to you.

**Will my taking part in the research be kept confidential?**

All the samples and information collected from you will be completely anonymised. The researchers will not be able to identify you from your sample once it has been collected. Your sample will have a unique code and it will be stored on a secure database but your name and date of birth will not be recorded or kept with the sample.

**What will happen to the samples**?

If you agree to donate samples to us, the project co-ordinators will be responsible for looking after them. You must understand that if eventually this research leads to any new diagnostic test you will not benefit financially.

The research we are carrying out is very new and changing all the time, so we will store some of the genetic material in the Royal Liverpool University Hospital tissue bank for use in future research projects. We may wish to share some of the samples or some of the information with other researchers who are working with us. If we do, the samples are anonymised and therefore it is impossible for us to share your personal information.

It may be beneficial, in the future, for us to work with a medical research company to develop new medical tests or treatments. We may wish to share some of the samples or some of the information with other researchers working for such a company. If we do so, the samples are anonymised and none of your personal information is available. Any additional studies will be approved by the relevant ethics committees.

What we will do with the information about your DNA**:**

We are asking you to give samples for genetic analysis only so that we can look at which genetic markers are present in normal saliva. We do not know the answer to this question yet, so we will not be able to tell you the results of what we find. We are not using your DNA sample to give you a 'genetic test' that will give you a diagnosis or predict your risk of oral disease, but we hope to use it to make such testing possible in the future.

We will not use any DNA sample given to this project for any purpose other than research.

**What will happen to the results of the research study?**

The results will be published in scientific journals and also presented at scientific meetings and any results published will not relate to individuals or mention any names. Once the research is completed the results of the study can be found at the website

www.headandneckcancer.co.uk. Individual test results will not be available as the samples collected will be anonymised.

**Who is organising and funding the research?**
The University of Liverpool, School of Dental Sciences.

**Who has reviewed the study?**
This study has been given a favourable ethical approval by Liverpool Research Ethics Committee.

**Contact Details for further information**

Miss April Matthews,
Mr Richard Shaw

School of Dental Sciences
Tel: 0151706 5275

**A1.3 Consent Form** Date:___/___/2008____

Cancer biomarkers in saliva: establishing a control group

**PLEASE READ CAREFULLY AND INITIAL EACH SECTION**

| | |
|---|---|
| I have read the attached information sheet on the use of samples for the above project and have been given a copy to keep. I have had the opportunity to ask questions about the project and understand why the research is being done and any foreseeable risks involved. | **YES/NO**…… |
| I understand how the saliva samples and swabs will be collected and that giving samples for this project is voluntary. | **YES/NO** …… |
| I agree to give saliva samples for this project | **YES/NO**…… |
| I agree to give oral swab samples for this project | **YES/NO**…… |
| I agree that the sample and the information I give will be looked after and stored at the University of Liverpool and maybe used for future molecular projects on oral cancer with ethical committee approval. | **YES/NO**……. |
| I understand that some of these projects may be carried out by researchers other than the current team, both within the University of Liverpool and elsewhere. | **YES/NO**……. |
| I understand that some of these projects may be carried out by researchers working for commercial organisations. | **YES/NO**……. |
| I confirm that I understand the purpose of this research study. I am aware that donation of samples is voluntary, and that my dental care will not be affected. | **YES/NO** ……. |

| ..... | …… | …….. |
|---|---|---|
| **Patient name** | **Date** | **Signature** |
| …… | …… | …….. |
| **Consented by (Block Capitals)** | **Date** | **Signature** |

**A1.4 Questionnaire**

**Cancer biomarkers in saliva: establishing a control group**

**Sample No _____**

**Age _____ years**

**Gender Male Female**

**Smoker Yes No**

**Number cigarettes smoked per day _____**

**Number of years you have smoked**

**Alcohol Yes No**

**Number of units per week _____**

**Reference guide:**

| | |
|---|---|
| **1 pint of lager** | **2-3 units** |
| **1 small glass wine** | **1 unit** |
| **1 single spirit** | **1 unit** |

**Have you ever been treated for oral cancer?**

**Yes No**

**Have you had any other type of cancer?**

**Yes No**

**If answered 'Yes' which type of cancer have you had? _____**

# Appendix 2: Commonly used reagents

| Reagent | Chemical composition |
|---|---|
| **1 X TE** | 10mM Tris-HCl, 1mM EDTA pH 7.5 |
| **Phosphate buffered saline** | NaCl 137mM 2.7mM K2HP04 10mM KH2PO4 2mM PH 7.4<br>Sigma http://www.sigmaaldrich.com/ **P4417** |
| **2xPK buffer** | 100mM Tris-HCl,2mM EDTA, 1% Tween 20 |

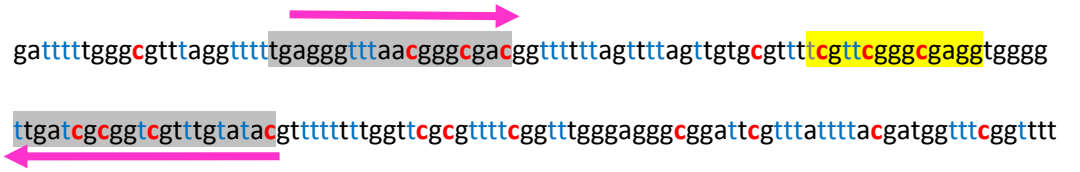
**Table A2.1: Chemical composition of commonly used reagents**

## Appendix 3: Target gene panel primer/probe sequences

gttttcggtaggcggcg gcgttgtttagtcgagcgtagac gggattttcgtagcgagatttttagcgatttttaaagttaa

aagttggcggcgggcgtcgggtttcgcgcgtttttacggtcgttgtttcgcgtcgtcg tcgtagttaaggagggtagga

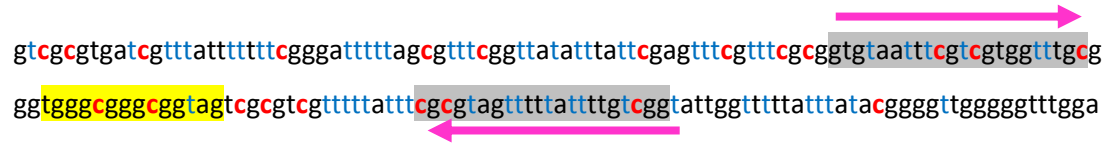
***ADAMTS9***



gattttgggcgtttaggttttgagggtttaacgggcgacggttttttagttttagttgtgcgtttttcgttcgggcgaggtgggg

ttgatcgcggtcgtttgtatacgttttttttggttcgcgttttcggtttgggagggcggattcgtttatttttacgatggtttcggtttt

***CCNA1***

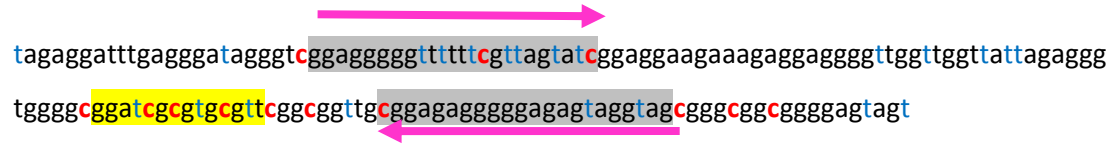


gtcgcgtgatcgtttatttttttcgggatttttagcgtttcggttatatttattcgagtttcgtttcgcggtgtaatttcgtcgtggtttgcg

ggtgggcgggcggtagtcgcgtcgttttttatttcgcgtagttttttatttttgtcggtattggttttttatttatacggggttgggggtttgga

gcggat

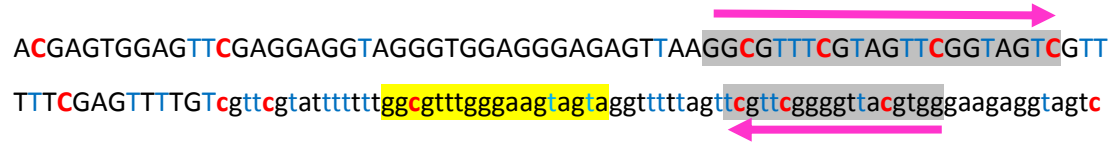
***CYGB***



tagaggatttgagggatagggtcggaggggttttttcgttagtatcggaggaagaaagaggaggggttggttggttattagaggg

tgggggcggatcgcgtgcgttcggcggttgcggagaggggggagagtaggtagcgggcggcggggagtagt

***P16***

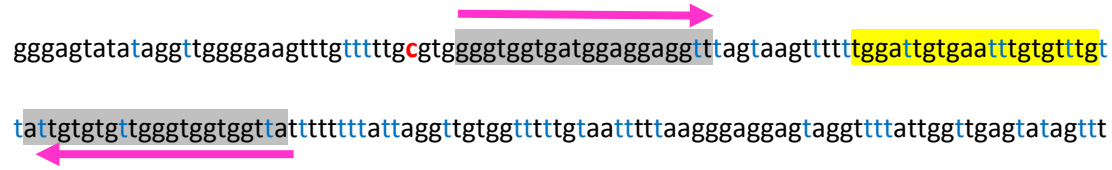


ACGAGTGGAGTTCGAGGAGGTAGGGTGGAGGGAGAGTTAAGGCGTTTCGTAGTTCGGTAGTCGTT

TTTCGAGTTTTGTcgttcgtatttttttggcgtttgggaagtagtaggttttttagttcgttcggggttacgtgggaagaggtagtc

gggttt

***TMEFF2***



gggagtatataggttggggaagtttgttttttgcgtggggtggtgatggaggaggtttagtaagttttttggattgtgaatttgtgtttgt

tattgtgtgttgggtggtggttatttttttttattaggttgtggttttttgtaattttttaagggaggagtaggttttattggttgagtatagttt

***ACTB***

**Figure 1: PCR primers and probes for the five target genes and beta-actin. The bases shaded in grey denote the forward and reverse primers as shown by the pink arrows and the bases shaded in yellow represent the probe sequence. The ts shaded in blue represent cytosines that will be converted to thymine residues by the bisulphite reaction. The CpGs are denoted by a red C and uppercase text denotes the first exon. ACTB is included for DNA input control and therefore is methylation-independent.**

# Appendix 4: Concordance between singleplex and duplex data

| | | Singleplex *ADAMTS9* | | Total |
|---|---|---|---|---|
| | | Neg | Pos | |
| Duplex *ADAMTS9* | Neg | 201 | 4 | 205 |
| | Pos | 0 | 1 | 1 |
| Total | | 201 | 5 | 206 |

McNemar Test p=0.125

**Table A4.1: Contingency table of duplex and singleplex *ADAMTS9* data**

| | | Singleplex *CCNA1* | | Total |
|---|---|---|---|---|
| | | Neg | Pos | |
| Duplex *CCNA1* | Neg | 196 | 3 | 199 |
| | Pos | 0 | 8 | 8 |
| Total | | 196 | 11 | 207 |

McNemar Test p=0.250

**Table A4.2: Contingency table of duplex and singleplex *CCNA1* data**

| | | Singleplex *CYGB* | | Total |
|---|---|---|---|---|
| | | Neg | Pos | |
| Duplex *CYGB* | Neg | 192 | 2 | 194 |
| | Pos | 1 | 9 | 10 |
| Total | | 193 | 11 | 204 |

McNemar Test p=1.0

**Table A4.3: Contingency table of duplex and singleplex *CYGB* data**

| | | Singleplex *P16* | | Total |
|---|---|---|---|---|
| | | Neg | Pos | |
| Duplex *P16* | Neg | 179 | 2 | 181 |
| | Pos | 9 | 15 | 24 |
| Total | | 188 | 17 | 205 |

McNemar Test p=0.65

**Table A4.4: Contingency table of duplex and singleplex *P16* data**

| | | Singleplex *TMEFF2* | | Total |
|---|---|---|---|---|
| | | Neg | Pos | |
| Duplex *TMEFF2* | Neg | 173 | 9 | 182 |
| | Pos | 0 | 11 | 11 |
| Total | | 173 | 20 | 193 |

McNemar Test p=0.004

**Table A4.5: Contingency table of duplex and singleplex *TMEFF2* data**

## Appendix 5: Demographic data

| | | Cohort | | Total |
|---|---|---|---|---|
| | | Normal | OSCC | |
| Age | 40.00 | 44 | 1 | 45 |
| | 60.00 | 29 | 28 | 57 |
| | 80.00 | 24 | 28 | 52 |
| Total | | 97 | 57 | 154 |

Fisher's exact (2-sided) p<0.001

**Table A5.1: Contingency table comparing age between normal and OSCC cohort. Age 40: equal to or < 40 years old, 60: 41-60 years old, 80: 61+years old**

| | | Cohorts | | Total |
|---|---|---|---|---|
| | | Normal | OSCC | |
| smoking | .0 | 38 | 8 | 46 |
| | 1.0 | 42 | 11 | 53 |
| | 2.0 | 17 | 28 | 45 |
| Total | | 97 | 47 | 144 |

Pearson chi Square (2-sided) p<0.001

**Table A5.2: Contingency table comparing smoking habits between normal and OSCC cohort. Smoking 0=non smokers 1=less than or equal to 20pack years 2=greater than 20 packyears.**

| | | Cohorts | | Total |
|---|---|---|---|---|
| | | Normal | OSCC | |
| Alcohol | .0 | 36 | 5 | 41 |
| | 1.0 | 59 | 25 | 84 |
| | 2.0 | 2 | 16 | 18 |
| Total | | 97 | 46 | 143 |

Fisher's exact (2-sided) p<0.001

**Table A5.3: Contingency table comparing alcohol intake between normal and OSCC cohort.**
**0=No alcohol 1=equal to or less than 28u/week 3= greater than 28u/week.**

| | | Cohorts | | Total |
|---|---|---|---|---|
| | | Normal | OSCC | |
| Gender | Male | 56 | 39 | 95 |
| | Female | 41 | 22 | 63 |
| Total | | 97 | 61 | 158 |

Pearson Chi square (2-sided)p=0.438

**Table A5.4: Contingency table comparing gender ratio between normal and OSCC cohort.**

| | | Cohorts | | Total |
|---|---|---|---|---|
| | | Normal | Dysplasia | |
| Age | 40.00 | 44 | 4 | 48 |
| | 60.00 | 29 | 31 | 60 |
| | 80.00 | 24 | 25 | 49 |
| Total | | 97 | 60 | 157 |

Fisher's exact (2-sided) p<0.001

**Table A5.5: Contingency table comparing age distribution between normal and dysplasia cohorts. Age 40: equal to or less than 40 years old, 60: 41-60 years old, 80: 61+years old**.

| | | Cohort | | Total |
|---|---|---|---|---|
| | | Normal | Dysplasia | |
| Smoking | .0 | 38 | 16 | 54 |
| | 1.0 | 42 | 15 | 57 |
| | 2.0 | 17 | 29 | 46 |
| Total | | 97 | 60 | 157 |

Pearson Chi square (2-sided) p<0.001

**Table A5.6: Contingency table comparing smoking habits between normal and dysplasia cohorts. 0: None smoker 1: less than or equal to 20packyears 2: greater than 20packyears**

| | | Cohort | | Total |
|---|---|---|---|---|
| | | Normal | Dysplasia | |
| Alcohol | .0 | 36 | 19 | 55 |
| | 1.0 | 59 | 32 | 91 |
| | 2.0 | 2 | 8 | 10 |
| Total | | 97 | 59 | 156 |

Fisher's exact (2-sided) p=0.023

**Table A5.7: Contingency table comparing alcohol intake between normal and dysplasia cohorts. 0=No alcohol 1 =equal to or less than 28u/week 2=greater than 28u/week**

| | | Cohort | | Total |
|---|---|---|---|---|
| | | Normal | Dysplasia | |
| Gender | Male | 56 | 34 | 90 |
| | Female | 41 | 26 | 67 |
| Total | | 97 | 60 | 157 |

Pearson Chi square (2-sided) p=0.896

**Table A5.8: Contingency table comparing age gender ratio between normal and dysplasia cohorts.**

| | | Cohort | | Total |
|---|---|---|---|---|
| | | Dysplasia | OSCC | |
| Age | 40.00 | 4 | 1 | 5 |
| | 60.00 | 31 | 28 | 59 |
| | 80.00 | 25 | 28 | 53 |
| Total | | 60 | 57 | 117 |

Fisher's exact (2-sided) p=0.455

**Table A5.9: Contingency table comparing age distribution between dysplasia and OSCC cohorts. Age 40 =equal to or less than 40 years old 60=41-60 years old 80= 61+years old**.

| | | Cohort | | Total |
|---|---|---|---|---|
| | | Dysplasia | OSCC | |
| Smoking | 0 | 16 | 8 | 24 |
| | 1 | 15 | 11 | 26 |
| | 2 | 29 | 28 | 57 |
| Total | | 60 | 47 | 107 |

Pearson Chi square p=0.418

**Table A5.10: Contingency table comparing smoking habits between dysplasia and OSCC cohorts. 0= Non smoker 1=equal to or less than 20packyears 2= 20packyears**

| | | Cohorts | | Total |
|---|---|---|---|---|
| | | Dysplasia | OSCC | |
| Alcohol | 0 | 19 | 5 | 24 |
| | 1 | 32 | 25 | 57 |
| | 2 | 8 | 16 | 24 |
| Total | | 59 | 46 | 105 |

Fisher's exact (2-sided) p=0.005

**Table A5.11: Contingency table comparing alcohol intake between dysplasia and OSCC cohorts. 0=No alcohol 1=less than or equal to 28u/week 2=greater than 28u/week**

| | | Cohort | | Total |
|---|---|---|---|---|
| | | Dysplasia | OSCC | |
| Gender | Male | 34 | 39 | 73 |
| | Female | 26 | 22 | 48 |
| Total | | 60 | 61 | 121 |

Pearson Chi square p=0.414

**Table A5.12: Contingency table comparing gender ratio between dysplasia and OSCC cohorts.**

# Appendix 6: Diagnosis groups and duplex qMSP data

| | | Diagnosis | | | | | |
|---|---|---|---|---|---|---|---|
| | | normal | mild dysplasia | moderate dysplasia | PVL | Severe dysplasia | MISCC |
| *CCNA1* | Neg | 97 | 15 | 20 | 5 | 14 | 5 |
| | Pos | 0 | 0 | 0 | 0 | 0 | 0 |
| Total | | 97 | 15 | 20 | 5 | 14 | 5 |

| | | Diagnosis | Total |
|---|---|---|---|
| | | OSCC | |
| *CCNA1* | Neg | 53 | 209 |
| | Pos | 9 | 9 |
| Total | | 62 | 218 |

Fisher's exact (2-sided) p=0.003

**Table A6.1: Contingency table of methylation status of *CCNA1* gene versus saliva sample diagnosis groups**

| | | Diagnosis | | | | | |
|---|---|---|---|---|---|---|---|
| | | normal | mild dysplasia | moderate dysplasia | PVL | Severe dysplasia | MISCC |
| *CYGB* | Neg | 97 | 15 | 20 | 5 | 13 | 5 |
| | Pos | 0 | 0 | 0 | 0 | 1 | 0 |
| Total | | 97 | 15 | 20 | 5 | 14 | 5 |

| | | Diagnosis | Total |
|---|---|---|---|
| | | OSCC | |
| *CYGB* | Neg | 52 | 207 |
| | Pos | 10 | 11 |
| Total | | 62 | 218 |

Fisher's Exact (2-sided) p=0.001

**Table A6.2: Contingency table of methylation status of *CYGB* gene versus saliva sample diagnosis groups**

| | | Diagnosis | | | | | |
|---|---|---|---|---|---|---|---|
| | | normal | mild dysplasia | moderate dysplasia | PVL | Severe dysplasia | MISCC |
| *P16* | Neg | 94 | 13 | 19 | 4 | 10 | 5 |
| | Pos | 3 | 2 | 1 | 1 | 4 | 0 |
| Total | | 97 | 15 | 20 | 5 | 14 | 5 |

| | | Diagnosis | Total |
|---|---|---|---|
| | | OSCC | |
| *P16* | Neg | 45 | 190 |
| | Pos | 17 | 28 |
| Total | | 62 | 218 |

Fisher's Exact (2-sided) p<0.001

**Table A6.3: Contingency table of methylation status of *P16* gene versus saliva sample diagnosis groups**

| | | Diagnosis | | | | |
|---|---|---|---|---|---|---|
| | | normal | mild dysplasia | moderate dysplasia | PVL | Severe dysplasia |
| *TMEFF2* | Neg | 96 | 15 | 19 | 4 | 13 |
| | Pos | 0 | 0 | 1 | 1 | 1 |
| Total | | 96 | 15 | 20 | 5 | 14 |

| | | Diagnosis | | Total |
|---|---|---|---|---|
| | | MISCC | OSCC | |
| *TMEFF2* | Neg | 4 | 52 | 203 |
| | Pos | 1 | 10 | 14 |
| Total | | 5 | 62 | 217 |

Fisher's Exact (2-sided) p<0.001

**Table A6.4: Contingency table of methylation status of *TMEFF2* gene versus saliva sample diagnosis groups**

## Appendix 7: Binned diagnosis and duplex qMSP data

| | | Binned diagnosis | | | | Total |
|---|---|---|---|---|---|---|
| | | Normal | Low risk | High risk | Malignancy | |
| *CCNA1* | Neg | 97 | 36 | 18 | 58 | 209 |
| | Pos | 0 | 0 | 0 | 9 | 9 |
| Total | | 97 | 36 | 18 | 67 | 218 |

Fisher's Exact (2-sided) p<0.001

**Table A7.1: Contingency table of methylation status of *CCNA1* versus binned diagnosis groups in saliva samples**

| | | Binned diagnosis | | | | Total |
|---|---|---|---|---|---|---|
| | | Normal | Low risk | High risk | Malignancy | |
| *CYGB* | Neg | 97 | 36 | 17 | 57 | 207 |
| | Pos | 0 | 0 | 1 | 10 | 11 |
| Total | | 97 | 36 | 18 | 67 | 218 |

Fisher's Exact (2-sided) p<0.001

**Table A7.2: Contingency table of methylation status of *CYGB* versus binned diagnosis groups in saliva samples**

| | | Binned diagnosis | | | | Total |
|---|---|---|---|---|---|---|
| | | Normal | Low risk | High risk | Malignancy | |
| *P16* | Neg | 94 | 33 | 13 | 50 | 190 |
| | Pos | 3 | 3 | 5 | 17 | 28 |
| Total | | 97 | 36 | 18 | 67 | 218 |

Fisher's Exact(2-sided) p<0.001

**Table A7.3: Contingency table of methylation status of *P16* versus binned diagnosis groups in saliva samples**

| | | Binned diagnosis | | | | Total |
|---|---|---|---|---|---|---|
| | | Normal | Low risk | High risk | Malignancy | |
| *TMEFF2* | Neg | 96 | 35 | 16 | 56 | 203 |
| | Pos | 0 | 1 | 2 | 11 | 14 |
| Total | | 96 | 36 | 18 | 67 | 217 |

Fisher's Exact (2-sided)p<0.001

**Table A7.4: Contingency table of methylation status of *TMEFF2* versus binned diagnosis groups in saliva samples**

| | | Binned diagnosis | | | | Total |
|---|---|---|---|---|---|---|
| | | Normal | Low risk | High risk | Malignancy | |
| Gene Panel | Neg | 94 | 32 | 11 | 38 | 175 |
| | Pos | 3 | 4 | 7 | 29 | 43 |
| Total | | 97 | 36 | 18 | 67 | 218 |

Fisher's Exact (2-sided)p<0.001

**Table A7.5: Contingency table of methylation status of gene panel versus binned diagnosis groups in saliva samples**

## Appendix 8: Age and qMSP duplex data

| diagnosis | | | Age groups | | | Total |
|---|---|---|---|---|---|---|
| | | | 40.00 | 60.00 | 80.00 | |
| Normal | *CCNA1* | Neg | 44 | 29 | 24 | 97 |
| | Total | | 44 | 29 | 24 | 97 |
| Low risk | *CCNA1* | Neg | 3 | 21 | 12 | 36 |
| | Total | | 3 | 21 | 12 | 36 |
| High risk | *CCNA1* | Neg | 1 | 7 | 10 | 18 |
| | Total | | 1 | 7 | 10 | 18 |
| Malignancy | *CCNA1* | Neg | 1 | 27 | 26 | 54 |
| | | Pos | 0 | 3 | 5 | 8 |
| | Total | | 1 | 30 | 31 | 62 |
| Total | *CCNA1* | Neg | 49 | 84 | 72 | 205 |
| | | Pos | 0 | 3 | 5 | 8 |
| | Total | | 49 | 87 | 77 | 213 |

High risk: Fisher's Exact (2-sided) p=0.745

Malignancy : Fisher's Exact (2-sided) p=0.194

Total: Fisher's Exact p=0.094

**Table A8.1: Contingency table of age against CCNA1 methylation stratified by diagnostic classification**

| Binned diagnosis | | | Age | | | Total |
|---|---|---|---|---|---|---|
| | | | 40.00 | 60.00 | 80.00 | |
| Normal | *CYGB* | Neg | 44 | 29 | 24 | 97 |
| | Total | | 44 | 29 | 24 | 97 |
| Low risk | *CYGB* | Neg | 3 | 21 | 12 | 36 |
| | Total | | 3 | 21 | 12 | 36 |
| High risk | *CYGB* | Neg | 1 | 7 | 9 | 17 |
| | | Pos | 0 | 0 | 1 | 1 |
| | Total | | 1 | 7 | 10 | 18 |
| Malignancy | *CYGB* | Neg | 1 | 25 | 27 | 53 |
| | | Pos | 0 | 5 | 4 | 9 |
| | Total | | 1 | 30 | 31 | 62 |
| Total | *CYGB* | Neg | 49 | 82 | 72 | 203 |
| | | Pos | 0 | 5 | 5 | 10 |
| | Total | | 49 | 87 | 77 | 213 |

High risk: Fisher's Exact (2-sided) P=1.0

Malignancy: Fisher's Exact (2-sided) p=0.77

Total: Fisher's Exact (2-sided) p=0.215

**Table A8.2: Contingency table of age against *CYGB* methylation stratified by diagnostic classification**

| Binned diagnosis | | | Age | | | Total |
|---|---|---|---|---|---|---|
| | | | 40.00 | 60.00 | 80.00 | |
| Normal | *P16* | Neg | 44 | 27 | 23 | 94 |
| | | Pos | 0 | 2 | 1 | 3 |
| | Total | | 44 | 29 | 24 | 97 |
| Low risk | *P16* | Neg | 3 | 19 | 11 | 33 |
| | | Pos | 0 | 2 | 1 | 3 |
| | Total | | 3 | 21 | 12 | 36 |
| High risk | *P16* | Neg | 1 | 6 | 6 | 13 |
| | | Pos | 0 | 1 | 4 | 5 |
| | Total | | 1 | 7 | 10 | 18 |
| Malignancy | *P16* | Neg | 1 | 21 | 27 | 49 |
| | | Pos | 0 | 9 | 4 | 13 |
| | Total | | 1 | 30 | 31 | 62 |
| Total | *P16* | Neg | 49 | 73 | 67 | 189 |
| | | Pos | 0 | 14 | 10 | 24 |
| | Total | | 49 | 87 | 77 | 213 |

Normal: Fisher's Exact(2-sided) p=0.159

Low risk: Fisher's Exact(2-sided) p=1.0

High risk: Fisher's Exact(2-sided) p=0.522

Malignancy: Fisher's Exact(2-sided) p=0.310

**Table A8.3: Contingency table of age against *P16* methylation stratified by diagnostic classification**

| Binned diagnosis | | | Age | | | Total |
|---|---|---|---|---|---|---|
| | | | 40.00 | 60.00 | 80.00 | |
| Normal | *TMEFF2* | Neg | 44 | 29 | 23 | 96 |
| | Total | | 44 | 29 | 23 | 96 |
| Low risk | *TMEFF2* | Neg | 3 | 21 | 11 | 35 |
| | | Pos | 0 | 0 | 1 | 1 |
| | Total | | 3 | 21 | 12 | 36 |
| High risk | *TMEFF2* | Neg | 1 | 7 | 8 | 16 |
| | | Pos | 0 | 0 | 2 | 2 |
| | Total | | 1 | 7 | 10 | 18 |
| Malignancy | *TMEFF2* | Neg | 1 | 27 | 25 | 53 |
| | | Pos | 0 | 3 | 6 | 9 |
| | Total | | 1 | 30 | 31 | 62 |
| Total | *TMEFF2* | .Neg | 49 | 84 | 67 | 200 |
| | | Pos | 0 | 3 | 9 | 12 |
| | Total | | 49 | 87 | 76 | 212 |

Low risk: Fisher's Exact (2-sided) p=0.417

Highrisk: Fisher's Exact (2-sided) p=0.542

Malignancy: Fisher's Exact (2-sided) p=0.549

Total: Fisher's Exact (2-sided) p=0.014

**Table A8.4: Contingency table of age against *TMEFF2* methylation stratified by diagnostic classification**

| Binned diagnosis | | | Age | | | Total |
|---|---|---|---|---|---|---|
| | | | 40.00 | 60.00 | 80.00 | |
| Normal | Gene Panel | Neg | 44 | 27 | 23 | 94 |
| | | Pos | 0 | 2 | 1 | 3 |
| | Total | | 44 | 29 | 24 | 97 |
| Low risk | Gene Panel | Neg | 3 | 19 | 10 | 32 |
| | | Pos | 0 | 2 | 2 | 4 |
| | Total | | 3 | 21 | 12 | 36 |
| High risk | Gene Panel | Neg | 1 | 6 | 4 | 11 |
| | | Pos | 0 | 1 | 6 | 7 |
| | Total | | 1 | 7 | 10 | 18 |
| Malignancy | Gene Panel | Neg | 1 | 17 | 19 | 37 |
| | | Pos | 0 | 13 | 12 | 25 |
| | Total | | 1 | 30 | 31 | 62 |
| Total | Gene Panel | Neg | 49 | 69 | 56 | 174 |
| | | Pos | 0 | 18 | 21 | 39 |
| | Total | | 49 | 87 | 77 | 213 |

Normal: Fisher's Exact p=0.159

Low risk: Fisher's Exact p=0.729

High risk: Fisher's Exact p=0.095

Malignancy: Fisher's Exact p=0.879

**Table A8.5: Contingency table of age against gene panel methylation stratified by diagnostic classification**

| Binned diagnosis | | | Age | | Total |
|---|---|---|---|---|---|
| | | | 41-60 | 60+ | |
| Normal | Gene Panel New | Neg | 27 | 23 | 50 |
| | | Pos | 2 | 1 | 3 |
| | Total | | 29 | 24 | 53 |
| Low risk | Gene Panel | Neg | 19 | 10 | 29 |
| | | Pos | 2 | 2 | 4 |
| | Total | | 21 | 12 | 33 |
| High risk | Gene Panel | Neg | 6 | 4 | 10 |
| | | Pos | 1 | 6 | 7 |
| | Total | | 7 | 10 | 17 |
| Malignancy | Gene Panel | Neg | 17 | 19 | 36 |
| | | Pos | 13 | 12 | 25 |
| | Total | | 30 | 31 | 61 |
| Total | Gene Panel | Neg | 69 | 56 | 125 |
| | | Pos | 18 | 21 | 39 |
| | Total | | 87 | 77 | 164 |

Normal: Fisher's Exact (2-sided) p=1.0

Lowrisk:Fisher's Exact (2-sided) p=0.464

High Risk:Fisher's Exact (2-sided) p=0.134

Malignancy: Fisher's Exact (2-sided) p=0.797

Total: Pearson Chi square (2-sided) p=0.361

**Table A8.6: Contingency table of age (patients over 40) against gene panel methylation stratified by disease groups**

## Appendix 9: Smoking and qMSP data

| Binned diagnosis | | | Smoking | | | Total |
|---|---|---|---|---|---|---|
| | | | None | =< 20PY | =>20PY | |
| Normal | *CCNA1* | Neg | 38 | 42 | 17 | 97 |
| | Total | | 38 | 42 | 17 | 97 |
| Low risk | *CCNA1* | Neg | 9 | 9 | 18 | 36 |
| | Total | | 9 | 9 | 18 | 36 |
| High risk | *CCNA1* | Neg | 5 | 5 | 8 | 18 |
| | Total | | 5 | 5 | 8 | 18 |
| Malignancy | *CCNA1* | Neg | 7 | 11 | 27 | 45 |
| | | Pos | 3 | 1 | 3 | 7 |
| | Total | | 10 | 12 | 30 | 52 |
| Total | *CCNA1* | Neg | 59 | 67 | 70 | 196 |
| | | Pos | 3 | 1 | 3 | 7 |
| | Total | | 62 | 68 | 73 | 203 |

Malignancy: Fisher's exact (2-sided) p=0.209

Total: Fisher's exact (2-sided) p=0.631

**Table A9.1: Contingency table of smoking habit against methylation of *CCNA1* stratified by diagnostic groups**

| Binned diagnosis | | | Smoking | | | Total |
|---|---|---|---|---|---|---|
| | | | None | =<20 PY | =>20PY | |
| Normal | *CYGB* | Neg | 38 | 42 | 17 | 97 |
| | Total | | 38 | 42 | 17 | 97 |
| Low risk | *CYGB* | Neg | 9 | 9 | 18 | 36 |
| | Total | | 9 | 9 | 18 | 36 |
| High risk | *CYGB* | Neg | 5 | 4 | 8 | 17 |
| | | Pos | 0 | 1 | 0 | 1 |
| | Total | | 5 | 5 | 8 | 18 |
| Malignancy | *CYGB* | Neg | 8 | 11 | 24 | 43 |
| | | Pos | 2 | 1 | 6 | 9 |
| | Total | | 10 | 12 | 30 | 52 |
| Total | *CYGB* | Neg | 60 | 66 | 67 | 193 |
| | | Pos | 2 | 2 | 6 | 10 |
| | Total | | 62 | 68 | 73 | 203 |

High risk: Fisher's exact (2-sided) p=0.556

Malignancy: Fisher's exact (2-sided) p=0.778

Total: Fisher's exact (2-sided) p=0.332

**Table A9.2: Contingency table of smoking habit against methylation of *CYGB* stratified by diagnostic groups**

| Binned diagnosis | | | Smoking | | | Total |
|---|---|---|---|---|---|---|
| | | | None | =<20 PY | =>20PY | |
| None | *P16* | Neg | 38 | 42 | 14 | 94 |
| | | Pos | 0 | 0 | 3 | 3 |
| | Total | | 38 | 42 | 17 | 97 |
| Low risk | *P16* | Neg | 8 | 9 | 16 | 33 |
| | | Pos | 1 | 0 | 2 | 3 |
| | Total | | 9 | 9 | 18 | 36 |
| High risk | *P16* | Neg | 2 | 3 | 8 | 13 |
| | | Pos | 3 | 2 | 0 | 5 |
| | Total | | 5 | 5 | 8 | 18 |
| Malignancy | *P16* | Neg | 7 | 8 | 24 | 39 |
| | | Pos | 3 | 4 | 6 | 13 |
| | Total | | 10 | 12 | 30 | 52 |
| Total | *P16* | Neg | 55 | 62 | 62 | 179 |
| | | Pos | 7 | 6 | 11 | 24 |
| | Total | | 62 | 68 | 73 | 203 |

**Normal: Fisher's exact (2-sided)** **p=0.005**

**Low risk: Fisher's exact (2-sided)** **p=0.796**

**High risk: Fisher's exact (2-sided)** **p=0.045**

**Malignancy: Fisher's exact (2-sided)** **p=0.613**

**Total: Pearson chi-square** **p=0.508**

**Table A9.3: Contingency table of smoking habit against methylation of *P16* stratified by diagnostic group**

| Binned diagnosis | | | smoking | | | Total |
|---|---|---|---|---|---|---|
| | | | .0 | 1.0 | 2.0 | |
| Normal | *TMEFF2* | Neg | 37 | 42 | 17 | 96 |
| | Total | | 37 | 42 | 17 | 96 |
| Low risk | *TMEFF2* | Neg | 8 | 9 | 18 | 35 |
| | | Pos | 1 | 0 | 0 | 1 |
| | Total | | 9 | 9 | 18 | 36 |
| High risk | *TMEFF2* | Neg | 5 | 4 | 7 | 16 |
| | | Pos | 0 | 1 | 1 | 2 |
| | Total | | 5 | 5 | 8 | 18 |
| Malignancy | *TMEFF2* | Neg | 7 | 9 | 28 | 44 |
| | | Pos | 3 | 3 | 2 | 8 |
| | Total | | 10 | 12 | 30 | 52 |
| Total | *TMEFF2* | Neg | 57 | 64 | 70 | 191 |
| | | Pos | 4 | 4 | 3 | 11 |
| | Total | | 61 | 68 | 73 | 202 |

Low risk: Fisher's exact (2-sided) p=0.5

High risk: Fisher's exact (2-sided) p=1.0

Malignancy: Fisher's exact (2-sided) p=0.85

Total: Fisher's exact (2-sided) p=0.798

**Table A9.4: Contingency table of smoking habit against methylation of *TMEFF2* stratified by diagnostic groups**

| Binned diagnosis | | | Smoking | | | Total |
|---|---|---|---|---|---|---|
| | | | None | =<20PY | >20PY | |
| Normal | Gene Panel | Neg | 38 | 42 | 14 | 94 |
| | | Pos | 0 | 0 | 3 | 3 |
| | Total | | 38 | 42 | 17 | 97 |
| Low risk | Gene Panel | Neg | 7 | 9 | 16 | 32 |
| | | Pos | 2 | 0 | 2 | 4 |
| | Total | | 9 | 9 | 18 | 36 |
| High risk | Gene Panel | Neg | 2 | 2 | 7 | 11 |
| | | Pos | 3 | 3 | 1 | 7 |
| | Total | | 5 | 5 | 8 | 18 |
| Malignancy | Gene Panel | Neg | 4 | 5 | 20 | 29 |
| | | Pos | 6 | 7 | 10 | 23 |
| | Total | | 10 | 12 | 30 | 52 |
| Total | Gene Panel | Neg | 51 | 58 | 57 | 166 |
| | | Pos | 11 | 10 | 16 | 37 |
| | Total | | 62 | 68 | 73 | 203 |

Normal: Fisher's exact (2-sided) p=0.005

Low risk: Fisher's exact (2-sided) p=0.342

High risk: Fisher's exact (2-sided) p=0.208

Malignancy: Fisher's exact (2-sided) p=0.152

Total: Fisher's exact (2-sided) p=0.552

**Table A9.5: Contingency table of smoking habit against methylation of gene panel stratified by diagnostic groups**

## Appendix 10: Alcohol consumption and duplex qMSP data

| Binned diagnosis | | | alcohol | | | Total |
|---|---|---|---|---|---|---|
| | | | .0 | 1.0 | 2.0 | |
| Normal | *CCNA1* | Neg | 36 | 59 | 2 | 97 |
| | Total | | 36 | 59 | 2 | 97 |
| Low risk | *CCNA1* | Neg | 10 | 22 | 3 | 35 |
| | Total | | 10 | 22 | 3 | 35 |
| High risk | *CCNA1* | Neg | 6 | 8 | 4 | 18 |
| | Total | | 6 | 8 | 4 | 18 |
| Malignancy | *CCNA1* | Neg | 5 | 24 | 15 | 44 |
| | | Pos | 2 | 3 | 2 | 7 |
| | Total | | 7 | 27 | 17 | 51 |
| Total | *CCNA1* | Neg | 57 | 113 | 24 | 194 |
| | | Pos | 2 | 3 | 2 | 7 |
| | Total | | 59 | 116 | 26 | 201 |

Malignancy: Fisher's exact (2-sided) p=0.455

Total: Fisher's exact (2-sided) p=0.383

**Table A10.1: Contingency table of alcohol consumption against *CCNA1* methylation stratified by diagnosis groups**

| Binned diagnosis | | | alcohol | | | Total |
|---|---|---|---|---|---|---|
| | | | None | =>28u | >28u | |
| Normal | *CYGB* | Neg | 36 | 59 | 2 | 97 |
| | Total | | 36 | 59 | 2 | 97 |
| Low risk | *CYGB* | Neg | 10 | 22 | 3 | 35 |
| | Total | | 10 | 22 | 3 | 35 |
| High risk | *CYGB* | Neg | 6 | 8 | 3 | 17 |
| | | Pos | 0 | 0 | 1 | 1 |
| | Total | | 6 | 8 | 4 | 18 |
| Malignancy | *CYGB* | Neg | 6 | 23 | 13 | 42 |
| | | Pos | 1 | 4 | 4 | 9 |
| | Total | | 7 | 27 | 17 | 51 |
| Total | *CYGB* | Neg | 58 | 112 | 21 | 191 |
| | | Pos | 1 | 4 | 5 | 10 |
| | Total | | 59 | 116 | 26 | 201 |

High risk: Fisher's exact (2-sided) p=0.222

Malignancy: Fisher's exact (2-sided) p=0.874

Total: Fisher's exact (2-sided) p=0.005

**Table A10.2: Contingency table of alcohol consumption against *CYGB* methylation stratified by diagnosis groups**

| Binned diagnosis | | | Alcohol | | | Total |
|---|---|---|---|---|---|---|
| | | | None | =<28u | >28u | |
| Normal | *P16* | Neg | 35 | 58 | 1 | 94 |
| | | Pos | 1 | 1 | 1 | 3 |
| | | Total | 36 | 59 | 2 | 97 |
| Low risk | *P16* | .00 | 10 | 20 | 3 | 33 |
| | | 1.00 | 0 | 2 | 0 | 2 |
| | | Total | 10 | 22 | 3 | 35 |
| High risk | *P16* | .00 | 4 | 6 | 3 | 13 |
| | | 1.00 | 2 | 2 | 1 | 5 |
| | | Total | 6 | 8 | 4 | 18 |
| Malignancy | *P16* | .00 | 6 | 20 | 12 | 38 |
| | | 1.00 | 1 | 7 | 5 | 13 |
| | | Total | 7 | 27 | 17 | 51 |
| Total | *P16* | .00 | 55 | 104 | 19 | 178 |
| | | 1.00 | 4 | 12 | 7 | 23 |
| | | Total | 59 | 116 | 26 | 201 |

Normal: Fisher's exact (2-sided) p=0.061

Low risk: Fisher's exact (2-sided) p=0.630

High risk: Fisher's exact(2-sided) P=1.0

Malignancy: Fisher's exact (2-sided) p=0.907

Total: Fisher's exact (2-sided) p=0.033

**Table A10.3: Contingency table of alcohol consumption against *P16* methylation stratified by diagnosis groups**

| Binned diagnosis | | | Alcohol | | | Total |
|---|---|---|---|---|---|---|
| | | | None | =<28u | >28u | |
| Normal | *TMEFF2* | Neg | 35 | 59 | 2 | 96 |
| | Total | | 35 | 59 | 2 | 96 |
| Low risk | *TMEFF2* | Neg | 10 | 21 | 3 | 34 |
| | | Pos | 0 | 1 | 0 | 1 |
| | Total | | 10 | 22 | 3 | 35 |
| High risk | *TMEFF2* | Neg | 5 | 8 | 3 | 16 |
| | | Pos | 1 | 0 | 1 | 2 |
| | Total | | 6 | 8 | 4 | 18 |
| Malignancy | *TMEFF2* | Neg | 5 | 23 | 15 | 43 |
| | | Pos | 2 | 4 | 2 | 8 |
| | Total | | 7 | 27 | 17 | 51 |
| Total | *TMEFF2* | Neg | 55 | 111 | 23 | 189 |
| | | Pos | 3 | 5 | 3 | 11 |
| | Total | | 58 | 116 | 26 | 200 |

Low risk: Fisher's exact (2-sided) p=1.0

High risk: Fisher's exact (2-sided) p=0.294

Malignancy: Fisher's exact (2-sided) p=0.662

Total: Fisher's exact (2-sided) p=0.364

**Table A10.4: Contingency table of alcohol consumption against *TMEFF2* methylation stratified by diagnosis groups**

| Binned diagnosis | | | alcohol | | | Total |
|---|---|---|---|---|---|---|
| | | | None | =<28u/wk | >28u | |
| Normal | Gene Panel | Neg | 35 | 58 | 1 | 94 |
| | | Pos | 1 | 1 | 1 | 3 |
| | Total | | 36 | 59 | 2 | 97 |
| Low risk | Gene Panel | Neg | 10 | 19 | 3 | 32 |
| | | Pos | 0 | 3 | 0 | 3 |
| | Total | | 10 | 22 | 3 | 35 |
| High risk | Gene Panel | Neg | 3 | 6 | 2 | 11 |
| | | Pos | 3 | 2 | 2 | 7 |
| | Total | | 6 | 8 | 4 | 18 |
| Malignancy | Gene Panel | Neg | 4 | 15 | 9 | 28 |
| | | Pos | 3 | 12 | 8 | 23 |
| | Total | | 7 | 27 | 17 | 51 |
| Total | Gene Panel | Neg | 52 | 98 | 15 | 165 |
| | | Pos | 7 | 18 | 11 | 36 |
| | Total | | 59 | 116 | 26 | 201 |

Normal: Fisher's exact (2-sided) p=0.061

Low risk: Fisher's exact (2-sided) p=0.647

High risk: Fisher's exact (2-sided) p=0.569

Malignancy: Fisher's exact (2-sided) p=1.0

Total: Fisher's exact (2-sided) p=0.005

**Table A10.5: Contingency table of alcohol consumption against gene panel methylation stratified by diagnosis**

## Appendix 11: Pathological stage and qMSP data

| | | pT | | | | Total |
|---|---|---|---|---|---|---|
| | | 1.0 | 2.0 | 3.0 | 4.0 | |
| *CCNA1* | Neg | 6 | 23 | 4 | 15 | 48 |
| | Pos | 0 | 4 | 1 | 3 | 8 |
| Total | | 6 | 27 | 5 | 18 | 56 |

Fisher's exact (2-sided) p=0.831

**Table A11.1: Contingency table of *CCNA1* methylation against tumour size**

| | | pN | | | Total |
|---|---|---|---|---|---|
| | | 0 | 1.0 | 2.0 | |
| *CCNA1* | Neg | 24 | 7 | 17 | 48 |
| | Pos | 3 | 1 | 3 | 7 |
| Total | | 27 | 8 | 20 | 55 |

Fisher's exact (2-sided) p=1.0

**Table A11.2: Contingency table of CCNA1 methylation against nodal status**

| | | ECS | | Total |
|---|---|---|---|---|
| | | No | Yes | |
| *CCNA1* | Neg | 9 | 15 | 24 |
| | Pos | 2 | 2 | 4 |
| Total | | 11 | 17 | 28 |

Fisher's exact (2-sided) p=1.0

**Table A11.3: Contingency table of *CCNA1* methylation against extracapsular spread**

| | | pStage | | | | Total |
|---|---|---|---|---|---|---|
| | | 1.0 | 2.0 | 3.0 | 4.0 | |
| *CCNA1* | Neg | 4 | 12 | 5 | 27 | 48 |
| | Pos | 0 | 0 | 2 | 6 | 8 |
| Total | | 4 | 12 | 7 | 33 | 56 |

Fisher's exact (2-sided) p=0.256

**Table A11.4: Contingency table of *CCNA1* methylation against pathological stage**

| | | pT | | | | Total |
|---|---|---|---|---|---|---|
| | | 1.0 | 2.0 | 3.0 | 4.0 | |
| *CYGB* | Neg | 6 | 24 | 4 | 13 | 47 |
| | Pos | 0 | 3 | 1 | 5 | 9 |
| Total | | 6 | 27 | 5 | 18 | 56 |

Fisher's exact (2-sided) p=0.297

**Table A11.5: Contingency table of CYGB methylation against tumour size**

| | | pN | | | Total |
|---|---|---|---|---|---|
| | | .0 | 1.0 | 2.0 | |
| *CYGB* | Neg | 22 | 7 | 18 | 47 |
| | Pos | 5 | 1 | 2 | 8 |
| Total | | 27 | 8 | 20 | 55 |

Fisher's exact (2-sided) p=0.869

**Table A11.6: Contingency table of *CYGB* methylation against nodal status**

| | | ECS | | Total |
|---|---|---|---|---|
| | | No | Yes | |
| *CYGB* | Neg | 8 | 17 | 25 |
| | Pos | 3 | 0 | 3 |
| Total | | 11 | 17 | 28 |

Fisher's exact (2-sided) p=0.050

**Table A11.7: Contingency table of *CYGB* methylation against extracapsular spread**

| | | pStage | | | | Total |
|---|---|---|---|---|---|---|
| | | 1.0 | 2.0 | 3.0 | 4.0 | |
| *CYGB* | Neg | 4 | 12 | 5 | 26 | 47 |
| | Pos | 0 | 0 | 2 | 7 | 9 |
| Total | | 4 | 12 | 7 | 33 | 56 |

Fisher's exact (2-sided) p=0.188

**Table A11.8: Contingency table of *CYGB* methylation against pathological stage**

| | | pT | | | | Total |
|---|---|---|---|---|---|---|
| | | 1.0 | 2.0 | 3.0 | 4.0 | |
| *P16* | Neg | 5 | 20 | 5 | 13 | 43 |
| | Pos | 1 | 7 | 0 | 5 | 13 |
| Total | | 6 | 27 | 5 | 18 | 56 |

Fisher's exact (2-sided) p=0.767

**Table A11.9: Contingency table of *P16* methylation against tumour size**

| | | pN | | | Total |
|---|---|---|---|---|---|
| | | .0 | 1.0 | 2.0 | |
| *P16* | Neg | 19 | 7 | 16 | 42 |
| | Pos | 8 | 1 | 4 | 13 |
| Total | | 27 | 8 | 20 | 55 |

Fisher's exact (2-sided) p=0.629

**Table A11.10: Contingency table of *P16* methylation against nodal status**

| | | ECS | | Total |
|---|---|---|---|---|
| | | .0 | 1.0 | |
| *P16* | Neg | 9 | 14 | 23 |
| | Pos | 2 | 3 | 5 |
| Total | | 11 | 17 | 28 |

Fisher's exact (2-sided) p=1.0

**Table A11.11: Contingency table of *P16* methylation against extracapsular spread**

| | | pStage | | | | Total |
|---|---|---|---|---|---|---|
| | | 1.0 | 2.0 | 3.0 | 4.0 | |
| *P16* | Neg | 3 | 8 | 7 | 25 | 43 |
| | Pos | 1 | 4 | 0 | 8 | 13 |
| Total | | 4 | 12 | 7 | 33 | 56 |

Fisher's exact (2-sided) p=0.409

**Table A11.12: Contingency table of *P16* methylation against pathological stage**

| | | pT | | | | Total |
|---|---|---|---|---|---|---|
| | | 1.0 | 2.0 | 3.0 | 4.0 | |
| *TMEFF2* | Neg | 6 | 23 | 2 | 17 | 48 |
| | Pos | 0 | 4 | 3 | 1 | 8 |
| Total | | 6 | 27 | 5 | 18 | 56 |

Fisher's exact (2-sided) p=0.036

**Table A11.13: Contingency table of *TMEFF2* methylation against tumour size**

| | | pN | | | Total |
|---|---|---|---|---|---|
| | | .0 | 1.0 | 2.0 | |
| *TMEFF2* | Neg | 24 | 8 | 16 | 48 |
| | Pos | 3 | 0 | 4 | 7 |
| Total | | 27 | 8 | 20 | 55 |

Fisher's exact (2-sided) p=0.486

**Table A11.14: Contingency table of *TMEFF2* methylation against nodal status**

| | | ECS | | Total |
|---|---|---|---|---|
| | | No | Yes | |
| *TMEFF2* | Neg | 9 | 15 | 24 |
| | Pos | 2 | 2 | 4 |
| Total | | 11 | 17 | 28 |

Fisher's exact (2-sided) p=1.0

**Table A11.15: Contingency table of *TMEFF2* methylation against extracapsular spread**

| | | pStage | | | | Total |
|---|---|---|---|---|---|---|
| | | 1.0 | 2.0 | 3.0 | 4.0 | |
| *TMEFF2* | Neg | 4 | 12 | 4 | 28 | 48 |
| | Pos | 0 | 0 | 3 | 5 | 8 |
| Total | | 4 | 12 | 7 | 33 | 56 |

Fisher's exact (2-sided) p=0.078

**Table A11.16: Contingency table of *TMEFF2* methylation against pathological stage**

| | | Stage | | | | Total |
|---|---|---|---|---|---|---|
| | | 1.0 | 2.0 | 3.0 | 4.0 | |
| Number genes positive | .00 | 3 | 7 | 3 | 19 | 32 |
| | 1.00 | 1 | 4 | 1 | 7 | 13 |
| | 2.00 | 0 | 1 | 3 | 3 | 7 |
| | 3.00 | 0 | 0 | 0 | 4 | 4 |
| Total | | 4 | 12 | 7 | 33 | 56 |

Fisher's exact (2-sided) p=0.557

**Table A11.17: Contingency table of number of methylated genes against pathological stage**

| | | pT | | | | Total |
|---|---|---|---|---|---|---|
| | | 1.0 | 2.0 | 3.0 | 4.0 | |
| Gene Panel | Neg | 5 | 15 | 2 | 10 | 32 |
| | Pos | 1 | 12 | 3 | 8 | 24 |
| Total | | 6 | 27 | 5 | 18 | 56 |

Fisher's exact (2-sided)p=0.551

**Table A11.18: Contingency table of gene panel methylation against tumour size**

| | | pN | | | Total |
|---|---|---|---|---|---|
| | | .0 | 1.0 | 2.0 | |
| Gene Panel | Neg | 13 | 6 | 13 | 32 |
| | Pos | 14 | 2 | 7 | 23 |
| Total | | 27 | 8 | 20 | 55 |

Fisher's exact (2-sided)p=0.321

**Table A11.19: Contingency table of number of gene panel methylation against nodal status**

| | | ECS | | Total |
|---|---|---|---|---|
| | | No | Yes | |
| Gene Panel | Neg | 7 | 12 | 19 |
| | Pos | 4 | 5 | 9 |
| Total | | 11 | 17 | 28 |

Fisher's exact (2-sided)p=1.0

**Table A11.20: Contingency table of gene panel methylation against extracapsular spread**

| | | pStage | | | | Total |
|---|---|---|---|---|---|---|
| | | 1.0 | 2.0 | 3.0 | 4.0 | |
| Gene Panel | Neg | 3 | 7 | 3 | 19 | 32 |
| | Pos | 1 | 5 | 4 | 14 | 24 |
| Total | | 4 | 12 | 7 | 33 | 56 |

Fisher's exact p=0.80

**Table A11.21: Contingency table of gene panel methylation against pathological stage**

## Appendix 12: Survival and qMSP data

| | | Fate | | | Total |
|---|---|---|---|---|---|
| | | Alive | DOD | DOC | |
| *CCNA1* | Neg | 30 | 11 | 4 | 45 |
| | Pos | 4 | 3 | 1 | 8 |
| Total | | 34 | 14 | 5 | 53 |

Fisher's exact (2-sided) p=0.499

**Table A12.1: Contingency table of *CCNA1* methylation against fate. DOD=died of disease, DOC=died of other cause**

| | | Fate | | | Total |
|---|---|---|---|---|---|
| | | Alive | DOD | DOC | |
| *CYGB* | Neg | 28 | 12 | 4 | 44 |
| | Pos | 6 | 2 | 1 | 9 |
| Total | | 34 | 14 | 5 | 53 |

Fisher's exact (2-sided) p=1.0

**Table A12.2: Contingency table of *CYGB* methylation against fate. DOD=died of disease, DOC=died of other cause**

| | | Fate | | | Total |
|---|---|---|---|---|---|
| | | Alive | DOD | DOC | |
| *P16* | Neg | 26 | 11 | 4 | 41 |
| | Pos | 8 | 3 | 1 | 12 |
| Total | | 34 | 14 | 5 | 53 |

Fisher's exact (2-sided) p=1.0

**Table A12.3: Contingency table of *P16* methylation against fate. DOD=died of disease, DOC=died of other cause**

| | | Fate | | | Total |
|---|---|---|---|---|---|
| | | Alive | DOD | DOC | |
| *TMEFF2* | Neg | 29 | 12 | 4 | 45 |
| | Pos | 5 | 2 | 1 | 8 |
| Total | | 34 | 14 | 5 | 53 |

Fisher's exact (2-sided) p=1.0

**Table A12.4: Contingency table of *TMEFF2* methylation against fate. DOD=died of disease, DOC=died of other cause**

| | | Fate | | | Total |
|---|---|---|---|---|---|
| | | Alive | DOD | DOC | |
| Gene Panel | Neg | 20 | 8 | 2 | 30 |
| | Pos | 14 | 6 | 3 | 23 |
| Total | | 34 | 14 | 5 | 53 |

Fisher's exact (2-sided) p=0.763

**Table A12.5: Contingency table of gene panel methylation against fate. DOD=died of disease, DOC=died of other cause**

## Appendix 13: Recurrence and duplex qMSP data

| | | Recurrence | | Total |
|---|---|---|---|---|
| | | No | Yes | |
| *CCNA1* | Neg | 28 | 18 | 46 |
| | Pos | 4 | 4 | 8 |
| Total | | 32 | 22 | 54 |

Fisher's exact (2-sided) p=0.702

**Table A13.1: Contingency table of *CCNA1* methylation against recurrence**

| | | Recurrence | | Total |
|---|---|---|---|---|
| | | No | Yes | |
| *CYGB* | Neg | 27 | 18 | 45 |
| | Pos | 5 | 4 | 9 |
| Total | | 32 | 22 | 54 |

Fisher's exact (2-sided) p=1.0

**Table A13.2: Contingency table of *CYGB* methylation against recurrence**

| | | Recurrence | | Total |
|---|---|---|---|---|
| | | No | Yes | |
| *P16* | Neg | 24 | 17 | 41 |
| | Pos | 8 | 5 | 13 |
| Total | | 32 | 22 | 54 |

Fisher's exact (2-sided) p=1.0

**Table A13.3: Contingency table of *P16* methylation against recurrence**

| | | Recurrence | | Total |
|---|---|---|---|---|
| | | No | Yes | |
| *TMEFF2* | .00 | 28 | 19 | 47 |
| | 1.00 | 4 | 3 | 7 |
| Total | | 32 | 22 | 54 |

Fisher's exact (2-sided) p=1.0

**Table A13.4: Contingency table of *TMEFF2* methylation against recurrence**

| | | recurrence | | Total |
|---|---|---|---|---|
| | | No | Yes | |
| Gene Panel | Neg | 19 | 12 | 31 |
| | Pos | 13 | 10 | 23 |
| Total | | 32 | 22 | 54 |

Pearson chi square p=0.724

**Table A13.5: Contingency table of gene panel methylation against recurrence**

# Appendix 14: Kaplan Meier survival curves and qMSP data

**Case Processing Summary**

| *CCNA1* | Total N | N of Events | Censored | |
|---|---|---|---|---|
| | | | N | % |
| .00 | 45 | 11 | 34 | 75.6 % |
| 1.00 | 8 | 3 | 5 | 62.5 % |
| Overall | 53 | 14 | 39 | 73.6 % |

**Overall Comparisons**

| | Chi-Square | df | Sig. |
|---|---|---|---|
| Log Rank (Mantel-Cox) | .434 | 1 | .510 |

**Table A14.1: Test of equality of survival distributions for the different levels of *CCNA1*.**

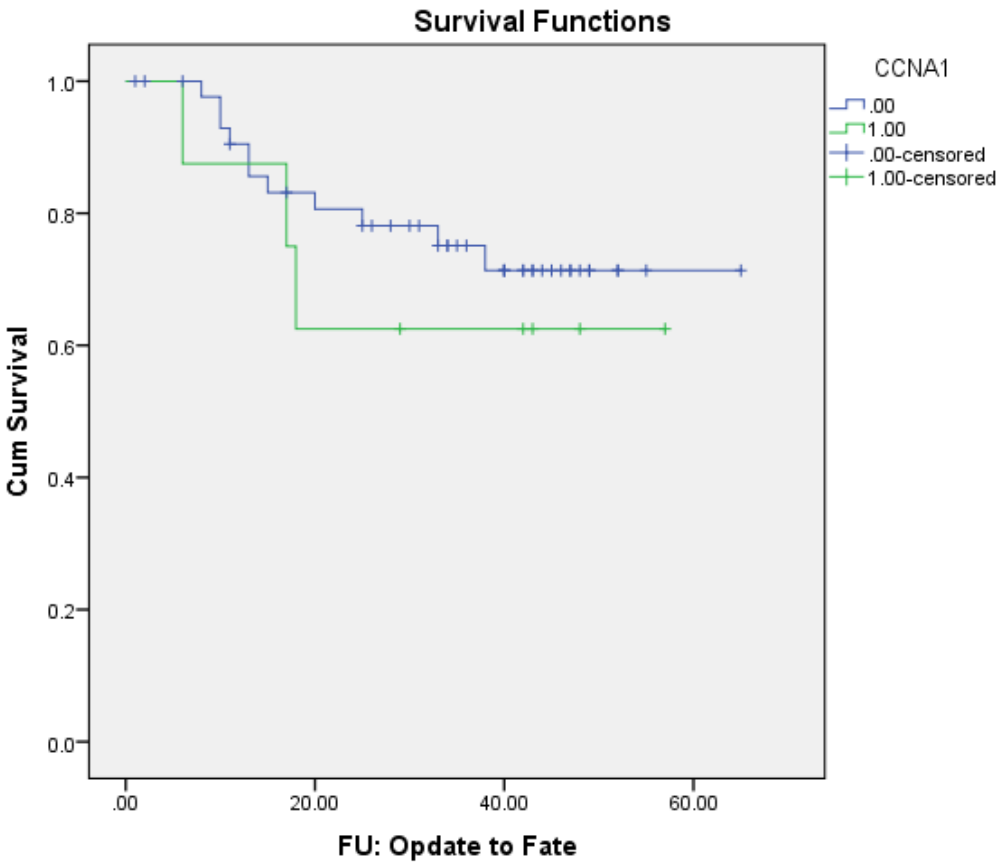



**Figure FI4.1: Kaplan Meier survival curve for *CCNA1* 0=unmethylated *CCNA1* and 1=methylated *CCNA1*. FU=operation date to fate (death of disease)**

**Case Processing Summary**

| *CYGB* | Total N | N of Events | Censored | |
|---|---|---|---|---|
| | | | N | % |
| .00 | 44 | 12 | 32 | 72.7 % |
| 1.00 | 9 | 2 | 7 | 77.8 % |
| Overall | 53 | 14 | 39 | 73.6 % |

**Overall Comparisons**

| | Chi-Square | df | Sig. |
|---|---|---|---|
| Log Rank (Mantel-Cox) | .290 | 1 | .591 |

**Table A14.2: Test of equality of survival distributions for the different levels of *CYGB*.**

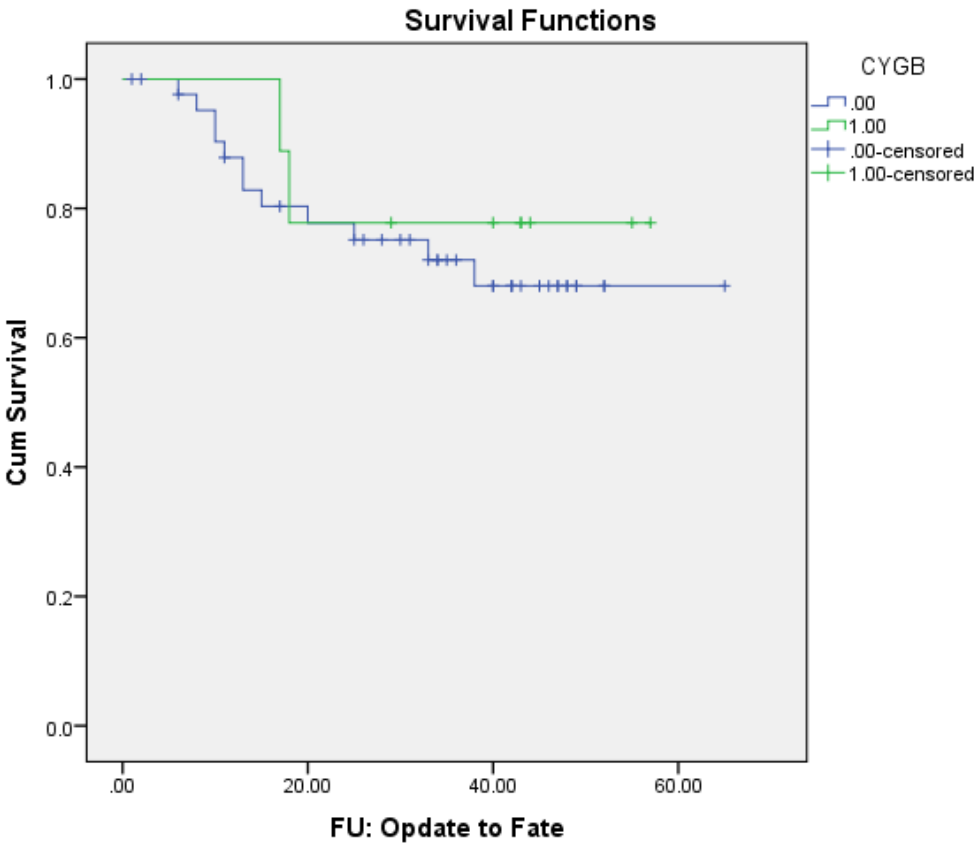



**Figure F14.2: Kaplan Meier survival curve for *CYGB*. 0=unmethylated *CYGB* and 1=methylated *CYGB*. FU=operation date to fate (death of disease)**

**Case Processing Summary**

| P16 | Total N | N of Events | Censored | |
|---|---|---|---|---|
| | | | N | % |
| Neg | 41 | 11 | 30 | 73.2 % |
| Pos | 12 | 3 | 9 | 75.0 % |
| Overall | 53 | 14 | 39 | 73.6 % |

**Overall Comparisons**

| | Chi-Square | df | Sig. |
|---|---|---|---|
| Log Rank (Mantel-Cox) | .135 | 1 | .713 |

**Table A14.3: Test of equality of survival distributions for the different levels of *P16***

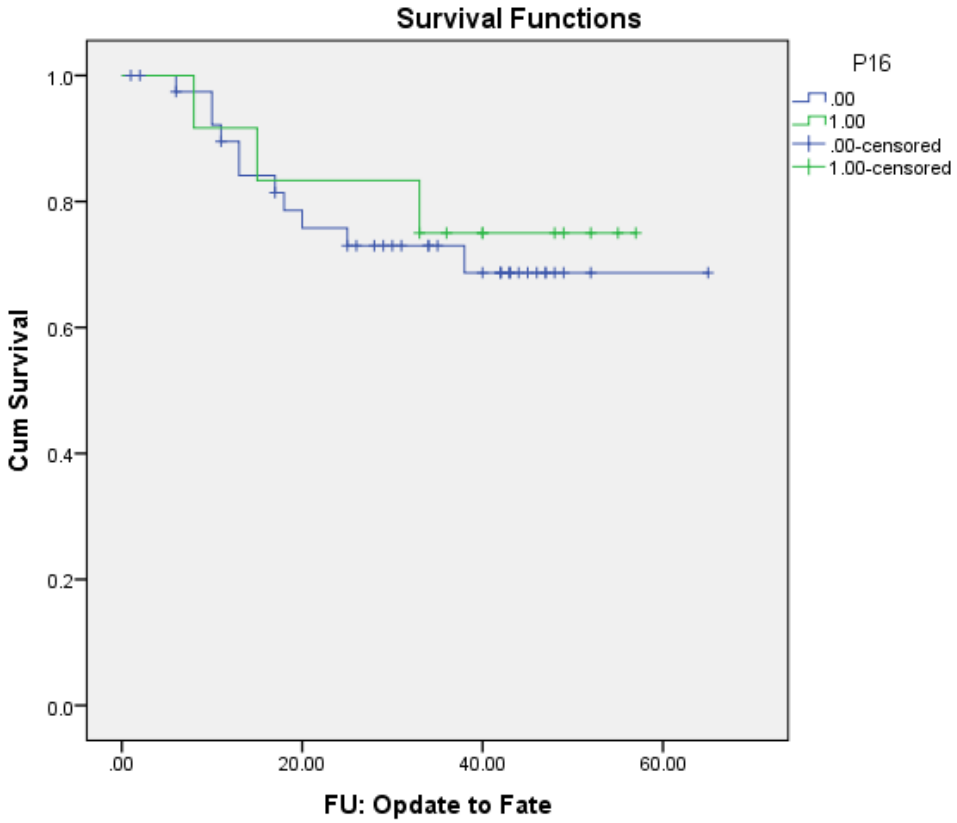



**Figure 14.3: Kaplan Meier survival curve for *P16*. 0=unmethylated *P16* and 1=methylated *P16*. FU=operation date to fate (death of disease)**

**Case Processing Summary**

| *TMEFF2* | Total N | N of Events | Censored | |
|---|---|---|---|---|
| | | | N | % |
| Neg | 45 | 12 | 33 | 73.3 % |
| Pos | 8 | 2 | 6 | 75.0 % |
| Overall | 53 | 14 | 39 | 73.6 % |

**Overall Comparisons**

| | Chi-Square | df | Sig. |
|---|---|---|---|
| Log Rank (Mantel-Cox) | .023 | 1 | .879 |

**Table A14.4: Test of equality of survival distributions for the different levels of *TMEFF2***

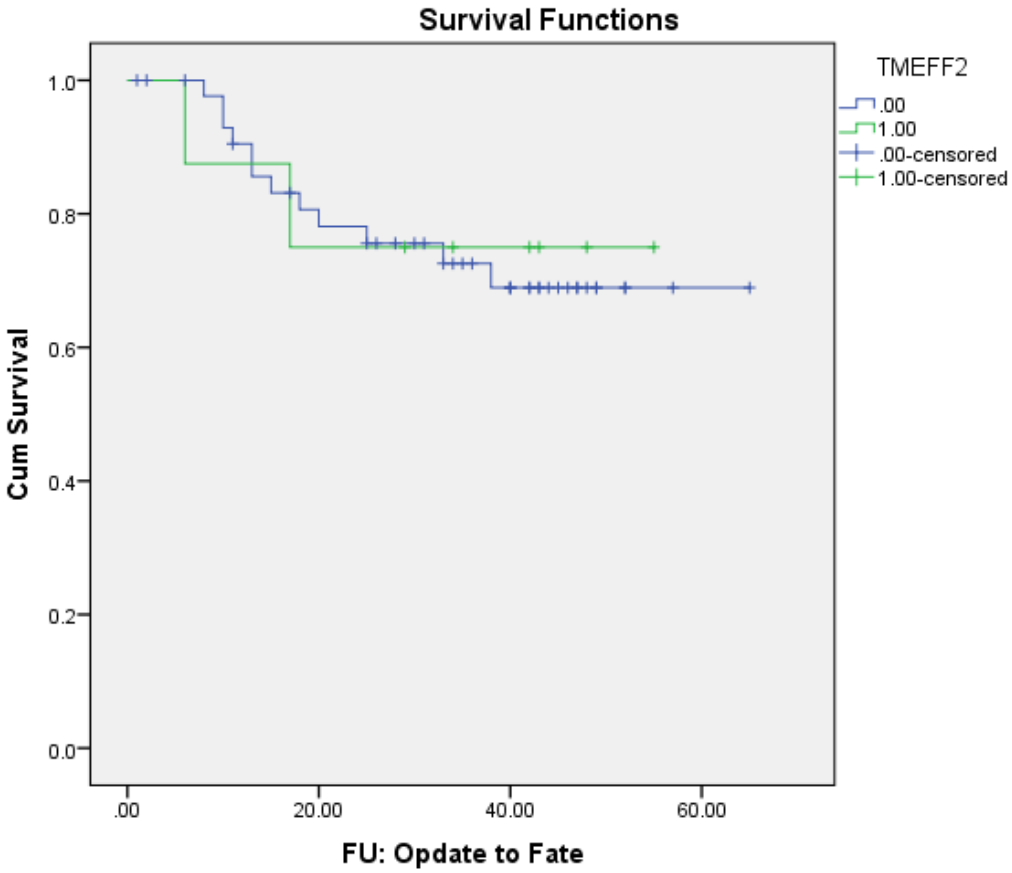



**Figure F14.4: Kaplan Meier survival curve for *TMEFF2*. 0=unmethylated *TMEFF2* and 1=methylated *TMEFF2*. FU=operation date to fate (death of disease)**

| Gene Panel | Total N | N of Events | Censored | |
|---|---|---|---|---|
| | | | N | % |
| Neg | 30 | 8 | 22 | 73.3 % |
| Pos | 23 | 6 | 17 | 73.9 % |
| Overall | 53 | 14 | 39 | 73.6 % |

**Overall Comparisons**

| | Chi-Square | df | Sig. |
|---|---|---|---|
| Log Rank (Mantel-Cox) | .067 | 1 | .795 |

**Table A14.5: Test of equality of survival distributions for the different levels of Gene Panel.**

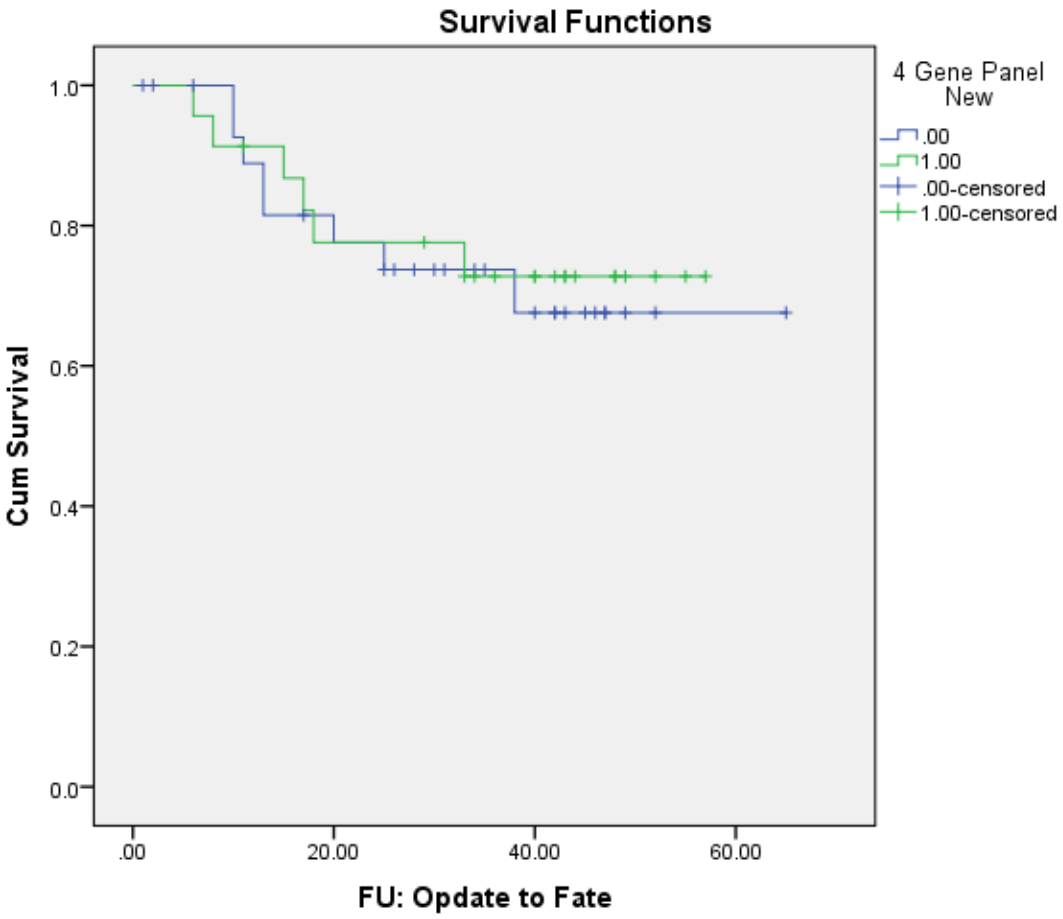



**Figure F14.5: Kaplan Meier survival curve for gene panel. 0=unmethylated gene panel and 1=methylated gene panel. FU=operation date to fate (death of disease)**