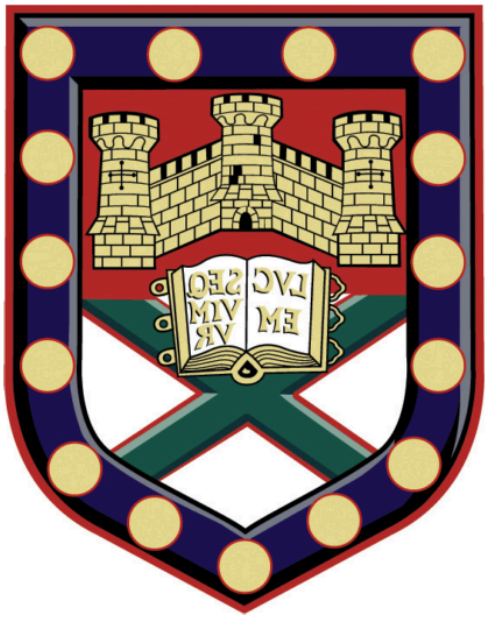
UNIVERSITY OF EXETER



VIBRATIONAL SPECTROSCOPY FOR THE ASSESSMENT OF VULVAL DISEASE

JONATHAN A FROST

PhD THESIS

SUPERVISORS: PROF. AC SHORE, PROF. N STONE, DR C KENDALL, DR G LLOYD, MISS K HILLABY AND MR R GORNALL.

August 2019

VIBRATIONAL SPECTROSCOPY FOR THE ASSESSMENT OF VULVAL DISEASE

Submitted by Jonathan Antony Frost to the University of Exeter

As a thesis for the degree of Doctor of Philosophy in Medical Studies

August 2019

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: ……..

Abstract

Vibrational spectroscopic diagnostic techniques have significant potential to improve the care of women with benign, premalignant and malignant vulval diseases by reducing the reliance on traditional biopsy and histopathology. These techniques also have the potential to augment clinicians' ability to differentiate different types of vulval disease at the time of surgery for neoplastic vulval disease. In addition, vibrational spectroscopic techniques offer the opportunity to assess molecular changes associated with the development of vulval cancer that are not apparent on routine histopathological assessment.

The work outlined in this thesis evaluates the role of emerging techniques in vibrational spectroscopy to address this need within three key themes:

1. Development of a vibrational spectroscopic diagnostic technique to reduce the reliance on traditional biopsy and histopathological diagnosis.

2. Development of a vibrational spectroscopic diagnostic technique for improving the delineation of disease margins at the time of surgery for pre-malignant and malignant vulval conditions.

3. Evaluation of a vibrational spectroscopic tool for augmenting and automating aspects of vulval histopathology.

Raman spectroscopic mapping of 91 fresh frozen vulval tissue sections combined with multivariate spectral analysis was used to demonstrate that malignant vulval disease could be differentiated from non-neoplastic and premalignant vulval disease with a sensitivity of 97% and specificity of 78%. The technique was then tested in experimental conditions closer to in-vivo application, measuring spectra from 91 whole fresh frozen tissue blocks using microscope and probe Raman systems. This

demonstrated the technique could differentiate malignant from non-neoplastic and premalignant vulval disease with sensitivities of 84% to 92% and specificities of 84% to 64% respectively. In a separate investigation vulval tissue blocks from 27 women with suspected lichen sclerosus underwent Raman spectroscopic point measurements. Multivariate analysis demonstrated Raman spectroscopy could be used to differentiate lichen sclerosus from other vulval disorders with a similar clinical appearance with a sensitivity of sensitivity of 91% and specificity of 80%.

Fourier transform infrared (FTIR) spectroscopic mapping of 93 fixed paraffin embedded tissue sections was used to demonstrate that malignant vulval disease could be differentiated from non-neoplastic and premalignant with vulval disease with an approximate sensitivity of 100% and specificity of 79%. In addition FTIR spectroscopy was used to differentiate molecular changes in vulval intraepithelial neoplasia (VIN) and lichen sclerosus (LS) found in association with vulval squamous cell carcinoma (SCC). Analysis of FTIR spectroscopic tissue maps from 48 patients demonstrated the technique could differentiate LS associated with SCC with a sensitivity of approximately 100% and specificity of 84% and VIN associated with SCC with a sensitivity of approximately 100% and specificity 58%.

This thesis demonstrates the considerable potential of vibrational spectroscopy in this clinical setting. The research has made significant progress in each of the three themes outlined above and indicates that further work is warranted to develop the techniques towards routine clinical application.

Acknowledgements

I would like to thank the Gloucestershire Hospitals NHS Foundation Trust Research and Innovation Forum, Cheltenham General Hospital Gynaecological Oncology Research Fund and the British Society for the Study of Vulval Disease whose funding made this study possible.

I would also like to thank everyone at the Biophotonics Research Unit at Gloucester Royal Hospital. Every member of the team has been a great source of help and support. Special thanks are due to Gavin Lloyd for his tireless efforts in teaching me how to code in Matlab® and for the many productive and enjoyable discussions that were essential throughout the project.

Thanks also to the physics departments at the University of Exeter and University of Bristol for the support with Raman probe development. In particular thanks are due to John Day, Alex Dudgeon and the physics engineering team.

I would like to thank the research and clinical teams at the Royal United Hospital in Bath, Musgrove Park Hospital in Taunton and Royal Devon and Exeter Hospital in Exeter for their support of this project.

I would also like to thank the clinical and nursing teams in Gloucester Royal Hospital and Cheltenham General Hospital who went out of their way to alert me to potential participants in the study and who facilitated the collection of clinical samples.

I am very appreciative of the efforts of the pathology department at Gloucestershire Hospitals NHS Foundation Trust, Jo Motte for her expert sample preparation and Linmarie Ludeman for screening and assessing the large volume of histological specimens required in this study.

My supervisors have been a delight to work with and I would like to thank each of them for their contribution to this project: Kathryn Hillaby and Rob Gornall for their much needed clinical experience and perspective; Catherine Kendall for her tireless support and attention to detail; Nick Stone for his vision and ideas and Angela Shore for the much needed experienced external view point.

I must finally thank my wife Lucy and children Alice and Rowan for their eternal patience and support.

This thesis is dedicated to Lucy, Alice and Rowan.

Contents

List of Figures

List of Tables

List of Abbreviations

ANOVA – Analysis of variance

AUC – Area under curve

CCD – Charge-coupled device

DNA – Deoxyribonucleic acid

EMSC – Extended multiplicative signal correction

FCM – Fluorescence confocal laser scanning microscopy

FTIRS - Fourier transform infrared spectroscopy

H&E – Haematoxylin and eosin

LDA – Linear discriminant analysis

LN – Lymph Nodes

LOOCV – Leave one out cross validation

LS – Lichen sclerosus

NICE – National Institute of Clinical Excellence

NIR – Near infra-red

OCT – Optical coherence tomography

PCA – Principal component analysis

PTFE – Polytetrafluoroethylene

RCM – Reflectance confocal microscopy

ROC – Receiver operator curve

RS – Raman spectroscopy

SCC – Squamous cell carcinoma

SNR – Signal to noise ratio

VIN – Vulval intraepithelial neoplasia

Section A: Literature Review

Chapter 1 The Vulva and Vulval Disease

1.1 The normal vulva

1.1.1 Anatomy of the vulva

The human vulva consists of the female genitalia including and immediately external to the hymen of the vagina. It is composed of the mons pubis, clitoris and external urethral meatus anteriorly; the labia minora and labia majora laterally and the fourchette posteriorly (Figure 1-1) (Standring, 2016).

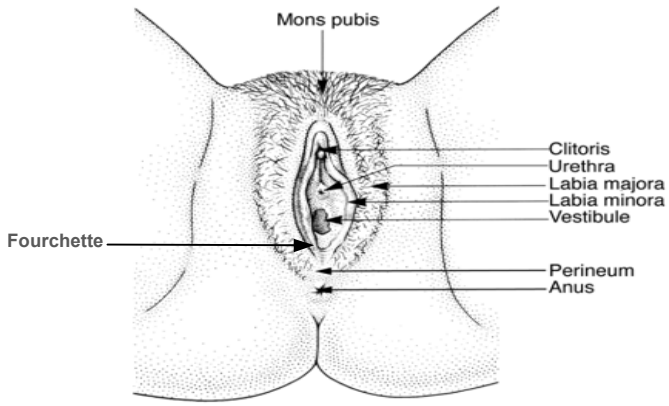


Figure 1-1 – Illustration of the anatomy of the vulva (Image courtesy of the The Association for Lichen Sclerosus and Vulval Health)

The epithelium covering the vulva displays significant regional differences in anatomical and histological structure due to the variation in embryonic derivation. Embryonic ectoderm forms the skin bearing areas of the mons pubis, labia majora and perineum whereas the vulval vestibule is formed of embryonic endoderm (Berek, 2012).

1.1.1.1 Vulval vestibule

The area between the hymen of the vagina, the labia minora, clitoris and posterior fourchette is termed the vulval vestibule. The vestibule is formed from the embryonic endoderm of the urogenital sinus. The area is non hair bearing and is covered in non keratinised stratified squamous epithelium containing several mucous secreting glands including Bartholin's, Skene's and periurethral glands as well as the urethral and vaginal orifices. The hymen forms the vaginal border of vestibule and consists of folds of fibrous tissue covered in non keratinised stratified squamous epithelium. The junction between the non keratinised epithelium of the vestibule and the medial surface of the keratinised epithelium of the labia minora is marked by Hart's line (Farage and Maibach, 2016; Standring, 2016).

1.1.1.2 Mons pubis and vulval labia

The mons pubis or mons veneris is a rounded protuberance anterior to the pubic symphysis comprising an assemblage of adipose tissue covered by keratinised stratified squamous epithelium. Superiorly the mons pubis borders the anterior

abdominal wall and inferiorly it joins the anterior labial commissure, the most anterior portion of the labia majora.

The labia majora extend posteriorly from the anterior commissure to form two prominent swellings on either side of the vulva. The prominence of the labia majora is created by subdermal fibro-adipose tissue deposition and this diminishes with age as the fat atrophies. The labia majora like the mons pubis are covered with keratinised stratified squamous epithelium containing sweat glands and sebaceous glands. The lateral surface of the labia majora are rugose, pigmented and contains hair follicles whilst the inner surface is smooth and hair free. The labia majora diminish posteriorly and merge with the surrounding skin below the level of the fourchette.

The labia minora are two lesser cutaneous folds lying medial to the labia majora extending laterally either side of the vaginal orifice. Unlike the labia majora they are composed of dense subdermal connective tissue without adipose deposition. Anteriorly each labia minora divides with the lower part forming the frenulum of the clitoris and the upper part forming the clitoral sheath. The posterior part of the labia minora surrounds the vulval vestibule and fuse at the fourchette.

The perineum lies between the vulval fourchette with its associated labia and the anus and consists of the hair bearing keratinised stratified squamous epithelium common to other parts of the vulva (Baggish and Karram, 2016; Farage and Maibach, 2016).

1.1.2 Histology of the normal vulva

1.1.2.1 Normal human skin

Normal human skin consists of three key layers. The most superficial layer is the epidermis, this covers the dermis, underneath which lies the hypodermis. The histological and biochemical structure of the epithelium covering the vulva closely resembles skin found elsewhere on the body with a few adaptive differences (Frienkel and Woodley, 2001).

1.1.2.1.1 Epidermis

The epidermis is the outmost layer of cells and is composed of stratified squamous epithelium. The epidermis varies in thickness depending on its position on the body from around 40 μm on the eyelid up to 1 mm in area of high mechanical stress (e.g. palms and soles of feet) (Jianhua Zhao *et al.*, 2008). The epidermis forms a barrier for water and protects the body from the external environment.

Keratinocytes are the principal cell type of the epidermis. These cells mature as they progress from the basal layer of the epidermis towards the surface of the skin. The epidermis is divided into four strata, the stratum germinosum, the stratum spinosum, the stratum granulosum and the stratum corneum (Figure 1-2). The stratum germinosum (single layer of cells) contains mitotically active cells necessary for the regeneration of the epidermis. After mitotic division the newly formed cells undergo progressive maturation as they migrate toward the surface of the skin. The stratum

spinosum (8-10 layers of cells) is the next layer the newly formed cells enter. The cells within this layer have many desmosomes covering their surface. These desmosomes anchor the cells together and contain thick tufts of keratinous filaments. Within the next most superficial layer (stratum granulosum, 3-5 layers of cells) cells begin to lose their nuclei and cytoplasmic organelles and develop into the keratinised squames of the stratum corneum. The stratum corneum is the outer most layer and contains non viable but biochemically active corneocytes densely packed with keratin (Frienkel and Woodley, 2001; Gartner, 2017).

The layers of the epidermis contain differing amounts of keratin. The stratum germinosum contains 30% keratin whilst the stratum corneum is 80% keratin. Studies have shown that epithelial keratins can be grouped into acidic (type I) or neutral-basic (type II) (Moll, 1991; Tobin, 2006). Cells express at least one type of type I keratin in combination with a type II partner keratin. Proliferative cells within the epidermis are associated with K5 and K14 keratin whilst maturing cells undergoing keratinocyte differentiation are typically associated with K1, K2, K10 and K11 keratin.

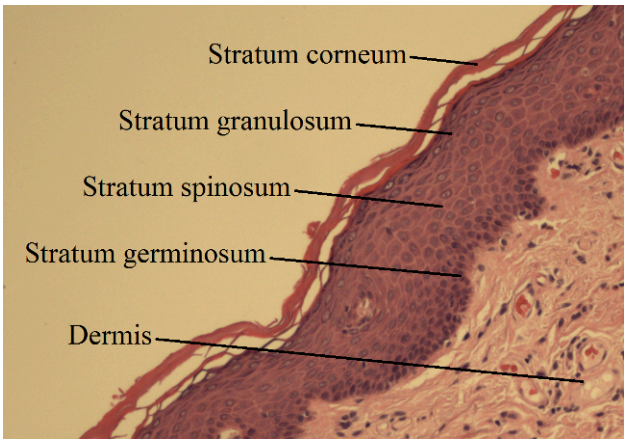


Figure 1-2 – H&E stained tissue section of normal vulval skin

The dermis is separated from the epidermis by a thin basement membrane. The epidermis contains extensions called Rete ridges that project into the dermis between dermal papillae. These dermal papillae form part of the more superficial of the two zones of the dermis, the papillary dermis. Beneath this lies the reticular layer that constitutes the bulk of the dermis.

1.1.2.1.2 Dermis

The dermis is a fibroelastic connective tissue which primarily contains fibroblasts which secrete collagen, elastin and other substances necessary for the support of the skin (Caspers, 2003). The dermis merges with the hypodermis that lies underneath the dermis. The hypodermis contains mostly adipose tissue and sweat glands.

The major structural proteins within the dermis are collagen and elastin (Table 1-1) and these are responsible for the tensile strength of human skin. Collagen represents around 75% of the dry weight of the dermis (Caspers *et al.*, 1998).

Dermal collagen is primarily type I (85-90%), type III (8-11%) and type V (3%) (Tobin, 2006). Collagens are crucial in ensuring the epidermis is anchored to the dermis. Collagen VII anchors the basement membrane to the extracellular matrix of the dermis and collagen XVII links basal keratinocytes with the basement membrane. The second major structural dermal protein is elastin. Elastin fibres can be extended by more than their full resting length without breaking and provide skin with its elasticity.

The portion of the extracellular matrix of the dermis not consisting of collagen or elastin fibres is termed the extrafibrillar matrix. This extrafibrillar matrix consists primarily of water, proteoglycans, glycoproteins and hyaluronic acid. Proteoglycans are formed by binding glycosaminoglycans to a protein. The main glycosaminoglycans in the dermis are hyaluronic acid, chondroitin sulphate, keratin sulphate, heparan sulphate, dermatan sulphate and heparin. An example of a key proteoglycan in the dermis is versican which contributes to skin tautness (Tobin, 2006).

Table 1-1 – Primary biomolecular constituents of human epidermis and dermis

Epidermis	Dermis
Keratin	Collagen (types I and III)
Urea	Elastin
Phenylalanine	Lipids
Ceramid	hyaluronic acid
Cholesterol	chondroitin sulphate
Lactate	keratin sulphate
Carotenoids	heparan sulphate
Lipids	dermatan sulphate
	heparan

1.1.2.1.3 Skin pigmentation

Melanin is the primary skin pigment responsible for variations in skin colour. Melanin synthesis occurs within melanosomes found within melanocytes in epidermis (Slominski *et al.*, 2004; Tobin, 2006). Melanocyte pigmentation is affected by various intrinsic and extrinsic factors which include body location, ethnicity, gender, age, chemical exposure and exposure to ultraviolet radiation.

1.1.2.2 Histology of the labia majora and mons pubis

The histological appearance of the epithelium within the labia majora and mons pubis closely resembles skin found elsewhere on the body with keratinised stratified squamous epithelium containing hair follicles with eccrine, apocrine and sebaceous glands (Burkitt, Young and Heath, 1995).

1.1.2.3 Histology of the labia minora, vulval vestibule and hymen

The epithelium of the labia minora, vulval vestibule and hymen is also a stratified squamous epithelium however it contains minimal keratinisation. This epithelium does not have hair follicles or underlying adipose tissue but still contains eccrine, apocrine and sebaceous glands.

1.1.3 Lymphatics drainage of the vulva

Vulval lymphatics drain to the superficial inguinal nodes and then to the deep inguinal lymph nodes (LNs) with the exception of clitoris that drains directly into the deep inguinal nodes. The deep inguinal nodes drain to the external iliac nodes and on to the paraaortic lymph nodes (Farage and Maibach, 2016).

1.1.3.1 The histology of lymph nodes

Lymph nodes are small bean shaped organs connected by lymphatic vessels. Lymph nodes have two main functional regions, the cortex and medulla, surrounded by an outer collagenous capsule. The capsule is punctured by several afferent lymphatic channels bringing incoming lymph. A single efferent channel at the hilum drains the lymph from the node. Inguino-femoral lymph nodes typically contain a large concentration of lipids compared with lymph nodes elsewhere in the body making them difficult to distinguish from surrounding fatty tissue (Gartner, 2017).

1.2 Benign diseases of the vulva

Benign vulval skin disorders are common and community based studies indicate that about one fifth of all women report significant vulval symptoms (ACOG, 2008). These symptoms and signs, including pruritis, pain and alterations in skin appearance are frequent reasons for presentation to primary and secondary care. Within secondary care the commonest causes of vulval symptoms are benign conditions such as vulval dermatitis, lichen simplex, vulval candidiasis, lichen sclerosus and lichen planus (ACOG, 2008).

Several of these benign vulval disorders have long term implications for the affected women both due to the direct effects of the disease and the risk of malignant transformation. In lichen sclerosus and lichen planus the risk of developing invasive vulval cancer is increased and long term surveillance is required (Carli *et al.*, 1995; Carlson *et al.*, 1998; Powell and Wojnarowska, 1999; RCOG, 2014).

1.2.1 Vulval lichen sclerosus

Lichen sclerosus (LS) also known as leucoplakia, lichen albus and lichen sclerosus et atrophicus is an inflammatory condition that can affect any skin site but is most commonly found in the vulva (Powell and Wojnarowska, 1999). The true prevalence of LS is unknown as however community studies in the elderly have demonstrated a prevalence as high as 3% (Leibovitz *et al.*, 2000). The aetiology of LS is unknown although there is some evidence to suggest that the lichen sclerosus has an autoimmune aetiology (Meyrick Thomas *et al.*, 1988). It has been observed that

auto immune conditions are more prevalent in women with lichen sclerosus compared unaffected women (Harrington and Dunsmore, 1981). In addition to this antibodies to extracellular proteins have been demonstrated in women with this condition (Oyama *et al.*, 2003).

Typical symptoms of lichen sclerosus include pruritis, irritation and soreness although the condition can be asymptomatic. Lichen sclerosus typically appears pale and atrophic although blistering, erosions, purpura and hyperkeratosis can be present. Vulval scarring leads to destruction of normal anatomical features and can result in fusion of the labia, stenosis of the interoitus as well as burying of the clitoris (Powell and Wojnarowska, 1999). This scarring can lead to significant physical, psychological and psychosexual morbidity in those affected (Haefner *et al.*, 2014). LS is complicated by development of squamous cell carcinoma in approximately 4% of women (Carlson *et al.*, 1998; Powell and Wojnarowska, 1999). This compares to a background risk of vulval cancer in the UK of 0.3% (Cancer Research UK, 2016).

Vulval lichen sclerosus may occur at any age but typically either appears in prepubertal girls or post menopausal women (Meffert, Davis and Grimwood, 1995). Diagnosis can be clinical although biopsy and histology is necessary if the diagnosis is uncertain or if there is a failure to respond to adequate treatment (Neill and Lewis, 2010). Histology typically shows a thinned epidermis, sub-epidermal hyalinisation with a deeper inflammatory infiltrate (Figure 1-3).

The mainstay of treatment is ultrapotent topical corticosteroids ointment applied regularly to the affected area. This treatment produces remission of symptoms in

ninety-five percent of women and maintenance treatment has the potential to the reduce the risk of progression to vulval cancer without significant adverse effects (Neill and Lewis, 2010; Chi *et al.*, 2012; Fistarol and Itin, 2013).

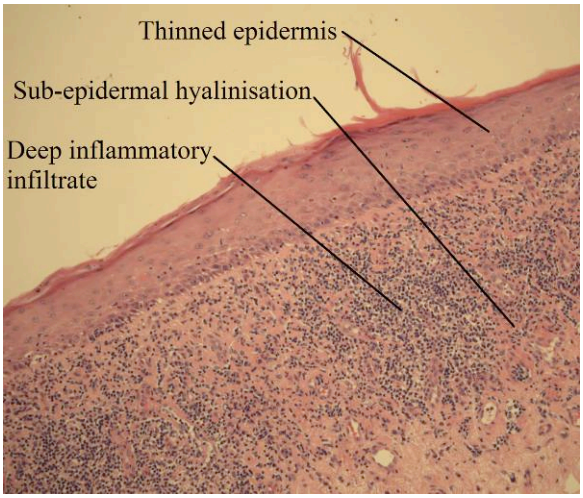


Figure 1-3 – H&E stained tissue section demonstrating histological changes in lichen sclerosus

After initial treatment patients require long term follow up either with regular review or self monitoring by the patient due the increased risk of progression to squamous cell carcinoma (RCOG, 2014).

1.2.2 Vulval lichen planus

Vulval lichen planus is a rare inflammatory condition of unknown aetiology that typically presents with vulval pruritis, irritation and soreness although similar to lichen sclerosus the condition can be asymptomatic. As with lichen sclerosus there is a small risk of progression to squamous cell carcinoma in patients with lichen planus (Cooper and Wojnarowska, 2006).

Diagnosis is based on the characteristic but varied clinical appearance. When the diagnosis is in doubt vulval biopsy is required. Histology shows irregular saw-toothed acanthosis, band like dermal infiltrate (primarily lymphocytic) and increased granular layer and basal cell liquefaction. Treatment is with topical corticosteroids ointment applied regularly to the affected area. As with lichen sclerosus patients often require long term follow up either with regular review or self monitoring by the patient (Cooper and Wojnarowska, 2006).

1.2.3 Vulval dermatitis

As with skin elsewhere in body the vulva can be affected by irritant, allergic, atopic or seborrhoeic dermatitis. Other, secondary causes of dermatitis include iron deficiency and infection. Diagnosis of vulval dermatitis is primarily based on clinical presentation and response to treatment. The condition may be mistaken for lichen sclerosus in some cases and clinical assessment can be supplemented by vulval biopsy if the diagnosis is not clear. After treatment long term follow up is not required provided the patient is asymptomatic as there is no significant increased risk of progression to vulval cancer (Bourke, Coulson and English, 2009).

1.3 Vulval intraepithelial neoplasia (VIN)

Vulval intraepithelial neoplasia (VIN) is a premalignant vulval skin condition with a reported risk of malignant transformation as high as 40-60% for untreated disease (Jones, Rowan and Stewart, 2005; Van Seters, Van Beurden and de Craen, 2005).

For this reason long term follow up of this condition is required (RCOG, 2014). VIN can affect women of any age, however a recent systematic review has suggested the mean age at diagnosis is 46 years (Van Seters, Van Beurden and de Craen, 2005). Women under 50 years comprise 75% of cases and the incidence in this age group is increasing (Jones, Rowan and Stewart, 2005; Judson *et al.*, 2006). Typical symptoms include pruritis and vulvodynia, which can be severe. VIN presents with multifocal disease in 80% of affected women and in these women premalignant neoplastic changes are more likely to be found elsewhere in the lower genital tract (van Beurden *et al.*, 1995).

The incidence of VIN is difficult to assess, as many women will not seek medical assessment of their vulval symptoms. In the United States the incidence has been estimated at 3 per 100,000 women (Judson *et al.*, 2006). VIN is often first detected alongside concurrent vulval SCC and it is postulated that the recent increase in vulval cancer in women under the age of 50 is linked to an increasing incidence of VIN (Jones, Baranyai and Stables, 1997; Eva *et al.*, 2009).

1.3.1 Classification of VIN

The categorisation of VIN has changed in recent years. Traditionally various names have been used for VIN type disease including Queyrat's erthroplasia, carcinoma simplex, early vulval cancer, carcinoma in situ (CIS), vulval atypia and dysplasia. The nomenclature was simplified in 1976 by the International Society for the Study of Vulvovaginal Diseases (ISSVD) and the terms vulvar atypia and carcinoma in situ were suggested (ISSVD, 1976). Following this, in 1986, the ISSVD introduced the

term VIN to replace the previous classifications. Initially the term VIN was split into: VIN 1 for mild dysplastic changes; VIN 2 for moderate dysplasia and VIN 3 to encompass severe dysplasia, carcinoma in situ and differentiated VIN. In light of a lack of evidence that VIN 1-3 exist in a biological or morphological continuum the classification was changed by the ISSVD in 2005 to encompass two types of VIN, usual type and differentiated type. These two types of VIN are thought to differ in pathogenesis, morphology and prognosis (Table 1-2). Both these types of VIN are considered to be high grade (VIN2-3) and use of the term VIN 1 is no longer supported. Newer terminology has now been recommended by the ISSVD in line with other HPV related anogenital lesions, however this terminology has not yet been widely adopted. Within the new terminology usual type VIN is renamed high-grade squamous intraepithelial lesion (HSIL) and the terminology for HPV independent (differentiated) VIN remains unchanged (ISSVD, 2015).

Table 1-2 – Features of the primary histological subtypes of VIN, adapted from Walker et al. 2012 (Walker and Mathews, 2012)

	Usual Type VIN (HSIL)	Differentiated VIN
Age group	Premenopausal	Postmenopausal
Overall Proportion of VIN (%)	95%	5%
HPV Association	Yes	No
Risk Factors	Smoking Immunosuppression	
Distribution	Often multifocal	Usually unifocal
Malignant Progression	Warty, basaloid SCC	Keratinising SCC

1.3.2 Diagnosis of VIN

Currently both the diagnosis of VIN and monitoring of patients with VIN are limited to visual assessment often supplemented by invasive tissue biopsies. The appearance

of VIN is highly variable and this can make it difficult to diagnose on clinical assessment alone. The lesions may appear as subtle pale or pigmented patches or more discrete papules and macules. There may be a single lesion or multiple lesions and they may be present in different anatomical locations. VIN on the keratinised outmost surface of the vulval typically appear as lichenified or hyperkeratotic plaque. By contrast lesions on the non-keratinised surfaces typically appear as pink or red macules. Pigmentation of lesions is also present in 10-15% of cases of VIN (Walker and Mathews, 2012).

Acetic acid applied to the vulva is often used as an adjunct to clinical examination. After application lesions appear pale due to the reversible precipitation of nuclear proteins. This technique lacks specificity as aceto-white changes may represent not only areas of VIN but also viral changes associated with HPV and Epstein-Barr virus, areas of tissue trauma, ulcers and erosions. The aceto-white reaction also varies in intensity in different anatomical locations and between patients making it an unreliable test in isolation (Maclean and Reid, 2011).

1% toluidine blue dye applied to the vulva has also been proposed as a diagnostic tool for the detection of premalignant and malignant vulval disease. Toluidine blue stains nuclei and increased staining is seen when it is applied to areas of neoplastic disease where cell nuclei are found closer to the surface. In one study a positive toluidine blue stain detected VIN with a sensitivity of 59% to 100% with specificity of 40% to 85% depending on the level of staining interpreted as a positive test (Joura *et*

al., 1999). The subjective nature of the test and the low positive predictive value limits is utility in routine clinical practice.

Cytology has been successfully employed on the cervix as a less invasive traditional technique for identifying those with premalignant conditions. Unfortunately as many vulval lesions have significant surface keratosis, cytological specimens from the vulval often consist of mainly enucleated squamous cells making accurate cytological diagnosis difficult. Studies assessing the use of cytology for the diagnosis of VIN reported that in 18% to 26% of cases the collected sample was inadequate for analysis (Levine *et al.*, 2001; Van den Einden *et al.*, 2012). The reported sensitivity of cytology for the detection of malignant and premalignant lesions varies from 32% to 96% with the study reporting the highest sensitivity also reporting a specificity of 30% (Levine *et al.*, 2001; Bae-Jump, Bauer and Van Le, 2007; Van den Einden *et al.*, 2012). In addition, only half of vulval carcinomas penetrate through the outermost dyskeratotic layer of skin present in both VIN and vulval carcinoma making differentiation between premalignant and malignant disease difficult (Nauth and Schilke, 1982). The low positive predictive value of the technique and its limitations in detecting invasive disease has meant cytology has yet to be established as a reliable technique for diagnosing vulval conditions (Smith, 2010).

In practice, as the appearance of VIN can be subtle and occult malignancy can be found in up to 22% of women with VIN, a high degree of suspicion and low threshold for tissue biopsy is usually applied (Modesitt *et al.*, 1998).

1.3.3 Usual type VIN

Usual type VIN (uVIN) includes warty, basaloid and mixed lesions and is associated with oncogenic human papillomavirus (HPV) infection, commonly HPV 16 and less often HPV 18 or 33 (De Vuyst *et al.*, 2009)(Hørding *et al.*, 1995). Risk factors for the development of uVIN include human immune deficiency virus infection and tobacco smoking. uVIN is found predominantly in younger women in their fifth decade and is more likely to present as multifocal vulval disease with concurrent intraepithelial neoplasia in the cervix, vagina or anus. A major concern in those with uVIN is the development of vulval SCC and uVIN is a precursor for the majority of basaloid and warty type vulval SCC (Hampl and Sarajuuri, 2006; Van der Avoort *et al.*, 2006; Reyes and Cooper, 2014). The rising incidence of vulval cancer is primarily due to an increase in vulval cancer in women aged 70 and below, suggestive of an increase in HPV related disease. This increase in HPV related disease may also account for the recent increase in the incidence of uVIN (Iversen and Tretli, 1998; Joura *et al.*, 2000). The recent introduction of routine immunisation against high-risk HPV subtypes may result in a decrease in the incidence of uVIN although the impact is not yet clear.

Low grade uVIN usually presents as white macules which may be multifocal whereas high grade uVIN usually presents as multifocal plaques. A minority of women present with hyperpigmented lesions.

Increase rates of progression to invasive disease and recurrence of uVIN have been linked to multifocal disease (Fehr *et al.*, 2013), large lesions (Wallbillich *et al.*, 2012)

and smoking (Fehr *et al.*, 2013). Immunocompromised women have also been reported to suffer from a higher rate of recurrence and progression to SCC (Wallbillich *et al.*, 2012; Fehr *et al.*, 2013).

1.3.3.1 Histological features of uVIN

Typical histological features of all types of uVIN include full thickness or near full thickness atypia with the loss of epithelial cell maturation associated with pleomorphism, cellular crowding, nuclear hyperchromasia, and abnormal mitotic figures.

The characteristic features of warty type VIN is a condylomatous appearance, hyperkeratosis, parakeratosis, cellular pleomorphism with abnormal cell maturation. Rete ridges are typically wide and deep. Multinucleation, acanthosis, koilocytosis, corps rounds and abnormal mitotic figure are also often found (Nucci, 2013; Gartner, 2017).

Basaloid VIN typically has a thickened epithelium which is flat and non papillomatous. There is a proliferation of undifferentiated cells with a uniform basaloid appearance and numerous mitotic figures. Corps rounds and koilocytotic cells may be present, however these are less prominent than with warty type VIN.

Both basaloid and warty VIN can extend into the sub epithelial skin appendages and this should not be classified as invasion. When warty and basaloid VIN are both present in a single lesion this is called mixed type VIN.

Immunostains for p16 and MIB-1 can be used to support the histological diagnosis of uVIN. P16 staining can be used as a surrogate marker for HPV infection and MIB-1 staining demonstrates the prevalence of the Ki-67 protein, a marker for cellular proliferation (Gartner, 2017).

1.3.4 Differentiated type VIN

Differentiated VIN (dVIN) accounts for less than 5% of VIN and in contrast to uVIN is not associated with HPV infection. It is more commonly found in older postmenopausal women. dVIN, is less commonly multifocal and is associated with mutations in the p53 tumour suppressor gene. dVIN typically develops in women with chronic vulval dermatoses typically lichen sclerosus and lichen simplex (Del Pino, Rodriguez-Carunchio and Ordi, 2013a)(Rolfe *et al.*, 2003). There is a higher risk of progression to vulval SCC in dVIN compared with uVIN. dVIN does not have a distinctive clinical appearance and it is not unusual for dVIN to be first diagnosed adjacent to a more clinically recognisable vulval SCC.

1.3.4.1 Histological features of dVIN

In dVIN cellular atypia is confined to the basal layers of the epithelium and may be difficult to distinguish. The characteristic histological features of differentiated type VIN include prominent eosinophilic cells within the parabasal and basal area usually associated with keratin formation and characteristic ‘pearl-like’ changes in the rete ridges. The premature differentiation of keratinocytes gives them a characteristic appearance with comparatively large vesicular nuclei with prominent nucleoli (Fox

and Wells, 2003). The extensive cellular differentiation makes this type of VIN difficult to recognise and it can be mistaken for more benign disease. When histological classification is difficult, staining for p53 protein can be helpful. P53 staining of an area of dVIN typically shows prominent, diffuse nuclear staining in basal and parabasal cells. This increasing in staining has been shown to be concomitant with somatic p53 mutations, highlighting the different pathogenesis of dVIN (Gartner, 2017).

1.3.5 Management of VIN

The majority of patients with VIN are affected by distressing vulval complaints in addition to being at increased risk of developing vulval SCC and for these reasons the VIN is usually actively managed (Shylasree *et al.*, 2008; Dominiak-Felden *et al.*, 2013). Traditionally the mainstay of therapy has been surgical excision although treatment is multimodal and topical medical therapy and laser ablation are accepted treatments.

Surgical excision can be mutilating and disfiguring and is associated with somatopsychic reactions and sexual dysfunction proportional to the size of the excision (Shylasree *et al.*, 2008). For this reason surgical excision of VIN remains standard management for small unifocal lesions (Edwards *et al.*, 2015). Disease excision in cases of multifocal disease requires extensive surgery and is not an optimal treatment due to the associated morbidity and adverse impact on quality of life (Aerts *et al.*, 2012; Edwards *et al.*, 2015). Following surgical excision reported recurrence rates are around 15% for women with negative surgical margins and 50%

for women with positive surgical margins (Jones, Rowan and Stewart, 2005; Van Seters, Van Beurden and de Craen, 2005).

Ablation of the vulval epithelium with for example, carbon dioxide laser vaporisation may reduce the morbidity of the treatment and produced preferable cosmetic results but a greater risk of treatment failure has been reported (Wallbillich *et al.*, 2012). In addition no tissue is available for histological analysis and multiple pre-procedure biopsies are required to exclude invasive disease.

Currently the extent of the surgical excision, or ablation, is guided by visual assessment and traditional tissue biopsies, as there is currently no established near-real-time non-invasive diagnostic technique for assessing vulval disease.

Historically numerous topical medical treatments have been used in the management of VIN. Agents including 5-flurouracil, bleomycin, trinitrochlorobenzene and alpha-interferon have all been used and for the most part disregarded either due to inefficacy or unacceptable side effects (Ta *et al.*, 2016). More recently imiquimod, an immune response modulator, has been demonstrated to be able to achieve complete response at six months after treatment, however the long term safety and efficacy has not been established (Kaushik *et al.*, 2015). Cidofovir a potent antiviral agent used to treat anal intraepithelial neoplasia has also been evaluated as a treatment of VIN. Initial studies have demonstrated comparable efficacy as imiquimod with a favourable safety and side effect profile, although long term data is awaited (Tristram *et al.*, 2014).

Photodynamic therapy (PDT) where light is used is conjunction with a photosensitising agent to elicit cell death has been used to treat VIN. The technique is well tolerated with good preservation of the vulval anatomy, however produces inconsistent results (Martin-Hirsch *et al.*, 1998; Hillemanns *et al.*, 2000; Winters *et al.*, 2008). A recent non randomised phase 2 trial evaluating imiquimod combined with PDT has indicated promising results with 64% of women displaying histological and clinical response to treatment (Winters *et al.*, 2008).

Vaccination with the bivalent or quadrivalent HPV vaccine has been demonstrated to reduce the risk of HPV related uVIN. Interestingly vaccination may also have a role in the treatment of uVIN. Initial studies have suggested that a synthetic peptide vaccination against oncogenic HPV 16 can clear HPV infection and produce a clinical and histological response in women with HPV 16 positive uVIN (Kenter *et al.*, 2009).

Amongst all the available treatment modalities there is no clear standard management of VIN as treatment choice must be tailored to the sub type of disease and the individual needs of the patient. Factors including the location and extent of the disease along with the patient's wishes will guide appropriate management. In many women a combination of treatment modalities will be core to ensuring remission of disease whilst minimising adverse effects of treatment, however if invasive disease is suspected clinically despite a biopsy indicating VIN then surgical excision is the standard management.

1.3.6 Guidance for the surveillance of VIN

Current guidance recommends that women with a history of VIN be seen in a specialist gynaecology clinic every six months indefinitely (RCOG, 2014). At each appointment symptoms are explored and a full vulval examination performed. If any suspicious lesions are seen then vulval biopsy may be performed.

1.4 Malignant disease of the vulva

Vulval cancer is primarily a disease of elderly women with the mean age at diagnosis being approximately 70 years. The incidence of vulval cancer is rising with an annual incidence rate of approximately 4 per 100,000 women within the UK (Office for National Statistics, 2013). Population based studies indicate that majority of vulval malignancies are squamous cell carcinomas (Platz and Benda, 1995). Other tumour types found on the vulva include malignant melanoma and rarely adenocarcinoma typically arising from Bartholin's glands.

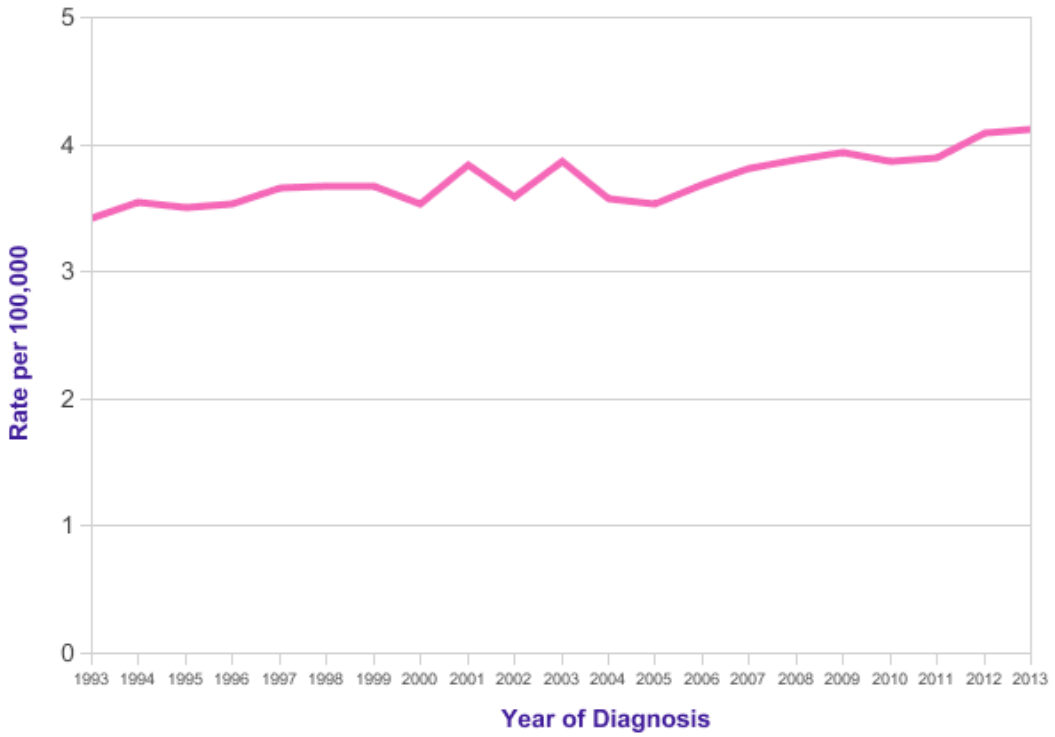


Figure 1-4 – Incidence of vulval cancer, Office for National Statistics. Cancer Statistics: Registrations Series. (Office for National Statistics, 2013)

1.4.1 Squamous cell carcinoma

Squamous cell carcinoma (SCC) is the commonest of all vulval malignancies accounting for more than 90% of all vulval malignancies and 4% of all gynaecological malignancies and there is evidence the incidence is rising (Judson *et al.*, 2006). The incidence of the disease increases with age and it is most common in women over 75 years old.

Women with vulval SCC typically present with an ulcerated, indurated or exophytic mass with pruritis, pain and/ or bleeding. Broadly vulval SCCs can be separated into two main groups, those associated with oncogenic HPV and those not associated with HPV infection. Tumours associated with HPV infection are more commonly

seen in younger women and SCC is often found in association with VIN or lichen sclerosus. The disease spreads by direct extension to adjacent structures in the vulva, vagina, rectum, deeper tissues or bone. Lymphatic spread is to the femoral and inguinal nodes.

Multiple morphological subtypes of invasive SCC are described, including non-keratinising, keratinising, basaloid, warty, spindled and verrucous. The main clinically relevant distinction is between verrucous and non-verrucous subtypes.

Verrucous carcinoma is rare and typically found in women in their ninth decade. It characteristically has low metastatic potential and a good prognosis (Nucci, 2013). Histologically verrucous carcinoma is well differentiated and with little atypia and a blunt margin of invasion. The pathogenesis of verrucous carcinoma is not well understood and does not appear to be associated with oncogenic HPV infection, however many cases are associated with lichen sclerosus or lichen simplex (Nascimento *et al.*, 2004; Nucci, 2013).

Basaloid and warty subtypes of SCC are associated with oncogenic HPV infection and keratinising tumours have typical keratin pearl formation not present in non-keratinising tumours.

1.4.1.1 HPV dependent carcinogenesis in the vulva

uVIN (HSIL) is a precursor for the majority of basaloid and warty type vulval SCC (Hampl and Sarajuuri, 2006; Van der Avoort *et al.*, 2006; Reyes and Cooper, 2014). The rising incidence of vulval cancer is primarily due to an increase in vulval cancer

in women aged 70 and below, suggestive of an increase in HPV related disease (Iversen and Tretli, 1998).

Malignant transformation of vulval epithelial cells by high-risk HPV subtypes is mediated through integration of the HPV DNA within the host genome leading to HPV induced expression of the E6 and E7 oncoproteins. These oncoproteins have numerous oncogenic effects including interference with control of the cell cycle, producing numeric and structural oncogenic chromosomal abnormalities (Vinokurova *et al.*, 2005; Zekan, Sirotkovic-Skerlev and Skerlev, 2011). In addition, methylation or mutations of the host's genome can occur that distorts transcriptional control, cell differentiation and viral gene expression resulting in altered molecular expression in tissues undergoing malignant transformation.

1.4.1.2 HPV independent carcinogenesis in the vulva

The typical precursor vulval disorders in cases of HPV independent vulval SCC are Lichen Sclerosus (LS) and differentiated VIN. LS and differentiated VIN develop into SCC through a HPV independent process which is not fully understood (Ueda *et al.*, 2011). Similar genetic and epigenetic changes have been noted to occur in areas of LS associated with SCC that are not present in LS found in isolation. These include mutation of TP53 and hypermethylation of MGMT and RASSF2A, suggesting a potential role for these genes in non HPV associated carcinogenesis however the exact mechanism is not well understood (Guerrero *et al.*, 2011; Trietsch *et al.*, 2015).

Overall the pathogenesis of vulval SCC is not as well described as that for more common cancers, however it is clear that in both the HPV dependent and HPV independent oncogenesis biomolecular changes occur in the vulval that lead to the development of invasive cancer.

1.4.1.3 Molecular field change in the development of cancer

The concept of field cancerisation was initially proposed by Slaughter et al. 1953 when studying abnormal tissue surrounding oral squamous cell carcinoma. The theory has been developed to explain the development of multiple primary tumours and local disease recurrence at sites separate from a primary invasive tumour. It has been proposed that vulval tumours also develop within preneoplastic fields of cancerisation containing histologically benign but genetically or epigenetically altered cells (Dakubo *et al.*, 2007). During the development of a preneoplastic field of cancerisation a stem cell acquires a genetic or epigenetic alteration that gives it a growth advantage over its neighbouring cells (Leemans, Braakhuis and Brakenhoff, 2011). Replication of the stem cell results in an expanding patch of altered daughter cells. With additional genetic or epigenetic alterations this patch replaces the normal epithelium and develops into a field of cancerisation. Within the abnormal preneoplastic field the epithelium may appear histologically benign however molecular alterations within cells result in an increased risk of malignant transformation. The development of this preneoplastic field change can be used to explain the multiple primary tumours and distant tumour recurrences that are often

observed in the vulva (Figure 1-5). The molecular basis for field cancerisation in the vulva has previously been explored by examining the distribution of X-chromosome inactivation, TP53 mutations and viral integration sites (Rosenthal *et al.*, 2002; Rolfe *et al.*, 2003; Vinokurova *et al.*, 2005). The data published demonstrated molecular monoclonality of anatomically distant vulval lesions supporting the concept of field cancerisation in the vulva. These studies of specific molecular changes in the vulva only show us part of the picture as the transformation of normal cells to cancer is a complex multistep process involving more that just a few molecular variations.

The development of fields of cancerisation and their persistence after surgery is problematic in the management of vulval SCC. Resection of the primary tumour including all histologically normal tissue may not remove all the molecularly abnormal cancer field leading to increased risk of local tumour recurrence and recurrence of cancer distant to the primary tumour (Figure 1-5). In the vulva, field cancerisation may explain the high incidence of disease recurrence at sites distant to the primary tumour sometimes termed second primary tumours (Van Seters, Van Beurden and de Craen, 2005). Observational studies have demonstrated that in those who have recurrent vulval SCC between 72% and 89% will have disease at a site distant to the primary tumour (Rouzier *et al.*, 2002; Woolderink *et al.*, 2006). Molecular changes associated with preneoplastic and neoplastic disease of the vulval are not well described although some studies have demonstrated clonal molecular changes at anatomically distant sites within the vulva supporting this concept of field cancerisation (Rosenthal *et al.*, 2002; Braakhuis *et al.*, 2003; Dakubo *et al.*, 2007).

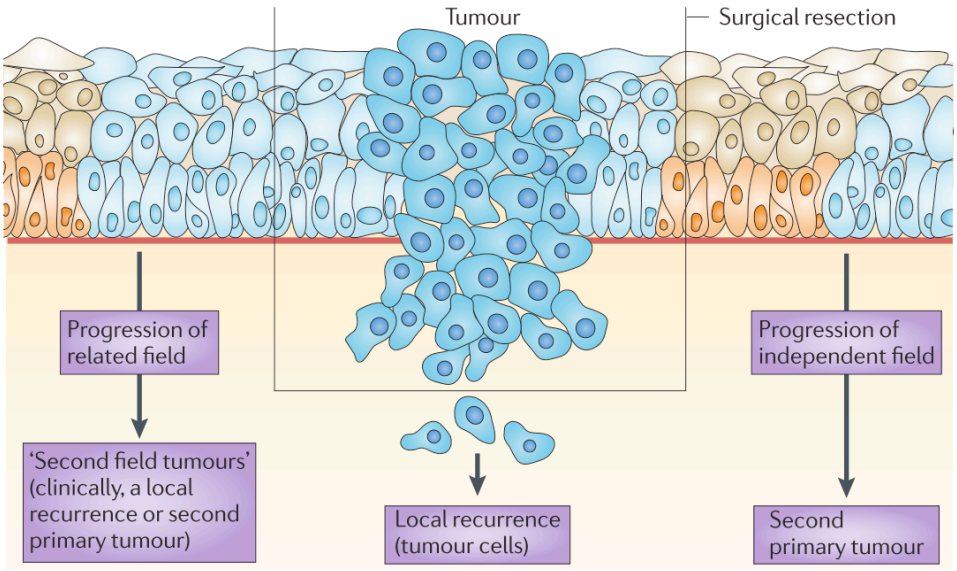


Figure 1-5 – Field cancerization and local relapse: The relationship between field cancerization and types of relapse is shown. Field cancerization is defined as the presence of one or more mucosal areas consisting of epithelial cells that have cancer-associated genetic or epigenetic alterations. A precursor field (shown in light blue) is monoclonal in origin and does not show invasive growth or metastatic behaviour, which are the hallmarks of an invasive carcinoma.

1.4.2 Vulval malignant melanoma

Vulval malignant melanoma is the second most common vulva malignancy accounting for approximately 9% of vulval malignancies. Malignant melanoma typically presents with the presence of a lump and/ or bleeding. Pre-existing pigmented lesions may develop into malignant melanoma and this change may be associated with pruritis, or an alteration in the shape size or colour of the lesion. These tumours can undergo early lymphatic or haematogenous dissemination and early detection is therefore paramount.

1.4.3 Basal cell carcinoma

Basal cell carcinomas of the vulva are uncommon and account for less than 5% of all vulval malignancies (Nucci, 2013). They occur predominantly in postmenopausal women with peak incidence in the eighth decade (Nucci, 2013). BCC may present as a deep ulcer with raised rolled margins, a superficial waxy slightly erythematous macule or as a polypoidal non-ulcerated lesion. These tumours primarily spread by direct invasion and rarely undergo lymphatic or haematogeneous spread.

1.4.4 Management of squamous vulval cancer

The primary treatment modality for vulval cancer is surgical excision of the disease. Treatment is individualised and in early stage disease usually consists of excision of the tumour with ipsilateral or bilateral inguinofemoral lymphadenectomy in all but FIGO stage Ia disease (Table 1-3) (RCOG, 2014).

Excision of the primary tumour should include the removal of an adequate margin of tissue without invasive disease. The rational for this is that the incidence of recurrence has been demonstrated to be directly related to the disease free surgical margin as assessed in the fixed histopathological specimen, with a decreased risk of recurrence with increased disease free margin.(Hacker *et al.*, 1981; Heaps *et al.*, 1990) Taking into account the contraction of tissue during excision and fixation, current guidelines recommend a 15 mm margin of non-invasive tissue is obtained on fresh surgical specimens (RCOG, 2014). Surgical excision of lichen sclerosus and

VIN in other parts of the vulva is considered during surgery to remove other possible foci of invasion and areas that pose an increased risk of recurrence (Regauer, 2011).

This treatment of vulval carcinoma is associated with significant physical, psychological, sexual and psychosexual morbidity both from the excision of the primary disease and from the inguinofemoral lymphadenectomy (Andersen *et al.*, 1988; Andersen, 1993). These adverse effects of the surgery are proportional to the extent of the surgery with the largest resections being associated with the highest morbidity.

The decision to undertake inguinofemoral lymphadenectomy is based on the risk of lymph node metastasis. In FIGO stage 1a disease (Table 1-3) the risk of lymph node metastasis is minimal and lymphadenectomy is not generally recommended (Hacker and Van der Velden, 1993; RCOG, 2014).

Inguinofemoral lymphadenectomy is therefore recommended for those women with FIGO stage 1b or above squamous cell carcinoma. Complete inguinofemoral lymphadenectomy is indicated in women with unifocal vulval tumours of more than 4 cm maximum diameter or where multifocal disease is present. The high rates of lymphoedema, infection and wound breakdown associated with complete nodal dissection has led to the development of sentinel node biopsy to avoid excision of disease free lymph nodes (Hauspy *et al.*, 2007). There is increasing evidence that women with a unifocal vulval tumour of less than 4 cm maximum dimension may be safely managed with excision of the sentinel lymph nodes before deciding on further management (Decesare *et al.*, 1997; Van Der Zee *et al.*, 2008; Levenback *et al.*,

2009). Currently assessment of sentinel node for the presence of micro metastasis is by ultrastaging with immunohistochemistry following fixation and sectioning of paraffin embedded lymph nodes. Currently the intraoperative assessment of lymph nodes is limited to visual inspection and palpation augmented with frozen section and cytological analysis although the diagnostic accuracy of these techniques is not well established.

Table 1-3 – International Federation of Gynecology and Obstetrics (FIGO) staging system

Stage I	**Tumour confined to the vulva**
Stage Ia	Lesions ≤ 2 cm in size, confined to the vulva or perineum and with stromal invasion ≤ 1 mm. No nodal metastasis
Stage Ib	Lesions > 2 cm in size or with stromal invasion > 1 mm confined to the vulva or perineum. No nodal metastasis
Stage II	**Tumour of any size with extension to adjacent perineal structures (lower 1/3 urethra; lower 1/3 vagina; anus) with negative node**
Stage III	**Tumour of any size with or without extension to adjacent perineal structures (lower 1/3 urethra; lower 1/3 vagina; anus) with positive inguinofemoral nodes**
Stage IIIa	(i) With 1 lymph node metastasis (≥ 5 mm), or (ii) 1–2 lymph node metastasis(es) (< 5 mm)
Stage IIIb	(i) With 2 or more lymph node metastases (≥ 5 mm), or (ii) 3 or more lymph node metastases (< 5 mm)
Stage IIIc	With positive nodes with extracapsular spread
Stage IV	**Tumour invades other regional (upper 2/3 urethra; 2/3 vagina) or distant structures**
Stage IVa	Tumour invades any of the following (i) Upper urethral and/or vaginal mucosa; bladder mucosa; rectal mucosa or fixed to pelvic bone, or (ii) Fixed or ulcerated inguinofemoral lymph nodes.
Stage IVb	Any distant metastasis including pelvic lymph nodes

The management of women found to have malignant disease within a sentinel lymph node is the subject of the ongoing GROINSS-V II trial (Eerens, 2010). Prior to this trial complete inguinofemoral lymphadenectomy was often performed when women were found to have metastatic disease within their sentinel lymph node. The

GROINSS-V II trial is an observational study assessing if women found to have sentinel lymph node metastases less than or equal to 2 mm in diameter could safely be spared the morbidity of full lymphadenectomy by treatment with primary radiotherapy (with or without chemotherapy). Within the GROINSS-V II trial those women with sentinel lymph node metastases greater then 2 mm with have complete inguinofemoral lymphadenectomy.

1.4.5 Management of non-squamous vulval cancers

Verrucous carcinoma and basal cell carcinoma of the vulva are rarely associated with lymph node metastasis and can usually be managed with wide local excision. In cases where resection would cause significant functional impairment basal cell carcinomas can also be successfully treated with radiotherapy.

Women with malignant melanoma also do not benefit from inguinofemoral lymphadenectomy and wide local excision of areas of disease is preferred. Relapse of malignant melanoma is common and there are limited strategies available to minimise the risk of relapse in this group.

1.5 Summary

The diagnosis and monitoring of vulval disease is undertaken by clinical evaluation supplemented by invasive tissue biopsies and histopathological analysis. Currently there is no established real time, objective technique for the assessment of vulval conditions. Vibrational spectroscopy, the topic of this thesis, offers the opportunity to

develop non-invasive probes for real time optical assessment, diagnosis and management of vulval conditions. The next sections of the thesis will detail the technique and progress to date in its use in the clinic.

Chapter 2 Vibrational Spectroscopy

2.1 Introduction

Light incident on matter may be absorbed, scattered or reflected by its constituent molecules. The interaction of the light with the molecular constituents of matter is defined by the structure of the molecules and wavelength of the incident light. Using the vibrational spectroscopic techniques of Raman spectroscopy and infrared spectroscopy we can measure the spectrum of the inelastic scattering or absorption of light from a tissue sample allowing us to analyse the molecular composition of the tissue.

2.1.1 Basic science and molecular vibration

Molecular vibration occurs when the atoms within the molecule are in periodic motion. Higher energy vibrational states can be excited when the molecule absorbs a photon or gains energy from an inelastic scattering event. It is this process of vibrational excitation that is exploited by vibrational spectroscopic techniques. The interaction between the light and a molecule is determined by both the composition of the molecule and the wavelength of the incident light. Photons striking a molecule may be absorbed, scattered or may not interact. Absorption of light by a molecule

only occurs at discrete wavelengths of light dependent on the composition and structure of the molecule. The wavelengths of light absorbed are dictated by the vibrational modes of the molecule. Molecules have a defined number of possible vibrational modes dependent on the number of atoms they contain (Equation 2-1) and the energy required to change the vibrational mode is different for each molecule (provided they are non-linear molecules, whereby they would have 3N-5 modes). Molecular vibrations can be excited by both the absorption of photons and scattering of photons. When the energy of the incident photon matches the difference between two molecular vibrational energy levels the photon is absorbed and the vibrational energy state of the molecule is increased.

$$Number\ of\ vibrational\ modes = 3N - 6$$

Equation 2-1

By contrast to absorption, scattering of light from a molecule can occur when the energy of the incident photon does not match the difference between two molecular vibrational energy levels.

Raleigh ‘elastic’ scattering occurs when the incident light causes the molecule to vibrate with the same frequency of the incident light, scattered light is emitted of the same wavelength as the incident light. This elastic scattering is the predominant form of scattering. In Raman scattering energy from the incident light is gained or lost by a change in molecular vibrational energy. This ‘inelastic’ process results in scattered light of a different energy, and therefore wavelength to the incident light.

Inelastic Raman scattering is a small effect and only accounts for 1 in 10^8-10^9 of all scattered photons.

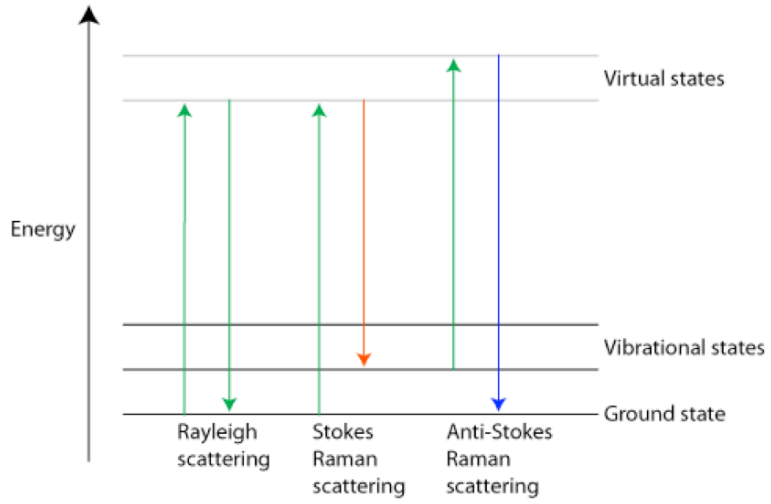


Figure 2-1 – Three forms of scattering

(Reprinted from http://www.doitpoms.ac.uk/tlplib/raman/raman_scattering.php, accessed March 2016)

In infrared spectroscopy a transition from one vibrational state to another only occurs in association with absorption of light. By contrast in Raman scattering there is change in the polarisation of the electron cloud associated with change from one vibrational energy state to another.

The vibrational energy change is different for each different molecule hence vibrational spectra can provide information about the molecules under analysis.

The two techniques of infrared spectroscopy and Raman spectroscopy are considered complementary as molecules that give strong bands in the IR spectrum may give weak bands in the Raman spectrum and visa versa. The electrical characteristics of the molecular vibrations are responsible for this complimentary

nature. Vibration of strong polar bonds (e.g. C-O, N-O and O-H bonds) will only have a small effect on polarisation and hence a relatively weak Raman effect. These strong polar bonds carry their charge during vibrational motion and can produce a large change in the dipole moment and strong IR absorption, unless neutralised by molecular symmetry. The converse is also true for weakly polarised bonds (e.g. C-C, C-H, C=C bonds). In simple symmetrical molecules such as carbon dioxide each vibrational mode will result in IR absorption or Raman scattering but not both (Figure 2-2). In complex asymmetrical molecules all the vibrational modes may be strongly absorptive in the IR region and Raman active.

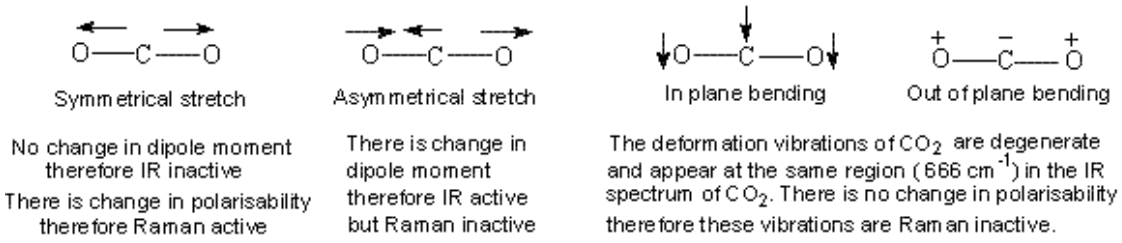


Figure 2-2 – Possible vibrational modes of carbon dioxide demonstrating complementary nature of IR and Raman spectroscopy.

(Reprinted from http://www.chemvista.org/ramanIR4.html, accessed March 2016)

2.2 Raman spectroscopy

2.2.1 Raman scattering

In Raman scattering the scattered photons can have a higher or lower energy than the incident light this is known as the ‘Raman shift’. Two types of Raman scattering have been described dependent on the vibrational energy state of the target

molecule when the scattering occurs. When light interacts with a molecule in its ground or unexcited state and the scattered photon has a lower energy than the incident photon this is termed Stokes Raman scattering. When the target molecule is already in a vibrationally excited state and the scattered photon has a higher energy than the incident photon this is termed anti-Stokes Raman scattering (Figure 2-1).

The change in energy in the scattered light can be detected, determining the difference in energy between the vibrational energy states of the molecules, by recording the intensity of the collected light against wavenumber shift to give a Raman spectrum. The Stokes and anti-Stokes scattered light undergoes an equal shift either side of the incident or elastically scattered light. This results in a symmetrical Raman spectrum apart from the difference in intensity of the scattered light. The number of molecules at each energy state, from thermal excitement, follows the Boltzmann distribution (Equation 2-2) so there are significantly fewer molecules in higher energy states. As a result the Stokes half of the spectrum, which doesn't require states to be pre-excited, is typically of greater intensity and this is the half routinely recorded in Raman spectroscopy.

$$N(state) \propto e^{-\frac{E}{kT}}$$

Equation 2-2 – Boltzmann distribution, where N is the number of molecules at a given energy state E and kT is the product of Boltzmann's constant and thermodynamic temperature.

2.2.2 The Raman spectra

A Raman spectrum is a plot of the intensity of the scattered the light (y axis) against the 'Raman shift' (x axis). By convention the 'Raman shift' is measured in wavenumber rather than wavelength and this is advantageous as it is directly proportional to the energy changed and it aids comparison between spectra produced with different wavelength of incident light. For a given material the position of the peaks on the Raman spectrum are the same irrespective of the wavelength of the incident light. The recorded spectrum may however be affected by variation of the incident light due to other interactions between the light and the material being examined, for example resonance Raman, fluorescence and absorption of light. In resonance Raman the wavelength of the excitation laser is close to that of electron transition in the compounds under examination, this can lead to greatly enhanced Raman scattering intensity. Conventionally the x axis varies from the positive wavenumber shifts produced by anti-Stokes Raman scattering on the left, through zero which represents the energy of the incident light, to negative wavenumber shifts produced by Stokes Raman scattering on the right.

Peaks within the Raman spectrum correspond to changes in vibrational energy of specific molecular bonds. Certain wavenumber shift spectral peaks can therefore be assigned to specific molecular bonds (Table 2-1). Furthermore the intensity of the

peak is proportional to the concentration of the molecular bond within the target material and within simple materials it is possible to produce a precise quantitative molecular fingerprint of the target. In molecularly complex material the wavenumber peaks may coalesce to form bands on the spectra and this can make the identification of individual peaks difficult. In this situation the overall shape of the spectra can be analysed as a molecular fingerprint of the target material.

Table 2-1 – Selection of Raman band assignments

(Reprinted from Applied Spectroscopy Reviews (Dieter Naumann, 2001), copyright 2001)

Frequency (cm^{-1})	Assignment[a]
~3059	$(C{=}C{-}H)_{(arom.)}$ str
~2975	CH_3 str
~2935	CH_3 and CH_2 str
~2870–2890	CH_2 str
~1735	>C=O ester str
~1650–1680	Amide I
~1614	Tyrosine
~1606	Phenylalanine
~1575	Guanine, adenine (ring stretching)
~1440–1460	C—H def
~1295	CH_2 def
~1230–1295	Amide III
~1129	C—N and C—C str
~1102	$>PO_2^-$ str (sym)
~1085	C—O str
~1061	C—N and C—C str
~1004	Phenylalanine
~852	"Buried" tyrosine
~829	"Exposed" tyrosine
~785	Cytosine, uracil (ring, str)
~720	Adenine
~665	Guanine
~640	Tyrosine (skeletal)
~620	Phenylalanine (skeletal)
~520–540	S—S str

The Raman spectroscopic analysis of known substances has allowed known biorelevant molecules to be mapped to Raman spectral peaks. This information can then be used to assess the molecular content of a sample under study or predict

expected changes in Raman spectra in different tissues (Stone *et al.*, 2007; Isabelle, 2009).

2.2.3 The Raman spectrometer

A basic Raman spectrometer requires a light source, collection optics, a filter to reject the excitation light and a spectrograph. Commonly the instrument will consist of a laser light source focused through a modified conventional microscope with a computerised stage. The scattered light is then collected through the microscope optics and directed through a notch filter onto a diffraction grating that splits the light onto a charge-coupled detector (Figure 2-3).

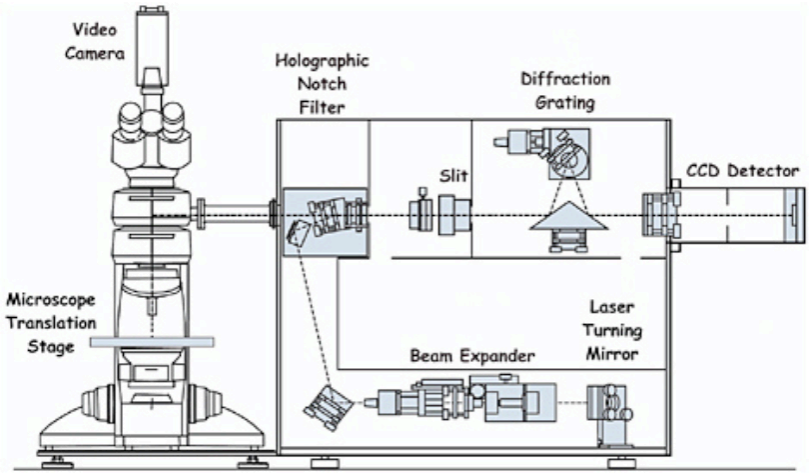


Figure 2-3 – Schematic representation of Raman microscope (Renishaw RM1000)

(Reprinted with permission from Scientific Examination of Art (Clark, 2005), copyright 2005)

2.2.4 Variants of Raman spectroscopy

In addition to standard dispersive Raman spectroscopy such as that described in 2.2.3 there are a number of variants that have been developed to overcome some of

the limitations of the standard technique. These include spatially offset Raman spectroscopy (SORS), coherent anti Stokes Raman spectroscopy (CARS), Hyper Raman, stimulated Raman spectroscopy (SRS), surface enhanced Raman spectroscopy (SERS) and Tip enhanced Raman spectroscopy (Begley, Harvey and Byer, 1974; Snow, Qian and Chang, 1985; Ziegler, 1990; Matousek *et al.*, 2005; Moskovits, 2005; Deckert, 2009). It is beyond the scope of this thesis to provide detail on these variants of the standard technique.

2.2.5 Raman spectroscopy of biological materials

With the development of improved laser technology Raman spectroscopy underwent a revival in the late 1960 and early 1970s. During this time there was increasing interest in the application of the technique to biological materials. Lord, Yu and Hartmann were pioneers in this area and their early work in the early 1970s measuring Raman spectra from native biological proteins and viral RNA laid the foundation for the modern application of Raman spectroscopy for the analysis of complex tissues (Lord and Yu, 1970; Hartman, Clayton and Thomas, 1973). Thanks to these early pioneers and those that followed the analysis of biological samples using Raman spectroscopy is now well established. Biological materials present numerous challenges to Raman spectroscopy. These include the relatively weak nature of the signal and the requirement for a highly sensitive detector making Raman spectrometers susceptible to interference from ambient light (Kendall *et al.*, 2009). In addition, visible wavelength light sources can elicit strong fluorescence in biological materials that may mask the much weaker Raman signal. Near infrared

wavelengths generally do not induce significant fluorescence, hence they are often used on biological applications.

As the vibrational modes of water, one of the primary constituents of tissue, produce only weak Raman signal the technique can be used to analyse tissue samples both in vivo and ex vivo without tissue preparation. This makes the development of Raman spectroscopic probes for in vivo diagnosis an attractive concept.

2.3 Fourier transform infrared spectroscopy

2.3.1 Infrared absorption spectra

The energy required to change the vibrational energy state of most molecular bonds within tissue corresponds to that of the mid-IR region of the electromagnetic spectrum (Haris and Chapman, 1992). Consequently infrared radiation is readily absorbed and the pattern of absorption can be used to produce an absorption spectrum on which the attenuation of infrared light can be demonstrated across a range of wavelengths. Similarly to the Raman spectrum this absorption spectrum can be used to analyse the molecular composition of tissue samples under study. In tissues many molecules will share the same molecular bond types so in complex materials infrared spectroscopy cannot be used to identify specific molecules but can be used to identify larger groups of molecules present (e.g. proteins and nucleic acids) creating a biomolecular finger print for the sample under analysis (Mackanos and Contag, 2010).

2.3.2 Fourier transform absorption spectra

In dispersive infrared spectroscopy the absorption spectra can only be measured for a narrow range of wavelengths at a time. This makes dispersive infrared spectroscopy a more time consuming technique and less practical for measuring clinically relevant biomaterials (Lasch and Kneipp, 2008). The use of array detectors including focal plane arrays and linear arrays to simultaneously detect multiple wavelengths increases the speed of spectral acquisition however Fourier transform infrared spectroscopy remains a quicker technique (Kazarian and Chan, 2013). In Fourier transform infrared spectroscopy, the wavelength of the light incident on a sample is varied within a set range though the use of an interferometer (Figure 2-4). This allows us to use Fourier transform infrared spectroscopy to measure the absorption of a wide range of wavelengths quickly. Using a discrete Fourier transform it is possible to then convert the raw data measured into a spectrum representing light transmitted at each wavelength. The transmission decreases exponentially with increasing absorbance by the sample under analysis according to the Beer-Lambert law (Equation 2-3). The absorbance can then be calculated from the measured transmittance using a logarithmic transformation.

Equation 2-3 – Beer-Lambert law, T is transmittance (relative amount of light passing through sample), ε is molar absorptivity, b is path length, c is molar concentration and A is absorbance.

$$T = e^{-\varepsilon bc} \quad (A = \varepsilon bc)$$

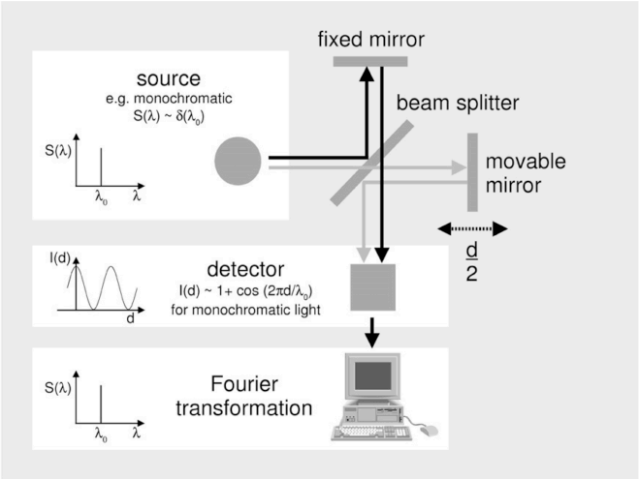


Figure 2-4 – Schematic of interferometer. (Reproduced from Biochimica et Biophysica Acta (Barth, 2007))

2.3.3 FTIR instrumentation

FTIR spectrometers require a polychromatic infrared radiation source, interferometer and detector. Light emitted from the infrared source is split by a beam splitter and approximately half directed towards a fixed mirror and the other half passes through to a moving mirror (Figure 2-4). The light incident on the fixed mirror is reflected back toward the beam splitter where approximately half passes through towards the detector. The light incident on the moving mirror is reflected back towards the beam splitter where approximately half is reflected towards the detector. The two beams passing from the beam splitter towards the detector combine and interfere. This interference is constructive or destructive depending on the optical path difference. The intensity of the light is measured at the detector relative to the position of the moving mirror and this creates the interferogram. The interferogram is the Fourier transform of the spectrum. By placing a sample of tissue between the interferometer and the detector we are then able to quickly measure an infrared absorbance

spectrum by converting the interferogram measured at the detector into a spectrum using a Fourier transform.

In contrast to dispersive infrared spectroscopy FTIR does not split the incident light into its spectral components but uses the interferometer to create light at different wavelengths whilst maintaining the intensity of the light passing through to the detector. This results in higher light intensity at the detector, faster data collection and higher signal to noise ratio (Barth, 2007). The majority of modern infrared spectroscopy is performed using the Fourier transform technique (Kendall *et al.*, 2009).

2.3.4 Modes of FTIR

There are three major modes of FTIR: transmission, transflection and attenuated total reflection (ATR). In transmission mode the sample is placed on a substrate that has low absorption of infrared radiation. The infrared radiation passing through the samples is then measured from the opposite side to the incident beam and the absorption inferred. In transflection mode the sample is placed on an infrared reflective slide and the transmission measurements are taken after the bean passes through the sample to the reflective slide and is reflected back though the sample. In ATR mode the beam is passed through an internal reflection element (IRE) with high refractive index. An evanescent wave is produced that extends beyond the IRE penetrating the sample that must be in contact with the element. The wave penetrates to a depth of 1-3 microns in biological substances for the 900 to 1,800 cm^{-1} spectral region (Baker *et al.*, 2014). This together with the high absorption of IR

radiation by water significantly limits the utility of the technique for in vivo diagnostics.

A comparison of the key differences between infrared spectroscopy and Raman spectroscopy is shown in Table 2-2.

Technical Considerations		
Property	Raman	FTIR
Commonly used wavenumber range (cm^{-1})	400-4000	800-4000
Light source	Monochromatic laser	Polychromatic
Detected light	Inelastic scattered	Mid infrared transmission
Molecular bond sensitivities	Non-polar bonds	Strong polar bonds
Signal to noise ratio over same timescale	Lower	Higher
Spatial resolution	<1 μm	≈1 μm with newer devices

Samples Considerations		
Restriction	Raman	FTIR
Sample contact	Required for in vivo probes	Not required in transmission
Sample thickness	Measurements possible	Thick samples can cause spectral saturation
Water content	Weak scattering signal	Strong absorption
Paraffin content	Strong scattering signal in fingerprint region	Narrow spectral bands in finger print region

Table 2-2 – Comparison of FTIR and Raman spectroscopy in tissue diagnosis

Chapter 3 Application of Vibrational Spectroscopy to Medical Diagnostics and the Clinical Need for Advanced Vulval Diagnostics

3.1 Vibrational spectroscopy of biological samples

The potential role of vibrational spectroscopy to accurately determine different histopathological groups has been well established at a number of different human anatomic sites. In the oesophagus Raman spectroscopy and FTIR have both been applied ex vivo and the ability of the technique to differentiate between normal, metaplastic, dysplastic and cancerous changes has been demonstrated (Shetty *et al.*, 2006; Almond, 2012; Almond *et al.*, 2014). Furthermore the ability of Raman spectroscopy to accurately distinguish different oesophageal tissue types in vivo has also been demonstrated (Bergholt *et al.*, 2011). Similarly studies on cellular changes in the stomach have demonstrated the ability of vibrational spectroscopic techniques to differentiate normal, metaplastic, dysplastic and cancerous changes both ex vivo and in vivo (Bergholt *et al.*, 2011, 2013; Li *et al.*, 2013; Huang *et al.*, 2014; Ito *et al.*, 2014). Raman spectroscopy has also been shown to be able to detect ductal carcinoma from normal breast tissue. Highlighting potential for the technique to be

used to assess the margins of surgical resection specimens from breast conserving surgical treatment of breast cancer (Kast *et al.*, 2008; Kong *et al.*, 2014). The potential for Raman spectroscopy to be used to assess tumour margins in surgical neuro-oncological treatments is currently the object of extensive investigation and studies to date have shown the technique has potential in this area (Toms *et al.*, 2006; Gajjar *et al.*, 2012; Ji *et al.*, 2013; Bahreini, 2015). Vibrational spectroscopic techniques have also been applied in the head and neck and studies have established the ability of the techniques to identify oral, oropharyngeal, laryngeal and para thyroid tumours (Stone *et al.*, 2000; Das *et al.*, 2006; Harris *et al.*, 2010; Singh *et al.*, 2012; Menzies *et al.*, 2014).

To date there has been very limited research published that has evaluated the role of vibrational spectroscopic techniques in the detection and assessment of vulval or vaginal diseases. In the search performed for this literature review no published papers in this area of study were identified. Due to the lack of previous study in this area it is necessary to assess the current evidence from other dermatological application of the techniques.

The epithelium of the vulva shares many characteristics with other skin on the human body, and hence vibrational spectroscopic research in this area is directly relevant to the vulva. It should be noted the spectrum of disease in the vulva is quite different from other skin sites with increased risk of SCC and decreased risk of BCC and melanoma (Gartner, 2017). The relevant research to date exploring the role of vibrational spectroscopy in the assessment of skin disease is reviewed below.

3.2 Vibrational spectroscopy for the characterisation of dermatological disease

3.2.1 Clinical diagnosis of dermatological disease and the role of advanced diagnostics

Traditionally the diagnosis of skin conditions has been by naked eye visual examination, magnified dermatoscopic examination and invasive biopsy where there is suspicion of malignancy or in cases of diagnostic uncertainty. Crucially this approach is dependent on the skill and experience of the clinician making the assessment. Variation in diagnostic accuracy amongst dermatologists has been demonstrated in numerous studies (MacKenzie-Wood, Milton and de Launey, 1998; Morton and Mackie, 1998; Lui *et al.*, 2012). Reported sensitivity for the diagnosis of melanoma by trained dermatologists varies from 49% to 81% with around a third of melanomas being inappropriately diagnosed as benign lesions (Lui *et al.*, 2012). Heal et al. found the sensitivity of clinical diagnosis of skin cancers was 63.9% for BCCs, 41.1% for SCCs and 33.8% for malignant melanomas (Heal *et al.*, 2008). Due to the inaccuracy of visual assessment the current gold standard for diagnosis of dermatological conditions is biopsy and histological examination. This process is invasive, remains subjective and slow as sample preparation and histology takes significant time. In addition a large number of benign lesions are needed to be excised to detect a few malignant ones. This issue was highlighted by a retrospective study of 4,761 excised pigmented lesions from patients attending 468

general practitioners. The study revealed that new practitioners excised 58 benign lesions per malignant lesion compared to 21 lesions excised by experienced practitioners per malignant lesion. The need for a near real time non-invasive technique for the diagnosis of skin conditions has driven research into the role of vibrational spectroscopy as a solution.

3.2.2 Vibrational spectroscopy for the assessment of dermatological conditions

Skin is a highly accessible area of the body and as such there is considerable interest in the use of vibrational spectroscopy as a near real time, non invasive, objective technique for the diagnosis of dermatological conditions (Gniadecka, Wulf and Mortensen, 1997; Nijssen *et al.*, 2002; Gniadecka *et al.*, 2004; Choi *et al.*, 2005). The structure of skin has been examined using both Raman and FTIR spectroscopy and current areas of research include: assessing drug absorption, solar radiation induced skin damage and the diagnosis of benign, premalignant and malignant skin diseases (Caspers *et al.*, 2001; Cartaxo *et al.*, 2010; Calin *et al.*, 2013; Cinotti *et al.*, 2013; S M Ali *et al.*, 2013).

3.2.3 Raman spectroscopy for the assessment of dermatological disease

Raman spectroscopy has some significant advantages over FTIR spectroscopy for the assessment of skin diseases. Raman spectroscopy produces spectra with a

high level of spatial detail that can provide a clearer demarcation between tissue layers (S. M. Ali *et al.*, 2013). As Raman spectroscopy only relies on the scattering of light, the excitation light and scattered light can be collected from the same tissue surface, allowing the technique to be applied in vivo (Barth, 2007). In addition, compared to FTIR spectroscopy Raman spectroscopy is relatively insensitive to the water present in skin.

Early studies by Gniadecka et al. that assessed a small number of biopsies from subjects with benign and malignant skin conditions highlighted Raman spectroscopy is able to detect differences in protein and lipid content of different skin conditions. These differences were seen in the spectral regions 1065-1094 cm^{-1} and 1243-1258 cm^{-1} (Gniadecka *et al.*, 1997; Gniadecka, Wulf and Mortensen, 1997). Knudsen et al. investigated the variation in Raman spectra taken from 13 different subjects at different bodily sites and at different times of the day (Knudsen *et al.*, 2002). The group found the relative intensities of amide bands I and III varied only slightly between individuals and on different days. The small spectral changes seen were thought to be due to changes in collagen structure and hydration status. The study concluded Raman spectra obtained from normal human skin are reproducible with limited variation. Caspers et al. demonstrated confocal Raman spectroscopy could be used to analyse the molecular composition and hydration status of different layers of the skin both in vitro and in vivo (Caspers *et al.*, 1998, 2001; Caspers, 2003; Caspers, Lucassen and Puppels, 2003). Raman spectroscopy has also been investigated as a potential tool for assessing the risk of developing skin cancer. Hata et al. validated the use of Raman spectroscopy for assessing carotenoid

concentrations and suggested those individuals with low skin carotenoid concentrations may be at increased risk of skin cancer (Hata *et al.*, 2000). The conclusions of the study were not confirmed by longitudinal data to assess causation and remain speculative.

These early studies were restricted to a small number of subjects; focused on analysing specific molecular changes and did not look to classify pathology based on spectral characteristics. The molecular changes observed indicated the technique had promise for the non-invasive diagnosis of skin conditions. Subsequent studies have applied statistical analytical techniques to automate the classification of disease from measured spectra.

One of the most investigated uses of Raman spectroscopy for disease classification is the discrimination of basal cell carcinomas and melanomas from normal skin (Gniadecka, Wulf and Mortensen, 1997; Nijssen *et al.*, 2002, 2014; Gniadecka *et al.*, 2004; C. a Lieber *et al.*, 2008; Cartaxo *et al.*, 2010; Ly *et al.*, 2010; Oliveira *et al.*, 2010; Bodanese *et al.*, 2012; Silveira Jr *et al.*, 2012; Philipsen *et al.*, 2013). In contrast to cancers of the vulva, basal cell carcinomas constitute over 80% of all non-vulval skin cancers, with squamous cell carcinomas and malignant melanomas only constituting 15% and 5% respectively. As a result of this the majority of research undertaken exploring the role of Raman spectroscopy in the assessment of skin cancer has focussed on BCCs rather than SCCs, which are more common in the vulva (Gniadecka, Wulf and Mortensen, 1997; Nijssen *et al.*, 2002, 2014; Choi *et al.*, 2005; C. a Lieber *et al.*, 2008; Ly *et al.*, 2010; Silveira Jr *et al.*, 2012; Philipsen *et*

al., 2013). Spectral markers associated with cancer have been identified using Raman spectroscopy and the diagnostic ability of the technique has been demonstrated, with sensitivities and specificities of greater than 90% and 80% respectively reported by many groups (Table 3-1) (Gniadecka, Wulf and Mortensen, 1997; Sigurdsson *et al.*, 2004; C. a Lieber *et al.*, 2008; Calin *et al.*, 2013). Sigurdsson et al. demonstrated Fourier transform Raman analysis of the skin surface in vitro combined with artificial neural networks could be used to discriminate BCC and melanoma from actinic keratosis, normal skin and benign pigmented naevi with a >94% sensitivity and >98% specificity (Sigurdsson *et al.*, 2004). Nijssen et al. utilised confocal Raman spectroscopy on 15 sections of basal cell carcinoma and were able to identify BCC from normal perilesional tissue with a sensitivity of 100% and specificity of 93% after leave one spectrum out cross validation (Nijssen *et al.*, 2002). It should be noted that in this study 18 BCC, 31 dermis and 10 epidermis spectra from 15 samples were included in the cross validated classification model increasing the probability of the result being due to a type one error (Nijssen *et al.*, 2002). Lieber et al. also evaluated the ability of Raman spectroscopy to discriminate cancerous lesions (9 BCCs and 4 SCCs) from adjacent normal tissue and reported a sensitivity of 100% and specificity of 91% (C. a Lieber *et al.*, 2008).

These early diagnostic studies are significantly limited by their small sample sizes and the resultant difficulty in producing reliable estimates of diagnostic accuracy. More recently, larger studies have focused on producing more reliable measures of diagnostic accuracy and the transition to in vivo application. Silveira et al. analysed 47 samples skin biopsies using a portable dispersive Raman system comprising 15

normal skin samples, 29 BCCs, and 4 melanomas (Silveira Jr *et al.*, 2012). They used a classification model applied to the relative contributions of collagen III, elastin, and melanin using Euclidean distance as a discriminator. This technique could differentiate normal from BCC and melanoma with a sensitivity of 98.2% and specificity of 96.6%.

The largest study published to date comes from Lui et al. and warrants detailed analysis (Lui *et al.*, 2012). The authors collected Raman spectra from 848 patients using a hand held optical fibre probe, coupled to a 785 nm excitation laser, spectrometer and charge coupled device. The effective spectral range was 500 to 1800 cm^{-1}, integration time 1 second and spectra were processed in real time. The spectral measurements were compared to a gold standard of either: expert clinical diagnosis by an experienced dermatologist (44%) or histopathological analysis (56%), in cases where biopsy was taken after Raman analysis. Only melanomas, SCCs, BCCs, actinic keratoses and lesions that visually mimic these conditions were included in the analysis. In total 518 lesions from 453 subjects were included in the analysis. This included 44 melanomas, 47 SCCs, 109 BCCs and 32 actinic keratoses. The reproducibility of the Raman measurements was assessed with repeated measurements of 30 different sites. Minimal variance was found between repeated measurements below 1055 cm^{-1}, with the higher wavenumber range being subject to more significant fluctuation in signal intensity. The spectra were analysed by full spectra and also by low variance (500 cm^{-1} – 1055 cm^{-1}) and high variance (1055 cm^{-1} – 1800 cm^{-1}) regions. Multivariate analysis for classification was by principal component analysis with generalised discriminant analysis (PC-GDA) and

concurrent partial least squares. Both analyses were validated with a leave one sample out cross validation. The study reported the diagnostic accuracy of Raman spectroscopy and multivariate analysis by three dichotomous groupings. The first grouping discriminated skin cancer and actinic keratosis from benign lesions with a sensitivity of 99%, specificity of 17% and corresponding area under the receiver operator curve of 0.879 ($p<0.0001$). The second grouping discriminated melanoma from benign pigmented lesions with a sensitivity of 99%, specificity of 15% and corresponding area under the receiver operator curve of 0.823 ($p<0.0001$). The third grouping discriminated melanoma from seborrheic keratoses with a sensitivity of 99%, specificity of 25% and corresponding area under the receiver operator curve of 0.898 ($p<0.0001$). In this study it appears the sensitivity of the classification model was optimised at the expense of specificity. This results in high negative predictive values of up to 0.99 with modest positive predictive values 0.49. The authors concluded that if implemented in the clinical environment Raman spectroscopy could reduce the need for invasive biopsy by around 50% whilst maintaining sensitivity for detection of malignant disease. These results are encouraging and highlight a real role for Raman spectroscopy in the assessment of dermatological conditions. The primary strengths of this study are the large numbers of samples included in the analysis and the application of a technology that could be used in a clinically relevant way, with short integration times and a handheld probe. There are some features of the study that should be noted when interpreting the results. The number of patients included in the study was reduced from the initial 848 patients recruited to 453 based on the benign lesion's visual similarity to premalignant or malignant disease. It is not

clear in the study how the included benign lesions were selected, however it appears selection was based on expert dermatological assessment. Excluded lesions not characterised in the study may appear to the inexpert clinician as visually similar to malignant lesions. This may limit the applicability of the results to expert dermatological practice, whereas the greatest reduction in the number of biopsies performed is likely to result from application to generalist clinical practice. In addition, the final data set consisted of 518 validated lesions from 453 subjects, therefore an unknown number of subjects in the analysis had multiple lesions assessed. This was not accounted for in the leave one lesion out analysis and may increase the likelihood of a type one error: if multiple benign or malignant lesions were included from individual subjects. Although the authors discuss the reproducibility of the spectral measurements for individual lesions the reproducibility between different spectrometer systems is not discussed. The ability to reproduce spectral measurements on different systems is crucial for the application of a Raman based diagnosis across a number of sites (Isabelle *et al.*, 2015).

This study combined with other work from the same research group has resulted in the production of the Verisante Aura, a commercial Raman spectroscopic medical device for the diagnosis of skin conditions in vivo (J Zhao *et al.*, 2008; Jianhua Zhao *et al.*, 2008; Zhao *et al.*, 2010; Wang *et al.*, 2012). This device has not been evaluated for the diagnosis of vulval skin conditions and the manufacturer advises against its use for vulval diagnostics. The reason for this is not stated, but as vulval skin has a significantly different spectrum of disease compared with other skin sites

and as this device is not validated on vulval skin samples, the current classification models are unlikely to be transferable to the vulva.

The studies discussed and those summarised in Table 3-1 evaluate the ability of Raman spectroscopy to replicate the existing diagnostic tests. As pathological molecular changes precede histological changes and the symptoms of disease Raman spectroscopy may also prove a useful tool for the early diagnosis of disease, although this has not yet been evaluated in clinical trials.

Table 3-1 – Studies assessing the role of vibrational spectroscopy in the diagnosis of malignant skin disease

Study	Type of Raman	In Vivo / In Vitro	Approach	Pathology	Comparator	Summary
Gniadecka et al. (1997)(Gniadecka, Wulf and Mortensen, 1997)	FT-Raman spectroscopy (1064 nm)	In vitro (punch biopsies)	Point measurement	BCC	Normal	Differences in Raman spectra between 16 BCC and 16 normal skin biopsies indicated alterations in protein and lipid structure in skin cancer samples.
Bodanese et al. (2010)(Bodanese *et al.*, 2010)	Dispersive Raman spectroscopy (830 nm)	In vitro	Unknown	BCC	Normal	Main spectral differences between these samples were in the region of 800 to 1000 cm^{-1} and 1200 to 1300 cm^{-1}. PCA and Mahalanobis distance applied to the scores of principal components could identify tissue with sensitivity and specificity of 89% and 93%, respectively, for the training group and 96% and 92% for the prospective group. A simplified biochemical model for collagen amount had sensitivity and specificity of 95% and 83% for the training group and 87% and 92% for the prospective group
Bodanese et al. (2012)(Bodanese *et al.*, 2012)	Fiber optic Raman probe (830 nm)	In vitro (biopsy fragments)	Unknown	BCC & Melanoma	Normal	PCA and Euclidean distance were employed to develop a discrimination model. Demonstrated that principal components (PCs) 1 to 4 accounted for 95.4% of all spectral variation. A total of 28 out of 30 spectra were correctly diagnosed as Normal, 93 out of 96 as BCC, and 13 out of 19 as Melanoma
Nijssen et al. (2002)(Nijssen *et al.*, 2002)	Dispersive Raman spectroscopy (850 nm)	In vitro (tissue sections)	2D Mapping	BCC	Normal from same samples	Cluster analysis of Raman maps of 15 tissue sections resulted in 15 clusters. Classification model based on linear regression resulted in sensitivity of 100% and specificity of 93% for identifying BCC from surrounding normal tissue.

Choi et al. (2005)(Choi *et al.*, 2005)	Confocal Raman spectroscopy (514.5 nm)	In vitro (tissue sections)	2D Mapping	BCC	Normal from same subjects	Analysis of biopsies from 10 patient demonstrated distinct Raman band differences between normal and BCC tissues in the amide I region of the spectra.
Lieber et al. (2008)(C. a Lieber *et al.*, 2008)	Portable confocal Raman system (825 nm)	In vivo	Point measurement	BCC/ SCC	Normal from same subjects	9 out of 9 BCC, 4 out of 4 SCC and 8 out of 8 inflamed scar tissues were correctly predicted by the diagnostic algorithm, and 19 out of 21 normal tissues were correctly classified.
Ly et al. (2009)(Ly *et al.*, 2010)	Polarized Raman microspectroscopy (785 nm)	In vitro (tissue sections)	2D Mapping	BCC	Normal from same subjects	Analysis of six samples of BCC demonstrated differences between the spectra of peritumoral and normal dermis.
Silveira et al. (2012)(Silveira Jr *et al.*, 2012)	Portable dispersive Raman system (830 nm)	In vitro (biopsy fragments)	Point measurement	BCC / Melanoma	Normal (unknown)	Analysis of 47 samples: 15 normal skin samples, 29 BCCs, and 4 melanomas. A classification model applied to the relative contribution of collagen III, elastin, and melanin using Euclidean distance as a discriminator could differentiate normal from BCC and MEL with a sensitivity of 98.2% and specificity of 96.6%.
Philipsen et al. (2013)(Philipsen *et al.*, 2013)	Fiber optic Raman probe (1064 nm)	In vivo	Point measurement	BCC / Melanoma	Normal (unknown)	Analysis of spectra from 55 healthy persons with different skin pigmentation, 25 BCCs, 41 pigmented nevi and 15 malignant melanomas revealed significant differences in the water band around 3250 cm^{-1}, the protein specific band around 1250 cm^{-1} (amide-III) and the amide-III ratio.
Nijssen et al. (2006)(Nijssen *et al.*, 2014)	Fiber optic Raman probe (720 nm)	In vitro (biopsy fragments)	Point measurement	BCC	Normal (perilesional skin)	Raman spectra (2800 to 3125 cm^{-1}) analysed from 19 BCC biopsy specimens and 9 biopsy specimens of perilesional skin. Consistent differences in the spectra of BCC and perilesional skin were identified.

Cartaxo et al. (2010)(Cartaxo *et al.*, 2010)	FT-Raman spectroscopy (1064 nm)	In vitro (biopsy fragments)	Point measurement	Melanoma	Pigmented nevi	10 samples of melanoma, 9 samples of pigmented nevi, and 10 samples of normal skin were analysed and the vibrational modes of the Polysaccharides, Tyrosine and Amide-I used to differentiate melanoma from the pigmented nevus
Lui et al. (2012)(Lui *et al.*, 2012)	Fiber optic Raman probe (785 nm)	In vivo	Point measurement	BCC / Melanoma / SCC / Actinic keratosis	Benign lesions visually similar to malignant lesions	Benign and malignant skin lesions (n=518) from 453 patients were measured, including melanomas, basal cell carcinomas, squamous cell carcinomas, actinic keratoses, atypical nevi, melanocytic nevi, blue nevi, and seborrheic keratoses. Lesion classification was made using a principal component with general discriminant analysis and partial least-squares. Comparisons included skin cancers and precancers from benign skin lesions, receiver operating characteristic (ROC)=0.879; melanomas from nonmelanoma pigmented lesions (ROC=0.823); and melanomas from seborrheic keratoses (ROC=0.898). For sensitivities between 95% and 99%, the specificities ranged between 15% and 54%.
Gniadecka et al. (2004)(Gniadecka *et al.*, 2004)	FT-Raman spectroscopy (1064 nm)	In vitro (biopsy fragments)	Point measurement	BCC / Melanoma	Normal/ pigmented nevi/ seborrheic keratosis	22 samples of melanoma, 41 of pigmented nevi, 48 of BCC, 23 of seborrheic keratosis and 89 of normal skin were analysed. Neural network analysis of Raman spectra achieved a diagnostic sensitivity of 85% and specificity of 99% for the diagnosis of Melanoma and for BCC diagnosis 97% and 98%, respectively.
Oliveira et al. (2010)(Oliveira *et al.*, 2010)	FT-Raman spectroscopy (1064 nm)	In vitro (biopsy fragments)	Point measurement	Melanoma	Normal/ metastatic melanoma in lymph Nodes	Analysis of 10 normal skin samples, ten melanoma samples and 9 lymph node metastasis allow correctly classification of the samples in 93.1% based on Phenylalanine, DNA, and Amide-I spectral features.

3.2.4 Fourier transform infrared spectroscopy for the assessment of dermatological disease

Unlike Raman spectroscopy the role of FTIR for the diagnosis of skin disease has been less well studied. It is likely that this is due to the challenges posed by strong absorption of mid-infrared radiation by water. This makes the technique difficult to apply in vivo, as penetration of infrared radiation using ATR FTIR is limited to 2-3 microns and transmission analysis would be limited to samples 20 microns thick and be impractical (Eikje, Aizawa and Ozaki, 2005). FTIR is well suited to the non-destructive analysis of dehydrated fixed cytological or histological tissue samples as it offers rapid analysis with high signal to noise ratio. The IR spectral characteristics of different components of human skin have been well described and an increasing number of studies have demonstrated the IR spectra contain sufficient information to offer valid and useful diagnostic information about a number of different types of skin tumour (McIntosh *et al.*, 1999; Skrebova, Aizawa and Aras, 2003; Eikje, Aizawa and Ozaki, 2005). FTIR spectra of benign (melanocytic naevi), premalignant (Bowen's disease, solar keratosis) and malignant (SCC, BCC and melanoma) disease indicate the most significant differences in the IR spectrum occur in the 800-1800 cm^{-1} region, when compared to normal skin tissue spectra (Skrebova, Aizawa and Aras, 2003). Spectra differences are thought to be due to variations in proteins, chromatin, and DNA peaks with spectral features due to DNA and Amide III at 965-970 cm^{-1}, 1071 cm^{-1}, 1084 cm^{-1}, 1095 cm^{-1}, 1220 $cm^{-1}$and 1245 cm^{-1} being modified and enhanced as tissues progress towards malignancy (Wong *et al.*, 1993; Skrebova,

Aizawa and Aras, 2003). In vivo investigations have been undertaken using ATR mode FTIR, however due to the limited penetration depth of the evanescent field, analysis is limited to the stratum corneum. Barry et al. and Brancaleon et al. investigated hydration and lipid concentration in the stratum corneum and identified differences in these parameters according to anatomical site (Barry, Edwards and Williams, 1992; Brancaleon *et al.*, 2001). Sukuta et al. investigated use of a modified form of ATR FTIR, Fourier transform infrared fibre optic evanescent wave (FTIR-FEW) spectroscopy and chemical factor analysis for the classification of melanoma and BCC from normal tissue. The study included 6 samples in each pathology group and demonstrated good agreement between pathological analysis and spectroscopic analysis for the differentiation of melanoma from normal tissue, highlighting the potential of FTIR as a diagnostic tool. Ex vivo studies using sections of routinely prepared paraffin embedded tissue have been performed to identify bimolecular markers of malignancy. Mordechai et al. analysed tissue from seven patients with melanoma and ten with cervical cancer using FTIR micro spectroscopy. Comparing spectra from malignant tissue with control tissue spectra from the same patients the study identified the RNA / DNA ratio as a possible common marker of malignant disease, thought to be due to a higher transcription rate during carcinogenesis (Mordechai *et al.*, 2004). Tfayli et al. undertook a feasibility study on 6 sections of paraffin embedded fixed skin from melanoma and benign epidermal naevi. The study again demonstrated the ability of FTIR micro spectroscopy and multivariate analysis for differentiating benign and malignant lesions (Tfayli *et al.*, 2005). This work was expanded in a study by Hammody et al. which analysed 55 sections of paraffin

embedded fixed skin containing melanoma using FTIR micro spectroscopy (Hammody *et al.*, 2008). In this study the authors identified eight sites of spectral variation between melanoma and adjacent unaffected tissue that were then used as spectral biomarkers. These corresponded to amide I and II, RNA/ DNA ratio; phosphate level; cytosine + guanine/ total absorbance intensities of all nucleic acid bases ratio; thyamine/ amide II ratio; tyrosine/ amide II ratio; DNA level and thiamine level. Using these biomarkers they were able to successfully differentiate malignant melanoma from normal skin with sensitivity of 86% and specificity of 82.5%. The authors proposed that these biomarkers could be used to determine the boundary of melanoma and normal tissue on the surface of the skin. If successful this technique could be used to determine resection margins at the time of surgery.

The studies discussed highlight the potential of FTIR for the ex vivo and in vivo characterisation and diagnosis of dermatological conditions. The studies are however limited by the small numbers of samples / patients included. To date there are no published studies examining the role of FTIR in the diagnosis of SCC on the vulval or other anatomical locations.

3.2.5 Alternative optical technologies for dermatological diagnosis

As previously discussed, the clinical diagnosis of dermatological disease is unreliable, necessitating frequent recourse to invasive tissue biopsy and histological analysis (chapter 3.2.1). Due to their accessibility skin disorders are amongst the

most investigated using optical diagnostic techniques. A range of optical techniques have been explored for the study of skin structures and diagnosis of skin disease. These include optical coherence tomography, fluorescence spectroscopy, diffuse reflectance spectroscopy, confocal microscopy and laser scanning microscopy.

3.2.5.1 Optical coherence tomography

Optical coherence tomography (OCT) has been investigated as a tool for the non-invasive diagnosis of skin lesions. OCT uses back scattered light to produce a cross sectional image of the skin up to 2 mm in depth. In the transition towards invasive cancer the epidermis thickens and the basement membrane becomes interrupted and discontinuous. It has been speculated that the appearance of the epithelium combined with a measurement of the epidermal thickness and the attenuation coefficient can be used to differentiate between healthy and premalignant vulval tissue. Wessels et al. performed in vivo OCT imaging in 16 patients with subsequently histologically confirmed VIN (Wessels *et al.*, 2012). The VIN OCT images were compared with that of adjacent normal skin. The authors report a sensitivity and specificity of 100% using the tomographically measured epidermal thickness and a sensitivity of 88% and specificity 94% using the attenuation coefficient, for differentiating VIN from normal skin. This study only compares VIN to normal skin and does not evaluate the efficacy of the technique for differentiating VIN from other causes of thickened skin or skin that may have an altered attenuated coefficient. The results are of interest but the comparison between VIN and normal skin is not one that is clinically useful.

Other studies have assessed the role of visual examination of the images produced by OCT to diagnose skin cancers. Olmedo et al. evaluated 27 patients with BCC and matched the images with the histopathology. They described the typical appearance of BCC OCT images and were able to correctly match 20 of the 27 images with the correct histopathology using the typical OCT features identified (Olmedo *et al.*, 2006). This study was not a blinded diagnostic test accuracy study and did not demonstrate the actual predictive value of OCT in diagnosing malignant disease. Further studies have shown the accuracy of OCT as a diagnostic tool varies with the experience of the person interpreting the images. Mogensen et al. compared OCT images of 104 skin cancers and 105 controls with histological diagnosis and found that the sensitivity and specificity achieved varies according to the experience of the observer (Mogensen *et al.*, 2009). Sensitivity varied from 57% to 94% and specificity from 43% to 96%, with the higher sensitivities and specificities being achieved by the most experience observers. This observation has led to research into machine learning analysis of OCT images. In a proof of concept study Bazant-Hegemark and Stone demonstrated an OCT prediction algorithm using principal component analysis combined with linear discriminant analysis could be used to differentiate different porcine tissue types with high classification accuracy (Bazant-Hegemark and Stone, 2008). Jogensen et al. implemented a machine learning approach for the differentiation of BCC from actinic keratosis. In this study blinded observers analysed the OCT images for the presence of set features including: layer thickness; position and extension of dense back scattering section; maximum depth of a local white spike; maximum depth of backscatter/ reflection; maximum depth of fading signal;

graininess; variation in penetration depth; appearance of dark circular lobules; number, depth and size of lobules; turbidity/ homogeneity; dark elongated vessel structures; reflectivity of vessel structures; intact epidermis, and separate top-layer. The set features from 78 lesions (41 BCC and 37 actinic keratosis) were fed into the machine-learning algorithm. With this approach the algorithm was able to correctly identify 81% of the actinic keratoses and 73% of BCCs (Jorgensen *et al.*, 2008).

In summary, several studies have demonstrated the ability of OCT to differentiate between benign, premalignant and cancerous skin lesions, using analysis of OCT images and determination of the optical attenuation coefficient. The technique is however largely limited by the need for subjective assessment of OCT images and the fact that classification is based on structural rather than biochemical analysis of tissues. Future work may enable the automation of image analysis to reduce the interoperator variability, leading to a more objective technique. In addition, OCT combined with other modalities such as Raman spectroscopy or fluorescence spectroscopy may enable structural information to be combined with biochemical information for multimodal diagnosis.(Robles *et al.*, 2011; Egodage *et al.*, 2015)

3.2.5.2 Fluorescence spectroscopy

Fluorescence spectroscopic techniques encompass light induced auto fluorescence and the use of exogenous agents to augment the light induced fluorescence. Auto fluorescent imaging is an appealing tool for optical diagnosis, as the technique allows differentiation of pathology on the basis of biomolecular composition of the tissue and does not require contrast agents (Bigio and Mourant, 1997; Kollias,

Zonios and Stamatas, 2002; Drakaki *et al.*, 2009; Borisova *et al.*, 2010). Different skin components fluoresce with different excitation wavelengths. Deep UV (260 to 300 nm) excites fluorescence from amino acids such as tyrosine, tryptophan and phenylalanine. 320 to 400 nm light excites fluorescence from structural proteins including collagen and elastin as well as nicotinamide adenine dinucleotide and flavins (Drakaki *et al.*, 2013). Cutaneous melanin fluoresces with excitation in the near infrared region (Han *et al.*, 2010). Researchers such as Panjehpour et al. have demonstrated laser induced (410 nm) fluorescence emission is reduced in SCC and BCC of the skin. This change in fluorescence profile can be used to classify SCC and BCC with an accuracy of 93%, however this is reduced to 78% in those with more pigmented skin due to the higher absorption of melanin (Panjehpour *et al.*, 2002). In addition to intensity shift in the fluorescence spectral profile can be used to characterise disease states. Drakari et al. observed BCC results in a red shift of the fluorescence spectrum produced with a 337 nm excitation source (Drakaki *et al.*, 2009). Exogenous florescent markers can be use when the endogenous fluorescence is weak or there is increased absorption such as that produced by melanin pigmentation (Borisova *et al.*, 2012). This photodynamic diagnostic approach has been investigated as a technique for the delineation of disease margins prior to surgical excision of skin cancer. Lee et al. applied aminolevulinic acid and protoporphyrin IX to 142 patients with facial SCC, BCC or Bowen's disease and used a Wood's lamp to identify the disease boundary, which was not visible with out the fluorescent imaging. The authors reported the technique was limited by the low specificity due to uptake of the dyes by scars and sites of inflammation or

infection. In addition, the fluorescence was found to vary according to the thickness and depth of the tumour, making identification of margins difficult (Lee, Kim and Kim, 2010). The complex, laminate and inhomogeneous nature of skin produces optical characteristics that confound the quantitative analysis of fluorescence spectra. As such the intensity of the fluorescence is not consistent between individuals with the same pathology and this limits the techniques ability to classify disease type in vivo.

3.2.5.3 Diffuse reflectance spectroscopy

Diffuse reflectance spectroscopy utilises the scattering and absorption properties of skin to produce a reflectance spectrum. This spectrum is dependent on diffuse scattering and the absorption from pigments within the skin (Drakaki *et al.*, 2013). Significant differences in the reflectance spectra have been demonstrated between SCC, BCC and actinic keratosis (Rajaram *et al.*, 2010). The reduction in scattering seen in malignant disease is thought to be due to breakdown of the dermal collagen matrix. Although scattering from the malignant cell proliferation is increased the overall scattering is reduced due to the change in dermal collagen (Rajaram *et al.*, 2010). Regrettably, the heterogeneity of human skin between individuals and by anatomical site leads to large variations in the reflectance spectra, making the technique difficult to apply to the classification of malignant disease.

3.2.5.4 Laser scanning confocal microscopy

Laser scanning confocal microscopy allows detailed imaging of the different cell layers within skin. This allows analysis of skin architecture without the need for

invasive biopsy, at resolutions comparable to histopathology (Astner *et al.*, 2008). In reflectance confocal microscopy (RCM) the illumination of the tissue and detection occurs at the same wavelength of light. This technique allows the skin to be explored in three dimensions by shifting the focal area. Characteristic RCM features have been described for SCC, BCC, melanoma and normal skin and the method has shown promise for the classification of malignant lesions and identification of tumour boundaries at the time of surgery (Gerger *et al.*, 2005; Horn *et al.*, 2007; Guitera *et al.*, 2012). RCM imaging is dependent on the difference in refractive index between tissue structures to produce images whereas fluorescence confocal laser scanning microscopy (FCM) allows the imaging of endogenous or exogenous skin fluorophore distribution. FCM can be used to target imaging at specific cellular or subcellular molecules using specific fluorescent markers. Astner et al. used sodium fluorescein injected into the epidermis to image the morphological characteristics of malignancy seen on histology, including enlarged nuclei, tumour nests, pleomorphism, nesting of cells and abnormal vascularisation (Astner *et al.*, 2008). The main limitations of RCM are the imaging penetration depth of 200 – 300 µm, the subjective nature of the image interpretation and deficiency of biochemical information. Confocal Raman laser scanning microscopy can be combined with RCM and FCM to provide additional biomolecular information from the tissue under investigation.

The alternative optical diagnostic techniques discussed all have a potential role in the management of vulval skin diseases and the advantages and disadvantages are summarised in Table 3-2.

Table 3-2 – Comparison of alternative optical diagnostic techniques

Modality	Data analysed	Advantages	Disadvantages
Optical coherence tomography	Tissue structure, attenuation coefficient	Defines histological structure	Insufficient accuracy, subjective
Fluorescence spectroscopy	Field changes	High sensitivity	High false positive rate
Diffuse reflectance spectroscopy	Tissue structure	Objective	Low diagnostic accuracy at present
Confocal laser microscopy	Tissue structure	Detailed cellular imaging possible	Subjective analysis, operator dependent
Raman Spectroscopy	Biochemical	Objective rapid assessment	Relatively weak Raman signal
Infrared spectroscopy	Biochemical	Objective assessment	Difficult to apply in vivo

3.3 Vibrational spectroscopy in the assessment of lymph nodes

The ability to detect metastatic disease in lymph nodes at the time of surgery has the potential to remove the need to perform a second operation should a sentinel lymph node be found to be diseased after histological examination. Various techniques have been used to facilitate the intraoperative assessment of lymph nodes including frozen tissue section and histology; touch imprint cytology and molecular assays. Frozen section is associated with a long processing time and the need for immediate access to an expert histopathologist. Touch imprint cytology is a rapid method for preparing a cytological sample for analysis. The technique involves pressing both cut

sides of a bisected lymph node onto a slide, staining, immediately followed by examination by an expert histopathologist. Das et al. applied vibrational spectroscopic techniques to the traditional touch imprint technique and demonstrated the feasibility of cytoimprint FTIR for intraoperative diagnosis (Das *et al.*, 2008). Molecular assays use nucleic amplification techniques to detect tumour specific RNA markers. This technique is destructive and impedes the subsequent histological confirmation of the lymph node disease status. In addition there are currently no available assays for the detection of metastatic vulval SCC.

Within the last decade there has been increasing interest in the use of Raman spectroscopy in the intraoperative assessment and classification of lymph nodes (Horsnell, 2012). Raman spectroscopy is an attractive solution as it offers the opportunity for non-destructive real time intraoperative assessment of the presence of cancer cells within nodes.

The ability of Raman spectroscopy to differentiate between normal lymph nodes and those with lymphoma or metastatic carcinoma in the head and neck has been demonstrated, with reported sensitivities and specificities of over 81% and 89% respectively (Lloyd *et al.*, 2013a). The technique has also been shown to be able to differentiate between axillary lymph nodes with and without metastasis in the working theatre environment, with sensitivities of 92% and specificity of 100% (Horsnell *et al.*, 2010, 2012; Horsnell, 2012). To date analysis of inguino-femoral lymph nodes from women with vulval cancer using vibrational spectroscopic techniques has not been reported.

3.4 Clinical need for advanced vulval diagnostics

As previously discussed, the diagnosis and management of vulval disease is based on clinical assessment and histopathological analysis (Chapter 1). Techniques enabling rapid molecular assessment of vulval disease could have wide ranging applications, from initial diagnosis and molecular risk assessment to helping delineate disease margins at the time of surgery. These potential applications and existing alternative techniques are discussed below.

3.4.1 Diagnosis of vulval disease

Currently, there are no established non-invasive techniques available for the objective diagnosis and monitoring of premalignant or malignant vulval skin disorders. At present diagnosis is based on clinical assessment supplemented by invasive tissue biopsy. Various non-invasive techniques have been proposed for the identification of neoplastic vulval disorders including Toludine blue staining of vulval skin and cytological examination of skin scrapings or swabs.

Toludine blue offers high sensitivity for the detection of VIN, however the technique has a prohibitively low positive predictive value, found to be 100% and 24% respectively in a study by Joura et al. (Joura *et al.*, 1999).

The utility of cytological examination of cells from the surface of the vulva is limited by the significant keratosis present in many lesions. This results in the majority of cells collected being anucleate squamous cells which offer limited information

(Smith, 2010). Reported sensitivities vary significantly from 32% to 97% for the detection of premalignant vulval lesions (Nauth and Schilke, 1982; Levine *et al.*, 2001; Bae-Jump, Bauer and Van Le, 2007; Van den Einden *et al.*, 2012). The higher reported sensitivities result from exclusion of non-diagnostic specimens, artificially elevating the reported diagnostic utility. Non-diagnostic specimens account for around one quarter of samples collected using a standard cytological brush collection system.

Neither of the techniques discussed can offer the diagnostic performance required to make them a viable adjunct or alternative to invasive biopsy. Vibrational spectroscopy offers an alternative to traditional tissue biopsy for the diagnosis of vulval disorders.

3.4.1.1 Raman spectroscopy for in vivo real time assessment of vulval disease

As discussed (Chapter 3.2.3) Raman spectroscopy shows significant promise in differentiating different dermatological disorders. Investigation into the role of Raman spectroscopy in the vulva has been very limited to date and the potential applications of the technique will now be discussed.

The most tangible application of the technique is as an investigation to be used alongside clinical assessment of the vulva to determine if further invasive biopsy is necessary. To be useful in this context the technique needs to be able to correctly identify those without clearly identifiable disease (test negative). If this is achieved

then those that test negative can have biopsies to explore the exact nature of the disease.

Using women with suspected LS as an example. If we have a technique that can detect those patients without LS with high specificity, then those individuals testing positive for LS could be safely treated without biopsy and the remainder further investigated with biopsy. If we have 100 women with suspected LS in whom biopsy is indicated and in whom the prevalence of LS is 60%, the sensitivity of Raman spectroscopy is 75% and the specificity 98% then we will be able to correctly identify 75% of those with LS and avoid the need for invasive biopsy in 45 women. This would reserve the majority of invasive biopsies (73%) for those women in whom LS is not the likely diagnosis and therefore those patients where biopsy is needed in order to obtain a diagnosis.

In other clinical situations the requirements of the investigation will be different. When assessing women with suspected precancer or cancer we need to ensure women with cancer have the nature of the disease explored with invasive biopsy. In this circumstance a high sensitivity is required. These varying requirements need to be accounted for in the diagnostic design of a Raman spectroscopic system.

A Raman diagnostic system also has the potential to guide where biopsies should be taken by identifying molecular changes associated with disease and increasing the chance of sampling the area of interest. This in turn could reduce the number of biopsies required to produce a definitive diagnosis.

Table 3-3 – Potential applications of Raman spectroscopy on the vulva

Application	Current gold standard	Putative advantages	Putative disadvantages
Clinical decision aid to guide need for invasive biopsy	Visual inspection/ magnified vulvoscopy	Avoidance of invasive biopsy Targeting of biopsies	
Optical biopsy (substitute to invasive biopsy)	Histopathology	Avoidance of invasive biopsy Faster assessment	Reduced sensitivity compared to gold standard
Molecular risk assessment	No established alternatives	Additional information to that provided by histopathology	Difficult to establish biochemical basis for assessment
Intra-operative guidance (detection of disease margins)	Visual inspection	Improved complete excision rates Reduced extent of surgery	

The ability to use the technique as a definitive optical biopsy and eliminate the need for invasive biopsy is currently limited, as Raman spectroscopy is unable to reproduce the full spectrum of information obtained by histopathological examination. The potential role of the technique is as an adjunct to standard assessment and to provide near real time assessment of pathology to guide further investigation and management.

3.4.1.2 Raman spectroscopy for intraoperative assessment of disease margins

During the surgical treatment of VIN or vulval cancer removing all cancer or VIN disease reduces the risk of the disease recurring. Current guidelines for the management of vulval cancer recommend the excision of primary tumours with a

minimum margin of 15 mm to ensure a histological margin of greater than 8 mm (RCOG, 2014). Inadequate excision margins have been repeatedly identified as an adverse prognostic factor and observational studies have identified the disease free margin to be an important predictor of primary tumour site recurrence. Heaps et al. study of 135 cases of vulval SCC reported that local control was achieved in 100% of cases where a histological margin of 10 mm or more was achieved and by contrast a margin of less than 8 mm was associated with a 47% local recurrence rate (Heaps *et al.*, 1990). The association between histological margin and disease recurrence has subsequently been demonstrated by other authors and reported recurrence rates vary from 20 to 50% in those with a pathological margin less than 8 mm and from 0 to 21% in those with a margin 8 mm or more (Burke *et al.*, 1995; De Hullu *et al.*, 2002; Chan *et al.*, 2007). More recent studies could not demonstrate the importance of a pathological margin >8 mm, however complete excision of the primary tumour remains an important prognostic factor (Groenen, Timmers and Burger, 2010; Höckel *et al.*, 2010; Woelber *et al.*, 2011). Observational studies have also demonstrated that despite the intention to attain a 8 mm pathological margin, this was only achieved in 14 to 40% of cases (Rhodes, Registrar and Lecturer, 1998; Rouzier *et al.*, 2002; Groenen, Timmers and Burger, 2010; Woelber *et al.*, 2011).

There is currently, no non-invasive near real time objective diagnostic technique that can be used for the identification of vulval tumour margins during surgery. A technique capable of determining tumour margins intraoperatively would, have the potential to reduce incomplete excision rates and in turn reduce the rate of local vulval cancer recurrence, whilst minimising the extent of surgery. Raman

spectroscopy has the potential to be developed as a valuable diagnostic aid in these circumstances.

In addition, as Raman spectroscopy elicits molecular information from tissue it may be possible to define a molecular surgical margin containing histologically benign tissue with molecular changes that predispose to the development of cancer. Although it is outside the scope of this thesis future studies may assess the prognostic value in defining surgical margins in this way.

3.4.2 Intraoperative assessment of lymph node metastasis

Vulval SCC is typically treated with excision and inguinofemoral lymphadenectomy, dependent on the size and position of the tumour (Chapter 1.4.4). The lymphadenectomy is associated with significant morbidity in 70% of cases: related to poor wound healing, infection, lymphoedema and lymphocyst formation (Lawrie *et al.*, 2014). This has led to significant interest in detection and assessment of sentinel lymph nodes in the treatment of vulval cancer. Sentinel nodes can be detected using a technetium tracer and blue dye injected around the primary lesion. Removal and analysis of the sentinel node can then be used to guide the need for further treatment. Currently the excised lymph node is either sent for immediate frozen sectioning and histopathological analysis or routine tissue fixation and paraffin histology. If the initial analysis of the sentinel node is negative then ultra-staging techniques can be used to detect micro-metastasis. If the node is found to contain cancer cells then further treatment with either full inguinofemoral lymphadenectomy or radiotherapy to the inguinofemoral lymph nodes is warranted. The safety of

omitting complete inguinofemoral lymphadenectomy in patients with a negative sentinel node is uncertain and currently the subject of the GROningen INternational Study on Sentinel nodes in Vulvar cancer (GROINSS-V) II observational study.

Currently there are no objective techniques for the assessment of sentinel lymph nodes at the time of initial surgery for vulval cancer. It is not possible to perform routine histopathology or ultra-staging at the time of surgery due to the need for tissue fixation. Frozen section analysis is dependent on the availability of an expert histopathologist and only assesses single sections through the nodes, so may miss metastases at other points within the node. Raman spectroscopy offers an objective real time adjunct to histopathological examination of lymph nodes and shows promise for the detection of diseased lymph nodes at the time of surgery (chapter 3.3). Such a technique could be performed in the absence of a histopathologist and would allow all surgery to be completed in one session without the need for a second operation, in situations where full inguinofemoral lymphadenectomy is required.

3.4.3 FTIR spectroscopy for spectroscopic aided detection of disease and augmented pathology

3.4.3.1 Spectroscopic aided detection of disease

When caring for women with suspected vulval cancer it is clearly important that histological diagnosis is accurate and delivered in a timely manner. There is recognition that delays in reporting pathological specimens and errors in reporting can negatively impact on patient care (Frable, 2006; Hirschowitz, Wells and Lowe,

2013; Volmar *et al.*, 2014). Studies analysing the rates of errors in histopathological practice by secondary review of the pathological specimens have revealed significant discrepancies exist between reports completed by different pathologists. The disagreement rate ranges from 9.0% to 18.7% for minor disagreements and from 1.4% to 6.8% for major disagreements (Kronz, Westra and Epstein, 1999; Weir, Jan and Colgan, 2003; Manion, Cohen and Weydert, 2008; Middleton *et al.*, 2014). Major discrepancies were defined as errors that affected or had the potential to affect patient care. These errors may lead to over or under treatment of significant disease and is an important concern for patients. Double reporting where two pathologists examine specimens has been proposed as a mechanism for reducing this diagnostic error, although it is not standard practice within the UK. Current recommendations suggest double reporting should only be used routinely in certain scenarios, such as some cases of high grade dysplasia in the gastrointestinal tract, melanomas and severely atypical naevi. Double reporting should also be considered for rare or uncommon tumours (Hirschowitz, Wells and Lowe, 2013). Whilst consultant histopathologists are considered independent practitioners, secondary review or double reporting of specimens is often used in general histopathological practice before a final report is issued. The current system relies on the individual pathologist's assessment of their own limitations and recourse to a second opinion when they judge this to be appropriate (Lind *et al.*, 1995).

With the increasing burden present on health services, an expanding repertoire of available tests, new clinical guidelines and fear of litigation, there is increasing pressure on the time available to histopathologists. In addition to this many

consultant histopathology posts and training posts within the United Kingdom are unfilled, further increasing time pressures (Centre for Workforce Intelligence, 2010). In light of these pressures a technique that would automate the process of pathological second opinion would be valuable. This would enable pathologists to obtain corroboration of their diagnosis and reduce the need for a second histopathologist. Automated analysis of digitised histopathological images, fluorescence imaging and vibrational spectroscopy have all been proposed as possible technological solutions (chapter 3.2.5). FTIR–S offers a computer aided detection technology based on the molecular composition of tissues and has real potential to offer a solution in this context (chapter 3.2.2). A FTIR–S aided detection system could be used by pathologists to help them reduce the occurrences of false positives and false negatives. Compared to chemical and immunohistochemical stains FTIR - S imaging has some distinct advantages, including minimal need for unusual sample preparation, no need for special stains or contrast agents and since the information is numeric it is highly amenable to objective computational analysis. Integrated into the pathology workflow spectroscopic analysis could sit alongside traditional histopathology to aid with histopathological decision making (Figure 3-1). The practical requirements of speed of data acquisition and processing have made the implementation of spectroscopic aided detection system in the healthcare setting difficult. These problems are being address and recent advances have made good progress towards enabling the technology to provide high throughput data acquisition within a compact unit at a realistic cost (Farries *et al.*, 2015).

When spectral analysis and data processing is quicker than traditional histopathology, the technique has the added advantage of being able to act as a triage mechanism, whereby those cases that are likely to contain an urgent diagnosis are flagged for priority pathological assessment. In the context of vulval disease the diagnosis of non-neoplastic, premalignant or malignant vulval disease could be confirmed using a spectroscopic aided detection technique. This has the potential to ensure malignant disease is not missed and that morbid surgical treatment of vulval cancer is not performed unnecessarily. An example of where the technique may be useful is in cases where a biopsy shows possible early invasive disease but the diagnosis cannot be clearly made. Currently in this situation a further biopsy and delay in diagnosis is usually necessary, however this may be avoided if spectroscopic aided detection of disease is available. In addition, FTIR spectroscopy has the potential to be used as a histopathological adjunct to improve the early diagnosis of SCC. FTIR spectroscopy has the potential to detect molecular changes, which identify those in whom further investigation is likely to find an occult malignancy. Early diagnosis would be advantageous as it is associated with a improved clinical outcome and survival (Schilder and Stehman, 2012).

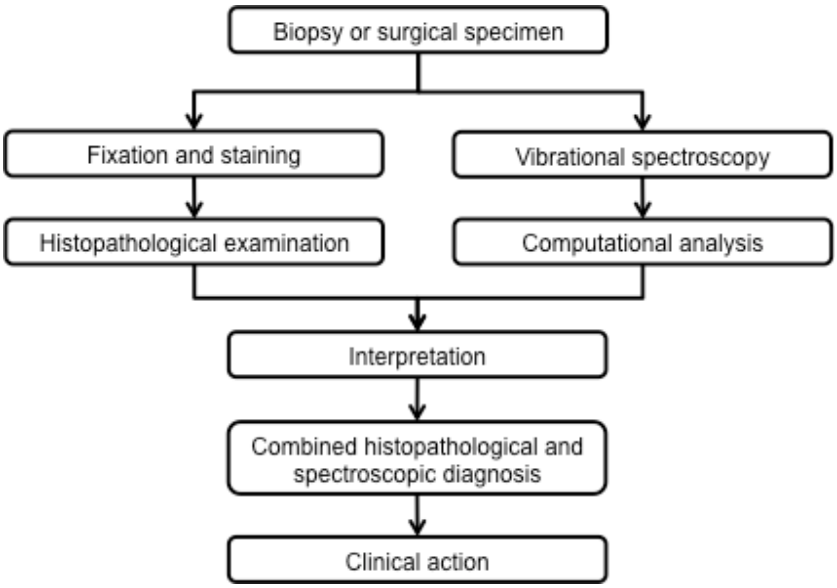


Figure 3-1 – Overview of the use of vibrational spectrscopy for augmenting histopathological diagnosis

3.4.3.2 FTIR spectroscopy for augmenting pathology

Another area in which molecular spectroscopic analysis may be advantageous is in providing prognostic information based on the chemical properties of tissue specimens. A recent study assessing the role of FTIR spectroscopy for the identification of prostate cancer at increased risk of recurrence demonstrated chemical changes that could be employed to outperform traditional recurrence prediction methods (Kwak *et al.*, 2015). Another study by Baker et al. investigated the use of FTIR spectroscopy in evaluating prostate cancer, and demonstrated the association between FTIR spectroscopic changes and clinically aggressive prostate cancer (Baker *et al.*, 2008). To date research into the role of vibrational spectroscopy for the identification of prognostic markers is limited and no studies assessing the role of the technique in vulval disease were identified.

There is a broad range of molecular changes in the vulval that contribute to the development of vulval cancer (chapter 1.4.1). Vibrational spectroscopy not only examines molecular changes resulting from known provocations (e.g. TP53 mutation and HPV integration) but also assesses other undescribed molecular changes (Ostrowska *et al.*, 2011). This gives a broader picture of the molecular changes present, but at the expense of detailed biomolecular causation, as vibrational spectroscopy does not allow us to determine the exact molecular alterations underlying the spectral differences identified.

Traditional histopathological analysis gives limited information about the biomolecular changes associated with an increased risk of malignant transformation. Vibrational spectroscopy has the potential to augment traditional histopathology, giving pathologists additional prognostic molecular information from the tissue under examination. As previously described women with LS and VIN are at increased risk of developing vulval SCC. Using vibrational spectroscopic analysis for molecular risk stratification of patients with VIN or LS may facilitate the identification of those at high risk of developing vulval SCC. In addition, the analysis of cancerous vulval disease has the potential to facilitate risk stratification, which could be used to modify management to prevent recurrence and ensure recurrences are detected early.

3.5 Conclusion

In conclusion, vibrational spectroscopy offers techniques for molecular assessment that have a relatively low cost; that only require a small amount of tissue and that

can be performed in a short amount of time (Old et al. 2014). Based on the findings from diagnostic investigation on skin and other areas of the body vibrational spectroscopic techniques have the potential to be usefully applied to both the initial diagnosis of vulval skin conditions in clinic and the histopathology laboratory as well as for the monitoring for malignant changes in women with precancerous vulval skin conditions.

In the management of vulval cancer spectroscopic probes offer a potential real time intraoperative diagnostic tool to detect non-visible disease margins and improve complete excision rates, and avoid disease recurrence, whilst minimising the extent of surgical excisions. In addition, vibrational spectroscopy offers a potential tool for intraoperative objective assessment of sentinel lymph nodes.

Vibrational spectroscopy also offers the potential to give biomolecular prognostic information that can be used to evaluate the risk associated with different types of vulval skin disease.

Section B: General Methods

Chapter 4 General Technical Considerations and Chemometric Analysis

This chapter outlines the general technical considerations to be considered when applying vibrational spectroscopy for the diagnosis of disease. This includes discussion of sources of noise and other variation in spectral data that is independent of signal from the tissue under analysis. The strategies for processing the spectral data to minimise the impact of non-tissue signal are discussed alongside the multivariate chemometric analytical techniques used for the classification of tissue type based on vibrational spectra.

4.1 Noise and interference

4.1.1 Sources of noise

The spectra obtained from both Raman and FTIR spectroscopy comprises the signal from the sample under analysis and noise. The main identifiable origins of noise in vibrational spectra include shot noise, dark noise, readout noise, quantisation noise and background light noise, these are summarised in (Table 4-1). This list is not exhaustive and there may also be other sources of noise whose origin is not readily identifiable.

The signal to noise ratio (SNR) is a useful measure of the intensity of the signal, relative to the noise. A high SNR is desirable in order to be able to identify the signal spectra from the noise (Lewis and Edwards, 2001; Brereton, 2003). Experimental protocols in this study have been designed to maximise the signal to noise ratio in the spectra recorded whilst maintaining spectral resolution and further detail on this is given in subsequent chapters.

The shot noise is a fundamental limit of photon spectroscopy. Even with a spectrometer that eliminates all other sources of noise the SNR will still be subject to the shot noise limit.

Type of noise	Summary	Mitigation	Spectroscopy technique affected
Shot	Noise resulting from the random nature of photons arriving at the spectroscope detector. It is proportional to the square root of the number of photons arriving at the detector.	The SNR due to shot noise increases with increased acquisition time or increased laser power.	Raman FTIR
Dark	Noise produced by the thermal generation of electrons within the spectroscope detector, independent of photons striking the detector. Dark noise is proportional to the duration of acquisition.	Cooling the detector reduces the dark noise. The Dark noise can also be measured and subtracted from the measured spectra.	Raman FTIR
Readout	Noise due to inaccuracies in measuring the number of electrons produced by the detector.		Raman FTIR
Quantisation	Noise relating to the resolution of conversion of the analogue signal from the detector into a digital signal.		Raman FTIR
Background light	Noise from light striking the detector that is not from the sample under analysis, for example light from the optical system or stay light from other sources.	Reducing the background light and using optics that do not scatter light will reduce the background light noise	Raman

Table 4-1 – Sources of noise in vibrational spectroscopy with principles for mitigation where applicable

4.1.2 Evaluating the signal to noise ratio

The SNR is calculated by dividing the relevant signal intensity by a measure of the noise in the spectrum. Noise can be estimated by assessing the variation in spectral intensity in a section of the spectrum where there is little or no signal from the sample under analysis. There is no standard method for calculating the signal to noise ratio. For example, the signal can be taken as the difference between the height of a peak of interest and the baseline of the spectra adjacent to the peak or the baseline of the whole spectrum. Likewise, the noise can be calculated as the simple range of an area of the spectrum with little or no signal or in more advanced ways such as the root mean squared of the difference from the mean in an area with no signal. It is important that the method used for calculating the signal to noise ratio is appropriate to the technique under investigation and consistent when applied to assess different spectra. The exact methods chosen for each application are discussed further in later chapters.

4.1.3 Other non-analytical interference

4.1.3.1 Cosmic rays

During the acquisition of Raman spectra high energy cosmic rays can strike the spectroscope detector, causing a peak in the spectra which is not as a result of signal from the sample under analysis. As the detector used in FTIR spectrometers is not sensitive to cosmic rays this type of interference is not relevant to FTIR spectroscopy. These unwanted signals can be avoided by taking the median spectra

from three consecutive acquisitions, as it is unlikely a cosmic ray will be present at the same place in two of the spectra. The median is taken rather than the mean because the mean can be significantly altered by the presence of a cosmic ray in one of the spectra whereas the median is unlikely to be altered. This triple measurement to avoid cosmic rays does however significantly increase the time needed to acquire the Raman spectra. An advantage of this median of three spectra approach is that taking the median counts at each wavenumber reduces the noise in the final spectrum. As three separate spectra are needed in this approach it takes longer to achieve the same signal to noise ratio as compared to a single acquisition (Figure 4-1 and Figure 4-2). With this cosmic ray reduction technique approximately double the acquisition time is necessary to gain spectra of similar quality to those obtained without this technique.

In spectral maps, where spectra are taken in a grid or line from adjacent sites close together, another type of median filter can be used to remove the effect of the cosmic rays. In a median filtered two dimensional Raman map the spectrum at the centre of a median filtering grid is defined as the median of the spectra at that point and all surrounding spectra depending on the median filter grid size (Figure 4-3). When spectra are not taken from sites close together neighbouring spectra may originate from molecularly different substances and so this approach can result in nonsense spectra.

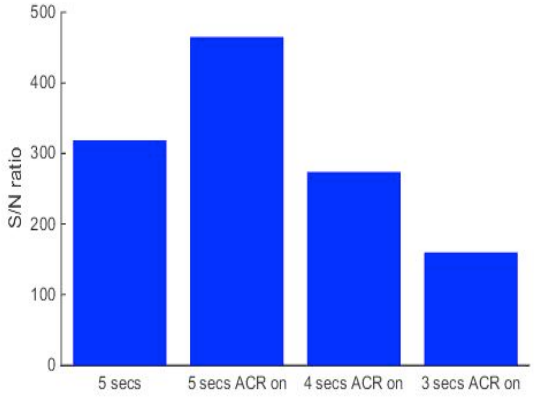


Figure 4-1 – Chart of signal to noise ratio of spectra from PTFE block with varying acquisition time with and without automatic cosmic ray removal (ACR)

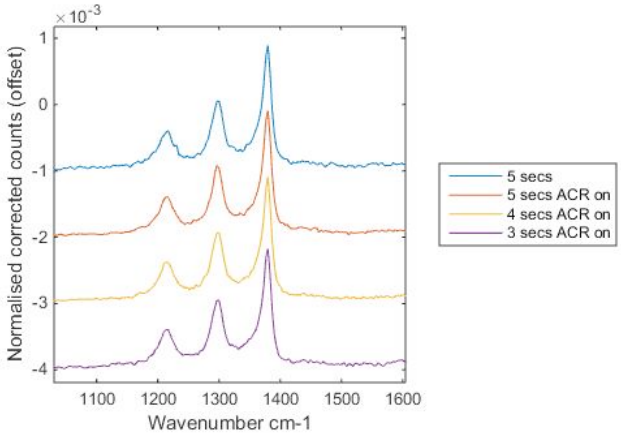


Figure 4-2 – Spectra from PTFE block with varying acquisition time with and without automatic cosmic ray removal (ACR)

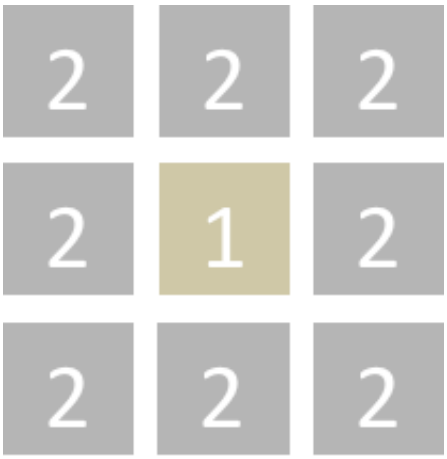


Figure 4-3 – Example of a three by three median filtering grid. The spectrum at 1 is the median of all spectra in the three by three grid (1 and 2)

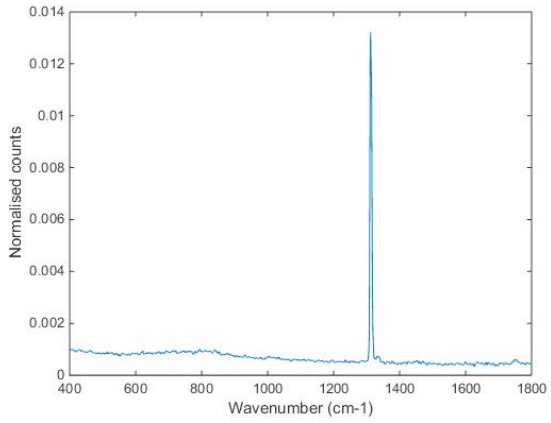


Figure 4-4 – Tissue spectrum with narrow cosmic ray

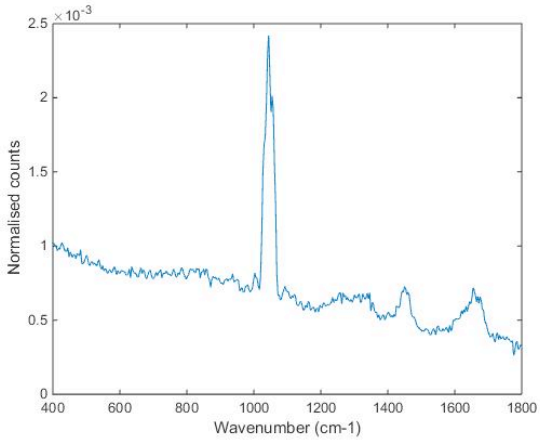


Figure 4-5 – Tissue spectrum with broad cosmic ray

Another approach to reduce the contribution from cosmic rays in spectra is to filter the measured spectra to remove the part of the spectrum consistent with cosmic rays. These areas of the spectrum distorted by cosmic rays can then be replaced with an approximation of what the spectrum is expected to be (Phillips and Harris, 1990; Cappel, Bell and Pickard, 2010). When this approach is used in this study the areas of the spectrum removed due to cosmic rays are replaced using a reconstructed spectrum based on PCA loadings from the entire data set and the predicted scores of the individual spectrum without the cosmic ray (chapter 4.2.1). When the cosmic rays are narrow with high detector counts (Figure 4-4) this approach works well as the sections of the spectra due to cosmic rays are readily identified. In addition when the cosmic ray is narrow the adjacent areas of the spectrum can be used to provide a good approximation of the expected spectrum where the cosmic ray is found. This approach is limited as some cosmic rays produce broad peaks with relatively low counts (Figure 4-5). Filtering these spectra to

remove the cosmic rays is more challenging. Algorithms to remove broad based rays also risk removing peaks of interest within the spectra.

4.1.3.2 Fluorescence

When a photon incident on a sample has sufficient energy to excite an electron between two energy states a photon will be emitted when the electron relaxes to its ground energy state. This process gives rise to fluorescence and can produce photons at wavelengths similar to those photons produced by Raman scattering. As fluorescence can result in a greater number of emitted photons than Raman scattering the Raman spectra can be concealed by the fluorescence. This effect is especially relevant in pigmented skin that contains large amounts of melanin which fluoresces intensely with near infrared excitation (Bodanese *et al.*, 2012; Silveira Jr *et al.*, 2012; Wang *et al.*, 2012).

This phenomenon can be avoided by selecting a laser excitation wavelength that does not have sufficient energy to excite molecular fluorescence. For this reason use of longer wavelength of laser should reduce the fluorescence from the samples. As the intensity of Raman scattering is inversely proportional to the excitation wavelength (to the fourth order), a longer excitation wavelength will not only reduce the background fluorescence but also the Raman signal. Therefore using a longer excitation wavelength requires a higher power laser source or longer acquisition time. It is for these reasons that an 830 nm laser is commonly selected for biophotonic applications (Stone, 2001; Butler *et al.*, 2016).

Another strategy for the reduction of fluorescence is the use of photobleaching. This technique is based on the principle that fluorescence from tissues decreases over time and that the Raman signal stays relatively constant. Exposing tissue to light that invokes fluorescence for a period of time prior to capture of the Raman spectrum can potentially reduce the contribution of the fluorescence in the final spectrum (Wang *et al.*, 2012).

If fluorescence cannot be reduced prior to measurement of spectra then it can be partially corrected after measurement by utilising a polynomial baseline correction. Assuming the fluorescence can be approximated with a low order polynomial fitted to the baseline of the spectra then the contribution from the fluorescence can be removed by subtracting this polynominal from the spectrum (Figure 4-6).

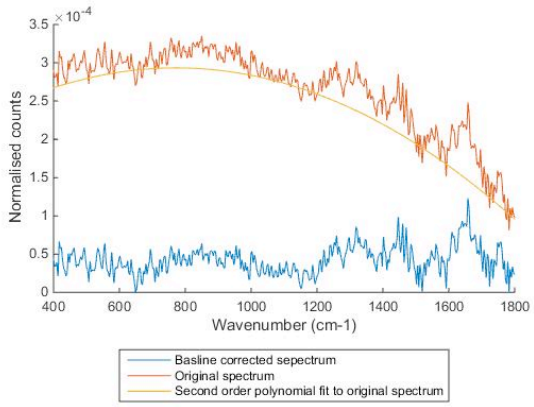


Figure 4-6 – Example of the effect of polynominal fit fluorescence reduction on a point Raman spectrum taken from an area of vulval melanosis (acquisition time 60 seconds)

4.2 Chemometric analysis

Vibrational spectra measured from biological samples contain contributions from a diverse range of different constituent molecules and as a result can appear complex. In addition, the spectral differences between different tissue types and pathology groups can be subtle and not obvious on visual inspection. Due to these factors multivariate analysis is required to explore the variance between spectra and develop classification models to identify subtle spectral differences. These classification models can then be used to determine the pathology present in tissue from the measured spectra (Defernez and Kemsley, 1997; Lloyd *et al.*, 2013b; Old *et al.*, 2014). From this the diagnostic performance can be assessed. Numerous statistical approaches have been used to interrogate spectral data in this way: including principal component analysis (PCA), linear discriminant analysis (LDA), partial least squares (PLS) and support vector machines (Lloyd *et al.*, 2012).

All these multivariate techniques used for the classification of tissue vibrational spectra require validation with an independent test set. This can be achieved using leave one sample out cross validation as this allows maximal use of the available data whilst minimising the risk of a type one error. Within this thesis spectral analysis was based on cross-validated PCA fed LDA and this is discussed in more detail below. All data analysis in this thesis was undertaken with Matlab® R2015b (Mathworks, USA) software using custom coded algorithms.

4.2.1 Principal component analysis

PCA is an unsupervised technique for multivariate pattern recognition which is well suited to handling large datasets (Brereton, 2009; Lloyd, 2009). PCA converts the spectral data into a set of uncorrelated variables called principal components (PC). This results in principle component scores (T) and loads (P) that represent the original data matrix (X) with error E as shown in Equation 4-1. When all principal components are included the error E will be zero.

$$X = TP + E$$

Equation 4-1

The principal components contain the main variation within the data set. The Eigenvalues calculated during PCA give a measure of the variation in the data accounted for by each PC and the PCs are ranked according to these values. The first principal component accounts for the greatest amount of variation within the data. Each PC is orthogonal to all other components and therefore represent independent sources of variation within the data, in data with only two sources of variation there will only be two principal components. In spectral data measured from tissue there are numerous sources of variation and these result in numerous principal components. For this reason, analysis is often limited to the PCs that account for the most variance within the data after the exclusion of other PCs that may describe noise or other variation that is not helpful in pathological group classification.

PCA is helpful in the analysis of spectral data as it reduces the dimensionality of the data without loss of important discriminating features. As PCA does not include information about pathology grouping, it is an unsupervised method and does not provide information on group membership. To classify spectra according to pathology group, a supervised analysis of relevant variation between spectra (principal components) is required. For the purposes of this thesis LDA is used to create the supervised classification model (Chapter 4.2.3).

4.2.2 Selection of principal components for LDA

Analysis of variance (ANOVA) can be applied to the principle component scores and the pathological groupings to assess which of the principal components are statistically significantly different between the pathological groups. These significant principal components can then undergo LDA to allow classification of pathology groups according to PC scores.

4.2.3 Linear discriminant analysis

Multiclass LDA can be used as a supervised method to maximise the variance in PC scores between pathology groups whilst minimising the within group variance. LDA analyses the PC scores of each of the pathology groups to produce LD scores and LDA loadings. This enables LDA to use the covariance of the whole dataset to define a linear discriminant border between the groups (Figure 4-7). This resultant model uses the distance of the LD scores to the mean of each group to predict to which pathological group each spectrum belongs. Each prediction is associated with a

given probability of that prediction being correct. This is based on the distance to the linear discriminant boundary and the mean of each group. Unlike histopathological analysis where a sample may be unclassifiable, LDA will always attribute a spectrum to a pathological group irrespective of how likely that classification is. This can lead to misclassification of data that does not fit well within the model as all data points will be assigned a probability of group membership based on the distance to the LD boundary (Figure 4-7). For this reason it is also important to consider the Mahalanobis distance from the mean of each group as greater the distance the poorer the fit to the model and the lower the probability that the predicted group is correct.

4.2.4 Group classification

Spectra measured from a sample of unknown pathology can be classified by a LDA classification model by estimating the PCA scores from the spectra and the PCA loadings (Equation 4-2). The predicted LD scores can then be estimated from the LDA loadings and the estimated PCA scores (Equation 4-3). These predicted LD scores can then be projected onto the LDA classification model covariance matrices. The LDA classification is then based on the distance of the predicted LD scores to the mean LD scores of the groups scaled by the intragroup variance. This results in group classification of the test (unknown) spectra (Figure 4-8).

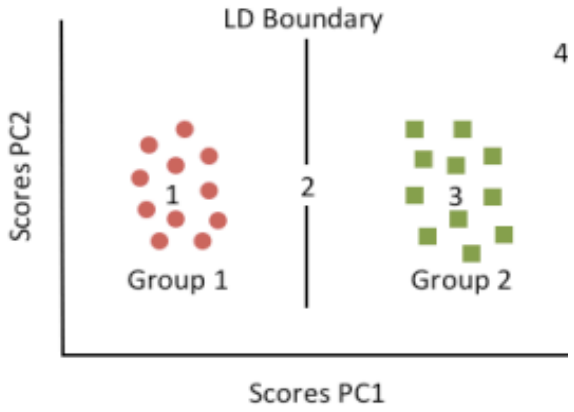


Figure 4-7 – Diagram demonstrating LDA classification of groups based on the scores of two PCs in a two group model. At point 1 (the mean of group 1) the predicted probability of membership of group 1 is 1 and the probability of membership of group 2 is 0; at point 2 on the LD boundary the probability of membership of either group is 0.5; at point 3 (the mean of group 2) the predicted probability of membership of group 1 is 0 and the probability of membership of group 2 is 1; and at point 4 the probability of membership of group 1 is 0 and the probability of membership of group 2 is 1 although the point is distant to the mean of group 2 and not a good fit to the LD classification.

$$X_{test}\, P = T_{estimated}$$

Equation 4-2 – Estimation of test set scores ($T_{estimated}$) from test spectra (X_{test}) and PCA loadings (P)

$$D = T_{estimated}\, (B)^{-1}$$

Equation 4-3 – Estimation of test set LD scores (D) from test spectra PCA scores ($T_{estimated}$) and LDA loadings (B)

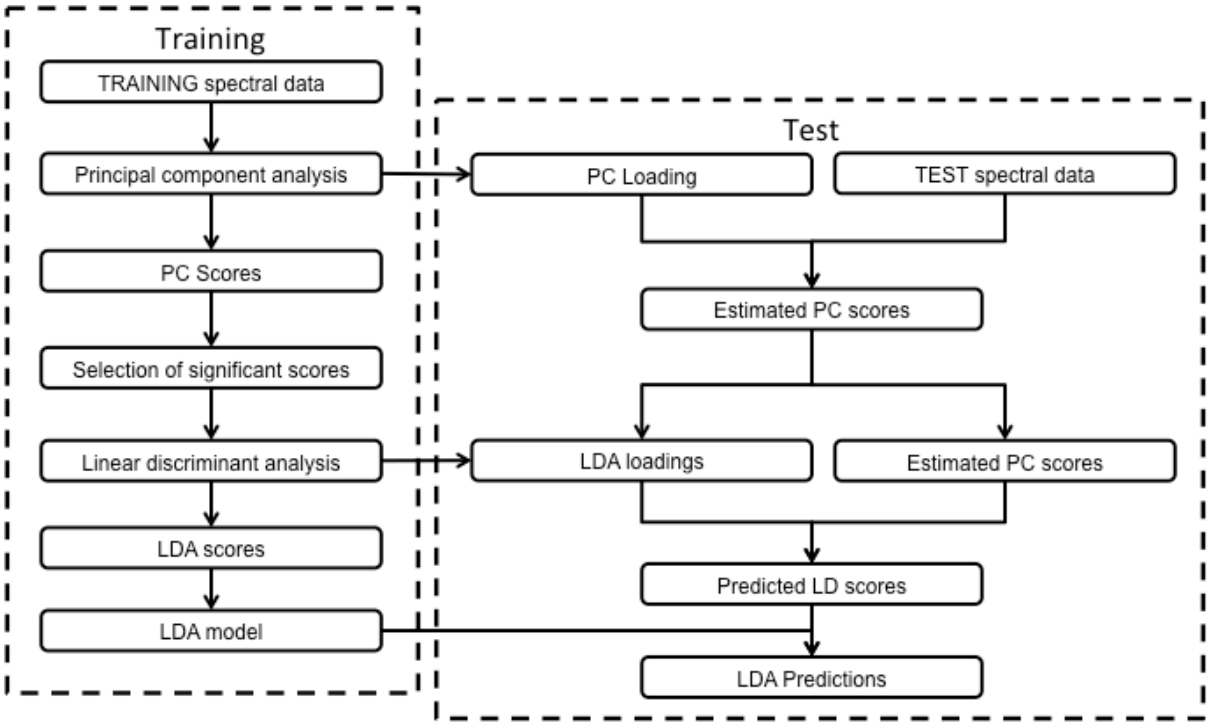


Figure 4-8 – PCA fed LDA classification model

This PCA fed LDA classification approach (Figure 4-8) can be used to predict the pathology group of each individual spectrum measured from a tissue sample. In LDA it is possible to move the discriminatory boundary to adjust the performance of the test to the application. For example if a high sensitivity is required a classification that favours the 'positive' group can be achieved by choosing a discriminant boundary further from the mean of that group. This would be at the expense of the specificity of the classification.

In addition, in clinical practice it is often the classification of the whole sample or patient that is important. Therefore if multiple spectra are measured from each sample, the group can be classified according to the proportion of spectra with each classification within the sample. This approach can also be individualised to the

clinical application. For example, when looking to identify cancer then a high sensitivity may be prioritised at the expense of specificity by using a lower threshold number of abnormal spectra to classify the sample as cancer.

4.2.5 Validation of results

It is important to validate predictions given by the PCA LDA classification model as testing the performance of the classification model using spectral data used to train the model may result in classification based on factors unrelated to the pathology groups of interest. Failure to adequately validate may result in an over estimation of the diagnostic performance and a type one error. If a large number of samples and spectral measurements were available then it would be possible to create a model from a training set of data and validate with data from an entirely different test data set. This approach requires a large number of participants and is not suitable when the number of participants is limited. When the number of participants is limited sequentially excluding some data as a test set and using the remaining data as a training set to create the classification model that is then tested on the excluded data is an alternative approach (Lloyd *et al.*, 2012). This approach allows most efficient use of the available data while reducing the risk of a type one error. We can use this principle in a leave one sample out cross validation loop in which each sample analysed is excluded in turn from the whole data set and tested against a classification model created from the remaining training set (Figure 4-9). By excluding all the data from a given participant in each loop then classification based on inter-participant variance can be avoided.

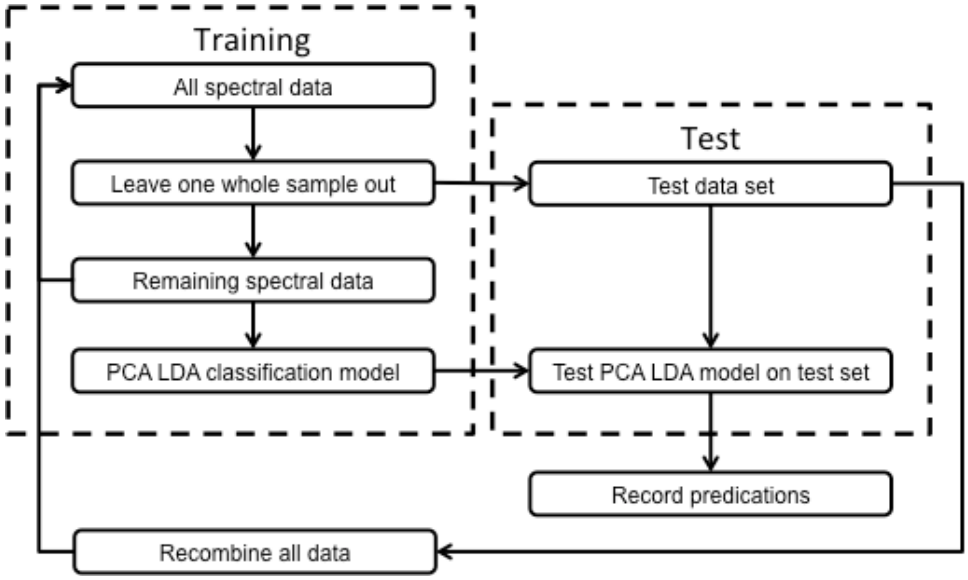


Figure 4-9 – Leave one sample out cross validation loop

4.2.6 Methods for the exclusion of outlying spectra

When multiple spectra or spectral maps are measured from a sample, the spectral data recorded may contain spectra that do not describe features typical of the type of sample under analysis and may not fit well within a PCA LDA classification model (Figure 4-7). These outlying spectra may result from non-tissue signal such as atmospheric water or contamination of the collection optics. It is important to exclude these outliers from PCA because if the spectral signal is significantly different from other tissue spectra this irrelevant variance will be described by the principal components. As this variance is not relevant to the classification of tissue groups it may impede the performance of the resultant LDA classification model.

4.2.6.1 D - and Q - statistic

To identify spectra for exclusion from the classification model that are not characteristic of tissue spectra a measure of the PC reconstruction error (Q statistic) and the difference from the mean spectrum (D statistic) can be used (Brereton, 2009; Brereton and Lloyd, 2016). Both these techniques are applied to spectral data after PCA has been performed (Figure 4-10). The identified outliers can then be excluded from the original spectral data set and PCA repeated for the purposes of the PCA LDA classification.

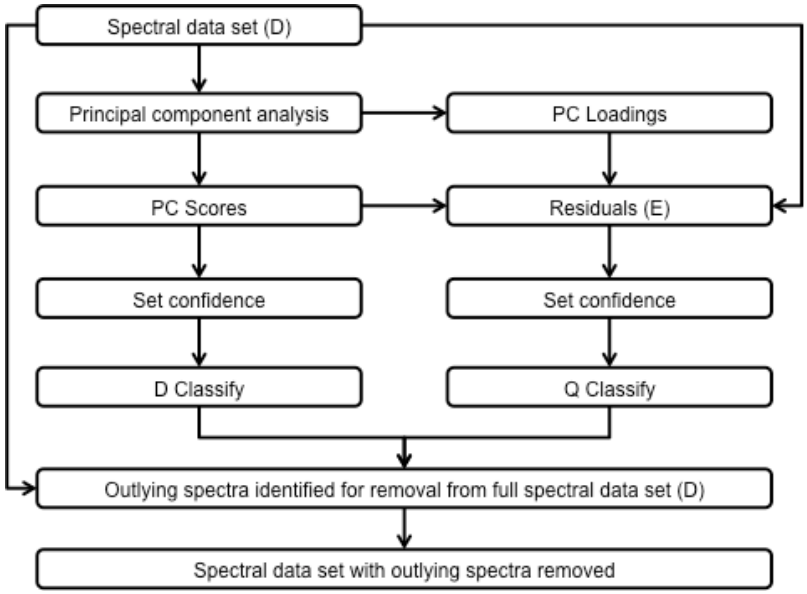


Figure 4-10 – Application of D and Q statistics for the exclusion of spectral data significantly different to all raw data

The application of the D statistic allows the calculation of the Mahalanobis distance from the mean of the PC scores and gives a specific confidence that a sample is a member of that group (e.g. a spectrum typical of the majority of spectra). The

spectra whose corresponding scores are outside of this distance can be excluded from the original data set (Figure 4-11).

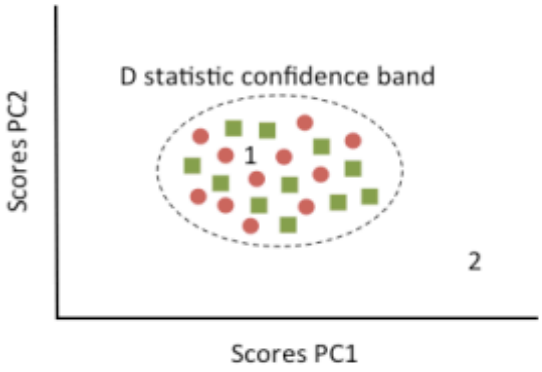


Figure 4-11 – Diagram showing D statistic confidence band where point 1 is within the confidence band and point 2 is outside the confidence band

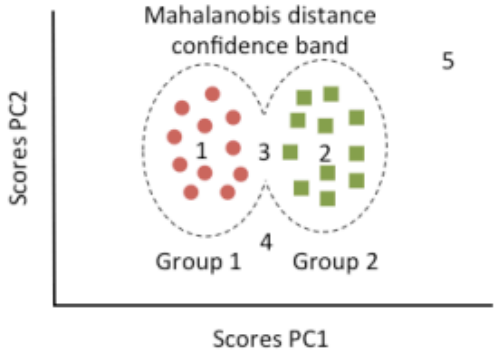


Figure 4-12 – Diagram showing the Mahalanobis distance confidence band from the means of the LD scores groups where points 1, 2 and 3 are included in the confidence band and points 4 and 5 are outside of the band.

The Q statistic or squared prediction error can be used to estimate how well a spectrum fits into a PC model of all data. It is calculated from the sum of squares of the residual error (E) obtained from the PCA model of the raw data. This is compared to a normal distribution of residual errors and threshold defined by the specific confidence that a sample is a member of that group or tissue type. When the Q statistic for a spectrum is outside of the set confidence limit this means the PCA model does not describe the spectrum well and that a different type of variation, not predominant in the raw data set, is present within the spectrum.

4.2.6.2 Mahalanobis distance

Once PCA LDA has been performed the resultant model will provide a prediction for all spectral data regardless of whether the data fits well within the model (Figure 4-7). A confidence interval derived from the combined Mahalanobis distance from the mean of each of the classification groups can be used to exclude predictions for spectra that do not fit well into the LDA model (Figure 4-12). By using this method we can restrict predictions to spectra that fit well within the LDA model and pathology groups present in the model.

4.2.7 Reporting diagnostic accuracy

Receiver operator curves are widely accepted as the standard method for reporting and comparing the accuracy of medical diagnostic tests (Zhou, Obuchowski and McClish, 2011). The performance of a diagnostic test is described by the sensitivity (i.e. ability to detect disease when present) and the specificity (i.e. ability to recognise disease absence when disease not present). The receiver operator curve (ROC) describes the sensitivity of the test over all possible false positive rates (i.e. 1-specificity). The ROC allows us to derive measures of accuracy that can be used to compare different diagnostic tests independent of the prevalence of the disease under investigation. The area under the ROC curve (AUC) combines both the key measures of accuracy (sensitivity and specificity) into a single measure of diagnostic accuracy and is the primary measure of diagnostic accuracy used in this thesis.

Section C: Raman Spectroscopy for the Assessment of Vulval Disease

Chapter 5 Introduction, aim, objectives, ethics and recruitment

5.1 Introduction

As discussed in chapter 3.4.1 there are currently no established alternatives to invasive biopsy for the assessment of women with vulval disease. The diagnosis of vulval disease can be clinical, although invasive tissue biopsy and histopathological evaluation is necessary if the diagnosis is uncertain; if there is suspicion of cancer or if there is a failure to respond to adequate treatment (Neill and Lewis, 2010). Invasive biopsy is unpleasant for women and its assessment relies on traditional tissue fixation and histopathological analysis. Although histopathological examination is considered to be the gold standard diagnostic procedure, the process is time consuming and expensive. In addition, invasive tissue biopsy followed by histopathology is subject to sampling errors as the number of areas of the vulva that can be mapped by this technique is limited by the morbidity of the biopsy procedure itself. In a review of 64 cases of stage I vulval cancer, Ross et al. demonstrated significant discordance between the depth of invasion on biopsy and that found in resection specimens (Ross and Ehrmann, 1987).

A reliable non-invasive diagnostic technique would reduce the need for invasive biopsy and allow a greater area of the vulva to be evaluated. Use of an objective technique for non-invasively assessing the vulva may in turn increase the sensitivity of vulval histopathology by targeting biopsies to the most abnormal areas.

During surgical management of vulval cancer and VIN it is important that all the disease is removed, when technically possible, as complete excision is associated with reduced disease recurrence (chapter 1.4.4). At present clinicians rely on visual assessment alone for the evaluation of disease margins at the time of surgery. A non-invasive tool for assessing the presence or absence of premalignant or malignant disease could aid the complete excision of disease whilst minimising the extent of the surgery necessary.

A non-invasive technique for objectively monitoring biochemical response to treatment may also allow clinicians to tailor management of precancerous vulval conditions to biomolecular disease response and reduce the risk of progression to vulval cancer. An example of where this approach may be useful is in the treatment of vulval lichen sclerosus with ultra potent topical corticosteroids (Chapter 1.2.1). Long term observational studies of women with vulval LS have not identified clinical features associated with malignant progression (Cooper *et al.*, 2004; A Lee, Bradford and Fischer, 2015). As treatment of LS is currently guided by reported symptoms and structural vulval changes, it is feasible that there is a cohort of patients who are undertreated, as these features have not been found to correlate well with a progression to malignant disease. Non-invasive biomolecular assessment of

response to treatment in LS may provide useful additional information when deciding on preventative treatment regimes.

In summary, Raman spectroscopy and multivariate analysis has the potential to reduce the need for invasive tissue biopsy for the assessment of vulval disease and to provide a valuable tool for the intraoperative assessment of disease margins during surgery for vulval cancer. Other potential applications of the technique include biomolecular assessment of response to non-surgical management of precancerous vulval conditions to allow treatment to be tailored to biomolecular rather than clinical response.

5.2 Aim

To investigate the ability of Raman spectroscopy to classify vulval pathology including vulval intraepithelial neoplasia, lichen sclerosus, vulval squamous cell carcinoma.

To achieve this aim the following objectives were set to progress experimentation with the technique from a purely scientific evaluation, under optimised laboratory conditions, towards application of the technique under conditions that more closely approximate in vivo use. This was designed to allow the subsequent development of a probe suitable for in vivo evaluation. The development of an in vivo probe is beyond the scope of this study.

5.3 Objectives

1. Evaluate the ability of Raman spectroscopy to classify vulval disease using Raman streamline mapping of transverse tissue sections at high resolution and explore underlying spectral differences.

    a. To investigate the ability of streamline Raman mapping at high resolution to classify non-neoplastic vulval skin, premalignant and malignant disease of the vulva.

    b. To identify the underlying spectral differences between the tissue types and explore the molecular basis for the classification of disease based on the Raman spectra.

2. Evaluate the ability of Raman spectroscopy applied to the surface of fresh and fresh frozen bulk tissue to classify vulval disease.

    a. Evaluate Raman spectroscopy applied to fresh vulval tissue for the diagnosis of clinically inconclusive vulval lichen sclerosus.

    b. Evaluate Raman spectroscopy applied to fresh frozen tissue for the classification of non-neoplastic vulval skin, premalignant and malignant disease of the vulva.

3. Evaluate the ability of a fibre optic probe and portable Raman spectrometer in classifying non-neoplastic vulval skin, premalignant and malignant disease of the vulva.

The practical assessment of objective 1 is addressed in Chapter 6 which evaluates the differentiation of vulval SCC, VIN and non-neoplastic vulval tissue using streamline Raman mapping at high resolution. Chapter 7 addresses objectives 2a and 2b by evaluating the ability to differentiate clinically inconclusive vulval lichen sclerosus, and separately non-neoplastic vulval skin, premalignant & malignant disease of the vulva using trans-epidermal micro-spectroscopic Raman spectral measurements. Objective 3 is assessed in Chapter 8 which assesses the efficacy of a fibre optic Raman spectroscopic probe for evaluating and differentiating non-neoplastic vulval skin, premalignant & malignant disease of the vulva.

5.4 Ethical approval and consent

Ethical approval for the collection and analysis of fresh and fresh frozen vulval tissue was granted by the National Research Ethics Service Committee South West – Exeter (REC reference - 14/SW/1077). Written informed consent was sought from each participant only after a full explanation had been given, an information sheet offered (Appendix C.1) and time allowed for consideration. Signed participant consent (Appendix C.2) was obtained prior to analysis of tissue samples.

Samples for all arms of this investigation were collected using the same ethical approval and consent procedure, except those samples obtained from the Imperial College Tissue Bank, London and the Exeter NIHR Clinical Research Facility, which were collected using each tissue banks general ethical approval.

5.5 Recruitment and sample processing

Fresh frozen samples obtained for this research study were analysed by micro spectroscopic streamline mapping, micro spectroscopic point measurements and probe measurements to fulfil objectives 1, 2b and 3. Fresh samples collected from participants undergoing diagnostic biopsy were analysed with micro spectroscopic point measurements to fulfil objective 2a. Outlines of both of these tissue-processing pathways are shown in Figure 5-1.

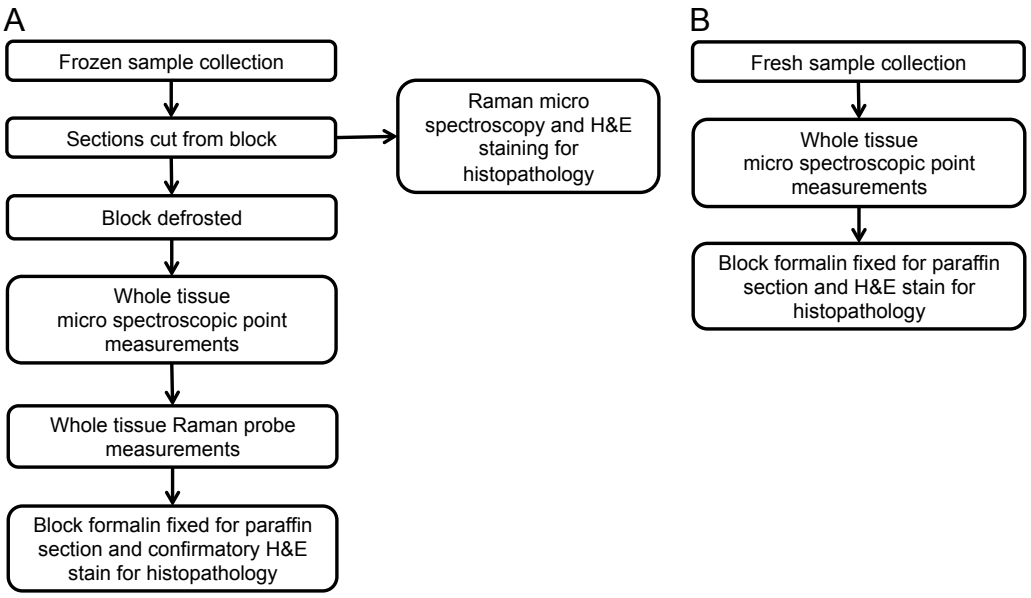


Figure 5-1 – Overview of sample processing for fresh frozen tissue (A) and fresh tissue (B)

Chapter 6 Ex Vivo Evaluation of Vulval Skin Disease Using Streamline Raman Spectroscopic Mapping

6.1 Introduction

In order to establish if Raman spectroscopy has the potential to differentiate non-neoplastic vulval skin, premalignant and malignant disease of the vulva it is desirable to obtain high quality Raman spectra from the area of the tissue that contains the molecular changes associated with the diseases under analysis. The primary cellular changes that occur with progression from normal or non-neoplastic vulval tissue to premalignant or malignant vulval tissue occur in the vulval epidermis (Walker and Mathews, 2012; Isaac and Young, 2013; Nucci, 2013). Analysing the Raman spectral changes isolated from these epidermal cells should demonstrate both the potential for the technique in differentiating the different pathologies and the spectral basis for differentiation.

Raman microspectroscopic mapping of tissue sections allows isolation of cellular layers that contain the pathological changes associated with disease (Lewis and

Edwards, 2001; Hutchings *et al.*, 2009). This is in contrast to point or probe measurements of bulk tissue samples where the Raman signal from the abnormal cellular layers is found in combination with signal from surrounding tissue structures. Using Raman microspectroscopic mapping disease classification can be related to specific cellular layers of the vulval epithelium. This should result in a clearer picture of the biomolecular changes that permit classification of disease from Raman spectra.

To date Raman microspectroscopic mapping based disease classification has been successfully applied to the analysis of multiple skin pathologies including basal cell carcinoma, solar radiation induced damage and malignant melanoma (Ly *et al.*, 2010; Ali *et al.*, 2012; S M Ali *et al.*, 2013; S. M. Ali *et al.*, 2013) There are no reports in the literature of the separation of skin layers for analysis and classification of SCC of the skin or pre malignant vulval disease using this technique.

6.2 Objectives

This investigation assesses objective 1a and 1b of the aims and objectives outlined in Chapter 5.

6.3 Overview of investigation

An outline of this investigation is shown in Figure 6-1 and the experimental methods for each step in the study are explained in more detail below.

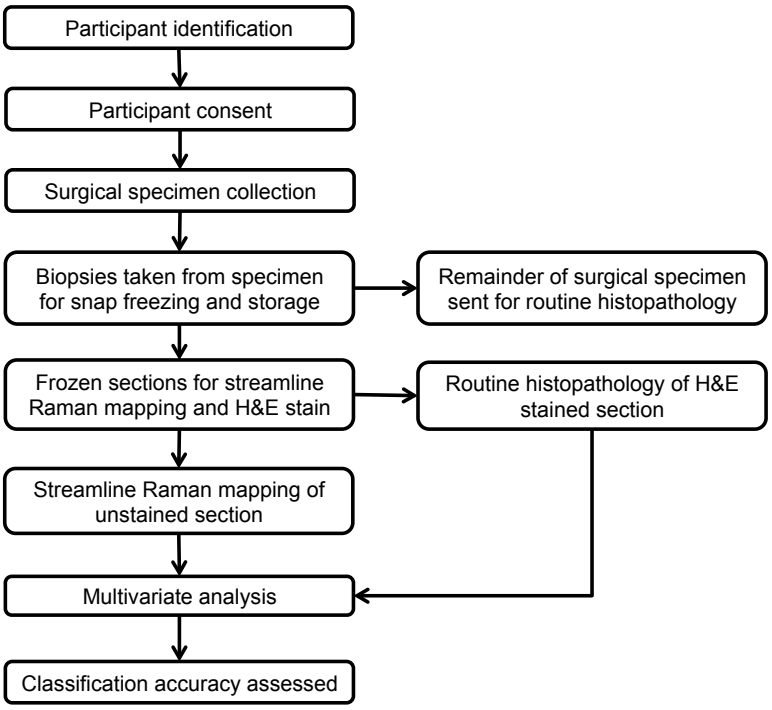


Figure 6-1 – Outline of the collection and analysis of vulval tissue for Streamline Raman mapping and analysis

6.4 Participant identification and sample collection

Tissue samples were collected from women undergoing excisional vulval surgery as part of their routine care in five hospitals across the south west of England (Figure 6-2) between November 2014 and December 2016.

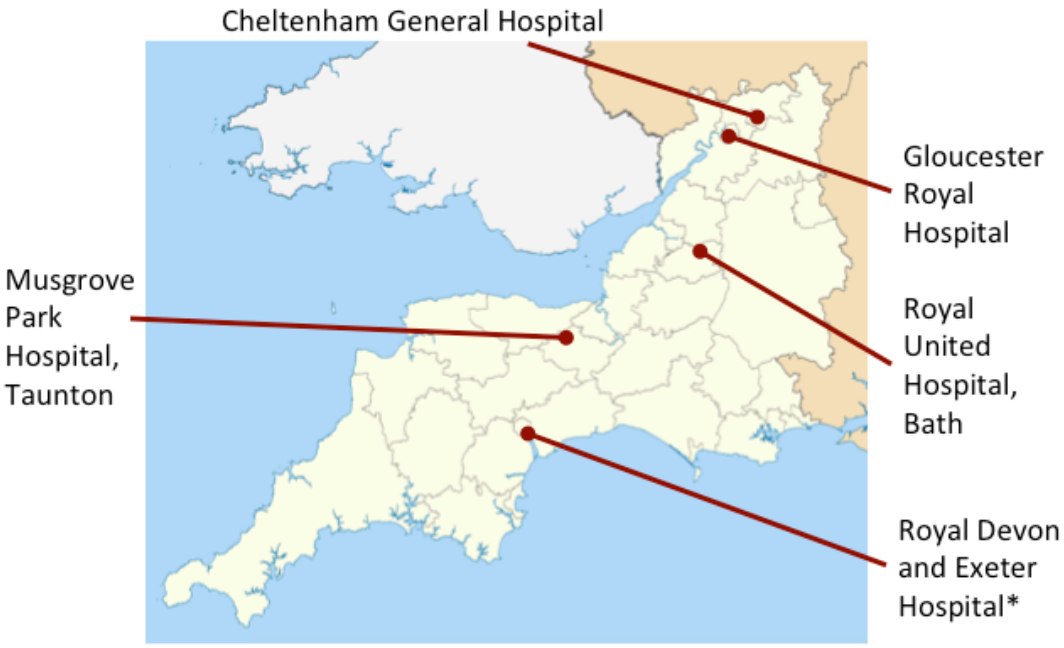


Figure 6-2 – Recruitment sites for the collection of vulval tissue.
*tissue bank collection

Patients with planned vulval excisional surgery were identified by clinicians providing specialist gynaecological care. Those identified gave written informed consent for inclusion into the study. Prior to surgical resection the participant's vulva was cleaned and prepared with 0.9% saline, 10% iodinated povidone solution or 0.05% chlorhexidine acetate solution according to the preference of the hospital and surgeon. Figure 6-3 shows the Raman spectra from each of these solutions placed on a calcium fluoride slide. All these solutions are water-soluble and can be readily washed from the tissue samples with 0.9% saline.

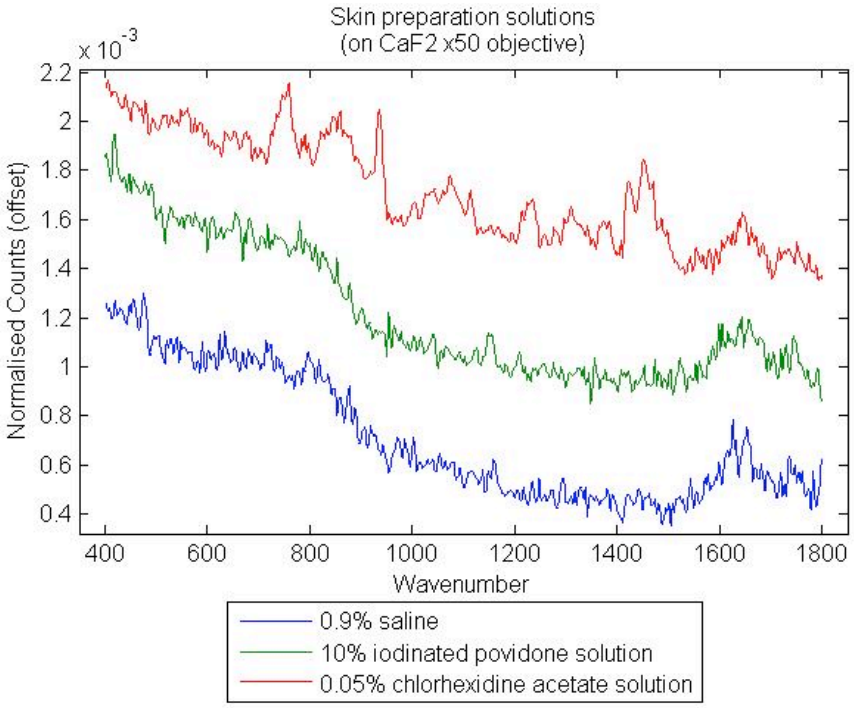


Figure 6-3 – Raman spectra of skin cleansing solutions used in vulval excisions and biopsies (preparations placed within CaFl2 vessel and spectra collected for 90 seconds through x50 HWD objective)

Immediately after the excisional surgery had been performed, research tissue samples were taken from the surgical resection specimen with a 6 or 8 mm Keyes biopsy punch. The size of the biopsies and the number of samples taken from each surgical specimen was dictated by the size of the specimen. It was necessary for adequate tissue to remain on the specimen to allow diagnostic histopathology for the patient's routine care. Key features reported by histopathologists that are used in determining the ongoing care of patients include the involvement of the margins of the excision specimen with tumour and the distance of the closest disease free tissue margin. Tissue samples collected therefore did not include the margin of any specimen and samples were not taken between the primary lesion and the closest

surgical margin to ensure reporting of key findings was not affected. For smaller specimens (≤3 cm max diameter) a single punch was taken from near the edge of the disease lesion. For larger specimens multiple punches were taken, including tissue from the primary lesion, tissue near the lesion and presumed non-neoplastic distant tissue, duplicated as possible (Figure 6-4).

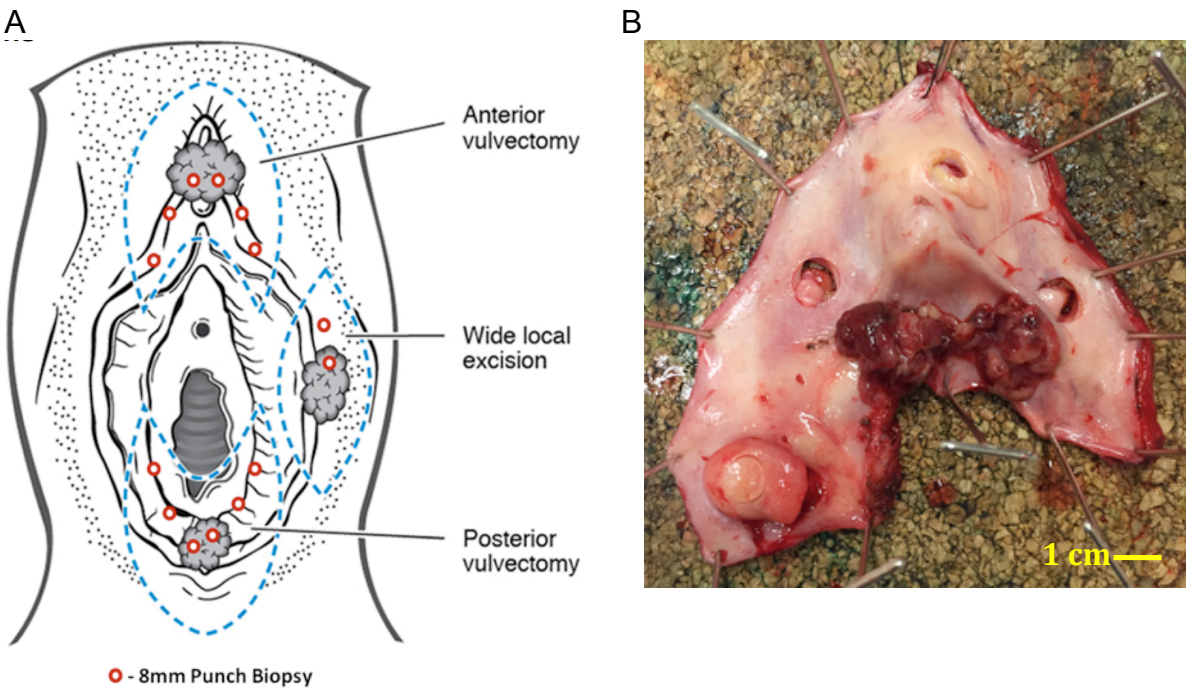


Figure 6-4 – A - Schematic of example tissue collection sites, B - Photograph of anterior vulvectomy specimen showing tissue collection sites

The samples taken were washed in 0.9% physiological saline; bisected; mounted onto cellulose acetate sheets (Sigma PC38378) to retain their correct orientation; placed into a 2ml Cryovial (Thermo Scientific) and snap frozen in liquid nitrogen prior to storage at -80^{O}C.

Tissue was also obtained from two tissue banks, the University of Exeter Clinical Research Facility and the Imperial College Hospitals Tissue Bank. Patients

contributing to these tissue banks were consented by the recruiting teams using the standardised nonspecific tissue bank consent procedure. Tissue collected by the University of Exeter Clinical Research Facility was collected according to the protocol above. Tissue samples collected by the Imperial College Hospitals Tissue Bank did not follow the same procedure. In these cases strips of tissue were cut from surgical resection specimens, wrapped in aluminium foil and snap frozen in liquid nitrogen prior to storage at -80^{O}C. These samples were transferred to the laboratory for processing frozen and packed in carbon dioxide pellets at a temperature of -79^{O}C.

6.4.1 Sample preparation and selection

Sections from the tissue samples collected were cut onto 1mm thick calcium fluoride windows using a standard low temperature cryostat. The cryostat mounting blocks were washed with detergent and rinsed with water prior to use and no cryo-embedding media (e.g. optimal cutting temperature compound) except water was used to mount the samples. The use of optimal cutting temperature compound (OCT) was avoided to minimise the biomolecular differences between the tissue samples under analysis and tissue in vivo. OCT has been shown to contribute significant Raman signal in the fingerprint region and this alters Raman spectrum measured from tissue sections (Shim and Wilson, 1996). The use of OCT was not necessary in order to obtain satisfactory tissue sections. This technique has been demonstrated to preserve the biochemistry of the tissue samples closer to that of live tissue (Smith, 2005).

Frozen tissue sections between 5 and 25 microns thick were evaluated to assess the impact of increasing section thickness on the signal to noise ratio (chapter 4.1.2) (Figure 6-5). The thicker 20 and 25 microns sections produced the highest SNR however these sections were technically difficult to cut and place on calcium fluoride with a high incidence of failure. For this reason section thickness of 15 microns was chosen for the analysis.

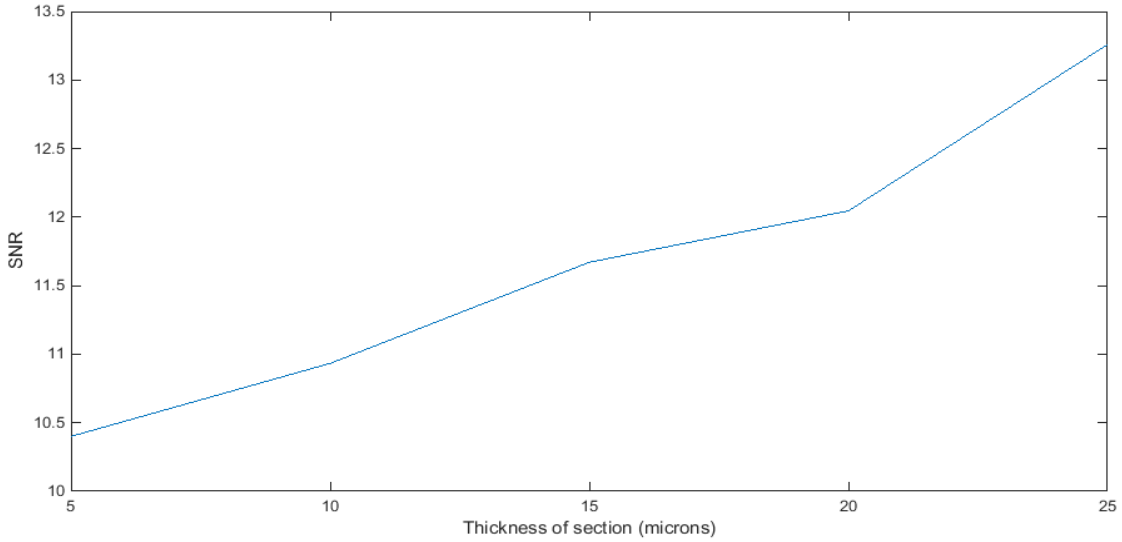


Figure 6-5 – Change in signal to noise ratio with changing frozen section thickness

Standard 4 micron tissue sections were cut on to glass before and after each section cut onto calcium fluoride. These sections underwent H&E staining by a biomedical scientist followed by histopathological analysis by a specialist gynaecological histopathologist to select samples of non-neoplastic vulval skin, VIN and SCC for further spectroscopic analysis. Vulval skin that did not display histological features of intraepithelial neoplasia or malignancy was designated as non-neoplastic.

In total 231 vulval tissue samples from 73 participants were assessed for inclusion in the study. As the intended mechanism for validation of the classification model produced from the spectral data was an internal leave one out cross validation (chapter 4.2.5) duplicates of the same pathology in multiple samples from the same participant were excluded to avoid a type one error. Samples containing malignant pathology not under investigation (e.g. adenocarcinoma) and samples where sections were of insufficient quality to allow histopathological diagnosis were also excluded from analysis. This left 91 samples from 64 participants that were suitable for inclusion, these are summarised in Figure 6-6.

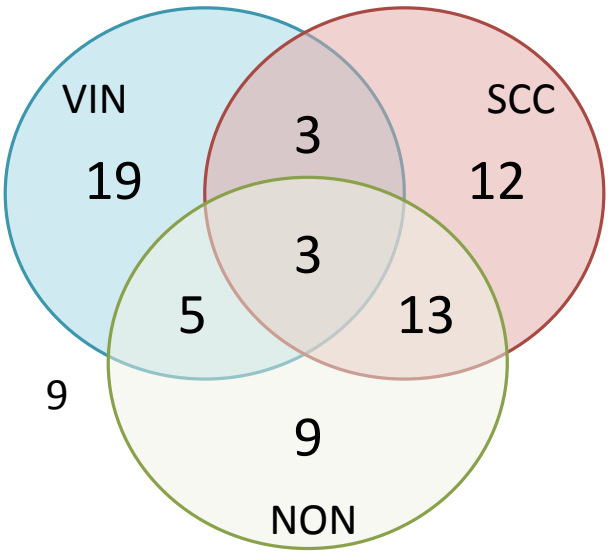


Figure 6-6 – Summary of the participants included in the investigation (VIN – vulval intraepithelial neoplasia, SCC – squamous cell carcinoma & NON – non-neoplastic vulval skin)

6.5 System instrumentation and experimental methods

A modified Renishaw system 1000 Raman microspectrometer adapted to facilitate streamline map acquisition was used for the measurement of the tissue. The

excitation laser was a Renishaw high power near infrared diode laser which supplies light at a wavelength of 830 nm. The laser generates 330 mW of power. The excitation light is focussed through a Leica microscope fitted with a Leica x50 long working distance lens giving 70 mW of power at the sample. The elastically scattered light was rejected with a dielectric coated edge filter and a 300 line/mm grating was used to disperse the light onto the CCD. When operated in streamline imaging mode a laser line is projected onto the sample and the motorised stage is used to step the sample in the direction of the laser line. The movement of the motorised stage is synchronised with the readout on the CCD and when the entire line has passed over a single point on the sample a Raman spectrum is recorded. The recorded spectrum comprises an accumulation of the entire signal recorded by the CCD as the laser line passes over the sample. The streamline technique allows high spatial resolution maps to be acquired faster than with conventional point mapping without a reduction in spectral intensity.

6.5.1 System calibration

The streamline Renishaw System 1000 dispersion micro spectrometer required calibration prior to each use. Calibration allowed accurate recording of the Raman shift allowing the comparison of spectra collected at different times. Without this calibration differences in the spectra may be present and these could be misinterpreted as clinically significant rather than as a result of changes in the instrument.

The calibration involved ensuring the grating of the detector is aligned so that the wavenumber shift recorded is accurate. This was achieved by measuring the spectrum of a silicon standard that has a distinct peak at 520.4 cm^{-1} and, if required, adjusting the offset of the spectrometer to align the peak correctly. The second calibration undertaken was an instrument response correction. For this correction, the Raman and fluorescence spectrum from a fragment of green glass (Figure 6-7) was measured. Illumination of the green glass with the 830 nm laser produced a broad fluorescence across the bio relevant wavenumber range (400 – 1800 cm^{-1}). The true or standard spectrum of the green glass sample is calculated from calibration against a tungsten filament lamp, which in turn is calibrated against a calibrated spectrometer at the National Physics Laboratory, Teddington, UK. The instrument response correction adjusts for the varying sensitivity of the spectrometer across the wavenumber range by applying a correction based on a ratio of the true or standard green glass spectrum and the measured green glass spectrum (Figure 6-8). The green glass is therefore a tertiary standard well suited to a regular instrument response correction, as a tungsten filament lamp is not required. In addition, the spectrometer laser beam can easily be focused on the glass and fluorescence recorded. This allows accurate instrument response correction under the same conditions as for the collection of Raman spectra from tissue samples.

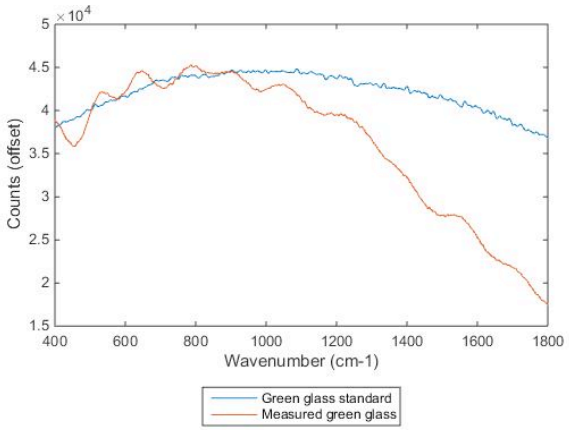


Figure 6-7 – Green glass standard spectrum and example of green glass calibration measurement

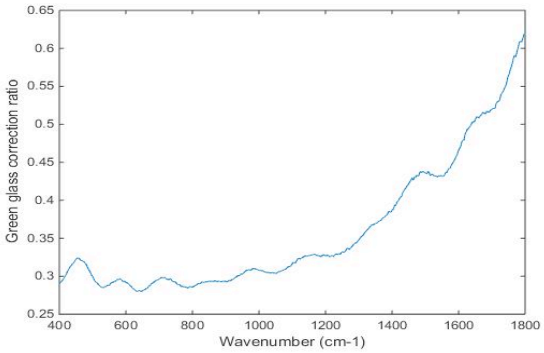


Figure 6-8 – Green glass correction applied to spectra to compensate for variation in spectrometer sensitivity across the wavenumber range

Further measurements of cyclohexane, a neon argon lamp and a polymer pipette tip were taken to cross check the calibration of the spectrometer. Characteristic wavenumber peaks are present in these standards across the bio relevant spectral range allowing identification of distortions in wavenumber spacing in the resultant spectra. The XYZ electronic stage was calibrated against a standard Vernier stage (Melles Griot, USA) to confirm accurate movement of samples under the microscope objective.

6.5.2 Acquisition parameters

The parameters of the system were optimised for the measurement of vulval tissue spectra. The smallest available streamline step size of 1.1 microns was chosen as this facilitated accurate correlation between the streamline Raman spectral maps and the H&E stained sections. The effect of changing the streamline composite acquisition time was assessed using a 15 µm section of histologically normal vulval

tissue this showed a linear relationship between increasing acquisition time and signal to noise ratio (Figure 6-9).

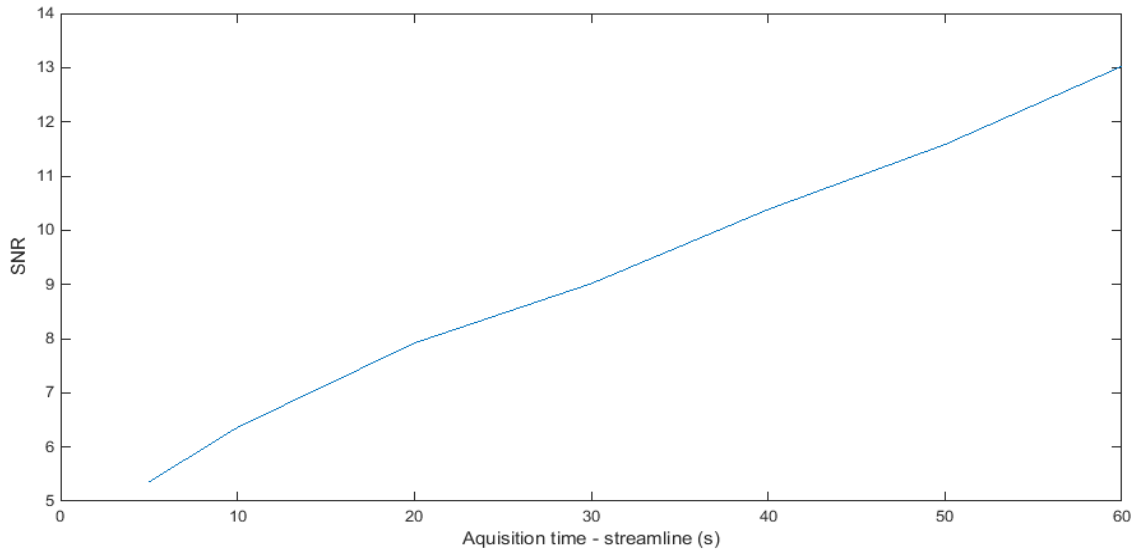


Figure 6-9 – Change in signal to noise ratio with changing spectral acquisition time

An acquisition time of 60 seconds was chosen to allow spectral mapping of regions of interest on tissue sections within a practical timeframe whilst delivering good SNR.

In the streamline mapping arrangement the spectral acquisition time equivalent to the acquisition time of static measurement (t_{eq}) is a function of the time it takes to translate the laser line across each pixel. The time taken (t) for the system to translate the laser line across a pixel of interest and the step size (Δx) is set when configuring the instrument to acquire a map. Using an x50 objective the laser line has a length of approximately 50 μm. Therefore for the x50 objective t_{eq} is given by Equation 6-1. Therefore the selected streamline acquisition time of 60 seconds is equivalent to a standard acquisition time of 1.32 seconds.

$$t_{eq} = t \times \frac{\Delta x}{50}$$

Equation 6-1 – Equation for calculating equivalent time for static acquisitions during streamline acquisitions

The size of the spectral map acquired varied according to the individual sample as the shape and distribution of the epidermal cells varied considerably. When setting the acquisition area the objective was to acquire spectra from an area of the epidermis in which the pathology could be identified on the corresponding H&E stained section.

6.5.3 Spectral preprocessing

When the spectrometer is centered on 1150 cm^{-1} the range of the spectrometer is from 456 cm^{-1} to 1790 cm^{-1}. This encompasses the majority of the biomolecular fingerprint region of the spectrum and the analysis used all available spectral data (Chapter 2.2.2). No adjustment was made to reduce the impact of cosmic rays during data acquisition however the spectral maps were median filtered after collection using a three by three filtering grid to reduce the impact of the cosmic rays (Chapter 4.1.3.1). All spectra underwent an instrument response correction using the spectrum measured from the green glass as previously described (chapter 6.5.1) (Grimbergen *et al.*, 2010). The spectra then underwent an extended multiplicative scatter correction (EMSC) to normalize the data to the mean of all the spectral data and correct the baseline created by auto fluorescence using a third order polynomial fit (Chapter 4.1.3.2) (Lieber and Mahadevan-Jansen, 2003).

6.5.4 Constituent analysis

The ability of Raman spectroscopy to identify the biochemical constituents of the tissue sections was assessed by comparison to standard commercially available compounds. The following constituents were selected as they represent the different areas within cells and the extracellular environment of vulval skin (Table 6-1). DNA is representative of the cell nucleus and the contribution of DNA is clearly seen in Raman spectra (Talari *et al.*, 2015). Other constituents present in the nucleus such as RNA and ribosomes have been demonstrated to only contribute minimally to the Raman scattering from the nucleus (Shafer-Peltier *et al.*, 2002). Cholesterol is required to form membranes within cells and is a key component of the extracellular lamellae in the stratum corneum. Cholesterol is also associated with tissue necrosis. Glycogen is present within metabolically active cells and has also be found to be diminished in malignant tissue (Hata *et al.*, 2000; Shafer-Peltier *et al.*, 2002). Collagen is an abundant structural protein in the extra cellular matrix of the dermis and basement membrane. Type 1 collagen accounts for 85-90% of all dermal collagen and it is distributed in a fibrillar pattern within the dermis providing tensile strength to the skin (Epstein and Munderloh, 1978; Nguyen *et al.*, 2013; Patterson, 2016).

The constituents were measured on the same modified Renishaw Raman system 1000 as the tissue maps and the same x50 microscope objective used to focus the laser light on the surface of the sample. The monochromator slit was set to 50 μm and the laser was set to full power for all the samples. The acquisition time was

optimised for each constituent in order to produce spectra with a high signal to noise ratio. The individual spectra were time normalised to match the intensity of the tissue spectra.

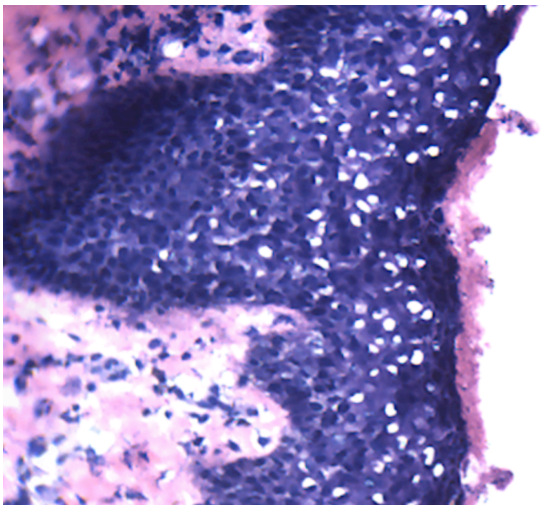
Constituent	Raman shift (cm^{-1}) (Talari *et al.*, 2015)	Commercial compound	Key physiology
Collagen type I	1447-1450	Sigma C774-5MG	85-90% of dermal collagen
Glycogen	1048, 1083, 1256 & 1333	Sigma G0885-5G	Present within metabolically active cells
Deoxyribonucleic acid	722,752,782 & 1663	Sigma D4522-5MG	Representative of the cellular nucleus
Cholesterol	699,882	Sigma C8667-1G	Required to form membranes in cells & extracellular lamellae in the stratum corneum

Table 6-1 – Summary of major bio molecular constituents of vulval skin

The estimated relative concentrations of the measured constituents were plotted onto spectral streamline maps measured from vulval tissue sections. The relative concentrations of the constituent were estimated by the following technique. A green glass instrument response correction was applied to the constituent and tissue spectra. Following this the baseline of the tissue spectra were corrected by the subtraction of a second order polynomial and all data was vector normalised. The relative concentration of the measured constituents within the tissue was then estimated using a non-negative least squares technique. Figure 6-10 shows the non-negative least squares coefficient plot for a section of normal vulval skin. This indicates that as expected the spectrum from the stratum germinosum and stratum spinosum demonstrates similarities with the spectrum from DNA, glycogen and cholesterol. The whole of the highly cellular epidermis shows similarities with cholesterol, consistent with the expected biochemistry. The spectra from the dermis, which is predominantly collagen, shows significant similarities with the spectrum of

type 1 collagen, the main type of collagen with the dermis. Also shown in Figure 6-10 is a plot of the non-negative least squares coefficient of the spectrum of a calcium fluoride slide showing as expected high correlation with the non-tissue bearing area of the slide. This analysis demonstrates that Raman spectra can be used to delineate the biomolecular differences within vulval skin and shows promise for the detection of the biomolecular changes associated with carcinogenesis such as increases in DNA concentration and decreases in glycogen and collagen concentration (Hata *et al.*, 2000; Shafer-Peltier *et al.*, 2002).

A



B

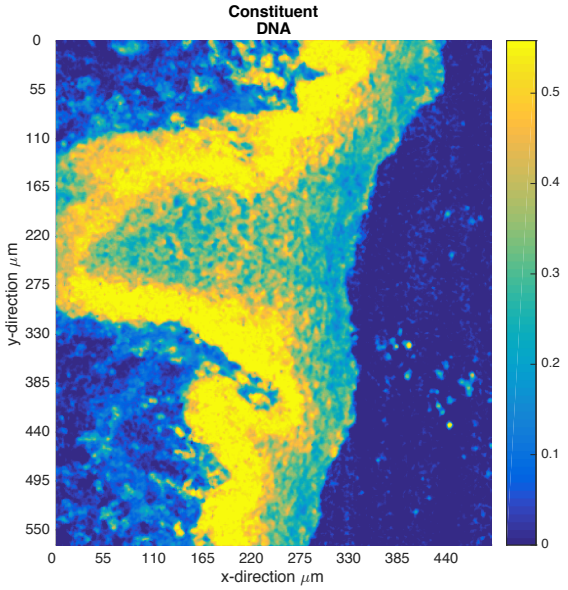


Figure caption on next page

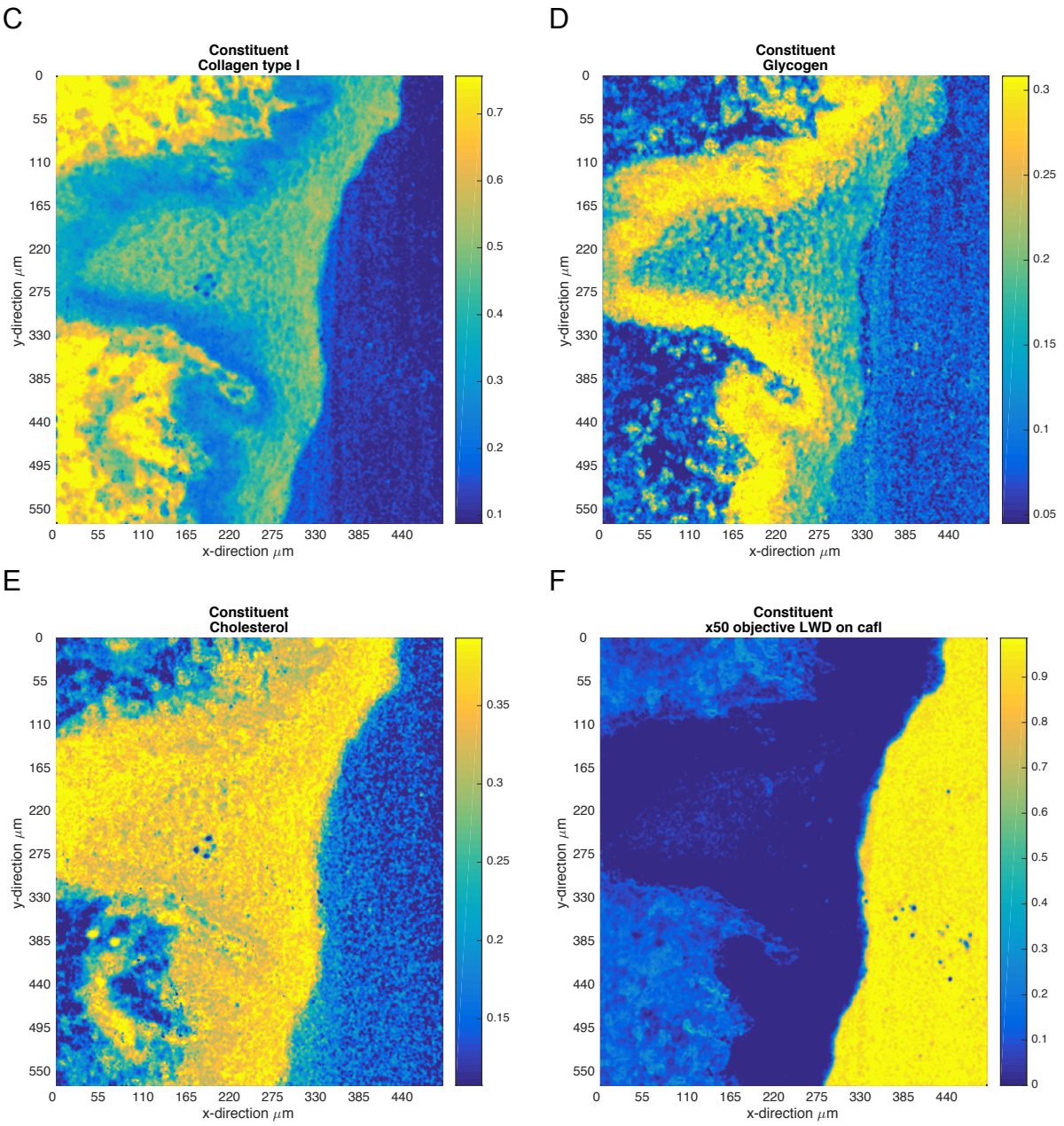


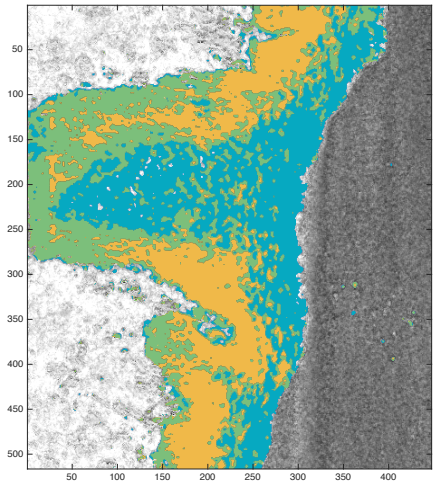
Figure 6-10 – Non-negative least squares coefficient plot of tissue spectra compared to the constituents of the tissue spectra. A – H&E stained tissue section; B - DNA spectrum; C – Type 1 Collagen spectrum; D – Glycogen spectrum; E – Cholesterol and F – Calcium fluoride slide spectrum.

6.5.5 Classification analysis

For the purposes of this study epidermal cells were selected as the primary site for investigation as these cells are the most likely to contain measurable biomolecular

changes that relate to the pathological processes under investigation. To select the spectra to be used for the classification analysis the data from each individual map were separated using k means clustering with eight centroid clustering with seeded spectra from manually selected epidermal, dermal and stratum corneum spectra and five other tissue spectra randomly selected from the spectral map. This enabled manual selection of the spectra measured from the epidermal cells by reference to the H&E stained section and white light image of the unstained tissue section (Figure 6-11). The spectra from the resultant k means relating to the epidermal cells were then saved for analysis.

A



B

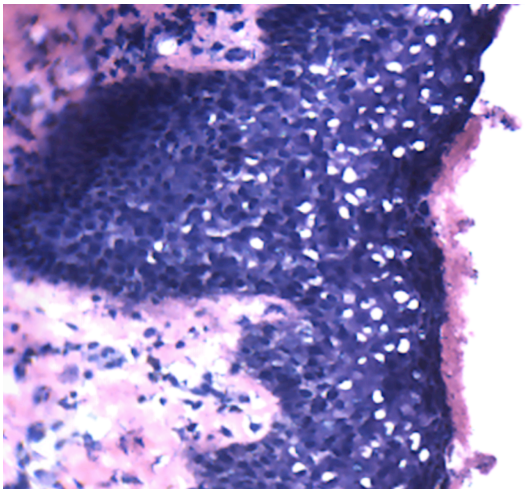


Figure 6-11 – Example of area of epidermal cells selected by k means for further analysis. A – The three k means corresponding to the the epidermal cells in green, yellow and blue; B – the corresponding H&E stained section.

Significant differences between those samples collected at remote sites was explored by applying ANOVA to the PC scores of the spectral data to assess if there were any significant spectral differences between groups. The F ratios for the individual PC scores did not reach significance for any of the external tissue sources

when compared to samples collected in the base hospitals of Cheltenham General Hospital and Gloucester Royal Hospital (Table 6-2). This suggested there is no significant difference in the spectra acquired from samples obtained from external sources compared with those collected locally.

Table 6-2 – ANOVA F ratios for scores of first 24 principal components for external sample sources when compared to the whole data set, F critical (95% Confidence) is 3.9469.

Principal component	Imperial College Hospitals Tissue Bank vs. all other spectral data	Musgrove park Hospital vs. all other spectral data	Royal United Hospital Bath vs. all other spectral data	Royal Devon and Exeter Hospital vs. all other spectral data
1	2.4535	0.0083	0.1903	0.1997
2	0.1243	0.3293	0.0639	0.5753
3	0.2642	0.0829	0.2932	0.0021
4	1.3797	0.0573	0.0305	0.1414
5	0.2067	0.0344	0.064	8.78E-04
6	0.0237	0.0572	0.6639	0.6147
7	0.0013	0.0092	0.0697	0.1126
8	2.04E-05	0.167	0.0255	0.0964
9	0.6778	0.0329	0.0963	0.0235
10	0.0213	8.90E-04	0.0063	0.0357
11	0.4312	0.0301	1.3483	0.3611
12	0.0232	0.0335	0.1424	0.1111
13	0.0126	0.0228	0.2391	0.0676
14	0.0019	0.2051	0.0597	0.7858
15	0.0536	0.0203	0.1535	0.3908
16	0.407	0.0557	0.3665	0.0152
17	0.0834	0.0039	0.0199	0.8154
18	0.2201	0.0056	0.0075	7.03E-04
19	0.5466	0.0528	0.8016	0.122
20	0.3656	0.0033	5.18E-04	1.3599
21	0.0131	0.0394	0.0643	0.0084
22	0.0012	0.0438	0.0026	0.9913
23	0.6755	0.001	0.0446	0.5609
24	0.0151	0.1951	0.0191	0.7171

The regions of interest as determined by k means clustering were then spatially binned to improve the signal to noise ratio of the spectra used for the analysis. The effect of the binning on the classification accuracy of the technique was assessed

using simple PCA LDA classification analysis and the increased binning was found to increase the performance up to bins of eleven beyond which there was only marginal benefit to increasing the binning further with the loss of spatial resolution (Figure 6-12).

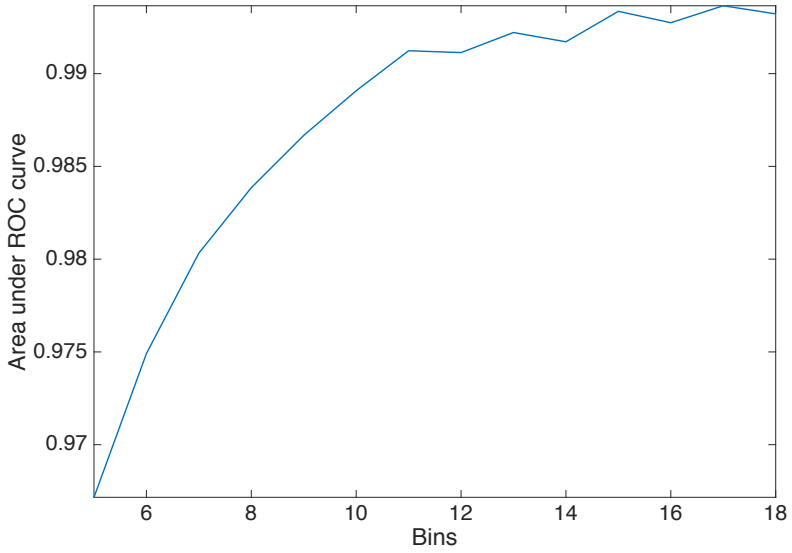


Figure 6-12 – Area under the curve for identifying SCC from VIN using PCA fed LDA with increasing bin size.

The tissue spectra were then mean centred and filtered for outlying spectra using the D and Q classifiers (Chapter 4.2.6.1). Analysis of the effect of filter in the data with the D and Q statistic shows a modest improvement in the diagnostic performance with exclusions (Figure 6-13 and Figure 6-14). In both the D and Q exclusions, the most significant improvement in diagnostic performance appears to occur up to a confidence threshold of 0.999 and for this reason, this threshold was used in the final analyses.

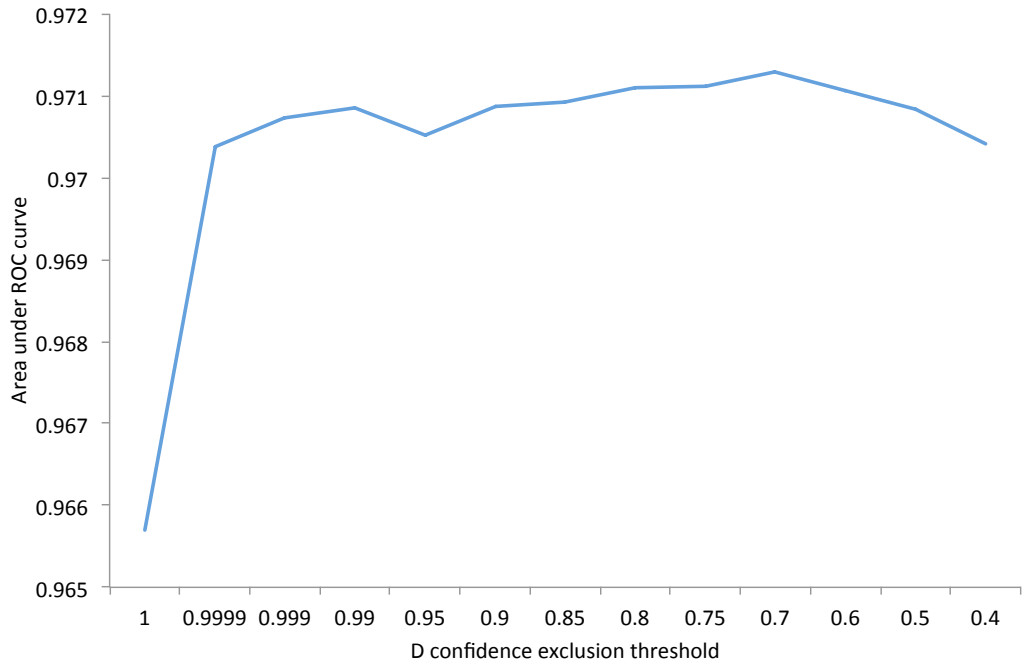


Figure 6-13 – Variation in the area under the ROC curve for identifying SCC from VIN using non-cross validated PCA fed LDA with increase exclusions using D statistic

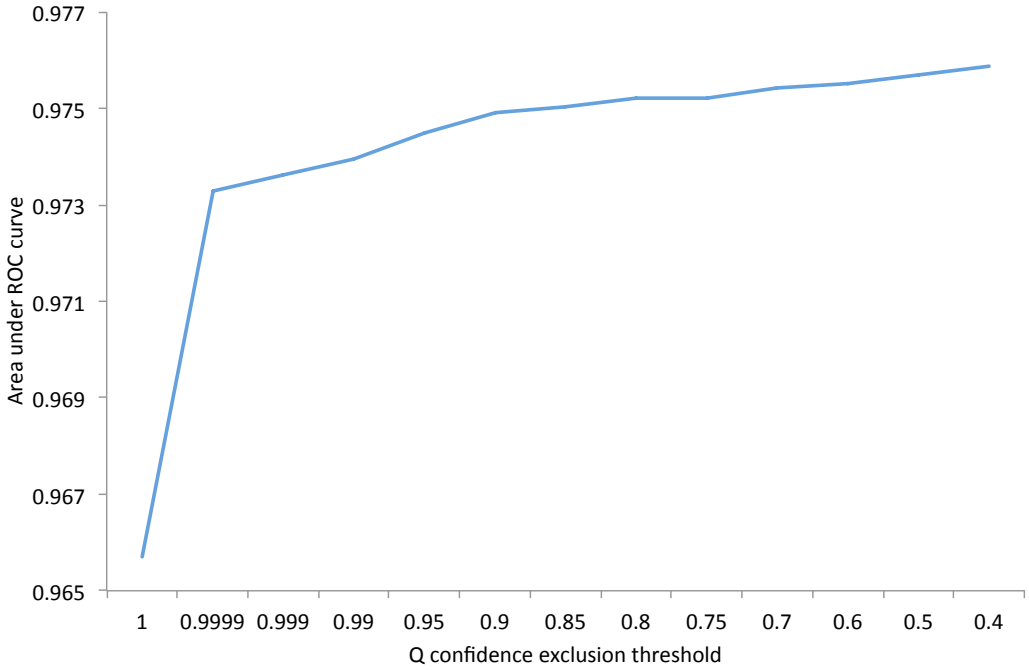


Figure 6-14 – Variation in the area under the ROC curve for identifying SCC from VIN using non-crossvalidated PCA fed LDA with increase exclusions using Q statistic

Classification performance was initially assessed using PCA LDA in a leave one sample out cross validation loop (Chapter 4.2.5). The data analysis performed is summarised in Figure 6-15 with details of the cross validation used given in Figure 6-16. Application of a final filter for spectra that do not fit well within the LDA classification model using the Mahalanobis distance with a confidence threshold of 0.95 also results in a modest increase in diagnostic performance. The area under the AUC curve for identifying non-neoplastic tissue from SCC and VIN using non-cross validated PCA fed LDA increased by 0.02 when excluding spectra with a confidence of 0.95. Final predictions were therefore filtered by Mahalanobis distance to the mean of each of the LDA groups (4.2.6.2). The classification performance of the PCA-LDA model was assessed on its ability to correctly classify the histological diagnosis of the whole sample test set and individual spectra.

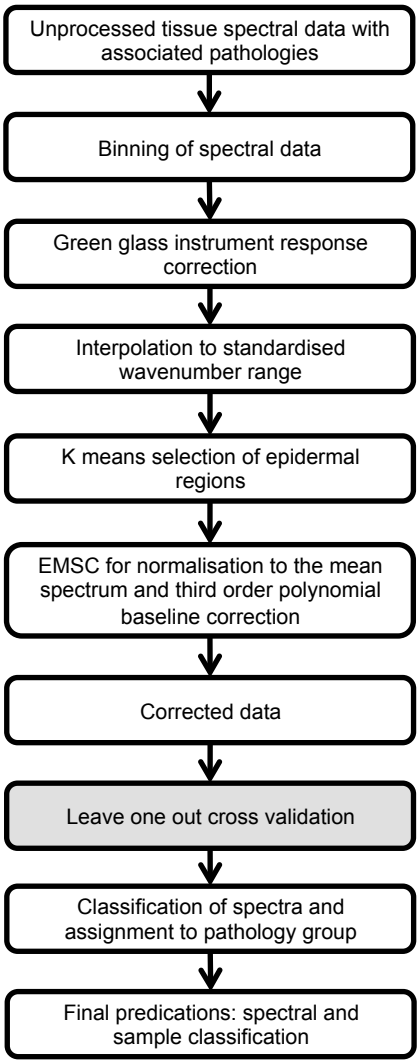


Figure 6-15 – Flow chart of data analysis of streamline Raman mapping measurements see Figure 6-16 for details of cross validation loop

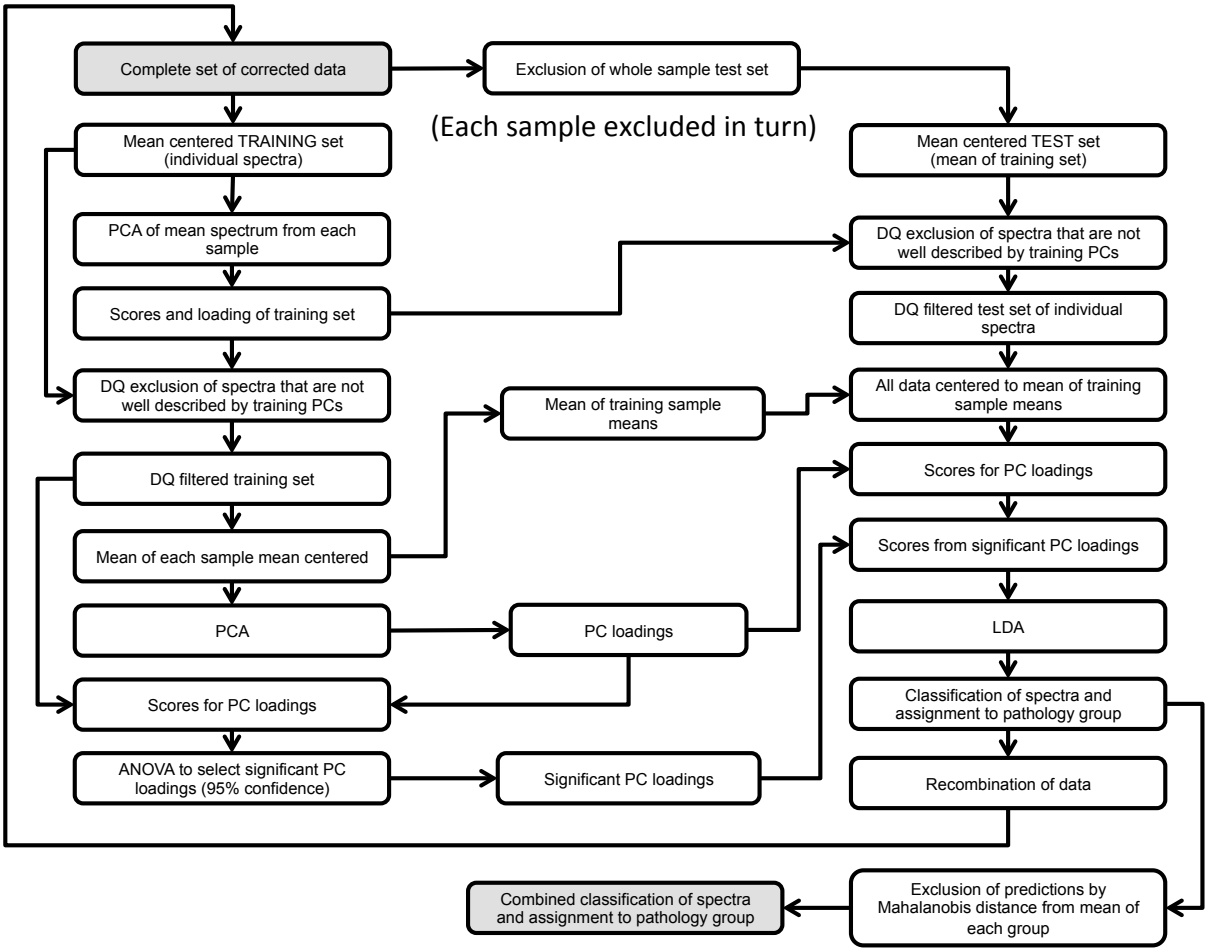


Figure 6-16 – Leave one whole sample out cross validation loop used for analysis of streamline spectral data

6.6 Results

A total of 91 different samples measured from 64 different participants underwent spectroscopic mapping and a total of 8,409,500 individual epidermal spectra were measured (Table 6-3). No participant contributed more than one sample of the same pathology, although some participants contributed to more than one pathology group (Figure 6-6).

Table 6-3 – Summary of mapping measurements recorded

Pathology group	Number of samples / maps measured	Total number of epidermal spectra measured
Non-neoplastic	30	2,190,221
Vulval intraepithelial neoplasia	30	2,327,556
Squamous cell carcinoma	31	3,891,723

To investigate the basis of the classification in cross-validation an initial PCA LDA analysis of the whole spectral data set was undertaken. This revealed that 90% of the variance in the data was accounted for in the first 12 PCs with 85% of the variance in the first 6 PCs (Figure 6-18). When ANOVA was applied to the PC scores and the known pathology groups of the whole data set, the significance of each PC for the classification of the disease state was assessed. The ANOVA F ratios for each PC are shown in Figure 6-19. Much of the variance in the data set is due to noise, variation within individual spectral maps or variation between spectral maps of the same pathology rather than due to pathological differences. As a result the most significant PCs as selected by ANOVA do not correspond to those that describe the most variance in the data.

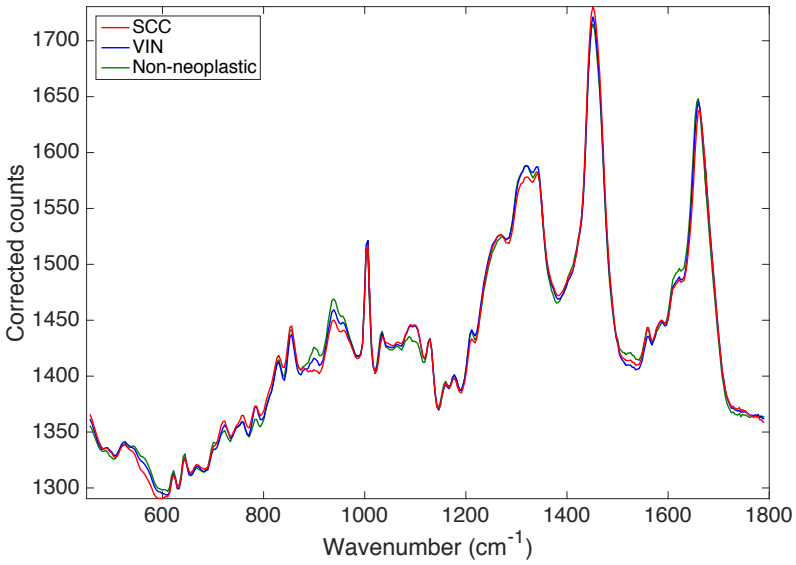


Figure 6-17 – Mean spectra of the pathology groups. Non-neoplastic vulval skin; VIN – vulval intraepithelial neoplasia and SCC – squamous cell carcinoma.

LDA of the significant PC scores demonstrated that the pathology groups could be separated by LD scores as shown in the histogram of the LD scores by pathology group (Figure 6-20).

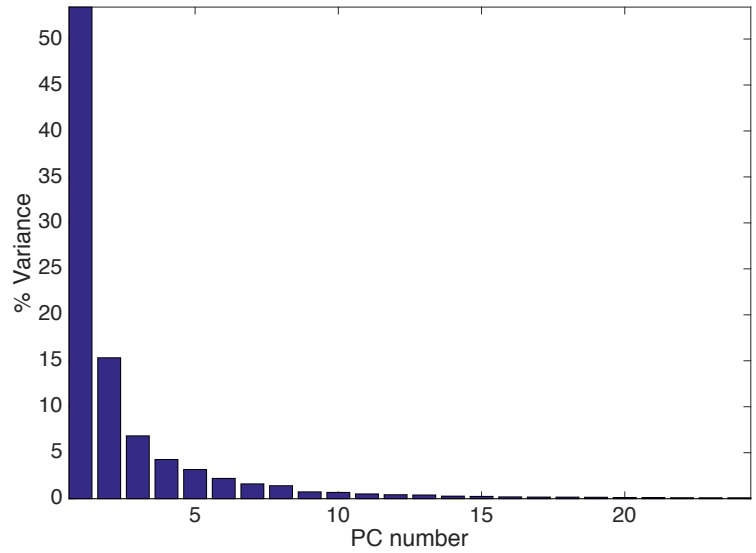


Figure 6-18 – Variance in data described by each principal component.

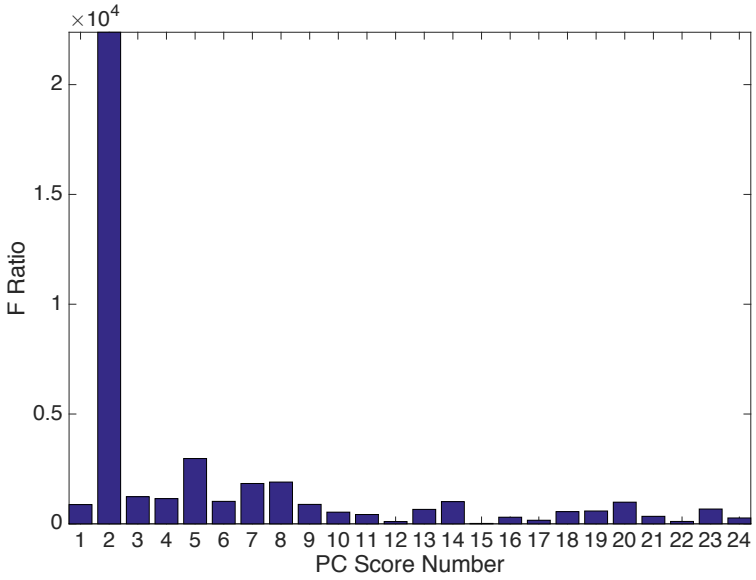


Figure 6-19 – ANOVA F ratios for principal component scores

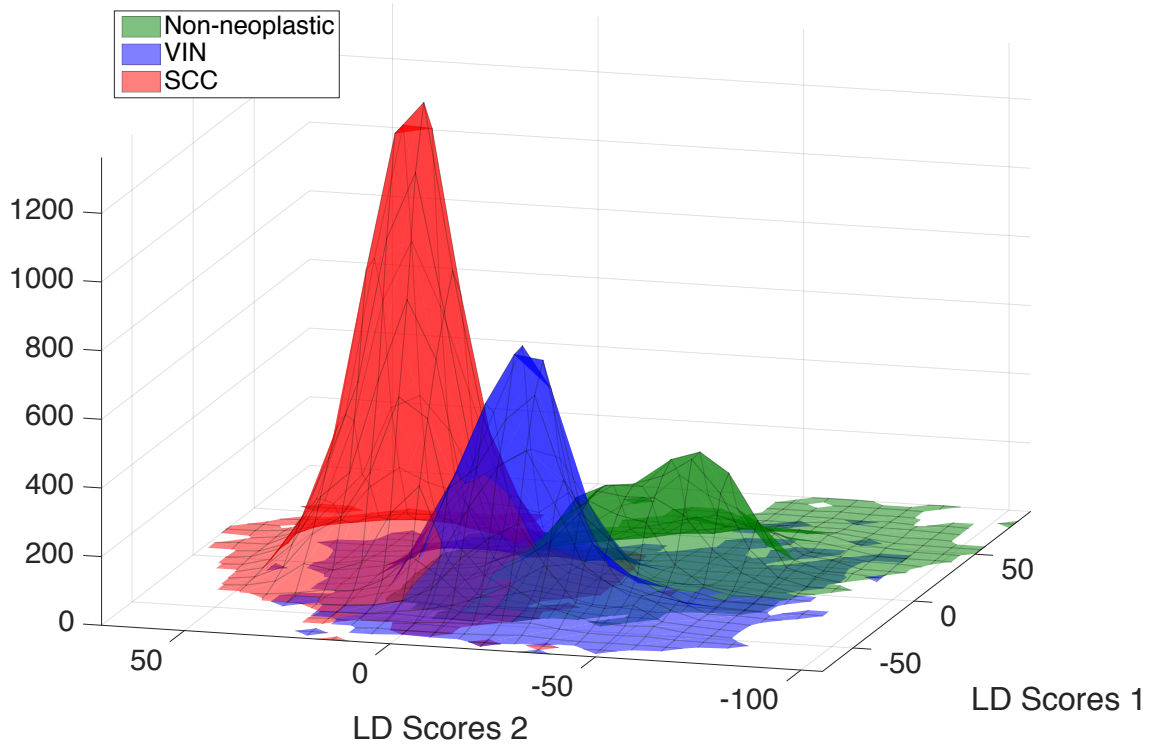


Figure 6-20 – Histogram of linear discriminant scores (VIN – vulval intraepithelial neoplasia, SCC – squamous cell carcinoma & non-neoplastic vulval skin)

The LD scores plot showed some overlap between the groups with more significant overlap between the non-neoplastic group and the VIN group. Of note the non-neoplastic group also appears to have bimodal appearance suggestive of the possibility of two subgroups. This is unsurprising given this group will encompass a range of benign vulval conditions such as lichen sclerosus.

The weightings given to each PC loading by the two linear discriminant functions are illustrated in Figure 6-23. There are several principal components with scores that show relatively small difference between the pathology groups using ANOVA but that are highly weighted in the linear discriminant model (e.g. PC score 20, Figure 6-24).

An explanation for this is that ANOVA is performing a univariate analysis of orthogonal principal components to assess which component scores contain the largest difference between the pathology groups. Therefore the loadings of the most significant ANOVA selected principal components are representative of differences in the spectra between the pathology groups. In addition as the principal components are orthogonal each loading represents a separate set of spectral differences between the pathology groups. The multivariate LDA does not treat the principal components as orthogonal and uses a combination of components to achieve separation of the pathology groups. In this process components are combined to achieve the maximum between group variance whilst minimising the intragroup variance. This combination of components means that assessing the loadings of individual components to understand the spectral differences between the groups is best achieved by examining the most significant loadings by LDA weighting rather than the most significant by ANOVA. To assess the overall spectral differences important in the LDA group classification we can consider the composite of all the PC loadings weighted by contribution to the LDA classification. In a three group model this results in two composite loadings to be interpreted alongside the distribution of the LD scores of the pathology groups, the mean spectra and the principal component loadings of the most significant scores (Figure 6-21, Figure 6-22).

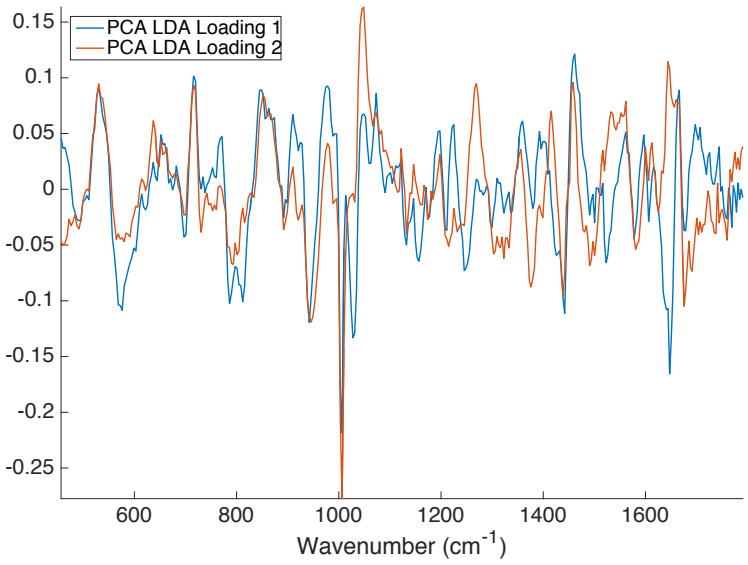


Figure 6-21 – Composite PCA LDA loading

The leave one out cross-validated PCA LDA classification model with DQ filtering of outlying spectra (Figure 6-15) demonstrated good classification of the three pathology groups. Table 6-4 shows the confusion matrix for the classification of the samples analysed with the linear discriminant boundary set midway between the pathology groups. Due to the larger number of spectra in the SCC and VIN groups the confusion matrix is likely to show unfavourable performance in the detection of SCC and to a lesser extent the VIN group. Importantly no samples of non-neoplastic tissue were misclassified as samples of SCC and no samples of SCC were misclassified as non-neoplastic tissue.

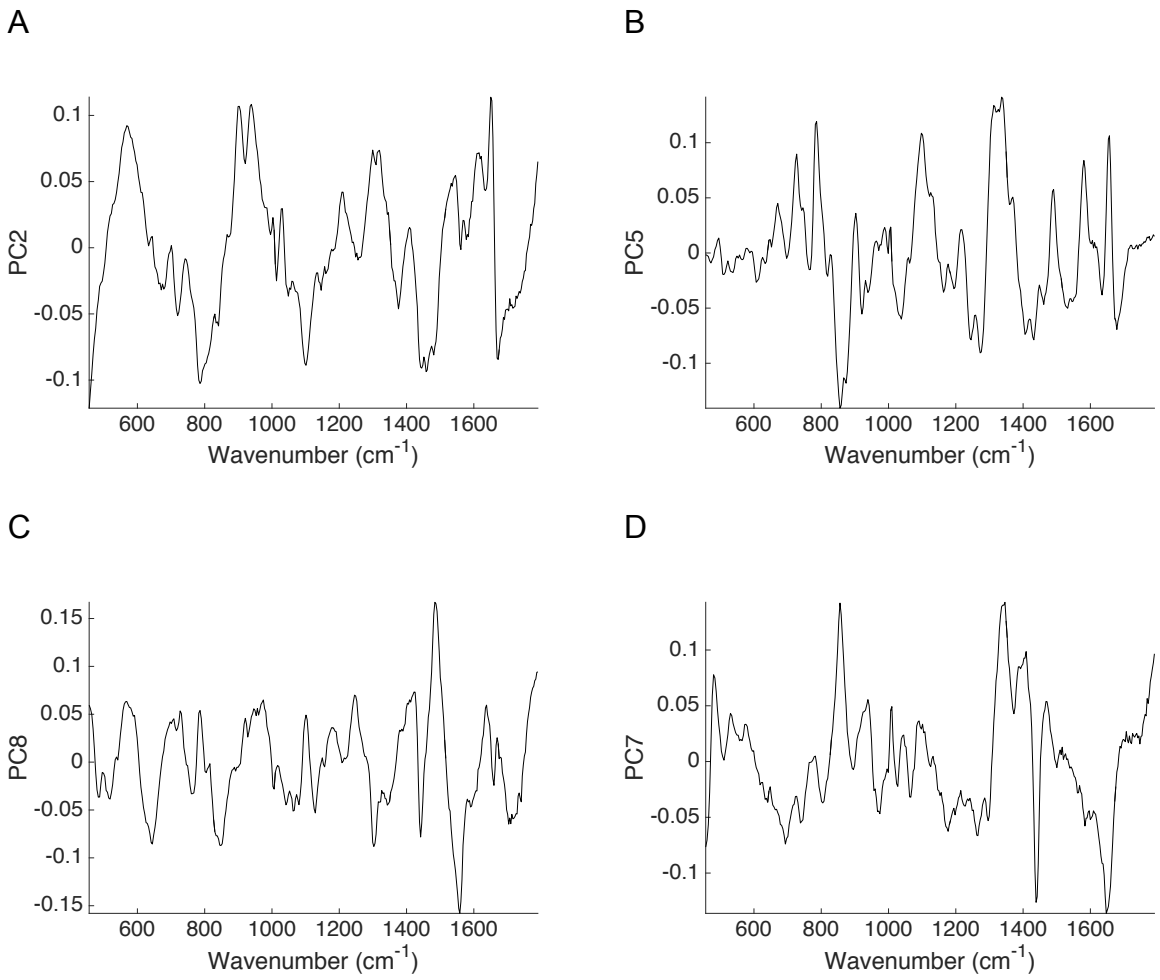


Figure 6-22 – Principal component loadings of four most significant scores as determined by ANOVA.

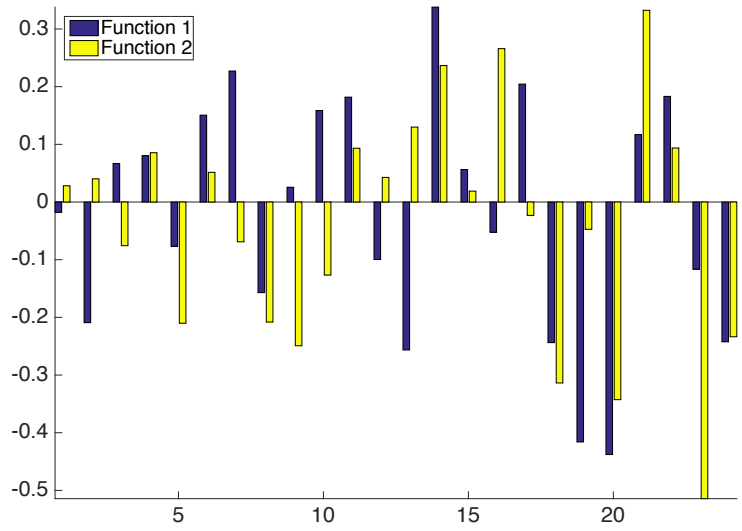


Figure 6-23 – Weight of the contribution of each principal component score to the linear discriminant model. In a three group model two linear discriminant functions are required to separate the groups and these are represented by function1 and function 2.

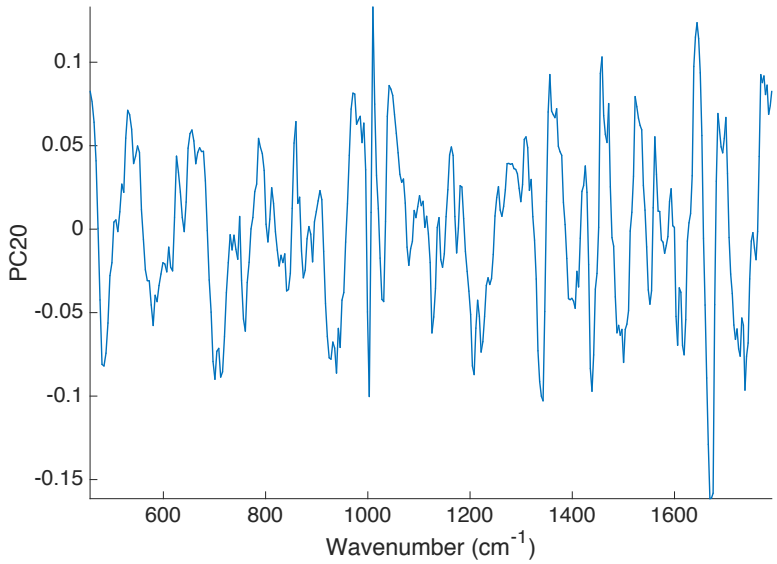


Figure 6-24 – PC loading 20

Table 6-4 – Confusion matrix for whole samples based on modal classification of spectra with linear discriminant boundary midway between the pathology groups, rows are true pathology group and columns are predictions.

Pathology group	Non-neoplastic	Vulval intraepithelial neoplasia	Squamous cell carcinoma
Non-neoplastic	23	7	0
Vulval intraepithelial neoplasia	6	20	4
Squamous cell carcinoma	0	2	29

The overall probabilities of group membership were calculated from the combined individual discriminant distances of the spectra in each group and plotted on a ternary plot to allow visualisation of the distribution of correctly and incorrectly classified samples (Figure 6-25). The ternary plot demonstrates that it is possible to define alternative discriminatory boundaries to tailor the classification to the clinical need. Table 6-5 demonstrates the alternative classification performance that can be achieved through alteration on the discriminatory boundary (Figure 6-25). With the modified boundary SSC is detected with a sensitivity of 100% and specificity of 93% with a positive predictive value of 0.89 and false positive rate of 0.07.

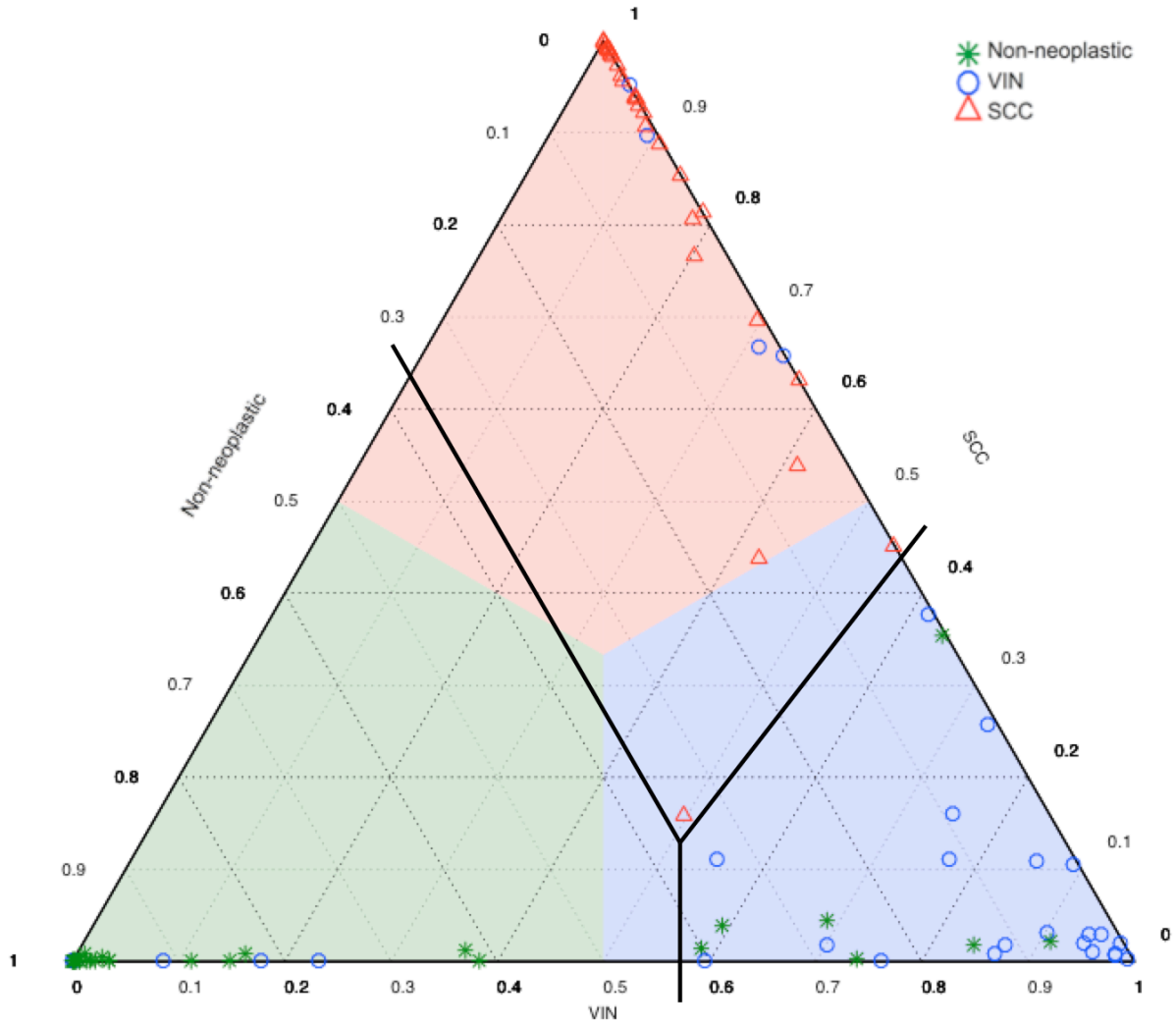


Figure 6-25 – Ternary plot of the predicted probabilities of a sample belonging to the three pathology groups under analysis. Non-neoplastic vulval skin, VIN - vulval intraepithelial neoplasia , SCC - squamous cell carcinoma. Legend indicates actual pathology of sample shaded areas show classification of samples based on discriminant boundary midway between groups. Solid lines indicates alternative discriminatory boundary that could be used for the purposes of classification.

Table 6-5 – Confusion matrix for whole samples based on combined probability of spectra with modified discriminatory boundary between the pathology groups, rows are true pathology group and columns are predictions.

Pathology group	Non-neoplastic	Vulval intraepithelial neoplasia	Squamous cell carcinoma
Non-neoplastic	23	7	0
Vulval intraepithelial neoplasia	6	20	4
Squamous cell carcinoma	0	0	31

To further assess the classification performance of the technique, receiver operator curves were plotted whilst varying the linear discriminant boundary. Comparisons were undertaken of the most clinically useful classifications as outlined in chapter 3.4.1. The first comparison looked at the ability of the technique to determine disease margins at the time of surgery for excision of VIN or SCC. The performance for the correct identification of non-neoplastic tissue from VIN or SCC using binned individual spectra is shown by the receiver operator curve in Figure 6-26. For individual binned spectra the area under the receiver operator curve is 0.95 and sensitivities of 82-98% can be achieved with specificities of 92-72%.

Each sample was also classified as SCC or VIN or non-neoplastic based on a threshold number of spectra that were classified as SCC or VIN. Using this approach and varying the threshold proportion of spectra used to classify the sample as SCC or VIN a receiver operator curve was created (Figure 6-27). The area under the curve is 0.89 indicating good performance for the identification of samples of SCC or

VIN from non-neoplastic vulval skin. Using this approach a sensitivity of approximately 97% can be achieved with a corresponding specificity of 71% with the number of abnormal spectra used to classify the sample of 3%.

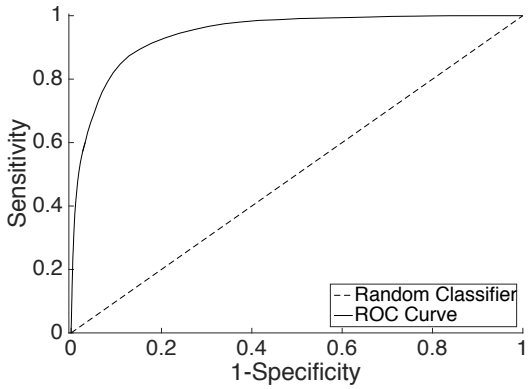


Figure 6-26 – Receiver operator curve for the detection of non-neoplastic tissue from SCC and VIN per spectrum. Area under curve, 0.95, Probability of falsely rejecting null hypothesis, <0.05

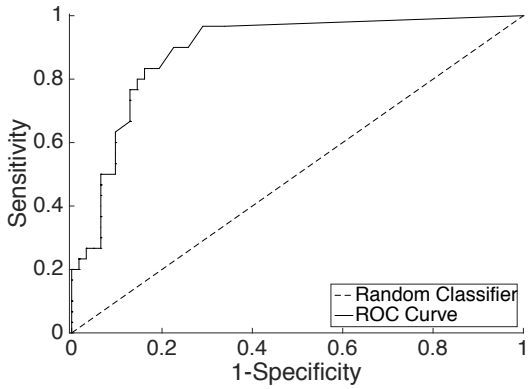


Figure 6-27 – Receiver operator curve for the detection of non-neoplastic tissue from SCC and VIN per sample. Area under curve, 0.89, Probability of falsely rejecting null hypothesis, <0.05

The second comparison undertaken assessed the role of correctly identifying SCC from VIN as this would be useful when monitoring women with VIN for progression to SCC. In this comparison the receiver operator curves for the technique are shown in Figure 6-28 and Figure 6-29. The results indicate that the technique shows promise for the identification of malignant change in premalignant vulval disease.

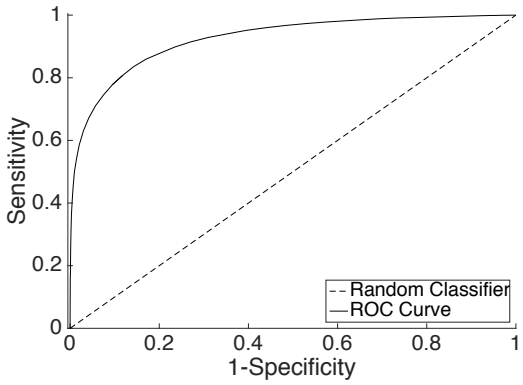


Figure 6-28 – Receiver operator curve for the detection of SCC from VIN per spectrum. Area under curve, 0.92, Probability of falsely rejecting null hypothesis, <0.05

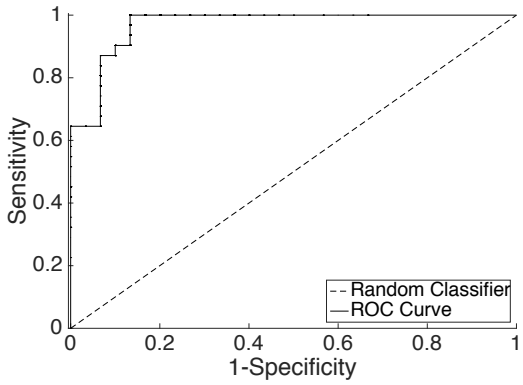


Figure 6-29 – Receiver operator curve for the detection of SCC from VIN per sample. Area under curve 0.97, Probability of falsely rejecting null hypothesis <0.05

Similarly the technique also has good performance in the separation of SCC from non-malignant vulval skin (non-neoplastic and VIN). This is illustrated in the ROCs in Figure 6-30 and Figure 6-31.

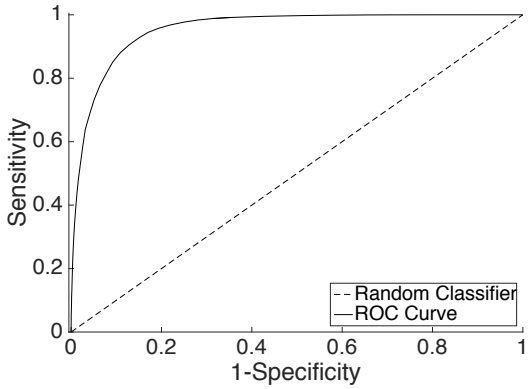


Figure 6-30 – Receiver operator curve for the detection of SCC from VIN and non-neoplastic tissue per spectrum. Area under curve, 0.95, Probability of falsely rejecting null hypothesis, <0.05

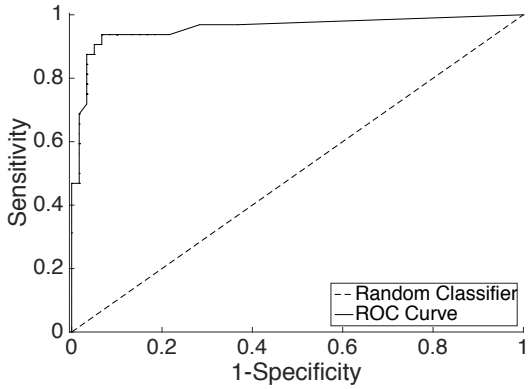


Figure 6-31 – Receiver operator curve for the detection of SCC from VIN and non-neoplastic tissue per sample. Area under curve, 0.96, Probability of falsely rejecting null hypothesis, <0.05

6.7 Spectral analysis

Analysis of the spectral differences between the non-neoplastic, premalignant and malignant vulva tissue was undertaken to understand the basis for disease classification using Raman spectra. Identification of specific biomolecular contributions to the spectral differences can be challenging, as at many wavelengths there is significant overlap between a number of possible bond vibrations and possible biomolecules present within the tissue. This overlap means that a given spectral appearance may have more than one possible molecular cause and hence assignment of spectral peaks to individual molecules is purely tentative. Understanding which biomolecules contribute to which peaks can be achieved though consulting the relevant reference literature and measurements of pure biomolecules made on the same Raman spectrometer (Barry, Edwards and Williams, 1992; Kendall, 2002; De Gelder *et al.*, 2007; Isabelle, 2009; Talari *et al.*, 2015).

Within the Raman vulval spectrum it is possible to tentatively identify some key molecular components. The major peaks present within the epidermal spectra can be assigned as follows: between 936 to 943 cm^{-1} there is a clear peak due to C-C stretching within proline and valine; from 1003 to 1005 cm^{-1} there is a peak due to a phenyl aromatic ring (breathing mode); between 1065 and 1094 cm^{-1} there is a combination of protein vibrations, phospholipids and nucleic acids; from 1243 to 1258 cm^{-1} there is phosphate from phospholipids and nucleic acids; The amide II protein band is found between 1267 and 1279 cm^{-1}; from 1301 to 1321 cm^{-1} there is

a peak caused by CH_2 in lipids; the peak between 1449 and 1453 cm^{-1} a peak is present comprising signal from protein and lipid and the Amide I is found between 1653 to 1669 cm^{-1} (Gniadecka *et al.*, 1997; De Gelder *et al.*, 2007). These spectral features are seen in the mean spectra of the pathology groups as shown in Figure 6-32.

Gniadecka et al. described changes in Raman spectra from whole non vulval skin found in association with the development of SCC of skin (Gniadecka *et al.*, 1997). These include a right shift in the Amide I band of proteins at 1661 cm^{-1}; decrease in the Amide III peak intensity at 1271 cm^{-1}; widening of the protein and lipid peak at 1451 cm^{-1}; an increase in the intensity of the phenylalanine peak at 1003 cm^{-1}; reduction in intensity of the peak at 939 cm^{-1} due to the C-C stretching mode of proline and valine and an increase in peak intensity at 1080 cm^{-1} due to phosphate in nucleic acids and phospholipids. This analysis was based on Raman spectra taken from whole tissue biopsies and was based on the analysis of just three participants with non-vulval SCC.

Within the spectral data from this study many of the features associated with malignancy reported by Gniadecka et al. can be found (Gniadecka *et al.*, 1997; De Gelder *et al.*, 2007). These include the right shift in the 1661cm^{-1} Amide I band, widening of the peak at 1451 cm^{-1}; a reduction in peak intensity at 939 cm^{-1}, an increase in peak intensity at 1080 cm^{-1} (Table 6-6).

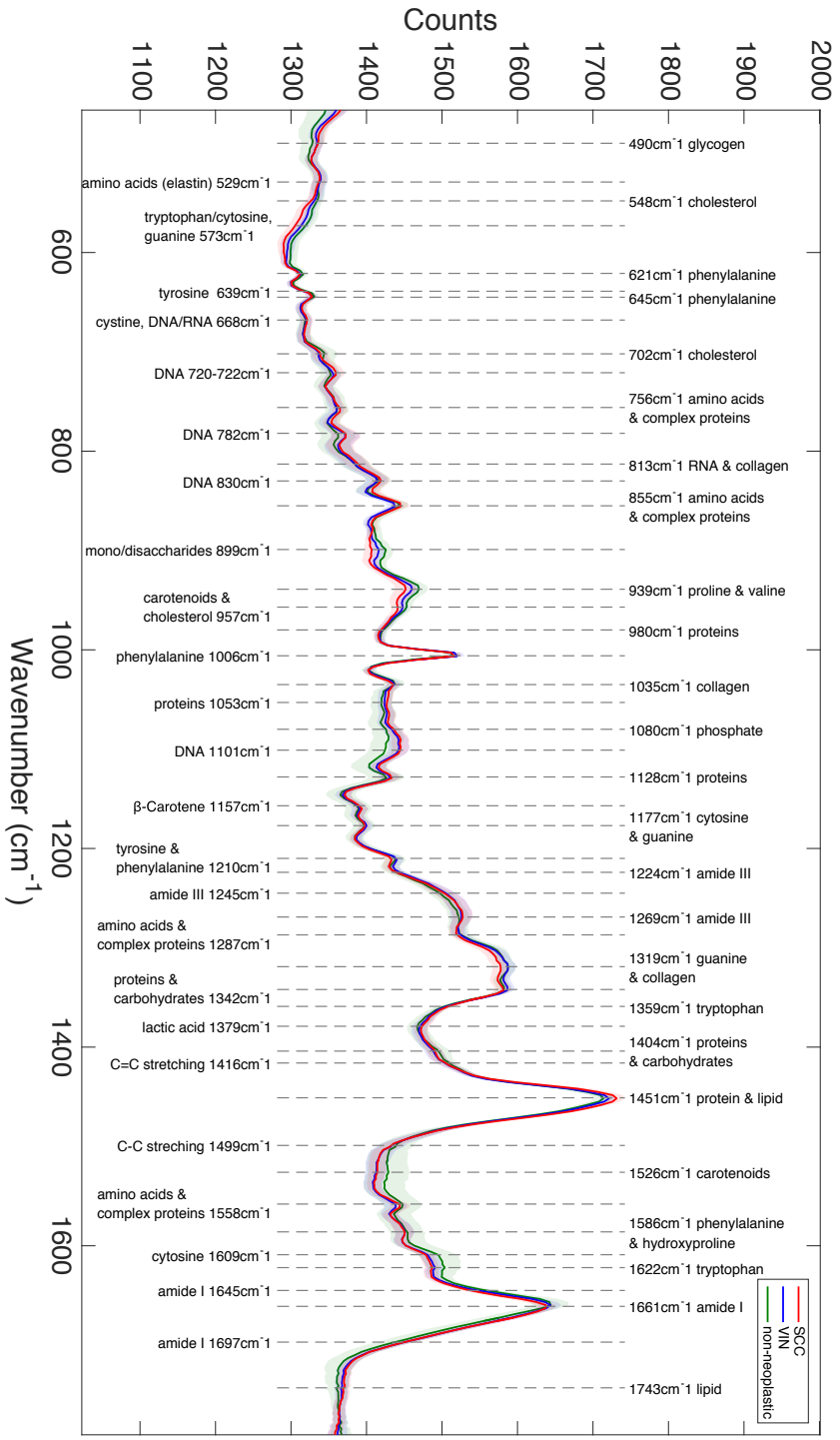


Figure 6-32 – Mean spectra of pathology groups with putative spectral peak assignments

In addition a number of further spectral differences can been observed when comparing neoplastic with non-neoplastic tissue spectra. These include increased in the prominence of the peaks at 723 cm^{-1}, 782 cm^{-1}, 830 cm^{-1}, 1101 cm^{-1} likely to represent DNA (720-722 cm^{-1}, 782 cm^{-1}, 1101 cm^{-1}), RNA (782 cm^{-1}) and nuclei acids (724 cm^{-1}, 830 cm^{-1}) and this would be consistent with the increase rate of cell proliferation in neoplastic tissue (Talari *et al.*, 2015). A reduction is seen in the mono/disaccharide peak at 899 cm^{-1} consistent with high metabolic activity in active neoplastic tissues. There is also an increase in the peaks at 756 cm^{-1}, 855 cm^{-1}, 1287 cm^{-1}, 1558 cm^{-1} and a reduction in the peak prominence at 1622 cm^{-1}. These peaks relate to amino acids including tryptophan, proline, tyrosine and cytosine as well as complex proteins including collagen. The 1622 cm^{-1} peak is most likely to represent tryptophan or the effects of hydrogen bond formation. Other identifiable changes in the spectra with malignant transformation are shown in Table 6-6.

Table 6-6 – Spectral peaks contributing to disease classification

Wavenumber (cm^{-1})	**Suggested biomolecular constituent allocation and change with malignant transformation where apparent**
529	Decrease in desmosine & isodesmosine (amino acids in elastin)
548	Decrease cholesterol
573	Decrease in tryptophan/cytosine, guanine
621	Decrease C-C twisting mode of phenylalanine (proteins)
639	Decrease in tyrosine ring breathing
645	Decrease C-C twisting mode of phenylalanine (proteins)
667-669	Increase C-S stretching mode of cystine (collagen type I), T, G (DNA/RNA)
702	Decrease cholesterol, cholesterol ester
720-722	Increased prominence of DNA peak
756	Increased prominence of amino acids including tryptophan, proline, tyrosine and cytosine peaks as well as complex proteins including collagen
782	Increased prominence of DNA peak
813	Increase RNA and collagen
830	Increased prominence of DNA peak
855	Increased prominence of amino acids including tryptophan, proline, tyrosine and cytosine peaks as well as complex proteins including collagen
899	Reduction in the mono/disaccharide peak

939	Reduction in intensity of the C-C stretching mode of proline and valine
957	Decrease carotenoid, cholesterol
980	Decrease in C-C stretching β-sheet (proteins)
1006	Increase phenylalanine, δ(ring)
1035	Decrease collagen
1053	Increase C-O stretching, C-N stretching (protein)
1080	Increase in peak intensity due to phosphate in nucleic acids and phospholipids
1101	Increased prominence of DNA peak
1128	Increase C-N stretching (proteins)
1157	Increase β-Carotene accumulation (C C stretch mode)
1177	Cytosine, guanine
1210	tyrosine and phenylalanine
1224	Amide III (β sheet structure)
1245	Amide III
1269	Increased prominence amide III band of proteins
1287	Increased prominence of amino acids including tryptophan, proline, tyrosine and cytosine peaks as well as complex proteins including collagen
1319	Guanine (B, Z-marker) CH_3CH_2 twisting (collagen assignment)
1342	CH deformation (proteins and carbohydrates)
1359	Increase in tryptophan
1379	Increase in δCH_3, δOH symmetric (lactic acid)
1404	CH deformation (proteins and carbohydrates)
1416	Decrease C=C stretching in quinoid ring
1451	Widening of the protein and lipid peak
1499	C-C stretching
1526	Decrease in carotenoids
1558	Increased prominence of amino acids including tryptophan, proline, tyrosine and cytosine peaks as well as complex proteins including collagen
1586	Phenylalanine, hydroxyproline
1608-10	Cytosine (NH2)
1622	Decreased prominence in tryptophan peak or decreased effects of hydrogen bond formation
1645	Right shift Amide I (α-helix)
1661	Right shift in the Amide I band
1697	Right shift Amide I (turns and bands)
1743	Lipid in stratum corneum

Wavenumber assignments from Kendall, Talari et al. and Nguyen et al. (Kendall, 2002; Nguyen *et al.*, 2013; Talari *et al.*, 2015)

The technique applied in this study allowed for the isolation of the epidermis, the region of vulval skin affected by neoplastic change and a more detailed analysis of the changes present within this skin layer. Compared to the currently published literature assessing whole samples of non-vulval skin this study has identified a number of additional spectral changes associated with malignant transformation in the vulval as outlined above. The relevance of these changes to malignant

transformation at other skin sites is less clear. The different pathogenic process that lead to the development of vulval SCC may explain the distinctive spectral changes identified in this study compare with the published studies assessing squamous neoplasia at other anatomical sites (C. A. Lieber *et al.*, 2008; J Zhao *et al.*, 2008; Zhao *et al.*, 2010; Lui *et al.*, 2012). Analysis of the role of the identified spectral changes in the PCA LDA classification model using the composite PCA LDA loadings demonstrates these changes appear to be important in the biomolecular classification of vulval disease using Raman spectroscopy (Figure 6-33).

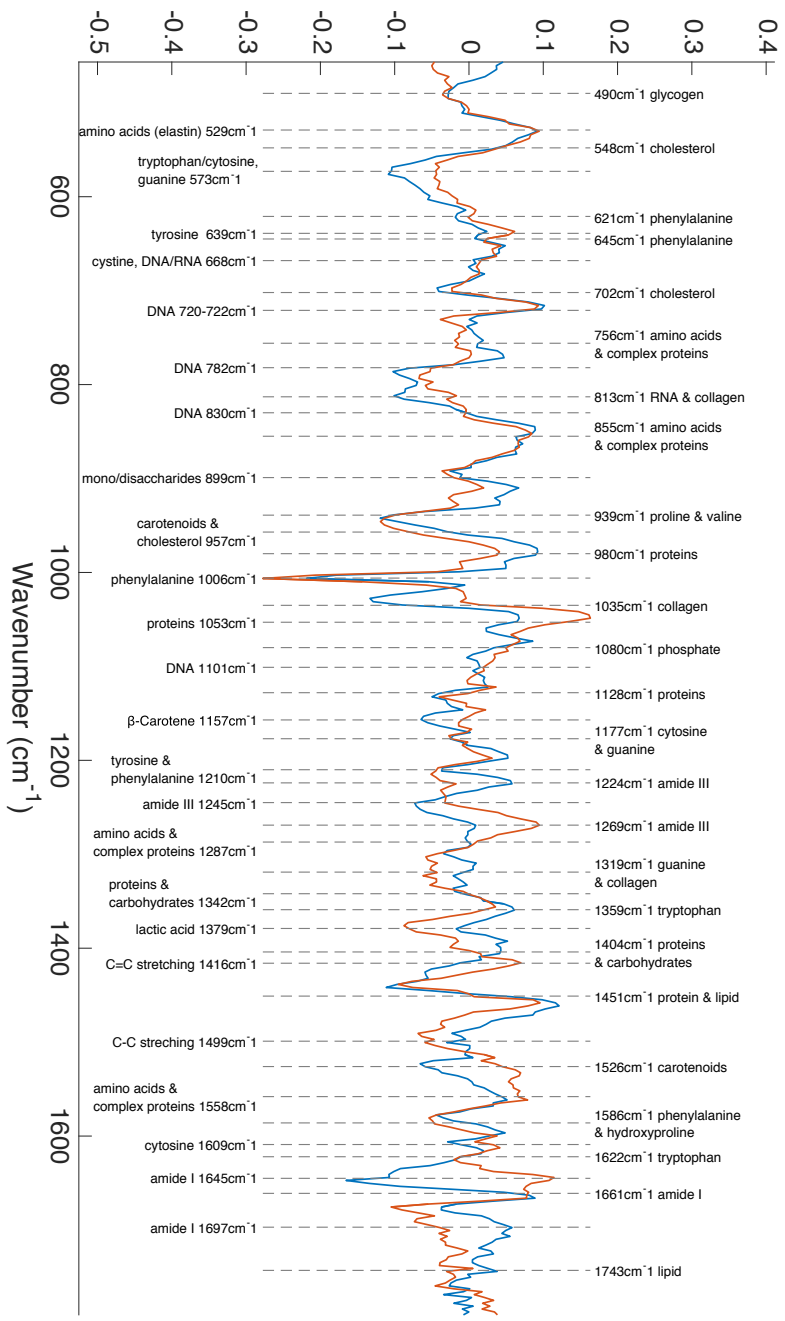


Figure 6-33 – Composite PCA/ LDA loadings with putative spectral peak assignments

6.8 Interim discussion

[Summative discussion and conclusions relating to the investigations outlined in Section C relating to Raman spectroscopy for the assessment of vulval disease can be found in Chapter 13]

The primary aim of this investigation was to investigate the potential for Raman spectroscopy to provide a valuable clinical tool in the management of premalignant and malignant vulval disease. This is the first reported study to analyse fresh frozen vulval tissue using streamline Raman mapping measurements. This study has demonstrated the ability of the technique to discriminate between non-neoplastic vulval tissue, vulval intraepithelial neoplasia and vulval squamous cell carcinoma with clinically relevant sensitivity and specificity. To our knowledge this is the first investigation to demonstrate this technique can be applied to these unique pathologies present within the human vulva. We have used novel methods of analysis developed for this study that combine classifications of grouped spectra across a spectroscopic map to give predictions for individual spectra and whole samples.

This study demonstrates that Raman spectra measured from samples of SCC can be differentiated from those from VIN or non-neoplastic vulval skin combined with a sensitivity of 97% and specificity of 78% (Figure 6-30) and that SCC can be differentiated from VIN alone with a sensitivity of 93% and specificity of 70% (Figure 6-28). This suggests the technique could deliver clinically useful diagnostic accuracy

for the purposes of diagnosis and monitoring disease progression. This study also demonstrates Raman spectra from neoplastic tissue and from non-neoplastic vulval skin can be classified with a sensitivity of 97% and associated specificity of 70% (Figure 6-26) suggesting the technique could deliver clinically useful diagnostic accuracy to determine disease margins in the surgical setting.

The spectral multivariate analysis performed in this investigation demonstrates the potential utility of Raman spectroscopy in the management of vulval disease. This investigation demonstrates that Raman spectroscopy has real promise for the classification of vulval disease and that the technique is able to provide clinically useful diagnostic performance within the laboratory setting and therefore warrants further investigation.

This investigation was performed under strictly controlled laboratory conditions with research grade spectrometers, selecting only epidermal cells for the multivariate analysis. The techniques utilised are of limited direct clinical application as the preparation of paraffin free tissue sections and streamline Raman mapping is technically challenging and time consuming. Analysis of tissue in this way could be used as an adjunct for routine histopathological analysis but it is likely that other techniques such as FTIRS would be better suited to this purpose (Chapter 11).

Although the approach used in this investigation yielded high diagnostic performance it is necessary, within the bounds of ex vivo experimentation, to evaluate the technique in conditions that are a closer approximation to in vivo application. The

experiments outlined in the following chapters are designed to test the technique under conditions that more accurately reflect the clinical environment.

Further discussion of future work and the steps necessary to translate these finding into clinical practice is located in Chapter 13.

Chapter 7 ExVivo Transepidermal Raman Spectroscopy for the Classification of Vulval Disease

7.1 Introduction

The investigation outlined in Chapter 6 demonstrated that Raman spectroscopy coupled with multivariate analysis is able to differentiate between Raman streamline micro spectra from the epidermis of non-neoplastic and neoplastic vulval skin. In order to be successful as a clinical diagnostic tool it is necessary to be able to differentiate between these tissue types using spectra collected in-vivo. Clearly it is not possible to perform mapping of tissue sections in-vivo so in working towards in-vivo application it is crucial to evaluate the classification accuracy of the technique using Raman spectra collected from the surface of vulval skin without isolation of the epidermal cell layers. As we have previously demonstrated in chapter 6.5.4 the biomolecular composition of the stratum corneum epidermis and dermis varies significantly. Due to this variation the Raman spectra obtained from vulval skin using surface (trans-epidermal) measurement is different from those obtained from tissue sections as there is a contribution from all the tissue layers. This variation may

impact on the classification accuracy of the technique as key features may be distorted by features from biomolecules more abundant in the stratum corneum or dermis for example type 1 collagen, elastin and keratin. In light of this anticipated complication of the Raman spectra the next stage in the investigation was to analyse tissue samples using high quality laboratory bench based spectrometers even though these cannot easily be applied to the in vivo setting. This was done to build a foundation for further experimentation with probe systems that can technically be applied to the vulva in vivo.

7.2 Objectives

This investigation assesses objective 2a and 2b of the aims and objectives outlined in Chapter 5.

To evaluate the ability of Raman spectroscopy applied to the surface of fresh and fresh frozen bulk tissue to classify vulval disease.

a) Evaluate Raman spectroscopy applied to fresh vulval tissue for the diagnosis of clinically inconclusive vulval lichen sclerosus.

b) Evaluate Raman spectroscopy applied to fresh frozen tissue for the classification of non-neoplastic vulval skin, premalignant and malignant disease of the vulva.

7.3 Overview of investigation

In this investigation the disease classification capability of Raman spectroscopy applied directly to the surface of blocks of vulval epithelium was evaluated. This was accomplished by the analysis of blocks of fresh vulval epithelium taken at the time of diagnostic biopsy or blocks of fresh frozen vulval tissue collected at the time of surgical excision. These blocks of tissue underwent Raman spectroscopy prior to routine histopathological analysis. An outline of the investigation is shown in Figure 7-1.

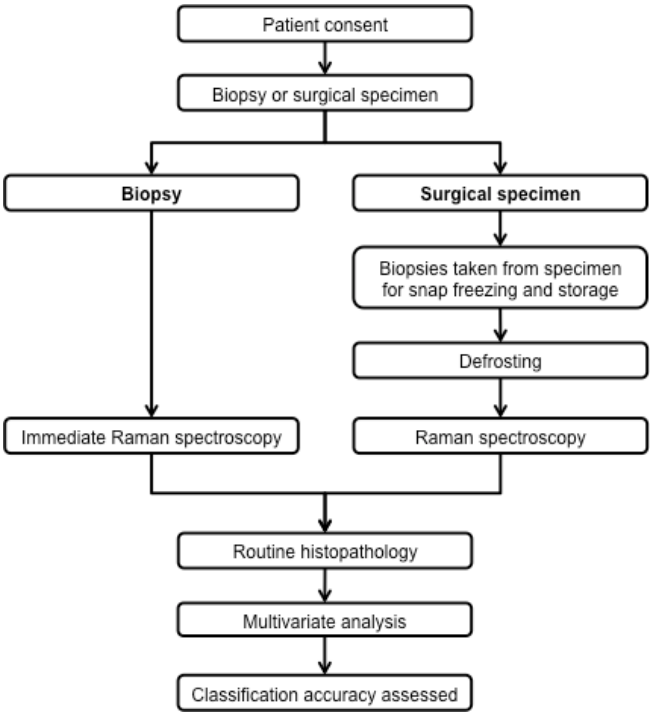


Figure 7-1 – Outline of the collection and analysis of vulval tissue for transepidermal analysis

Two separate comparison groups were analysed, simulating two different clinical scenarios where there is the potential for Raman spectroscopy to perform an important role as a tool capable of classifying the disease present within the tissues.

7.3.1.1 Comparison one

In the first comparison the role of Raman spectroscopy in the assessment of women suspected to have a diagnosis of LS, in whom a clinical diagnosis could not be established with confidence, was explored. National guidelines suggest when LS is suspected and a diagnosis cannot be made on clinical grounds alone then invasive biopsy is indicated (Edwards *et al.*, 2015). Biopsy samples taken from women with suspected lichen sclerosus were analysed in this comparison. The main differentials when a patient is undergoing biopsy for suspected LS are non-specific inflammation due to other causes or variants of normal vulval skin. It is important to establish the correct diagnosis as this dictates the most appropriate management. Furthermore, LS is associated with an increased risk of malignant transformation and as such requires surveillance (Edwards *et al.*, 2015). Comparison of spectral data collected from these tissue types allows us to investigate the role of Raman spectroscopy in reducing the need for invasive biopsy for the diagnosis of LS, a common vulval condition. The spectral characteristics of LS are explored and the diagnostic performance assessed.

In addition to analysis of the study group it would have been desirable to investigate LS in patients in whom clinical diagnosis is possible in order to characterise the full range of the disease. In this preliminary ex vivo study this was not possible as taking

invasive biopsies for research purposes alone was reported to be unacceptable to women during initial consultation with patient representatives. This also meant that obtaining tissue for ex vivo analysis of a normal control group was unacceptable to women, limiting the experimental design.

7.3.1.2 Comparison two

The second comparison assesses the role of Raman spectroscopy in the identification of non-neoplastic, premalignant and malignant vulval disease. This comparison simulates the role of Raman spectroscopy for both the diagnosis of disease and the identification of disease margins at the time of surgery. In this comparison samples of tissue taken from women at the time of biopsy or surgery for VIN or SCC were analysed, comparing non-neoplastic, pre-malignant and malignant vulval tissue. The specimens used were from the same cohort of patient as the specimens used in the investigation outlined in Chapter 6.

7.4 Participant identification and sample collection

As shown in Figure 7-1 two types of sample collection were used in this study: diagnostic biopsy specimens and samples collected from surgical resection specimens. Only biopsy specimens were used in the analysis for comparison one whereas samples from surgical resection specimens were included for analysis in comparison two.

7.4.1 Participant identification and sample collection – diagnostic biopsies

Women referred to specialist gynaecology services at Gloucestershire Hospitals NHS Foundation Trust for assessment of vulval complaints were considered for recruitment to the study. Clinicians providing specialist gynaecological care identified women with a suspected diagnosis of lichen sclerosus who required invasive biopsy due to diagnostic uncertainty as part of their routine care. The women identified gave written consent for recruitment to the study. Vulval biopsies were collected using a 4 mm Keye's biopsy punch with physiological (0.9%) saline skin preparation and 2% lidocaine subdermal infiltration for anesthesia. 2% lidocaine solution was used as it has a less significant Raman profile compared to other commonly used local anesthetic solutions (Figure 7-2). The Raman spectra from commonly used local anaesthetics: 2% lidocaine hydrochloride with adrenaline 1:80,000, 3% prilocaine hydrochloride with felypressin 0.03IU/ml and 2% lidocaine are demonstrated in Figure 7-2.

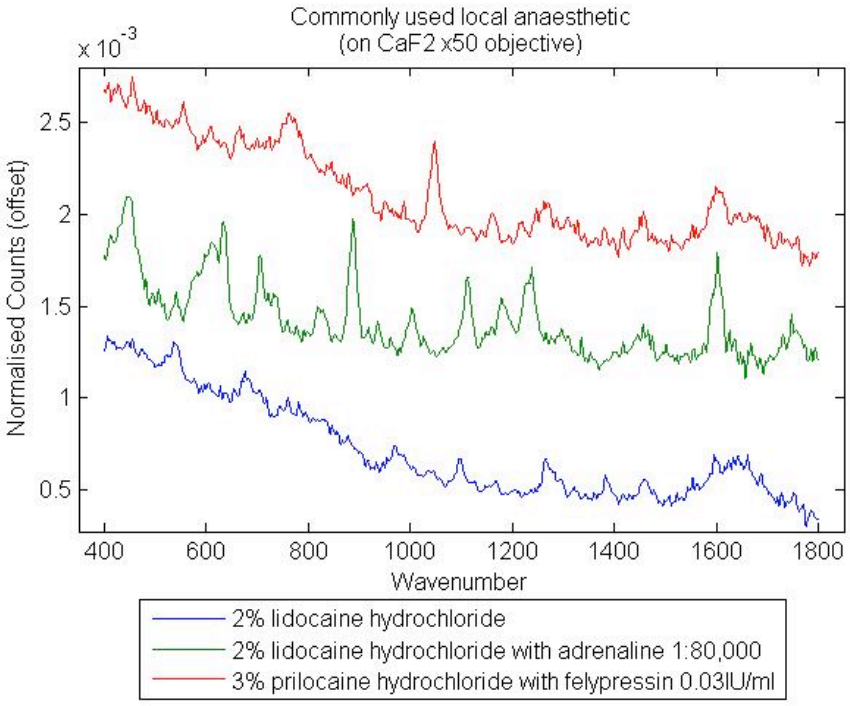


Figure 7-2 – Raman spectra of commonly used local anaesthetic preparations (preparations placed within CaFl2 vessel and spectra collected for 90 seconds through x50 HWD objective)

This tissue collection procedure results in a circular disc of vulval epithelium 3-4 mm in diameter, comprising the full thickness of the epidermis with underlying dermal tissue. The location and number of biopsies collected was directed by clinical need rather than the requirements of the research. Biopsies were washed and transported in physiological (0.9%) saline to prevent dehydration to the on-site spectroscopy laboratory for immediate analysis. After spectral analysis samples were marked with dye to allow orientation and fixed in formalin for paraffin embedding, sectioning, H&E staining and histopathological analysis. No additional tissue was taken from patients solely for the purposes of the research.

7.4.2 Sample collection – surgical resection specimens

Samples were collected from women undergoing vulval excisional surgery as described in chapter 6.3.

7.5 Instrumentation and experimental methods

The Raman spectrometer used for the collection of the transepidermal spectra from the vulval tissue was a modified Renishaw system 1000 dispersion micro spectrometer optimised for tissue analysis in the near infrared range. The spectrometer excitation beam came from an 830 nm laser with a power output of 330 mW. This was connected to a Leica confocal polarise light microscope with a motorised stage and selection of objective lenses, chosen for their minimal fluorescence and Raman signature. The elastically scattered light was rejected with a dielectric coated edge filter and the inelastically scattered light passed through a monochromator slit set to 50 μm onto a 300 line/mm grating which was used to disperse the light onto the CCD. The spectrometer and stage were controlled via a computer interface running Renishaw WIRE software (version 2.0, service pack 9).

7.5.1 System calibration

As with the streamline Renishaw system 1000 microspectrometer the system required calibration prior to each use with silicon standard as previously described

(chapter 6.5.1). A green glass and other standards were also measured to enable correction for variations in instrument response (chapter 6.5.1).

7.5.2 Spectroscope objective characteristics

The choice of microscope objective determines the size of the acquisition volume and the quality of the spectra produced. An objective with a strong Raman signal may mask subtle tissue spectra. The Raman spectra of the standard objectives used with this system were tested by focussing the objective on a polished chromium surface in order to measure the Raman signal from light traveling in both directions through the objective, these are shown in Figure 7-3. Both the x2.5 and the x5 objectives have significant Raman signal, which could mask the tissue spectra. The x50 and x80 objectives have relatively low Raman signal and were found to be suitable for the collection of tissue spectra. A larger beam size makes histopathological correlation more practical as the Raman signal is collected from a larger area and is more likely to be representative of the tissue as a whole. As the x50 objective had the largest beam size of the two lowest Raman signal objectives it was chosen for the analysis. Advancing the beam produced by the x50 objective into a silicon wafer whilst recording the resultant Raman spectra was used to assess the depth of the laser spot size which was found to be in the order of 17 microns (Figure 7-4).

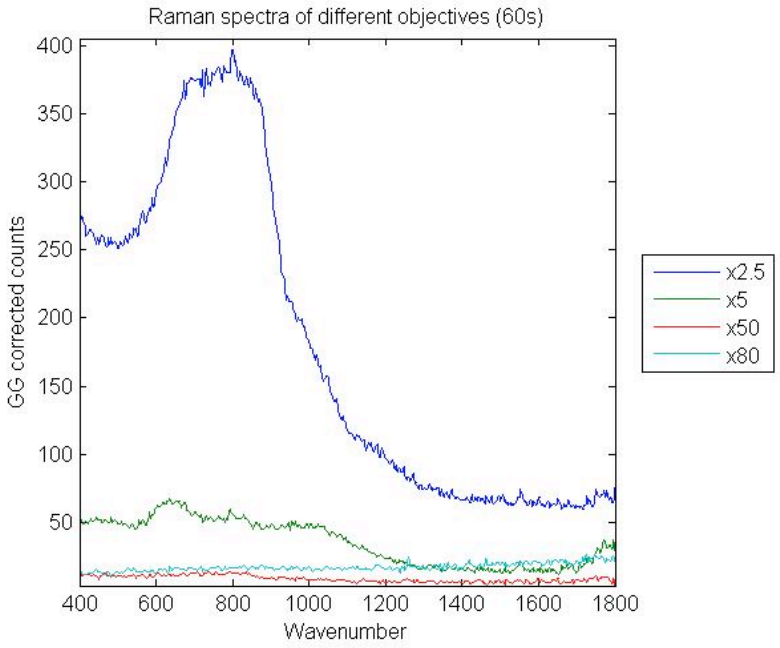


Figure 7-3 – Raman spectra of different microscope objectives with 60 second acquisition time.

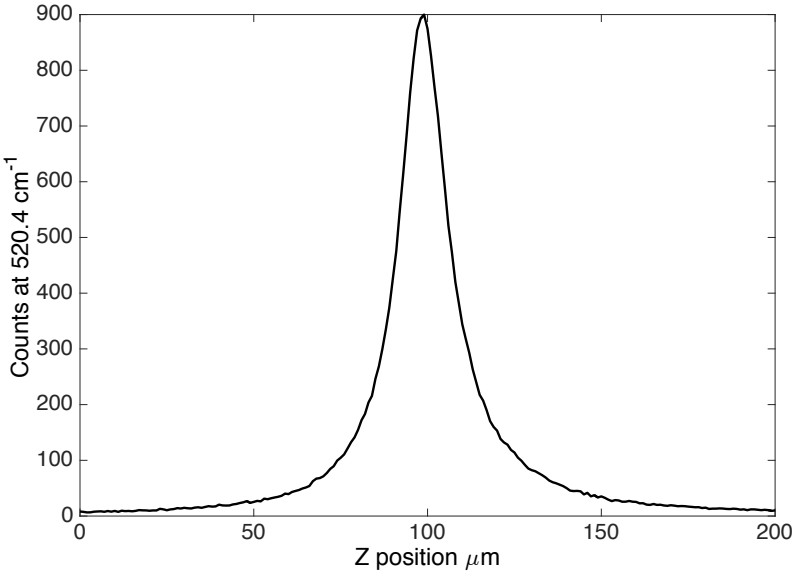


Figure 7-4 – Counts at 520.4 cm^{-1} as silicon wafer is advanced towards x50 HWD objective, full width half maximum of peak is 17 microns.

7.5.3 Sample mounting

Samples needed to be securely mounted to facilitate reproducible spectral measurements across the sample and to prevent the sample drying out. In order to

allow reproducible standardised collection of spectra at set depths through the vulval skin samples they needed to be mounted perpendicular to the beam from the spectroscope. Initial experiments with samples placed directly onto calcium fluoride slides within a droplet of 0.9% saline resulted in drying of the tissue. This drying led to disorientation of the sample surface during spectral acquisition. Samples were also mounted within a polymer cone (pipette tip) and this allowed appropriate sample orientation, however samples were still subject to drying over the long acquisition times. In addition the surface of the samples was not level and this made acquisition of adjacent spectra challenging. The addition of a 250 micron calcium fluoride cover slip both retained the moisture within the samples and produced a level surface for the acquisition of multiple spectra across the sample (Figure 7-5). Figure 7-6 shows the spectra acquired from the materials that made up the trial sample mount (calcium fluoride, polymer cone, polymer black foam and coated alloy). A laser power meter was used to assess the impact of the calcium fluoride cover on the laser power. With the x50 objective the laser power at the sample is on average 5.9 mW, when a 0.25 mm calcium fluoride window is placed in front of the power meter the power remains at 5.9 mW, therefore the window does not adversely affect the intensity of the light incident on the tissue.

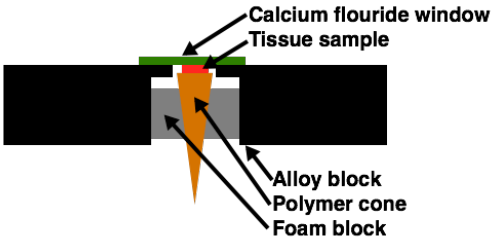


Figure 7-5 – Prototype tissue mount for Raman spectroscopy (not to scale)

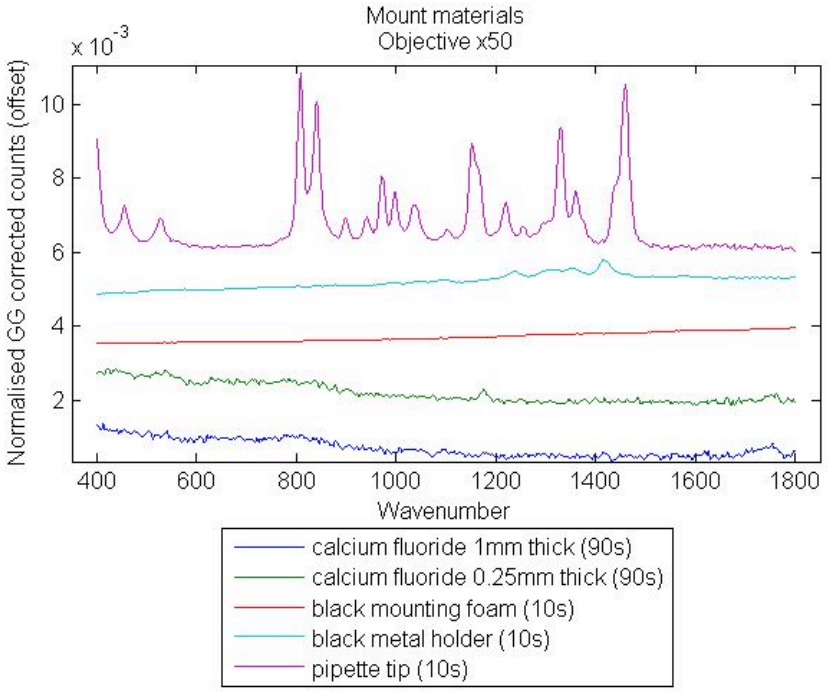


Figure 7-6 – Raman spectra of materials comprising prototype mounting system

This mounting system was tested using chicken skin biopsies, Figure 7-7 shows point spectra from a 4 mm punch biopsy in the prototype mounting without cover slip, in the prototype mounting with 0.25 mm calcium fluoride cover slip and placed on a calcium fluoride slide. The primary drawback with this mounting method was the lack of a mechanism for ensuring the proximity of the sample and the calcium

fluoride window. When the sample was not in close contact with the calcium fluoride window the spectra acquired from the tissue are significantly affected (Figure 7-8).

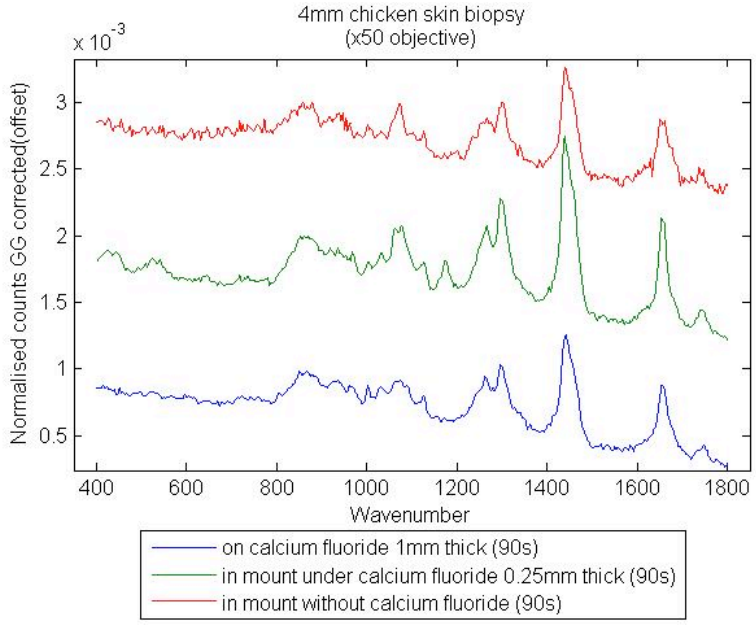


Figure 7-7 – Raman spectra of 4 mm chicken skin biopsy using different mounting methods

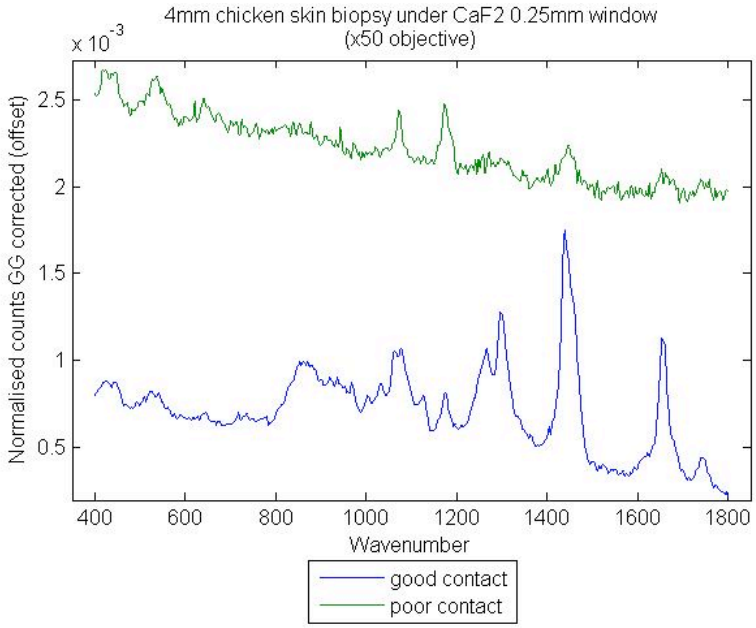


Figure 7-8 – Raman spectra of 4 mm chicken skin biopsy in good and poor contact with cover window

To overcome the problem of consistent contact of the sample with the calcium fluoride window a second mounting system was designed to allow increased and even pressure of the sample on the window (Figure 7-9). The threaded plunger and secured calcium fluoride window allowed the sample to be pressed against the window, reducing poor contact between the two.

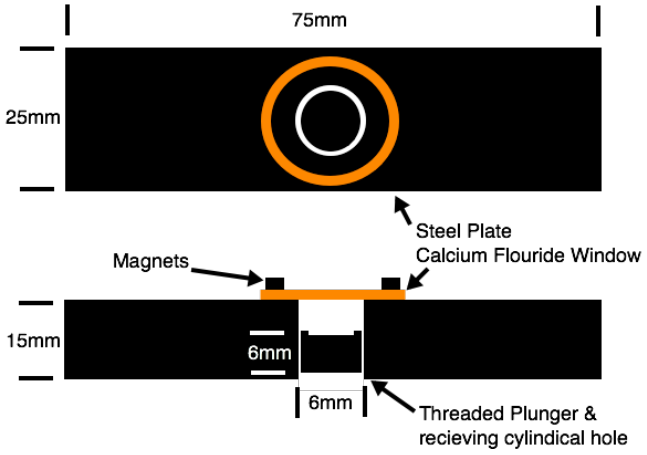


Figure 7-9 – Design for sample mount (not to scale)

7.5.4 Optimisation of parameters for spectral measurement of fresh bulk tissue

The signal to noise ratio (SNR) of measured tissue spectra was used to optimise the experimental setup and acquisition parameters. The signal was calculated as the height of the main tissue peak between 1420 and 1480 cm^{-1} above the median value of an area of adjacent spectrum between 1400 and 1420 cm^{-1}. The noise was calculated from the median of the absolute difference between the actual spectrum and a median filtered spectrum using a section of the spectrum with little biomolecular signal (1700-1800 cm^{-1}). The median values of the absolute difference

were used, as it was necessary to remove the effect of cosmic rays, which are beyond the control of instrument optimisation.

For fresh vulval tissue one of the key considerations in setting measurement parameters is the time taken to collect the spectra. If a typical 4 mm punch biopsy is under Raman analysis for greater than three and a half hours there is a risk of the sample drying and this risked adversely affecting the quality of the histopathological assessment of the tissue. The spectral collection parameters were therefore optimised to provide adequate number of spectra with good spectral quality within three and a half hours per sample.

7.5.4.1 Spectral acquisition time

The time taken to acquire a spectrum directly affects the SNR of the resultant spectrum. The gain in SNR is most significant between 0 and 30 seconds, above 20 seconds there is only a modest increase in the SNR, which appears to plateau at 50 seconds (Figure 7-10). A shorter acquisition time is desirable to enable the analysis to be undertaken in a time frame that minimises drying & distortion of the tissue and reduces the chance of thermal damage from the laser. For these reasons an acquisition time of 20 seconds was selected for the analysis.

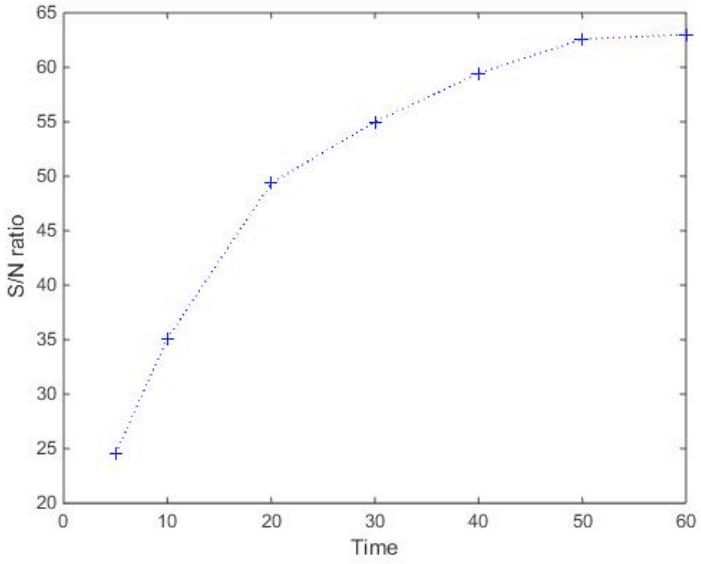


Figure 7-10 – SNR for increasing acquisition time for spectra taken from normal vulval skin

7.5.4.2 Fluorescence

To combat problematic fluorescence (chapter 4.1.3.2) found in skin with high concentration of melanin the effects of photo bleaching were investigated. Using the 830 nm laser to photo bleach the tissue did not seem to have any significant effect on the amount of fluorescence from the tissues. For this reason photo bleaching was not used during spectral collection. An iterative polynomial curve fitting baseline correction was applied to all spectra to remove the effect of fluorescence present in the spectra (Lieber and Mahadevan-Jansen, 2003).

7.5.4.3 Cosmic ray removal

From the measurement of fifty 20 second spectra the chance of a cosmic ray being present in a spectrum was estimated as 1 in 20. As affected spectra can be

identified and corrected during analysis and as the total time fresh samples can undergo analysis is limited, it was concluded more useful data could be gathered without using the triple measurement median spectrum removal technique (chapter 4.1.3.1). Cosmic rays were therefore identified and removed during data analysis, as described later.

7.5.4.4 Spectral acquisition depth

Histopathological and therefore biochemical changes associated with vulval disease do not typically occur at the surface of epithelium where the stratum corneum is found (Chapter 1). Pathological changes in benign and malignant vulval disease occur primarily at the level of the junction between the dermis and the epidermis (Nucci, 2013). Within vulval skin this area is situated a varying distance from the superficial surface of the stratum corneum. It is therefore sensible to acquire Raman spectra at different depths within the samples. This approach is limited by the reduced signal to noise as the laser spot is focused into the sample and experiments revealed a significant drop off in the signal to noise ratio when the focus was advanced beyond 150 microns into the tissue.

7.5.4.4.1 Tissue modelling

To further gain insight in to the depth of tissue a Raman signal can be acquired from, an experimental phantom was produced to simulate the signal returned from skin at different depths. For light of wavelength 830 nm the reduced scattering coefficient of skin was calculated to be 22.39 cm^{-1} and the absorption coefficients of the epidermis

and dermis were calculated to be 1.715 cm^{-1} and 0.249 cm^{-1} respectively. These values were calculated using the methodology and experimental tissue properties from Jacques et al. (*Skin Optics*; Jacques, 2013). The reduced scattering coefficient was matched to a 2.48% solution of Intralipid® (Sigma-Aldrich).

To investigate the depth at which spectra could be acquired a cuboid piece of polytetrafluoroethylene (PTFE) was place in a petri dish on the stage of the spectroscope and the petri dish filled with 2.48% intralipid solution until the PTFE was entirely covered. Due to the highly hydrophobic nature of the PTFE this required a layer of intralipid approximately 1400 μm thick to cover the PTFE. This layer could be reduced to 500 μm before the surface of the intralipid broke over the PTFE. This allowed analysis of the signal from the PTFE block at varying focal depths through varying thickness of intralipid skin phantom. The resultant spectra were green glass corrected and normalised and the height of the largest PTFE peak at 732 cm^{-1} (Figure 7-11) was isolated to represent the changing signal from the PTFE block (Figure 7-12). The results of this show that even when the depth of the intralipid is over 1 mm the PTFE Raman signal can be detected. This depth is equivalent to the thickest combined stratum corneal and epidermal layer in vulval skin. An adjustment is required for the absorption of light within skin, which is not satisfactorily accounted for in this model.

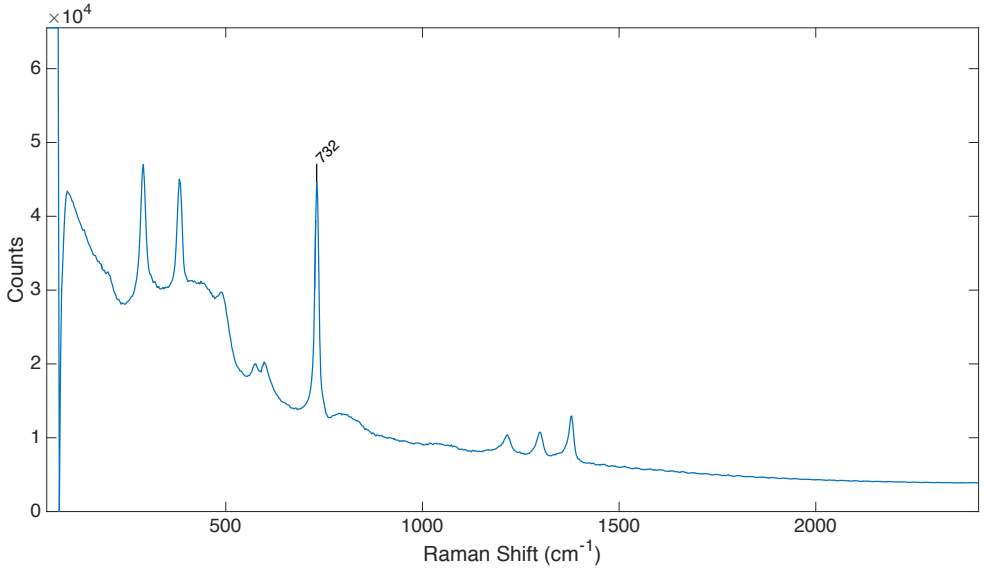


Figure 7-11 – PTFE spectrum, x50 objective, 10 sec integration time

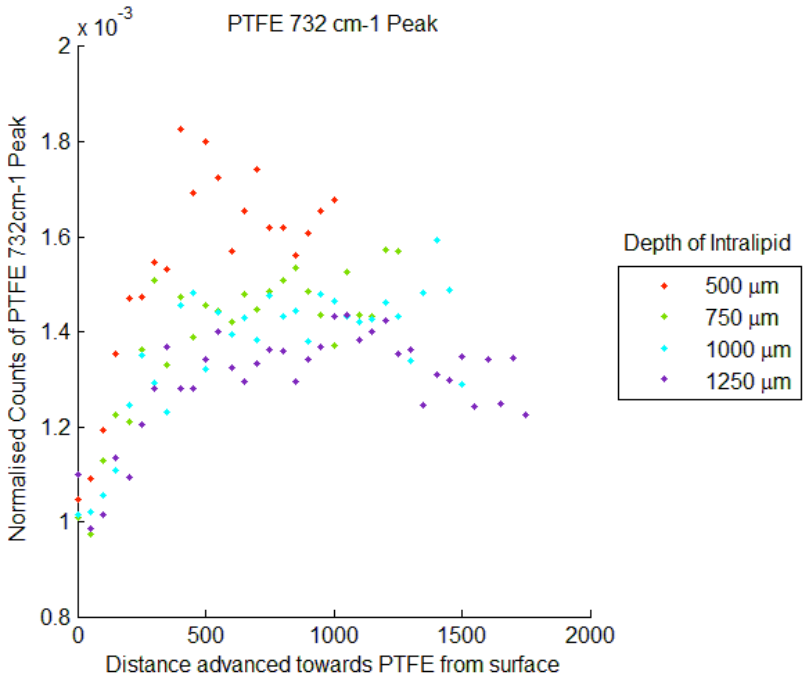


Figure 7-12 – Normalised counts of the 732 cm^{-1} peak as the microscope focus advances towards the PTFE through different depths of Intralipid 2.48%, x50 objective, 60s integration time.

7.5.4.5 Final acquisition parameters

A two-stage approach was chosen for the acquisition of spectral measurements from blocks of fresh/ fresh frozen vulval tissue. The first stage was applied to spectral data collected for comparison one and the second stage applied to the additional spectral data collected for comparison two.

7.5.4.5.1 Acquisition parameters – comparison one

Comparison one investigates both the spectral differences between different vulval pathologies and the depth of laser focus most likely to produce diagnostic spectral information. Spectra were acquired in a line across the central 2 mm of the fresh vulval 4 mm punch biopsy samples at intervals of 50 μm with the laser focused on the surface of the sample and an acquisition time of 20 seconds. In order to assess the optimal focal point for acquiring spectra in this type of tissue this process was repeated after the microscope stage was advanced towards the objective. The stage was advanced in steps of 15 μm and spectra measured at each position until the sample had been advanced 150 μm towards the objective (Table 7-1). An average of 440 spectra were measured from each sample.

Table 7-1 – Acquisition parameters for comparision one

Parameter	Value
Acquisition time	20 secs
Stage z increments	15 microns
Total z displacement	150 microns
Lateral increments	50 microns
Lateral displacement	2000 microns
Total Spectra	440
Total acquisition time	3 hours

7.5.4.5.2 Acquisition parameters – comparison two

The second stage of the data acquisition for comparison two was based on the results of the first comparison between LS and other inflammatory conditions 7.6.1. As the optimum depth for spectral acquisition for the purposes of classification (Chapter 7.6.1) was found to be 15 microns, spectra were only collected at this depth in comparison two (Table 7-2).

Table 7-2 – Acquisition parameters for comparision two

Parameter	Value
Acquisition time	20 secs
Total z displacement	15 microns
Lateral increments	50 microns
Lateral displacement	2000 microns
Total Spectra	40
Total acquisition time	15 minutes

7.5.5 System calibration

System calibration was carried out as described in chapter 6.5.1.

7.5.6 Histopathological correlation

Histopathological correlation was achieved with an agreed protocol for marking of punch biopsies. Designated histopathologists familiar with the protocol cut all samples sent to the histopathology laboratory. Samples underwent Raman spectroscopy along a line through the sample and were marked with ink on one side of the skin surface with a black dot. This allowed the pathologist to cut the sample into two parts (Figure 7-13) and cut sections for staining and histological analysis

along the line explored with the spectroscope. All samples underwent histopathological assessment by a specialist gynaecological pathologist blinded to the results of the spectroscopic analysis. The results of this assessment were then used in building and validating a PCA LDA based classification model (chapter 4.2).

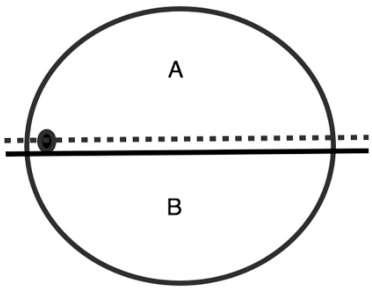


Figure 7-13 – Marking of punch biopsies (ink dot on left of the biopsy allows the pathologist to cut the specimen along the solid line, mount section A in paraffin and cut sections for staining and histology at the level marked with the dashed line)

7.5.7 Spectral preprocessing and analysis

Spectral data points outside of the 400 to 1800 cm^{-1} biomolecular fingerprint range were excluded from the analysis and all spectra underwent an instrument response correction using the spectrum measured from the green glass as described in chapter 6.5.1 (Grimbergen *et al.*, 2010). Autofluorescence was removed using a second order polynomial fitting and subtraction (chapter 4.1.3.2) (Lieber and Mahadevan-Jansen, 2003) and the spectra were then vector normalized. Cosmic rays were identified by isolating spectral points where the difference between the spectra and the median spectra for that sample at the given depth deviated from the median absolute difference by greater than 4 standard deviations (chapter 4.1.3.1).

Where cosmic rays were identified the affected region of the spectrum was removed and substituted by a reconstruction using principal component loadings from analysis of spectra free from cosmic rays. The resultant spectra were then analysed with principal component analysis (PCA) to reduce the dimensionality of the data in preparation for linear discriminant analysis (LDA).

As expected on visual inspection the spectral data obtained using the transepidermal approach appeared to contain a more significant contribution from sources of noise than the data collected through streamline Raman spectroscopy median filtered and binned. Analysing this spectral data using the cross validated classification approach outlined in chapter 6.5.5 resulted in the inclusion of multiple principal components that were found to be significantly different between the pathology groups using ANOVA but that on inspection appear to contain predominantly noise with little evidence of biological Raman signal. In order to select principal components for LDA that were biological relevant, alternative data analysis methodologies were developed and evaluated and compared to the method outlined in chapter 6.5.5.

The first alternative analytical method used a sub cross validation loop to determine the number of PCs to use in each cross validation loop. In this approach, after exclusion of the test set in cross validation loop and filtering by the D and Q statistic a further analytical sub loop was undertaken to assess the number of principal components that contribute to improved disease classification (Figure 7-14). Within the sub loop each sample in the training set was excluded in turn and a further round of PCA and LDA undertaken using the remaining data as the sub training set and the

excluded sample as the sub test set. The diagnostic performance of the classification model was then assessed after adding each PC in turn starting with the most discriminatory and ending with the least discriminatory (Figure 7-15). The number of principal components that resulted in the maximum number of correct predictions in the sub test set was then noted and the sub loop was then repeated for the next sub test sample and the mean number or PCs that resulted in the maximum number of correct predictions applied to the test sample in the primary cross validation loop.

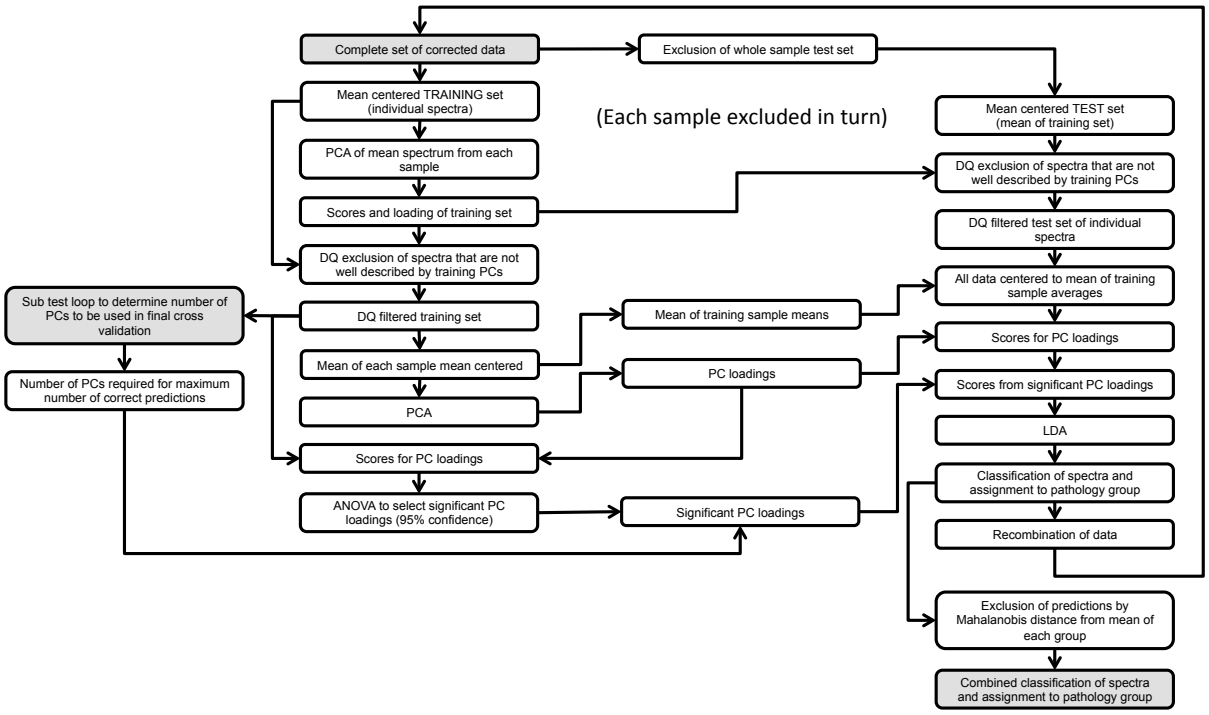


Figure 7-14 – First alternative method for selecting principal components within leave one sample out cross validated PCA LDA classification

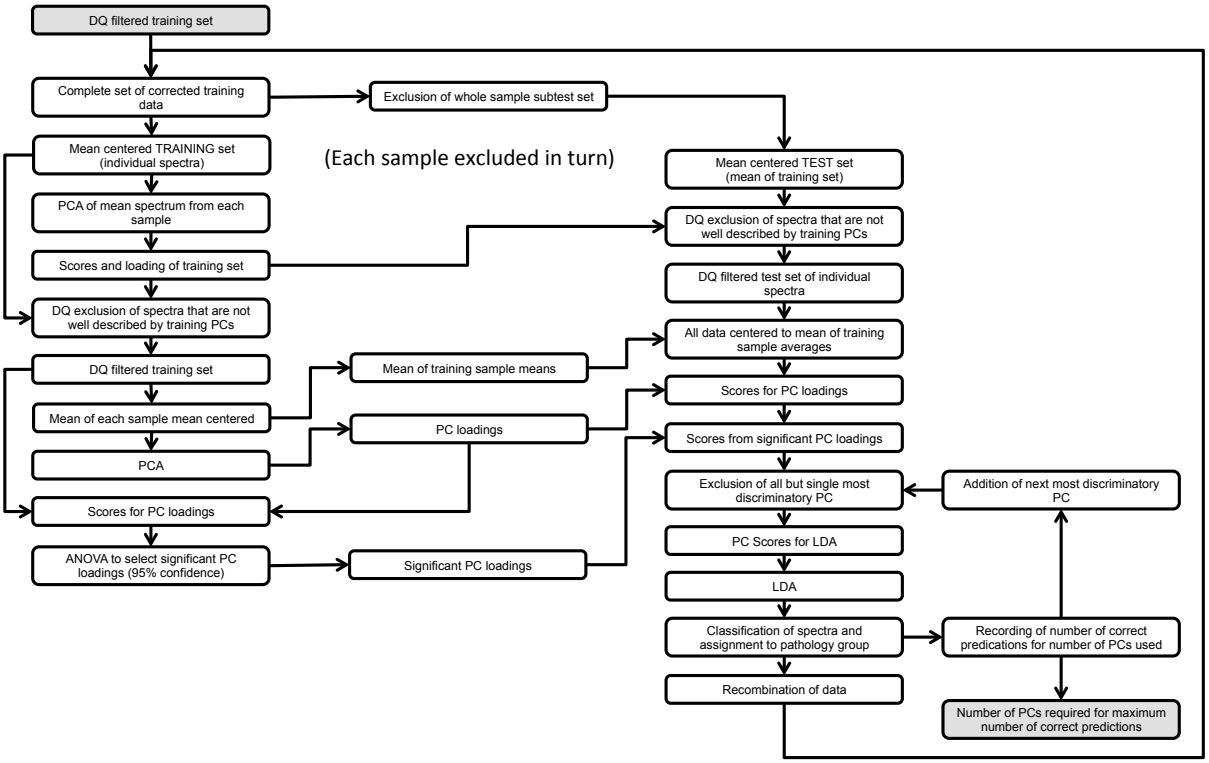


Figure 7-15 – Sub loop for determining the number of biological relevant PCs in the first alternative method for selecting PCs within leave one sample out cross validated PCA LDA classification

The second alternative cross validation methodology uses a sub loop and stepwise selection of the actual principal components to be used in the final analysis to reduce the significance of noise in the classification model and reduce the risk of a type 2 error (Figure 7-16).

In the stepwise cross validation cycle the data from a whole tissue sample (test set) was removed from the complete data set. After this a randomly selected 25% of the remaining full spectra were then removed from the data set as a stepwise sub test set. PCA fed LDA was then performed on the remaining data. The resultant model was then tested on the stepwise sub test set using each principal component (PC) in turn to identify the PC that resulted in the highest proportion of correct predictions (Figure 7-17). This PC then remained in the PCA-LDA model and the remaining PCs

were tested alongside this PC in the same manner to identify the next best performing PC. This process was repeated adding the next best performing PC to the model each time. Only those PCs that improved the performance of the PCA-LDA model were then selected for classification of the whole sample test set. The classification performance of the PCA-LDA model was assessed on its ability to correctly classify the histological diagnosis of the whole sample test set. This leave one out stepwise cross validation loop was then repeated evaluating each sample in turn.

This method had the advantage of removing principal components that described variation in the data not useful for classification such as those that described noise. This approach allowed improved classification performance without compromising the integrity of the leave one sample out cross validation and without increasing the chance of a type 1 error.

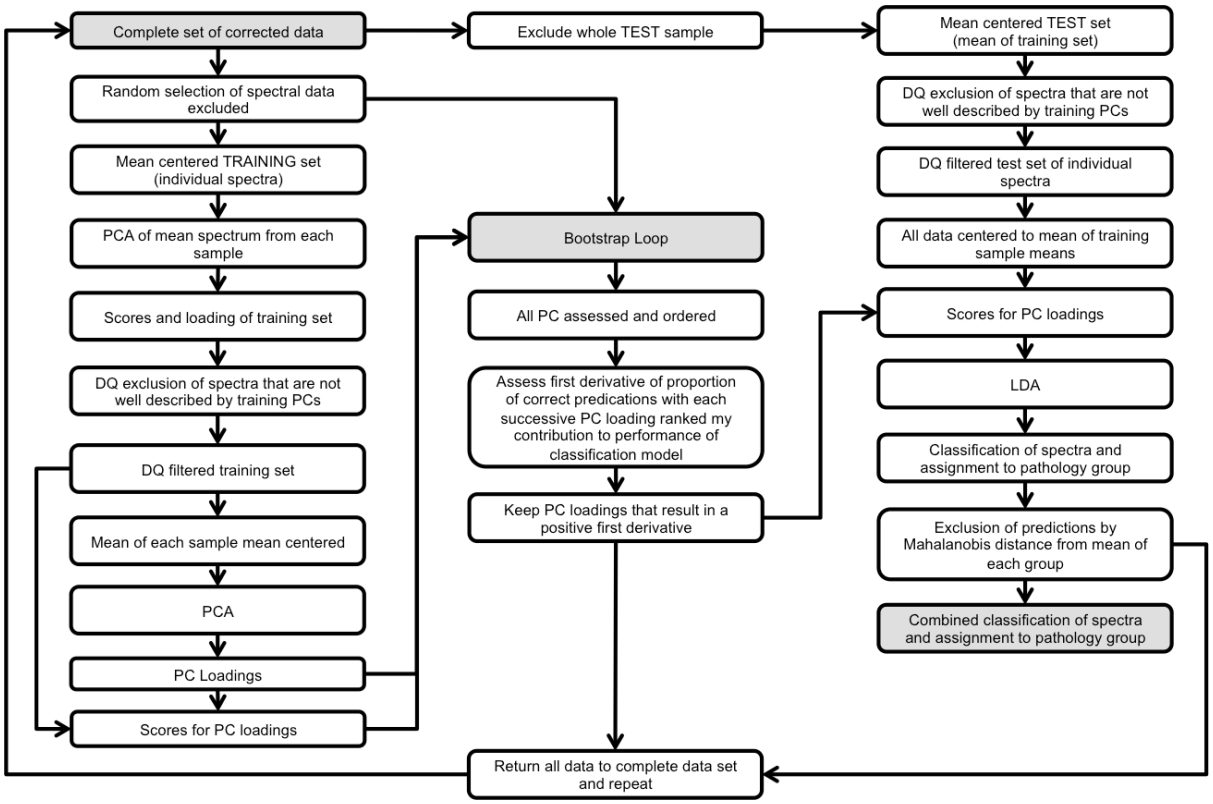


Figure 7-16 – Second alternative method for selecting principal components within the leave one sample out cross validated PCA LDA classification

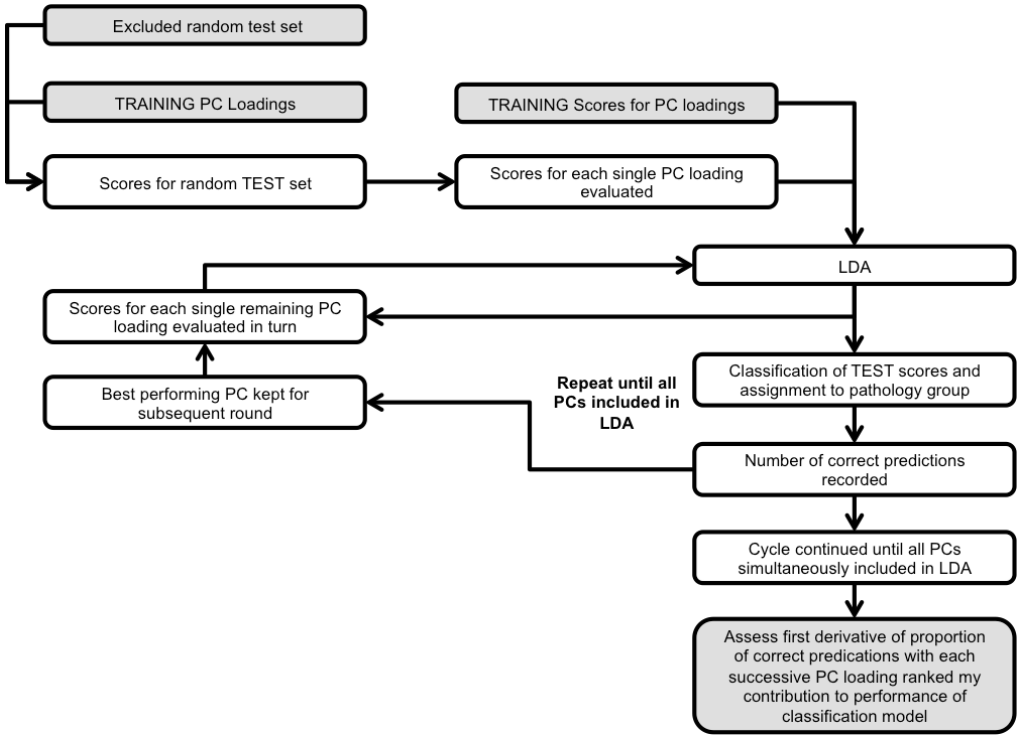


Figure 7-17 – Sub loop for identifying the biological relevant PCs in the second alternative method for selecting PCs within leave one sample out cross validated PCA LDA classification

The classification performance of individual spectra was assessed using the receiver operator characteristic to simulate measurements that could be acquired in a clinical setting. The probability of obtaining a false positive result was evaluated using the method described by Obuchowski et al. (Obuchowski and McClish, 1997; Obuchowski, Lieber and Wians, 2004).

Comparing the three different methods for cross validation in the context of spectra with significant noise using the data collected using the transepidermal approach demonstrated the effect on the diagnostic performance of applying the different analytical approaches. A sample comparison of the areas under the ROC curve (AUC) when differentiating SCC from VIN using the three approaches using all the data was undertaken. Improved classification performance was seen with the second alternative (boot strap) classification approach (AUC 0.88) when compared with both the approach applied in chapter 6.5.5 (AUC 0.86) and the first alternative approach (AUC 0.85). The second alternative (boot strap) classification approach was therefore adopted to assess the diagnostic performance of the technique.

7.6 Results

7.6.1 Comparison one - diagnosis of clinically inconclusive vulval lichen sclerosus

7.6.1.1 Summary of measurements

Samples from 27 women with suspected LS were subject to spectroscopic and histopathological analysis. Of these, 15 women had a histological diagnosis of LS. Of the remaining 12 women, the majority had findings of nonspecific inflammation only, some had no specific histological abnormality and there was one case of mild dysplasia (Table 7-3).

Table 7-3 – Histopathological diagnoses of study samples

Histological Diagnoses	**N**	**%**
Lichen sclerosus	15	56
Other diagnoses:	12	44
Non specific inflammation	(8)	
Mild Dysplasia	(1)	
No specific histological abnormality	(3)	

Spectra were measured from all samples between 30 minutes and 7 hours of the samples being collected. A mean of 440 spectra were collected from each sample. A two group analysis of the data was performed to differentiate those samples from women with LS from those with other diagnoses as this was the most clinically relevant comparison possible using the available data.

7.6.1.2 Diagnostic classification

The initial PCA LDA analysis demonstrated that the technique was able to correctly classify the majority of samples based on the spectra acquired using the linear discriminant scores (Figure 7-18). This result was replicated in the leave one out stepwise cross validation. The bimodal distribution of linear discriminant scores in the lichen sclerosus group (Figure 7-18) suggests there may be separate biomolecular subtypes of the disease although no correlation between the scores and histological features could be identified.

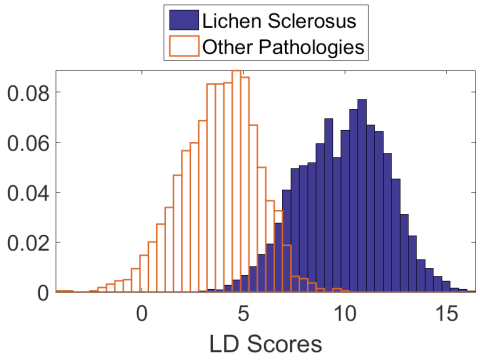


Figure 7-18 – Histogram of score values from linear discriminant function

Analysis of the spectra from different focal points within the samples revealed the diagnostic performance was optimal when the focal point was advanced 15 microns into the samples (Figure 7-19). Only data from this depth (40 spectra from each sample) were used in the final analysis.

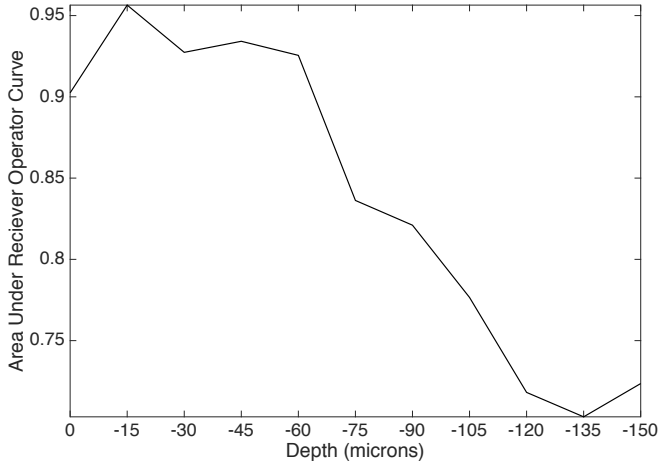


Figure 7-19 – PCA LDA classification performance (area under receiver operator curve) using spectra obtained from different depths of focus.

Defining the biochemical constituent peaks responsible for discrimination within the LDA model is challenging, as Raman spectroscopy does not allow simple biochemical quantification of target tissues. This is due to the compound effect on the Raman spectra from the numerous molecular constituents of the tissue.

The dominant histological feature of LS is hyalinisation of the dermis typically seen alongside epidermal atrophy and subdermal inflammatory cell infiltrate. The dermis primarily consists of collagen fibres with smaller quantities of elastin fibres and other extracellular proteins. In human skin type 1 collagen accounts for 80% of all dermal collagen, in LS significant biochemical alterations occur around the zone of hyalinisation with changes in the organisation and expression of type I, III collagen fibers, a relative increase in the expression of type V collagen and a decrease in the expression of extracellular matrix protein 1 and elastin. These changes may form the molecular basis for the differentiation of the two groups in this study.(Godoy *et al.*, 2015)

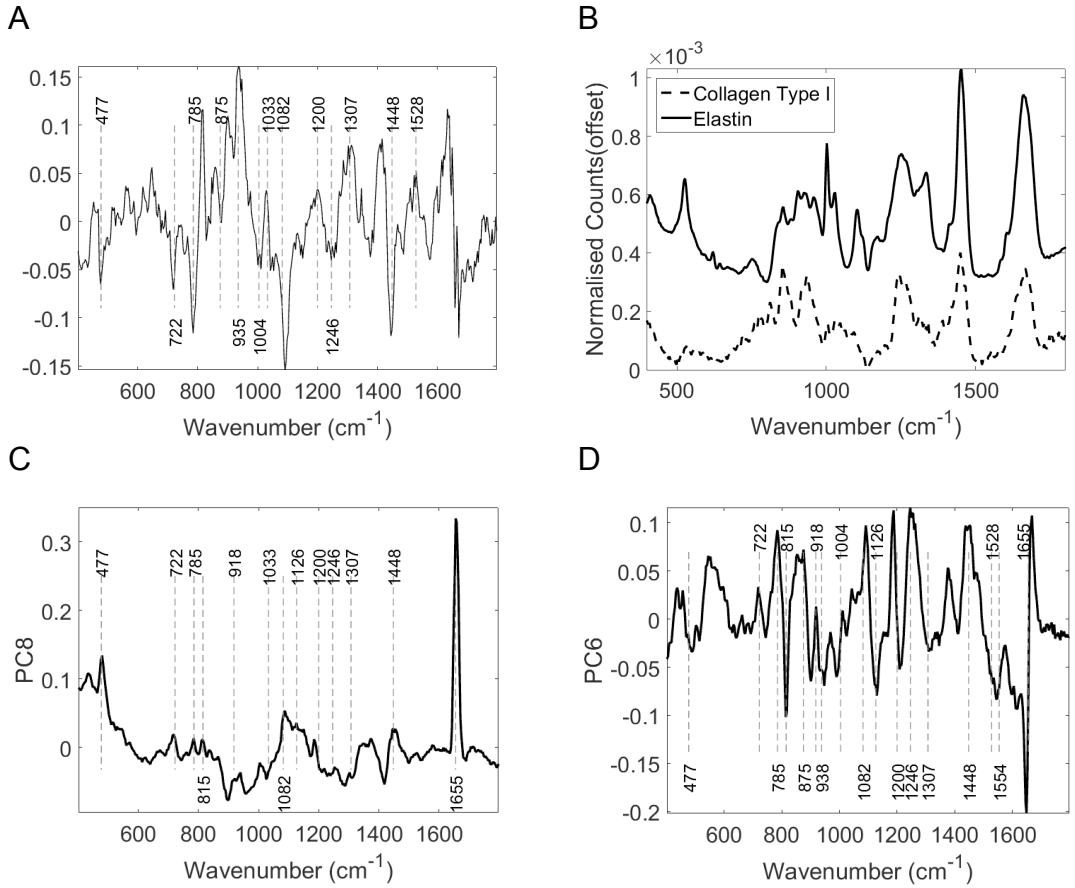


Figure 7-20 – A - Composite of all principal component loadings weighted by contribution to the linear discriminant classification model; B – Collagen type I (Sigma C7774) and elastin (Sigma E1625) spectra; C - Most significant principal component loading; D - Second most significant principal component loading.

Examining the two most significant principal component loadings (ANOVA) and the composite of all principal component loadings weighted by contribution to the linear discriminant classification model (Figure 7-20) spectral differences between the two groups (Figure 7-21) can be seen to occur in the regions associated with collagen, elastin and other structural proteins (Table 7-4). Changes are also seen in the lipid, fatty acid, nucleic acid and carbohydrate regions of the spectra (Table 7-4). These changes may well be associated with epidermal atrophy, vacuolar degeneration and the lichenoid lymphocytic inflammatory infiltrate seen in LS.

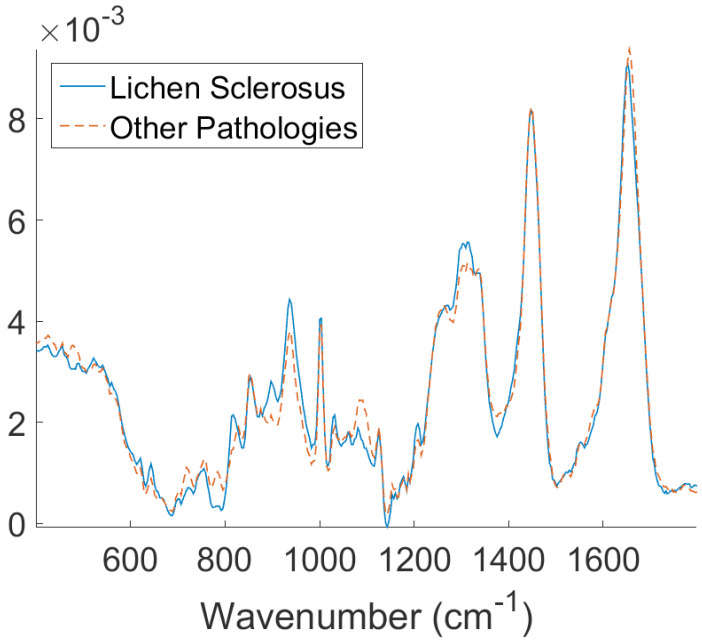


Figure 7-21 – Mean Raman spectra of pathology groups, lichen sclerosus and other pathologies (Table 7-3).

Analysis of diagnostic performance after cross validation demonstrated Raman spectroscopy combined with multivariate analysis was able to differentiate vulval samples from women with LS from those with other diagnoses with a sensitivity of 91% and specificity of 80%. The area under the receiver operator curve was 0.95 with a false positive rate of <0.05 (Figure 7-22).

Table 7-4 – Putative assignment of key Raman spectral bands seen in vulval skin.

Wavenumber (cm^{-1})	Suggested biomolecular constituent allocation
Type 1 Collagen	
815	C-O-C stretching
722	C-S stretching
856 and 920	Proline ring
875	Hydroxyproline ring
938	C-C stretching
1004 & 1033	Phenyalanine
1043	Proline
1082	Carbohydrate residues of collagen
1668	Amide I
1246	Proline rich region
1271	Proline poor region
1448	CH deformation
Nucleic acids	
722	DNA
785	DNA/ RNA
1325-1330	Nucleic acids
Carbohydrates	
477	Polysaccharides
Other protein and amino acids	
918	Proline, hydroxyproline
935	C-C stretching mode of proline and valine and protein backbone / glycogen
1000	Phenylalanine
1053	C-O stretching, C N stretching (protein)
1126	C-N stretching vibration (protein vibration)
1200-1300	Amide III (proteins)
1307	CH3/CH2 twisting, wagging, and/or bending mode of collagens and lipids
1528	Carotenoid
1554	Amide II

Wavenumber assignments from Talari et al. and Nguyen et al. (Nguyen *et al.*, 2013; Talari *et al.*, 2015)

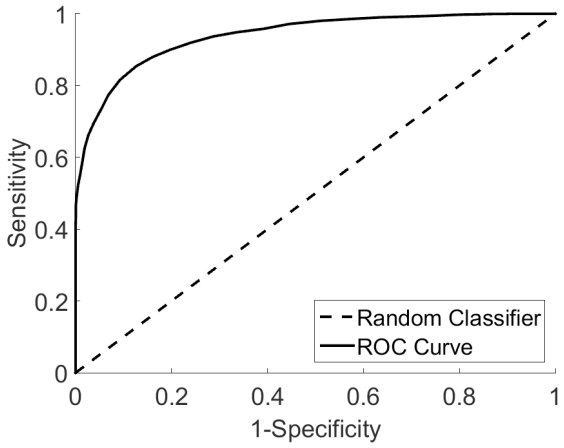


Figure 7-22 – Receiver operator curve for the detection of LS from other pathologies, area under curve 0.95, type 1 error rate <0.05

7.6.2 Comparison two – identification of premalignant and malignant disease of the vulva

7.6.2.1 Summary of measurements

In this comparison 91 tissue samples from the 64 participants previously identified for streamline Raman (chapter 6.4) mapping were measured using the standardised parameters (Figure 6-6). The histopathological classification of the samples measured are summarised in Table 7-5.

Table 7-5 – Samples measured for comparison two

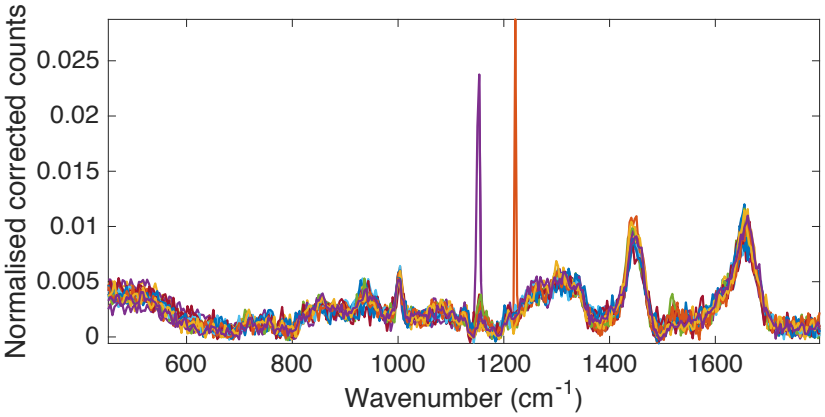
Histological Diagnoses	Number of samples measured	Total number of spectra measured
Non-neoplastic vulval tissue	30	1,200
Vulval intraepithelial neoplasia	30	1,200
Squamous cell carcinoma	31	1,240

40 Spectra were collected from each of the of the 91 tissue samples giving a total of 3,640 spectra.

7.6.2.2 Spectral processing and exclusion

All spectra underwent an instrument response correction, baseline correction and cosmic ray removal (chapters 6.5 and 7.5.7). Potential cosmic rays were identified and corrected in 22% of the 3,640 measured spectra. An example of a single sample's spectra before and after cosmic ray removal is given in Figure 7-23 demonstrating effective removal of easily identifiable cosmic rays. 35% of spectra were excluded by PC reconstruction error (Q statistic) and difference from the mean spectrum (D statistic) combined (chapter 4.2.6.1). This resulted in the complete exclusion of four whole samples from the analysis using the D and Q statistic methods.

A



B

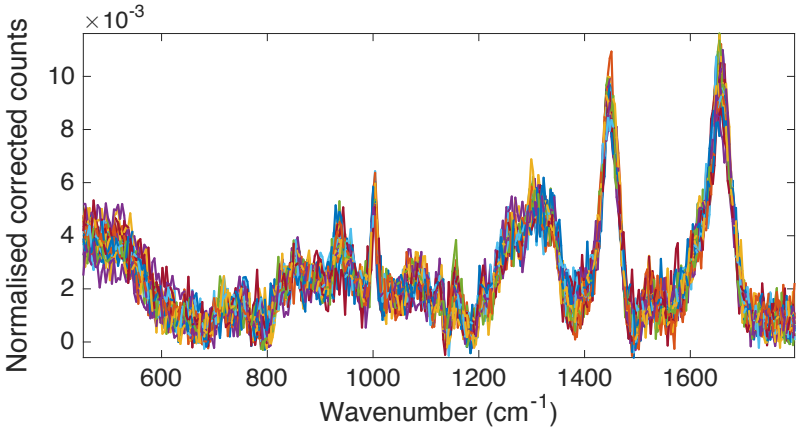


Figure 7-23 – Raman spectra from a tissue block before (A) and after (B) cosmic ray removal

7.6.2.3 Diagnostic classification

A three group analysis of the data was performed to assess the performance of the technique for differentiating between vulval SCC, VIN and non-neoplastic vulval disease as outlined in Table 13-1 as this mirrored the clinical needs identified.

Initial PCA demonstrated that 78% of the variance in the spectral data could be explained in the first four principal components (Figure 7-24). In the bootstrap cross

validation loops between 12 and 21 principal components were ultimately selected for use in the LDA classification model in each cycle of the loop (Figure 7-25)

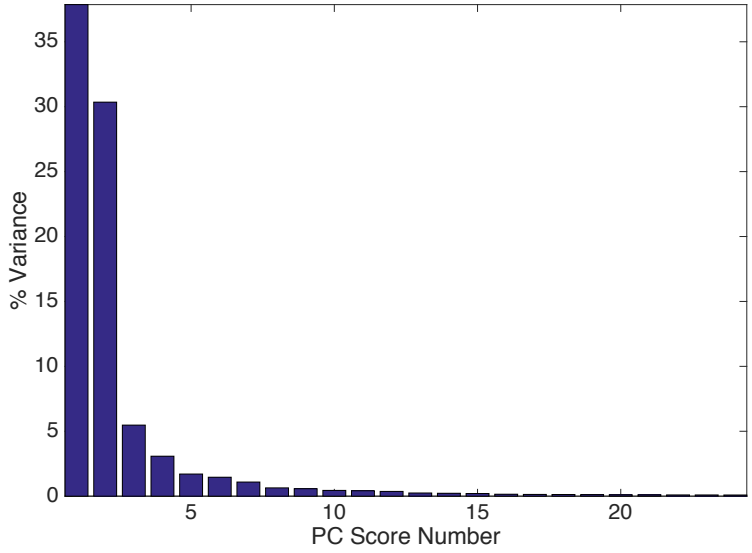


Figure 7-24 – Variance in data described by each principal component, 78% of variance described in the first four principal components.

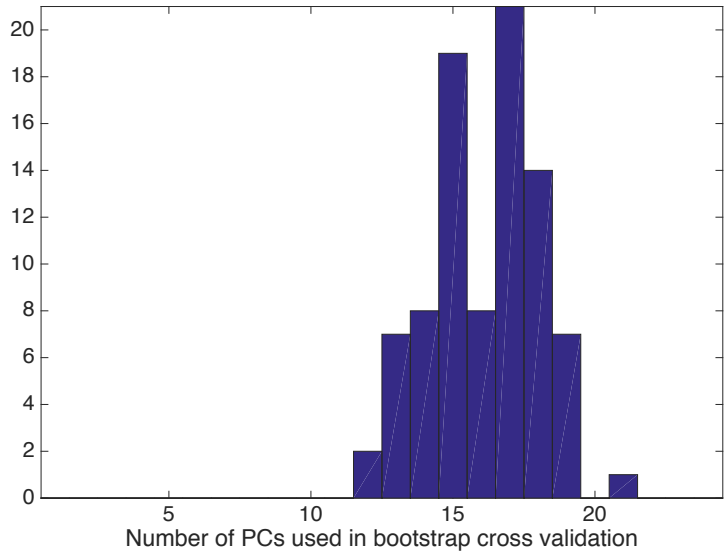


Figure 7-25 – Number of principal components used in each of the bootstrap cross validation loops

Figure 7-26 illustrates the distribution of the LD scores for each classified spectrum. As expected, compared to the LD scores from the spectra of the highly selected epidermal streamline mapping (Figure 6-20) there is significantly more overlap between the pathology groups. Greater overlap between the groups would be expected as the trans epidermal approach to spectrum acquisition and inability to select spectra from solely the area of interest means there is a great contribution to the spectra from less relevant tissue areas such as the stratum corneum and dermis.

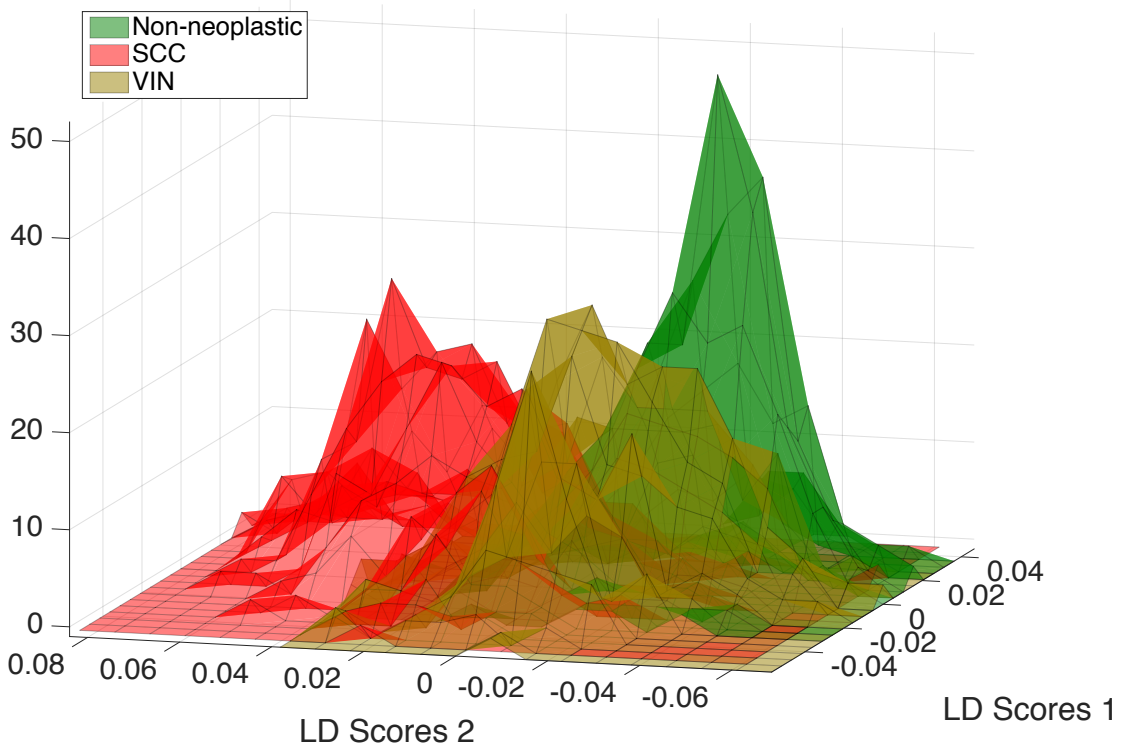


Figure 7-26 – Histogram of linear discriminant scores. Non-neoplastic vulval tisse, SCC – squamous cell carcinoma, VIN – vulval intraepithelial neoplasia

The principal components that are used to distinguish the pathology groups are not those that account for the greatest variability between the spectra (Figure 7-28) and

the importance of more subtle spectral difference in the classification may in part be responsible for the significant overlap between the groups.

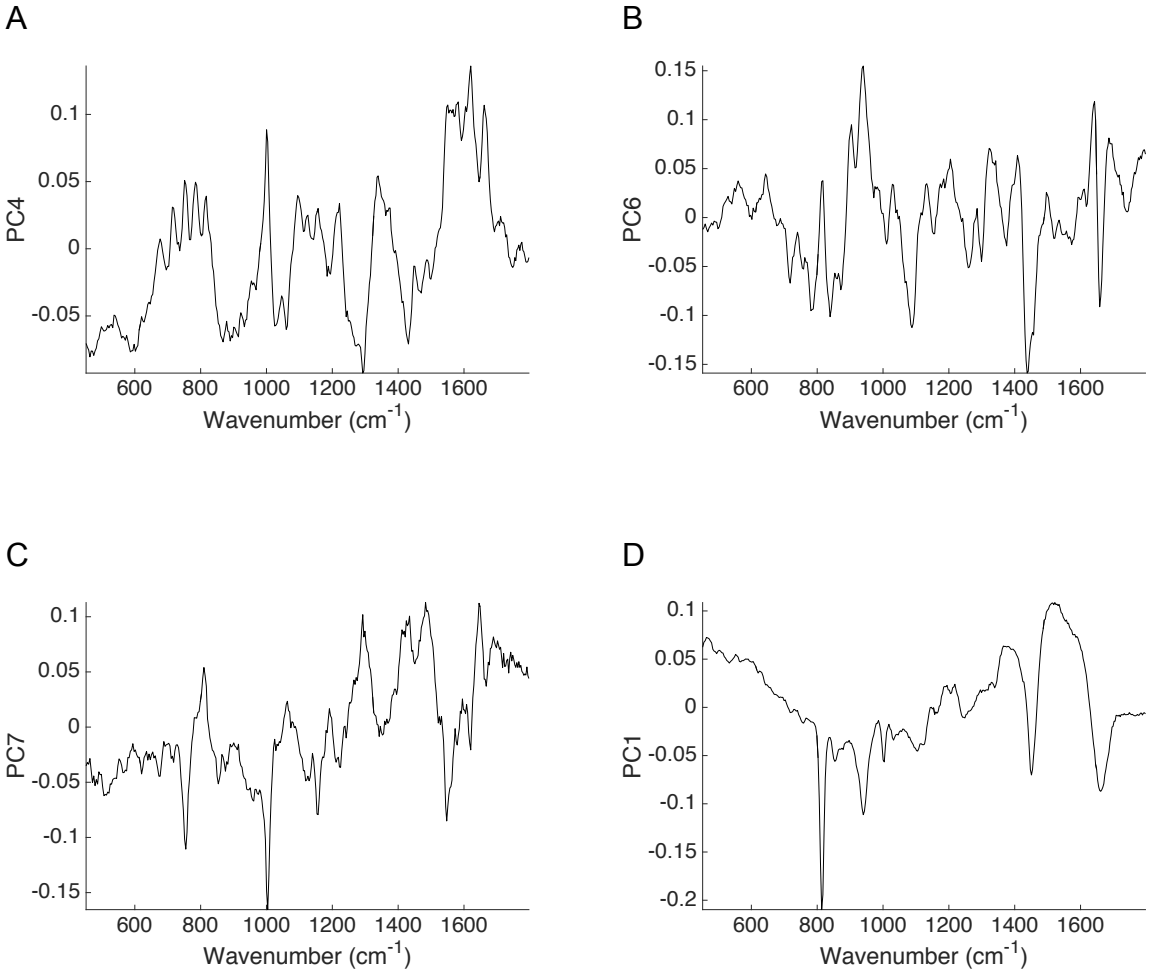


Figure 7-27 – Principal component loadings of four most significant scores as determined by ANOVA.

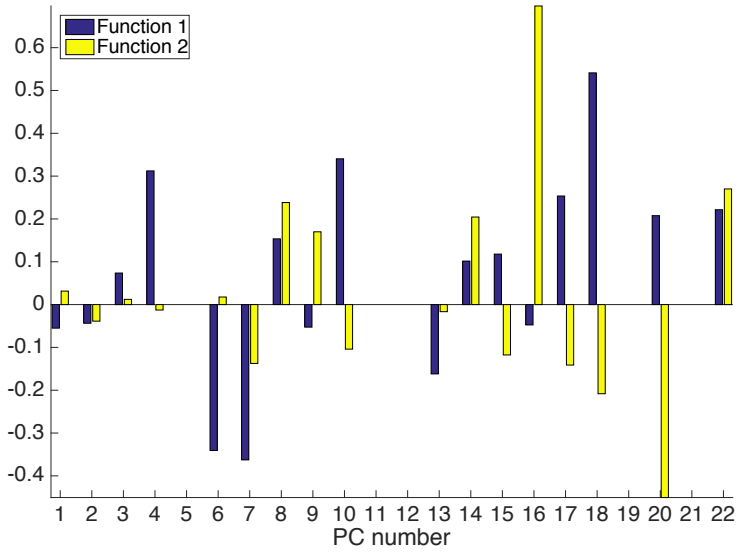


Figure 7-28 – Weight of each principal component score within the linear discriminant model created from non-cross validated analysis.

Table 7-6 shows the results of the cross-validated classification model when analysed by individual spectra. The classification performance is significantly poorer than the results obtained by streamline mapping tissue sections with overall sensitivity for the detection of VIN or SCC from non-neoplastic vulval skin of 75% with a specificity of 67% when the linear discriminant boundary lies midway between the pathology groups. When analysing by sample (Table 7-7) the specificity is improved but the sensitivity is approximately the same, giving 76% sensitivity and 74% specificity.

Table 7-6 – Confusion matrix for spectra based on modal classification of spectra with linear discriminant boundary midway between the pathology groups, rows are true pathology group and columns are predictions.

Pathology group	Non-neoplastic	Vulval intraepithelial neoplasia	Squamous cell carcinoma
Non-neoplastic	602	209	86
Vulval intraepithelial neoplasia	422	388	173
Squamous cell carcinoma	82	215	723
Sensitivity	67	39	71
Specificity	75	78	86

Table 7-7 – Confusion matrix for whole samples based on modal classification of spectra with linear discriminant boundary midway between the pathology groups, rows are true pathology group and columns are predictions.

Pathology group	Non-neoplastic	Vulval intraepithelial neoplasia	Squamous cell carcinoma
Non-neoplastic	22	5	2
Vulval intraepithelial neoplasia	13	10	5
Squamous cell carcinoma	2	7	21
Sensitivity	76	36	70
Specificity	74	80	88

Figure 7-29 shows a ternary plot of the predicted probabilities of a sample belonging to each of the three pathology groups. There appears to be separation between the neoplastic and non-neoplastic tissues. No samples of non-neoplastic tissue have a high predicted probability of being cancer with most samples having a predicted probability of belonging to the cancer group of less than 5% and all samples having a probability of belonging to the cancer group of 41% or less. Conversely no samples

of SCC had a high predicted probability of belonging to the non-neoplastic group. This indicates the technique has promise in identifying SCC from non-neoplastic vulval skin. The distinction between the VIN and cancer groups and the VIN and non-neoplastic groups is significantly less clear. Some samples of VIN have high predicted probability of being in the non-neoplastic or SCC group; some samples of non-neoplastic tissue have high predicted probability of being in the VIN group and some samples of SCC have high predicted probability of being in the VIN group. This is seen in the inverted clustering in the ternary plot (Figure 7-29). This would suggest that the technique is able to detect a continuum of physiological changes that lead to the development of cancer (chapter 1.4.1). This may account for the poor classification ability using standard linear discriminant boundaries in the three group model.

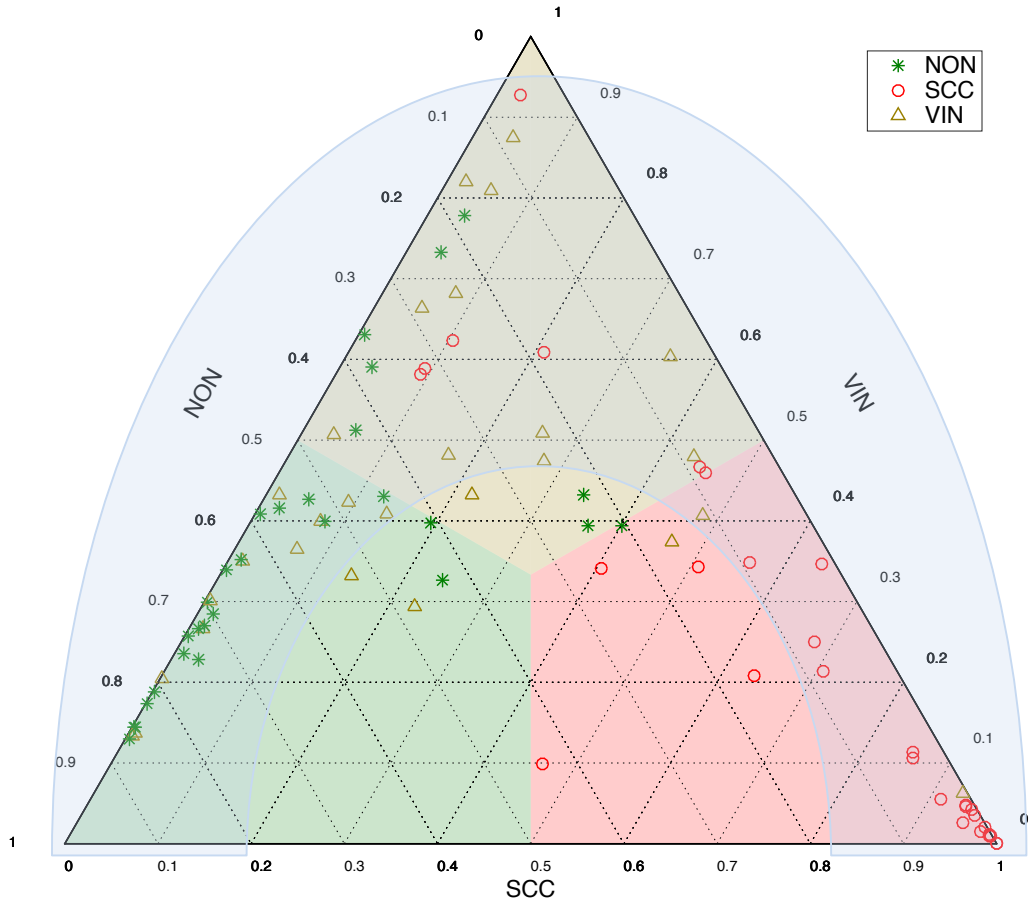


Figure 7-29 – Ternary plot of the predicted probabilities of a sample belonging to each of the three pathology groups under analysis. NON – non-neoplastic perilesional vulval skin, VIN – vulval intraepithelial neoplasia, SCC – squmous cell carcinoma. Legend indicates actual pathology of samples, axes are probability of group membership based on linear discriminant distance. The shaded area indicates the clustering of probabilities in a pattern that suggests the detection of biomolecular changes associated with carcinogenesis.

By altering the linear discriminant boundaries that define the predicted pathology within the model we can assess what diagnostic performance can be achieved with this approach. The receiver operator curve for the detection of non-neoplastic tissue from SCC and VIN is shown in Figure 7-30, this shows that by varying the linear discriminant boundary sensitivities between 80 to 90% can be achieved with corresponding specificities of 66 to 54% with an area under the curve of 0.85. For

distinguishing SCC from VIN and non-neoplastic tissue the performance is better and sensitivities of 85 to 90% can be achieved with specificities of 78 to 68% and an area under the curve of 0.90. In the identification of SCC from VIN sensitivities of 84 to 88% for specificities of 72 to 65% can be achieved with an area under the curve of 0.88 (Figure 7-32). Although the threshold of acceptability of diagnostic performance will of course vary by the application these results show promise for the technique in the applications previously described.

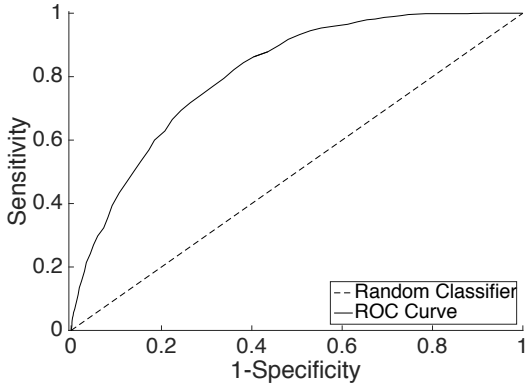


Figure 7-30 – Receiver operator curve for the detection of non-neoplastic tissue from SCC and VIN per spectrum. Area under curve, 0.85, Risk of falsely rejecting null hypothesis, <0.05

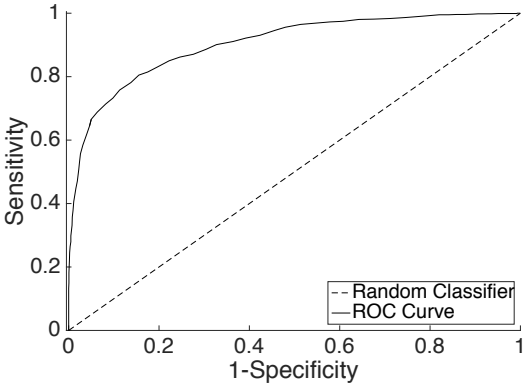


Figure 7-31 – Receiver operator curve for the detection of SCC from VIN and non-neoplastic vulval tissue per spectrum. Area under curve, 0.90, Risk of falsely rejecting null hypothesis, <0.05

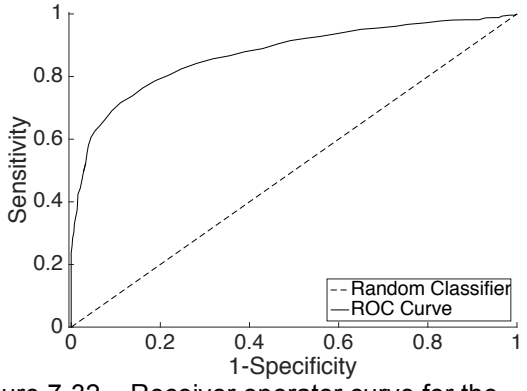


Figure 7-32 – Receiver operator curve for the detection of SCC from VIN per spectrum. Area under curve, 0.88, Risk of falsely rejecting null hypothesis, <0.05

7.6.2.4 Spectral analysis of biomolecular differences between pathologies

As previously explored (chapter 6.7) the identification of specific biomolecular changes in the spectra between the different groups is challenging. In the streamline model several spectral changes were seen in line with the published literature. In this model the previously reported changes seen with malignant transformation, including a right shift in the 1661 cm^{-1} Amide I band a widening of the peak at 1451 cm^{-1} and an increase in peak intensity at 1080 cm^{-1} are again seen (Gniadecka *et al.*, 1997; De Gelder *et al.*, 2007).

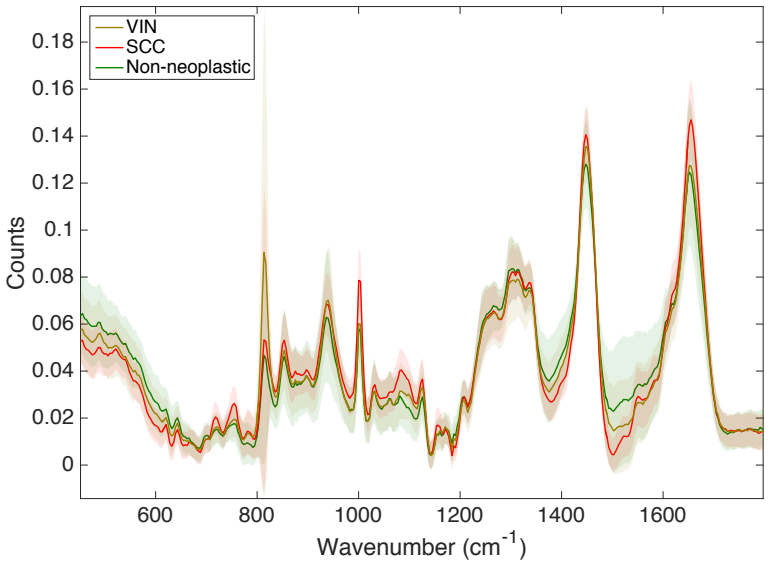


Figure 7-33 – Mean spectra of the pathology groups. Non-neoplastic, VIN – vulval intraepithelial neoplasia and SCC – squamous cell carcinoma. +/- one standard deviation marked by shading either side of line.

Similarly to the streamline model a number of other additional spectral differences can be identified when comparing neoplastic to non-neoplastic tissue. These include

increased prominence of the peaks at 723 cm^{-1}, 782 cm^{-1}, and 1101 cm^{-1} likely to represent DNA (720-722 cm^{-1}, 782 cm^{-1}, 1101 cm^{-1}), RNA (782 cm^{-1}) and nucleic acids (724 cm^{-1}, 830 cm^{-1}) and again this would be consistent with the increase rate of cell proliferation in neoplastic tissue (Talari *et al.*, 2015).

As expected there is some additional signal seen in the spectra from the wider sampling area from the trans epidermal approach. In Figure 7-34 the position of the main peaks of the type one collagen are shown on transepidermal Raman spectra (Appendix B). In these spectra we can see increased prominence of some of the key collagen peaks at 749 cm^{-1}, 814 cm^{-1}, 1245 cm^{-1}, 1266 cm^{-1} and 1419 cm^{-1}. This is consistent with a sampling volume that includes the dermis of the vulval tissue.

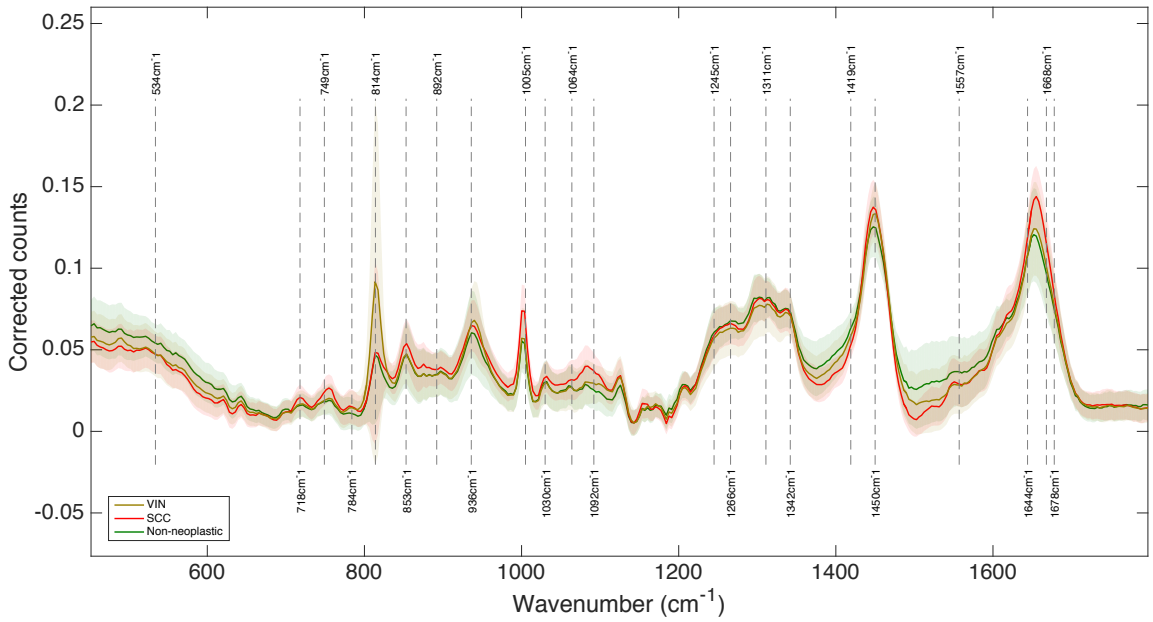


Figure 7-34 – Mean Raman spectra of pathology groups using a transepidermal approach showing the position of peaks relating to collagen type I.

The previously seen increased prominence in the peak at 830 cm^{-1} is not clearly seen in this data due to the presence of a new large peak at 814 cm^{-1} likely from the

collagen within the sampling volume obscuring the smaller peak seen at 830 cm^{-1}. In addition the previously seen reduction in the mono/disaccharide peak at 899 cm^{-1} is not clearly identifiable. The increased prominence in the amino acid and protein peaks at 756 $cm^{-1,}$ 855 cm^{-1} and 1558 cm^{-1} is again seen alongside the reduction in peak prominence at 1622 cm^{-1}. It is not clear what specific biomolecular changes are responsible for these spectral changes seen in association with malignant transformation however they appear consistently in spectra collected using the streamline and trans epidermal technique as demonstrated in Figure 7-35.

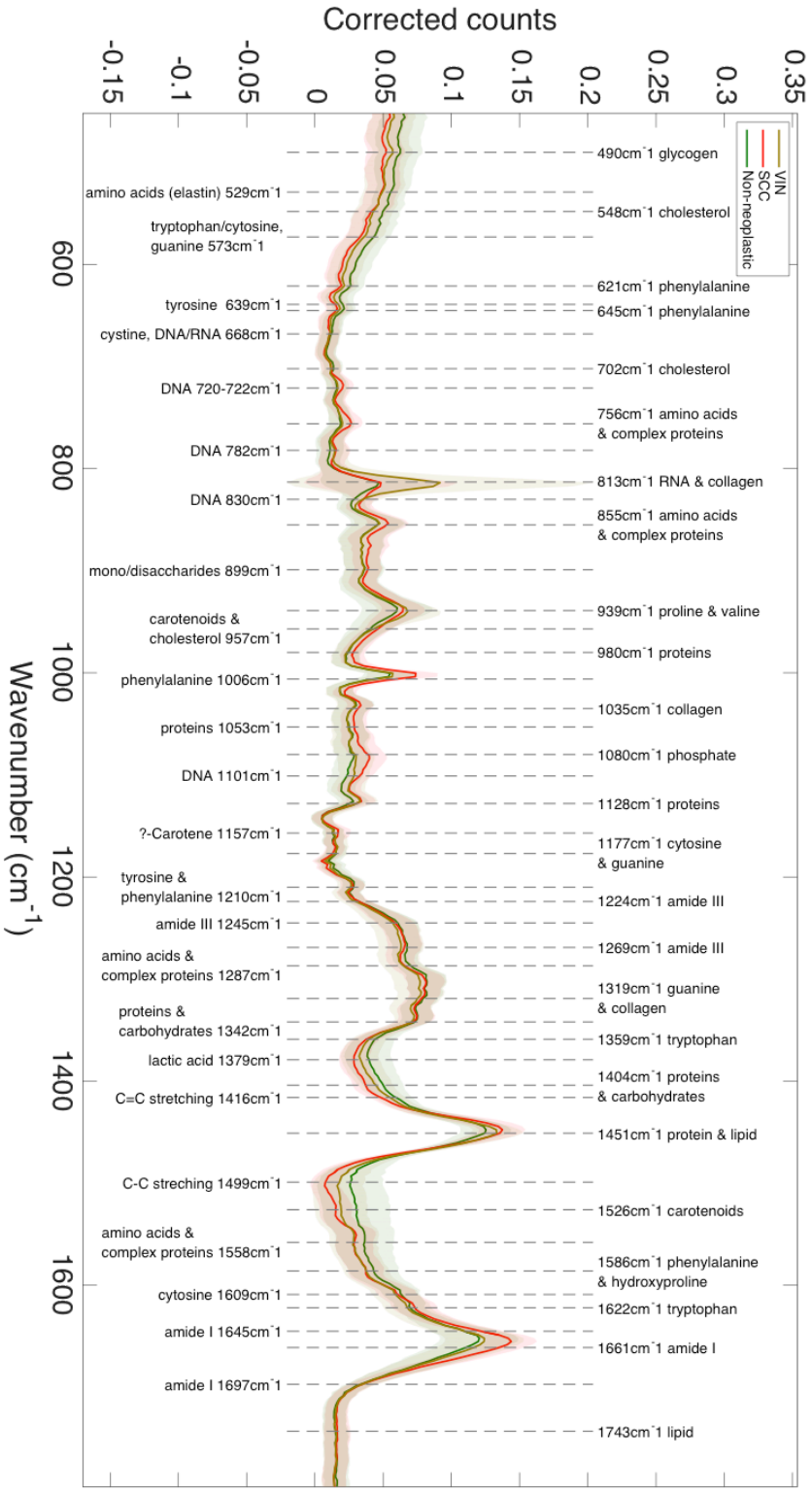


Figure 7-35 – Mean spectra of pathology groups with putative spectral peak assignments

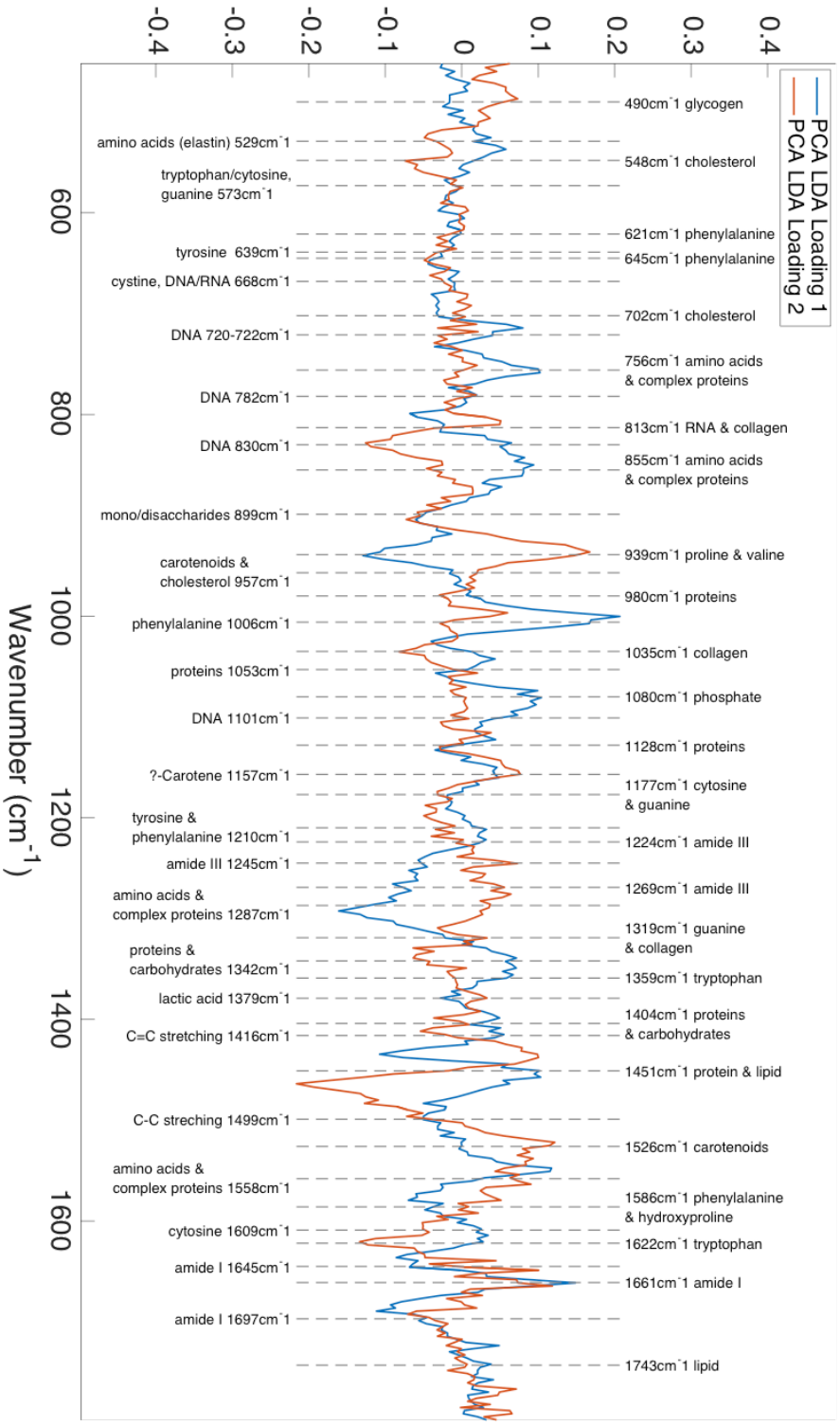


Figure 7-36 – Composite PCA LDA loading with putative peak assignments

These identified spectral differences are also reflected in the composite PCA LDA loading (Figure 7-36) and the changes identified appear to be significant in the predicted classification of the pathology groups as determined by deviation in the combined PCA LDA loading.

7.7 Interim discussion

[Summative discussion and conclusions relating to the investigations outlined in Section C relating to Raman spectroscopy for the assessment of vulval disease can be found in Chapter 13]

7.7.1 Comparison one - diagnosis of clinically inconclusive vulval lichen sclerosus

The aim of this investigation was to use Raman spectroscopy to differentiate lichen sclerosus from other vulval pathologies in the women where a diagnosis of lichen sclerosus is suspected but where biopsy is required to make the diagnosis. This study is the first reported application of this technique for the diagnosis of lichen sclerosus at any body site. Other studies using Raman spectroscopy to assess dermatological conditions have not included this important pathological group in their analyses.

The results demonstrate Raman spectroscopy combined with multivariate analysis has real potential as a useful tool in the diagnosis and monitoring of women with lichen sclerosus, a common vulval skin condition. In this study of patients who

required invasive tissue biopsy for diagnosis, Raman spectroscopy was able to correctly diagnose LS with clinically applicable performance.

Further research is required to validate this technique using probe-based systems (Chapter 8). The design of probes for in vivo trials should consider that laser focus beneath the surface of the skin will likely yield improved diagnostic performance, compared to probes that focus at the surface.

7.7.2 Comparison two – identification of premalignant and malignant disease of the vulva

In this comparison the aim was to evaluate Raman spectroscopy applied to whole blocks of fresh frozen tissue for the purposes of classifying non-neoplastic vulval skin, premalignant and malignant disease of the vulva. This is the first reported study to attempt classifying these pathologies using Raman spectra collected from whole vulval tissue blocks. In this investigation it was possible to use individual Raman spectra to classify the pathology present within the tissue with clinically relevant classification accuracy. This replicated the results of the experiment outlined in Chapter 6. Comparing the streamline spectroscopic mapping of epidermal cells and transepidermal methods a predictable decrease in the diagnostic accuracy was demonstrated as shown in Table 7-8. In this analysis the Raman spectra also appears to define a continuum of biomolecular changes from non-neoplastic tissue through VIN to frank SCC (Figure 7-29). This fits with our understanding of the carcinogenic processes leading to the development of vulval SCC (chapter 1.4.1.3).

Table 7-8 – Comparision of the diagnositc accuaracy of streamline spectroscopic mapping of epidermal cells and transepidermal Raman measurements

Diagnostic comparison	Comparator 1		Comparator 2	Area under ROC curve streamline mapping	Area under ROC curve Point measurement
1	SCC and VIN	from	non-neoplastic vulva	0.95	0.85
2	SCC	from	VIN	0.92	0.88
3	SCC	from	VIN and non-neoplastic vulva	0.95	0.90

Although the trans-epidermal approach used in this investigation yielded good classification performance it is necessary, within the bounds of ex vivo experimentation, to evaluate the technique in conditions that are a closer approximation to in vivo application. The bench based microscope spectrometer system used in this analysis is capable of producing high quality spectra but cannot be used in vivo as the objective cannot easily be applied to the vulva. In the following chapter the technique is evaluated under conditions that would be easier to apply in the clinical environment.

Further discussion of future work and the steps necessary to translate these finding into clinical practice is located in Chapter 13.

Chapter 8 Evaluation of Vulval Skin Using Fibreoptic Raman Spectroscopic Probe

8.1 Introduction

In the investigations outlined in Chapter 6 and Chapter 7 it was demonstrated that Raman spectroscopy coupled with multivariate analysis is capable of objectively identifying malignant from premalignant and non-neoplastic vulval disease with clinically acceptable diagnostic performance. Both of these investigations were performed using large microscope based laboratory spectrometers. This approach is not suitable for in-vivo use as it would be very technically challenging to collect Raman spectra in the clinic or theatre environment using these laboratory bench based systems (Kallaway *et al.*, 2013; Old *et al.*, 2014).

To apply the technique in-vivo it is necessary to have a portable spectroscopic system with excitation and collection optics that can be positioned flexibly and freely on the patients' vulva to improve access to tissues. To apply a fibre optic Raman probe to the classification of disease in the vulva it is necessary overcome a number of key challenges. The Raman signal is comparatively weak, and only 1 in 10^6 to 10^9 photons undergo inelastic scattering and fibre optic probes used in vivo need to

overcome the fundamental signal to noise challenge of Raman spectroscopy. Clinical environments are not currently optimised for Raman spectroscopy and environmental light may produce unwanted spectral noise (Kendall *et al.*, 2009). In addition the photoluminescence from the optics and delivery fibres of the probe itself can contribute to further undesirable spectral noise (Day *et al.*, 2009). The use of fibre optic probe based spectroscopic systems has been reported for the classification of skin disease but these studies have not demonstrated the use of probes for disease classification for premalignant or malignant vulval disease. The largest study to date to use a fibre optic probe was that of Lui et al. and is discussed in detail in chapter 3.2.3 (Lui *et al.*, 2012). In this study the authors use a fibre optic probe optimised for diagnosis rather than margin assessment. The probe illuminates and collects Raman signal from a 3.5 mm diameter spot on the surface of the skin. While this is appropriate for the diagnostic purposes it is not optimal to have such a large spot size when assessing surgical excision margins as more precise delineation of the margin is required if the size of the excision and therefore surgical morbidity is to be minimised. In addition these larger probes only allow the assessment of superficial surface surgical margins and do not permit the insertion of the probe through the skin to assess deep margins. Nijssen et al. evaluated a simplified Raman probe design with a single 200 micron excitation and collection fibre for the classification of BCC from perilesional skin taken from 17 patients (Nijssen *et al.*, 2014). This study demonstrated that the small diameter probe was able to correctly classify tumour with reported 100% accuracy and non-involved tissue with reported 99% accuracy. In the study Nijssen *et al.* coupled the probe to a

modified Renishaw 1000 spectrometer like that used in the study in Chapter 7. This bench based system is not portable and not suitable for clinical use. To demonstrate Raman spectroscopy can be applicable to the clinical purposes outline in Chapter 5 it is necessary to reproduce the classification accuracy demonstrated in Chapter 6 and Chapter 7 using a probe system capable of taking measurements in the clinic or theatre environment with the spatial and classification accuracy required to determine disease margins. Such a system requires narrow gauge fibre optic probes to be used alongside a portable spectrometer system. In this investigation the disease classification capability of a Raman spectroscopic probe applied directly to the surface of vulval epithelium was evaluated. This was accomplished by the analysis of blocks of vulval epithelium taken at the time of surgical excision as outlined in chapter 6.4.

8.2 Objectives

This investigation assesses objective 3 of the aims and objectives outlined in Chapter 5.

To evaluate the ability of a fibre optic probe and portable Raman spectrometer in classifying non-neoplastic vulval skin, premalignant and malignant disease of the vulva.

8.3 Overview of investigation

In this investigation the disease classification capability of Raman spectroscopy applied directly to the surface of blocks of vulval tissue was evaluated. This was accomplished by analysing blocks of fresh frozen vulval tissue collected at the time of surgical excision. These tissue blocks underwent Raman spectroscopy prior to routine histopathological analysis. An outline of the investigation is shown in Figure 8-1.

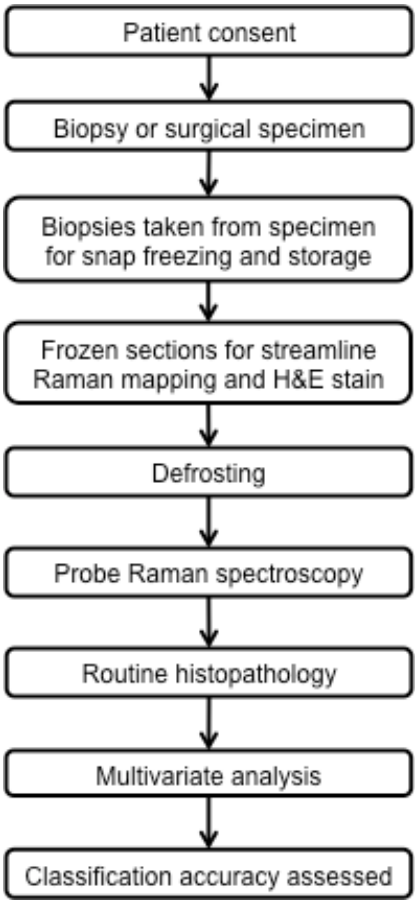


Figure 8-1 – Investigation flowchart for the evaluation of Raman spectroscopic probe

8.3.1 Participant identification, sample collection and sample preparation

Participants were identified and samples were collected from women undergoing vulval excisional surgery as described in chapter 6.4. Prior to analysis all samples were fully defrosted at 21^{o}C and washed in 0.9% physiological saline to remove any residues present on the surface of the epithelium.

8.4 Instrumentation and experimental methods

8.4.1 Spectrometer

A Headwall (Diegem, Belgium) Explorer Raman Spectrometer coupled to an Andor (Belfast, Northern Ireland) iDus CCD was used for this study. In compact spectrometers geometric distortions such as keystone can lead to spatial misregistration on the detector. The spectrometer uses an aberration-corrected retro-reflective design to reduce the effects of keystone and to provide a short focal length spectrometer with a compact and portable design (Figure 8-2). The iDus DU420A-BR-DD CCD used for this investigation has a back illuminated deep depletion CCD with anti-fringing air cooled to -80^{O}C with a near infrared enhanced wedged fused silica window.

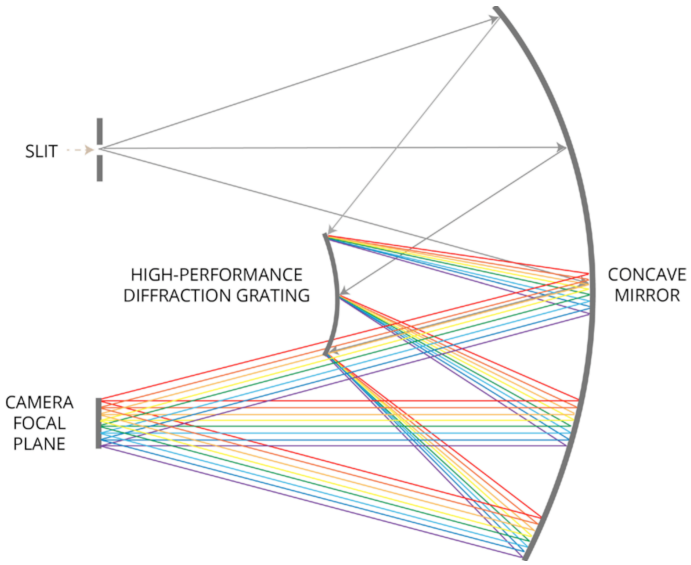


Figure 8-2 – Aberration-corrected retroreflective spectrometer (reproduced from Andor inc. technical specifications)

The back illumination of the CCD allows photons to directly strike the photosensitive part of the detector without having to first pass through the polysilicon electrodes (Figure 8-4). This configuration increases the quantum efficiency of the detector but does make the detector susceptible to constructive and destructive optical interference when illuminated with coherent photons. This results in optical etaloning that is especially problematic in the near infrared wavelengths used for this study, this will be described in more detail in later chapters.

The CCD was controlled using the Andor Solis for Spectroscopy software version 4.27 2014 (Belfast, Northern Ireland)

This spectrometer configuration results in a portable closed system suitable for use and transportation within the clinical environment (Figure 8-3).

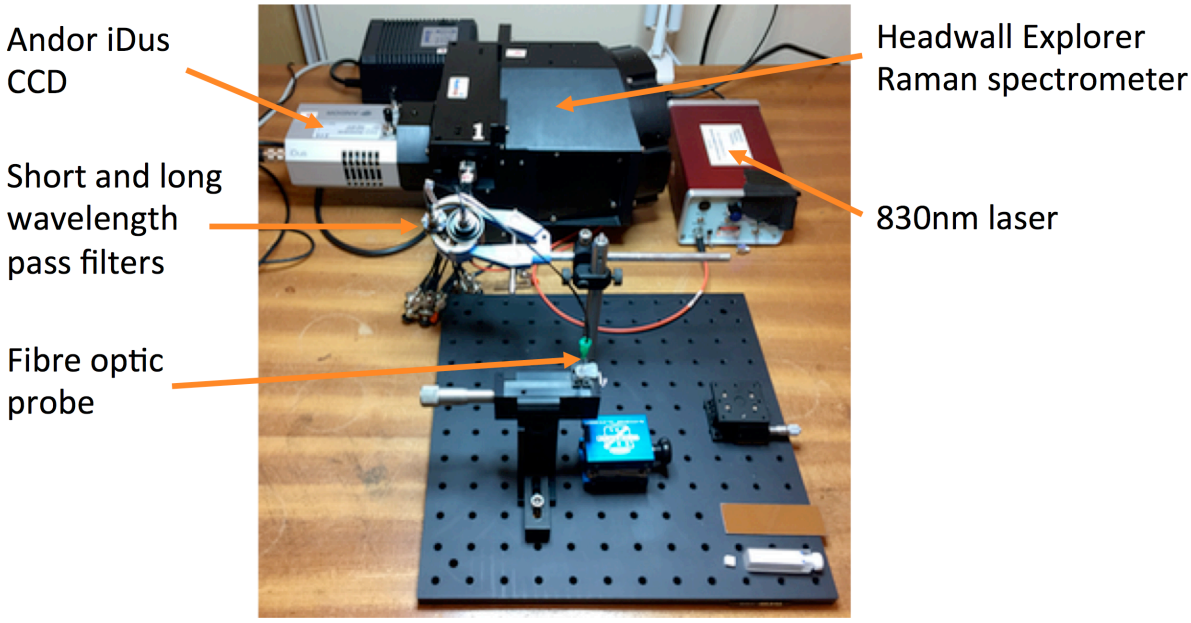


Figure 8-3 – Probe spectrometer system (Headwall Explorer Raman spectrometer coupled to an Andor iDus CCD and fiber optic probe)

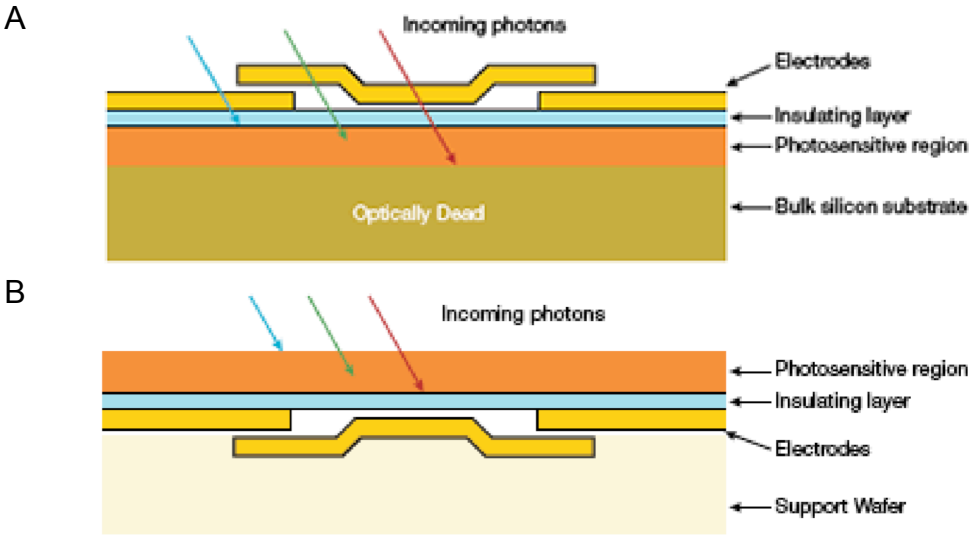


Figure 8-4 – A - Typical front illuminated CCD (cross section), B - Typical back illuminated CCD (cross section) (reproduced from Andor inc. technical specifications)

8.4.2 Probe design

In selecting an appropriate design of fibre optic probe for the purposes outlined in Chapter 5 the optical characteristics of the probe including the Raman scattering and photoluminescence created within the probe were considered alongside the potential clinical applications. In this study the objective was to assess the classification accuracy of the technique using measurements taken from the surface of the vulva skin with a spatial accuracy that allows the determination of disease margins at the time of surgery. In practice vulval cancer not only affects the surface of the skin but also invades deep into the subcutaneous tissues. The design of a Raman probe needs to consider the likely need to assess disease margins at depth through the vulval skin although evaluating this is outside the scope of this thesis. Considering the need for spatial precision and the potential future need to take invasive measurements a fine bore needle probe similar to those proposed by Day and Stone and Komachi et al. suits both these purposes (Komachi *et al.*, 2009; Day and Stone, 2013).

The probe evaluated comprised two 62.5 μm core 15 cm lengths of metal-coated, fibre with low concentrations of hydroxyl groups (Oxford Electronics, Oxford, UK). These two lengths of fibre were fixed adjacent to each other at the distal tip of the probe and polished at 90^{O} to the fibres. This assembly is narrow enough to fit within a 20 gauge needle as is suitable for the identification of disease margins and subcutaneous application. One of these fibres was connected to the spectrometer

and the other to an 830 nm stabilised semiconductor laser (Innovative Photonics Solutions, Monmouth Junction, NJ, USA).

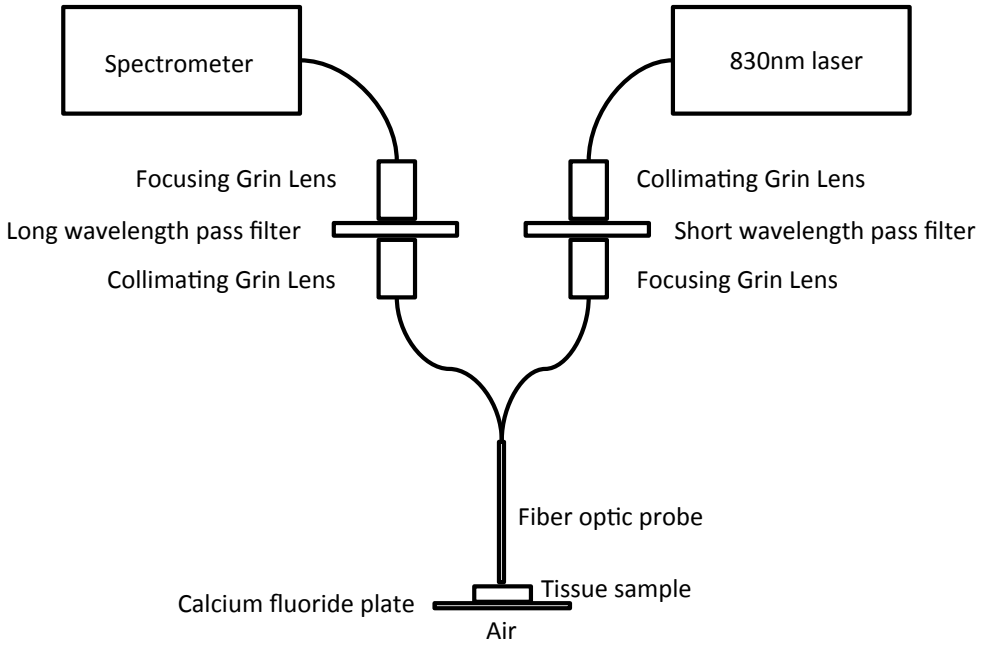


Figure 8-5 – Optical layout of probe system

To reduce the Raman scattering and photoluminescence in probe systems it is usual to incorporate interference filters as close to the signal source as possible. As it is not possible to incorporate these filters within the tip of such a narrow probe the filters were inserted close to the tip of the probe at 10 cm from the distal tip. At 10 cm from the tip of the probe the filters would not interfere with the operation of the probe in vivo. The excitation path consisted of a single 62.5 μm core fibre patch lead connected to a gradient-index (GRIN) lens, short wavelength pass filter and second, identical GRIN lens which was coupled to the excitation fibre of the probe. The collection path had the identical arrangement of GRIN lenses with a long wavelength pass filter (Figure 8-5).

8.4.3 Probe evaluation and tissue modelling

The laser power at the distal end of the excitation fibre sample was 60 mW. The signal background of the entire experimental setup was measured by recording a 20 second spectrum from the surface of a polished stainless steel slide. This demonstrated significant background signal in the bio-relevant region of the Raman spectrum that is especially pronounced at 436 cm^{-1} 490 cm^{-1}, 603 cm^{-1}, 803 cm^{-1}, 1049 cm^{-1}, 1184 cm^{-1} and 1605 cm^{-1} (Figure 8-6).

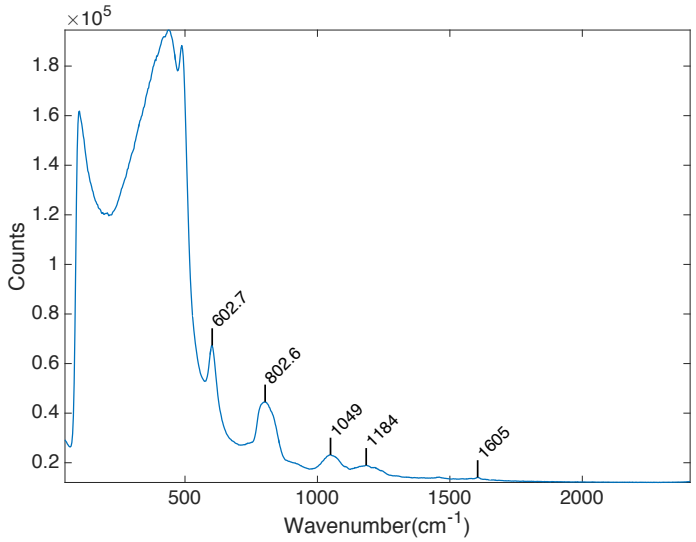


Figure 8-6 – 20s spectrum from polished stainless steel slide

The sampling volume of the probe in tissue was explored using a two-layer tissue phantom consisting of a 3 mm thick block of polytetrafluoroethylene (PTFE) fixed to the bottom of 5 ml glass beaker filled with a scattering and absorbing liquid placed on a micrometer stage (Figure 8-8). This arrangement allowed the Raman signal acquired from tissue layers at different depths in tissue to be simulated. PTFE was chosen to provide the test signal as it has a refractive index and scattering properties similar to those of biological tissues (Iping Petterson *et al.*, 2015). The PTFE also

provides a distinctive Raman signal in the biorelevant region of the spectrum and is therefore well suited to simulate tissue signal in phantom simulations (Iping Petterson *et al.*, 2015).

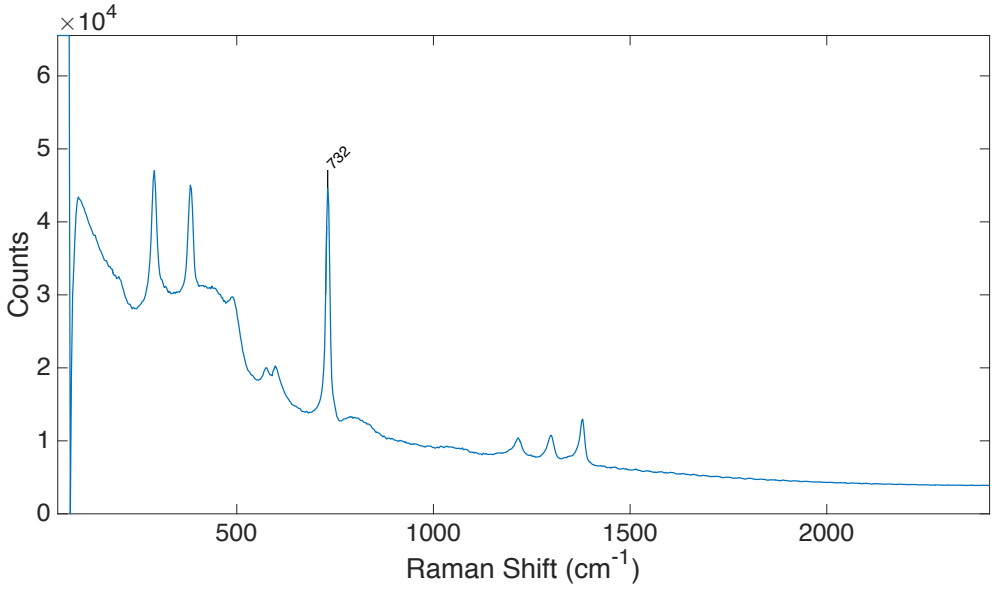


Figure 8-7 – 10 second PTFE Raman spectrum acquired using probe system

The second layer of the tissue phantom was a solution of Intralipid® (Sigma-Aldrich, 20 %), India ink (American MasterTech Scientific, Lodi, CA, USA) and water matched to the scattering and absorption properties of human skin. Intralipid was used at the scattering agent as it has well characterised optical properties in the near infrared spectral region (Pogue and Patterson, 2006; Ninni, Martelli and Zaccanti, 2011; Aernouts *et al.*, 2014). India ink is a well suited phantom as it has well characterised optical properties, a broadband absorption spectrum and is easily incorporated into Intralipid solution (Madsen, Patterson and Wilson, 1992; Pogue and Patterson, 2006; Di Ninni, Martelli and Zaccanti, 2010). The ink used came from a solution used by Iping Petterson et al. who had previously characterised the ink by

UV-vis absorption spectroscopy (Iping Petterson *et al.*, 2015). The absorption and scattering characteristics of human skin at 830 nm were calculated from reference values and methodology described by Jacques et al. (Jacques and Pogue, 2008; Jacques, 2009, 2013). For the purposes of this simulation the optical properties of the dermis and epidermis were considered to be the same within the phantom. Whilst it is technically feasible produce a phantom with multiple layers with different optical properties for simplicity a uniform liquid phantom was used. The scattering coefficient of epidermis was estimated to be 22.39 cm^{-1} and absorption coefficient estimated to be 1.715 cm^{-1}. It should be noted the absorption coefficient of dermis is estimated to be 0.249 cm^{-1} using S Jaques methodology with tissue properties from Jacques, (2013). Optical properties of biological tissues: a review. Physics in Medicine and Biology, 58(11), R37–61. The probe was fixed from above allowing the distances from the probe tip to the PTFE block to be adjusted using the micrometre stage.

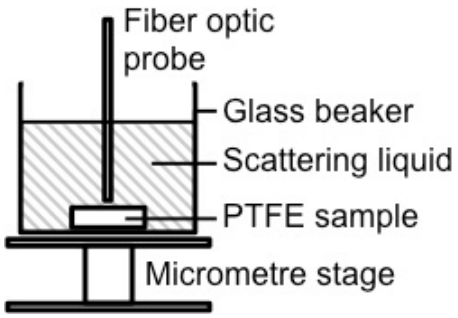


Figure 8-8 – Tissue phantom experimental setup

Ten second spectra were taken at 50 μm intervals between 0 and 1500 μm from the PTFE within the phantom liquid. The counts at the prominent 732 cm^{-1} peak (Figure 8-7) were plotted against depth and are shown in Figure 8-9. This demonstrated that in the tissue phantom model the signal from the PTFE was detected at depths of up to 1000 μm. This demonstrates that if the optical characteristics of the vulval tissue are similar to that of the phantom model then the probe should be able to acquire Raman signal from the full thickness of the epidermis in almost all tissue samples. The probe should therefore be able to detect the tissue area of interest within the samples.

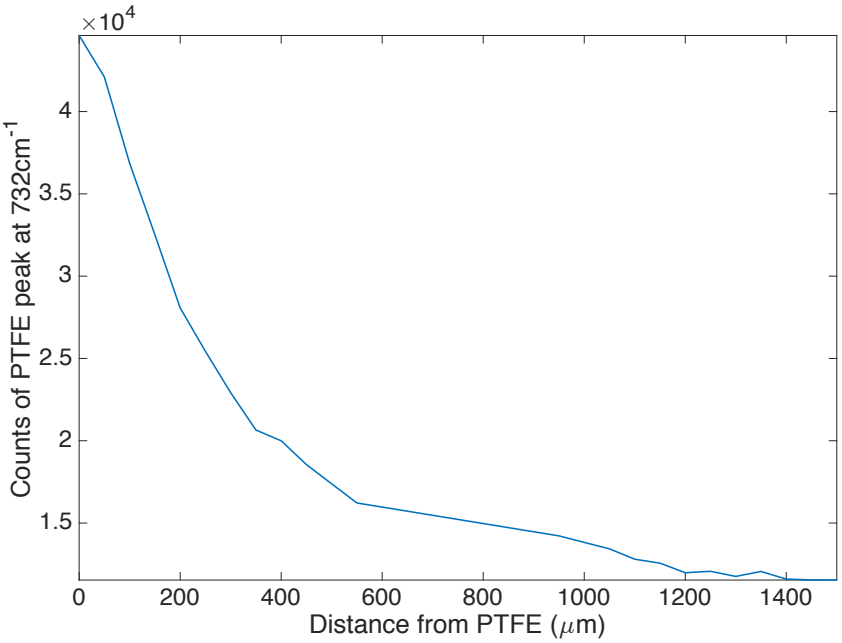


Figure 8-9 – Counts at 732 cm^{-1} peak from 10s acquisition with distance from PTFE in intralipid and India ink phantom.

The acquisition time for the probe measurements was assessed by evaluating the performance of the probe on a sample of chicken skin used as a biological model for

human skin. The signal to noise ratio was assessed in a range of acquisition times between one and twenty-five seconds. Above twenty-six seconds, detector saturation occurred across the wavenumber range rendering longer acquisition times unsuitable. The spectra measured were normalised and the signal was determined from the CH_2 bending peak at 1440 cm^{-1} and the noise from the signal poor area of the spectrum between 2309 cm^{-1} and 2369 cm^{-1} using the methodology described in chapter 4.1. The results of this evaluation are shown in Figure 8-1 where a rapid increase in the signal to noise ratio is seen that plateaus between acquisition times of 17 and 23 seconds. Based on this and the need to avoid detector saturation an acquisition time of 20 seconds was selected for the measurements of the vulval tissue.

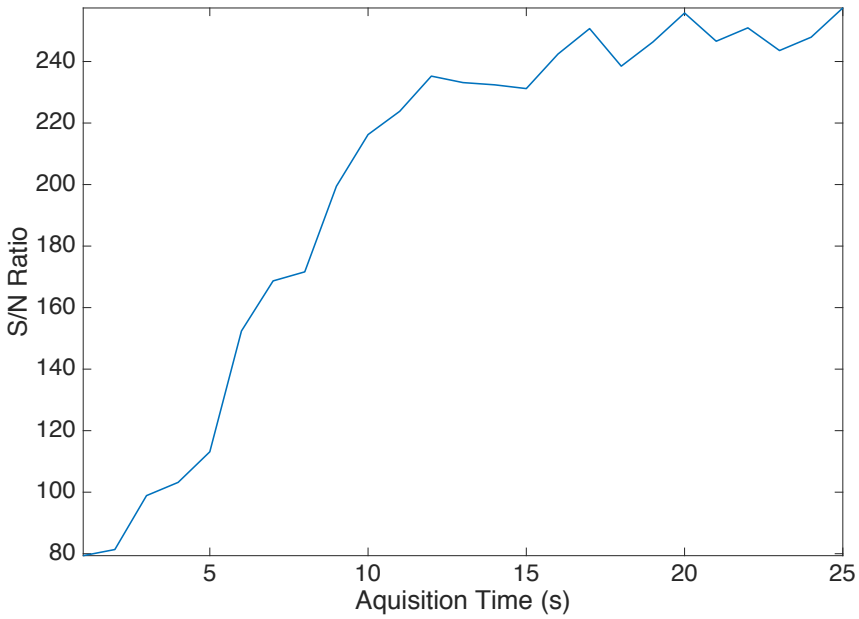


Figure 8-10 – Signal to noise ratio with increasing acquisition time for speactra measured from chicken skin using the experimental probe system

8.4.1 Calibration

Instrument calibration was performed on each use of the spectrometer by measuring a 0.2 s spectrum from crushed aspirin tablet in a quartz cuvette. This gave a reliable standard to calibrate the Raman shift measured by the instrument using set peaks within the aspirin Raman spectrum (Figure 8-11) calibrated against the spectrometer used in Chapter 7. The Raman shift was then corrected to the CCD pixel position by the Solis software using third order polynomial. An instrument response correction was performed using a spectrum measured from a sample of green glass during calibration as previously described (chapter 6.5.1).

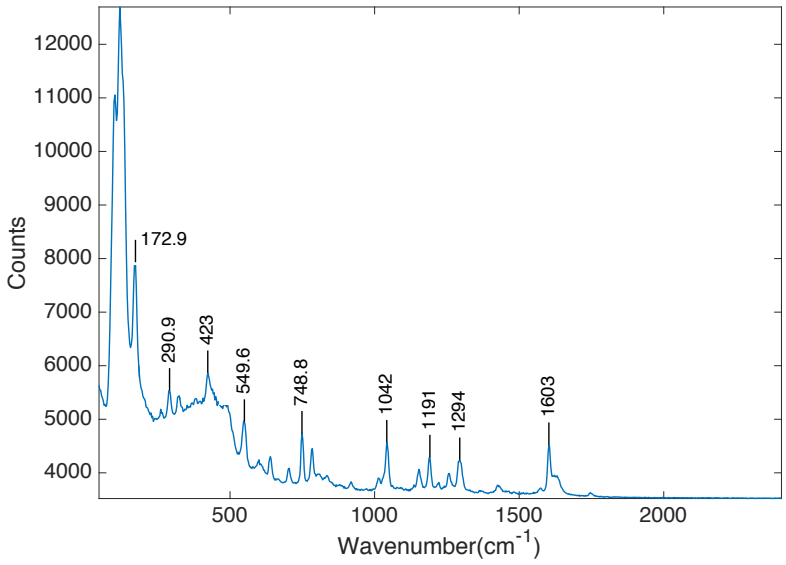


Figure 8-11 – 0.2 s Aspirin spectrum demonstrating calibration peaks

8.4.2 Sample measurement

For the purposes of measuring tissue samples the probe was mounted on an optical bench within a hypodermic needle to allow control of the angle of the probe tip

against the surface of the tissue samples (Figure 8-12). An adjustable stage was placed beneath the probe to allow the specimen to be advanced toward the probe until contact was made between the tissue and the probe. Tissue samples were placed on a calcium fluoride window to prevent unwanted interference with the bio-relevant area of the Raman spectrum from surrounding material. When positioning the samples, the stage with the calcium fluoride slide and specimen resting on top was advanced toward the probe until contact was made and the probe displaced ensuring good contact between the specimen and probe. During measurements the probe tip was allowed to rest on the surface of the tissue with no additional pressure other than the mass of the probe to maintain contact between the tissue and the probe. Raman spectra were measured at ten positions on each of the vulval samples. The spectral acquisition time was set at 20 s and to avoid the need for cosmic ray removal in the analysis the median of three measurements was taken at each point on the vulval sample

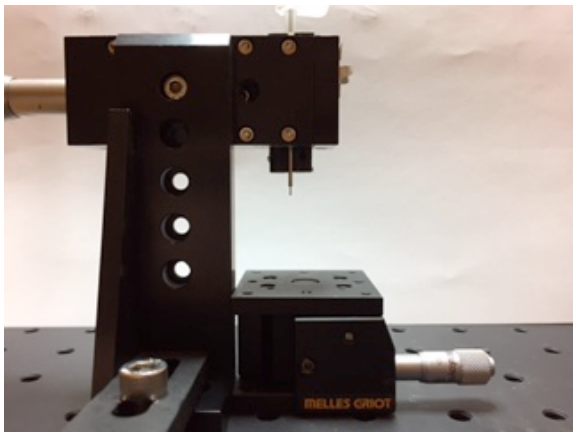


Figure 8-12 – Fibre probe mounting system

8.4.3 Data preprocessing and analysis

After spectral acquisition the data were optimised to reduce the contribution of non-tissue signal within the spectra. The optical etaloning from the back illuminated CCD was found to be especially apparent in the tissue samples measured due to the relatively low signal obtained from biological tissues. This etaloning alongside the background signal from the apparatus and fluorescence was addressed through a combined correction applied to all the measured spectra.

In order to identify the contribution of the etalon to the data a 3rd order polynomial fitted to the green glass spectrum was subtracted from the green glass spectrum (Figure 8-13). This gave and approximation of the etalon spectrum present within the tissue spectra (Figure 8-14).

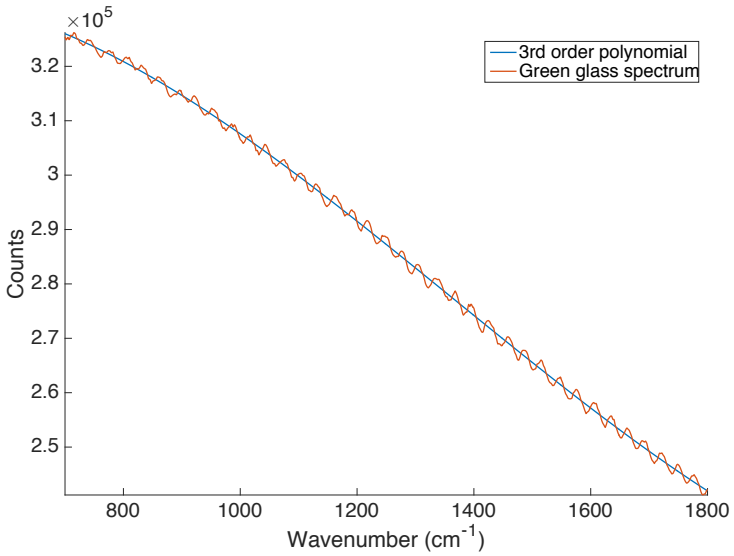


Figure 8-13 – Spectrum measured from a sample of green glass and 3rd order polynomial fitted to the spectrum.

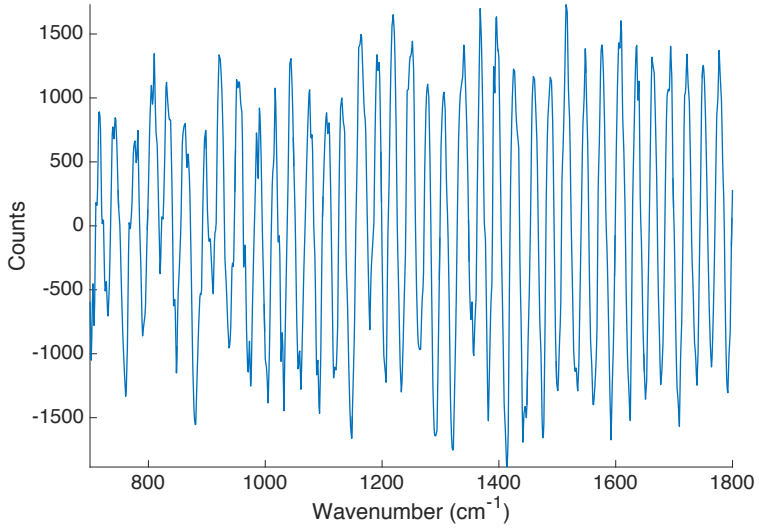


Figure 8-14 – Etalon spectrum after correcting for instrument response.

The background spectrum as measured from the stainless steel slide and the etalon approximation was considered to account for the main sources of non-tissue signal within the raw data. In order to identify the tissue signal from the background, etalon and fluorescence an iterative least squares fit was performed using a modification of the technique described by Lieber et al. (Lieber and Mahadevan-Jansen, 2003). The

background, etalon and a third order polynomial fitted to the spectral data underwent the modified iterative least squares fit to the measured spectral data. The resultant fitted spectrum was then taken as an approximation of the contribution of the background, etalon and fluorescence to the measured data and subtracted from the data to isolate the underlying tissue signal. An example of a corrected and uncorrected spectrum is given in Figure 8-15. Correction removes the variation in the data caused by the background and etalon and allows clearer visualisation of tissue peaks, for example the amide I band at 1661 cm^{-1}, the protein and lipid band at 1451 cm^{-1} and the amide II protein band at 1273 cm^{-1}.

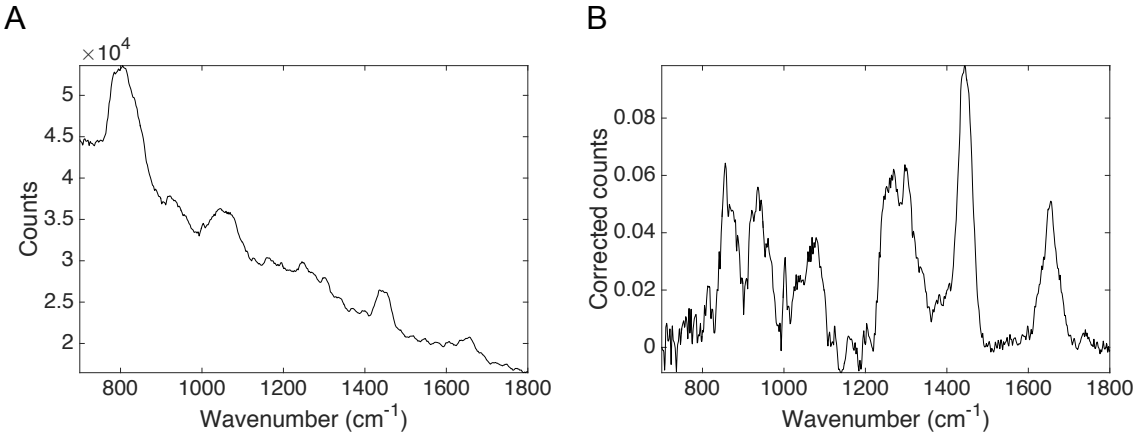


Figure 8-15 – Example uncorrected probe spectrum (A) and the same spectrum after baseline, instrument response and etalon correction (B)

After this combined correction the spectral data was filtered for outlying spectra using the D and Q statistic technique described in chapter 4.2.6.1. The schema for the preprocessing steps applied to the raw spectral data is outlined in Figure 8-16.

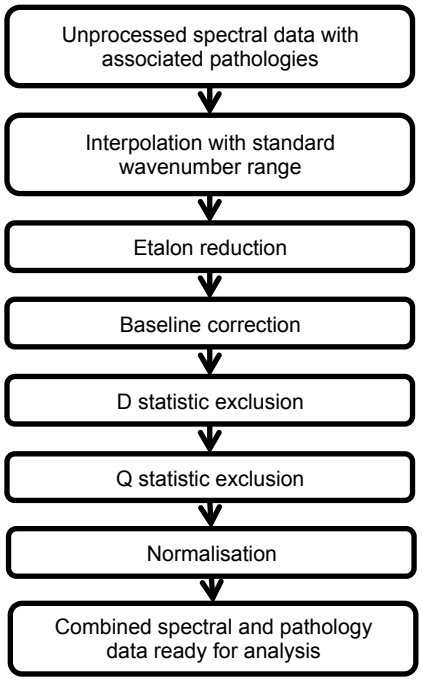


Figure 8-16 – Preprocessing algorithm for probe spectral data

After pre-processing the classification accuracy achievable was assessed using the boot strap cross validation PCA LDA method outlined in chapter 7.5.7, Figure 7-16 and Figure 7-17.

8.5 Results

8.5.1 Summary of measurements

In this comparison 910 spectra were measured from 91 tissue samples from the 64 participants previously identified for transepidermal Raman analysis. The histopathological classification of the measured samples is summarised in Table 8-1. In nine instances (1% of spectra measured) individual spectral measurements needed to be repeated to obtain ten spectra per sample. The most common reasons

for repeat measurements were software failure and accidental movement of the apparatus.

Table 8-1 – Samples measured for probe investigation

Histological Diagnoses	Number of samples measured	Total number of spectra measured
Non neoplastic vulval epithelium	30	300
Vulval Intraepithelial Neoplasia	30	300
Squamous cell carcinoma	31	310

8.5.1 Spectral processing and exclusions

In the cross validation classification analysis a mean of 34.18% (± 42.02, 95% CI) of spectra were excluded by the DQ statistic from the training set and 33.38% (± 66.01, 95% CI) of spectra were excluded from the test sets in each loop. This resulted in the complete exclusion of 5 whole samples from the test analysis using the D and Q statistic methods. In addition, detector saturation occurred in a small number of spectra below 700 cm^{-1} and for this reason the spectra were truncated for analysis and the spectral data below 700 cm^{-1} was excluded from the analysis.

8.5.2 Diagnostic classification

A three group analysis of the data was performed to assess the performance of the technique for differentiating between vulval SCC, VIN and non neoplastic vulval

disease as outlined in Table 13-1 as this mirrored the clinical needs identified and the aim of the investigation.

Initial PCA demonstrated that 77% of the variance in the spectral data could be accounted for by the first four principal components (Figure 8-17). In the boot strap cross validation analysis (chapter 7.5.7) the minimum number of principal components used was 4 and the maximum was 17 (Figure 8-18).

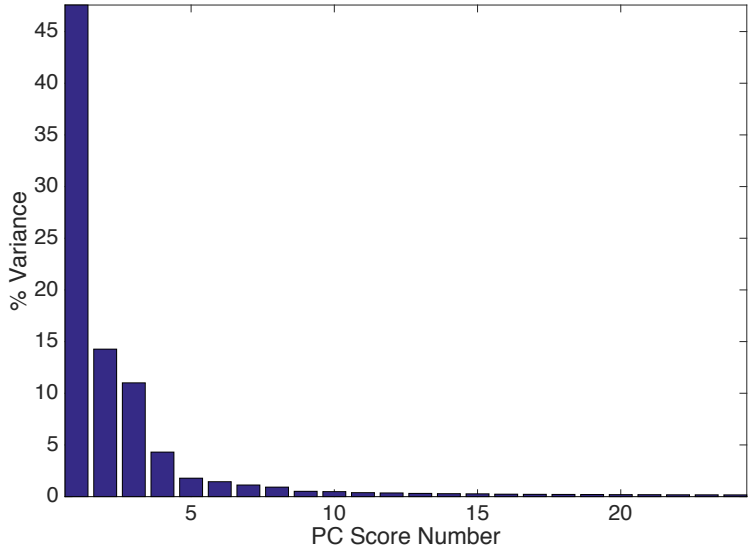


Figure 8-17 – Variance in data described by each principal component, 77% of variance described in the first four principal components.

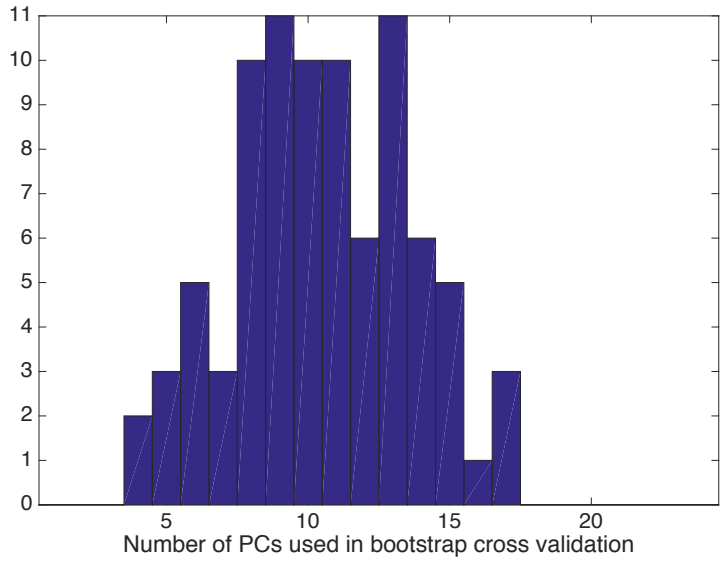


Figure 8-18 – Number of principal components used in each of the bootstrap crossvalidation loops

The distribution of the LD scores for each of the classified spectra is illustrated in the LD scores plot in Figure 8-19. There does appear to be separation of the three groups but as expected this is less pronounced than with the investigations in Chapter 6 and Chapter 7 given the lower signal to noise in the spectra measured from the Raman probe and the inevitable contribution of signal from less relevant tissue areas such as the epidermis and dermis.

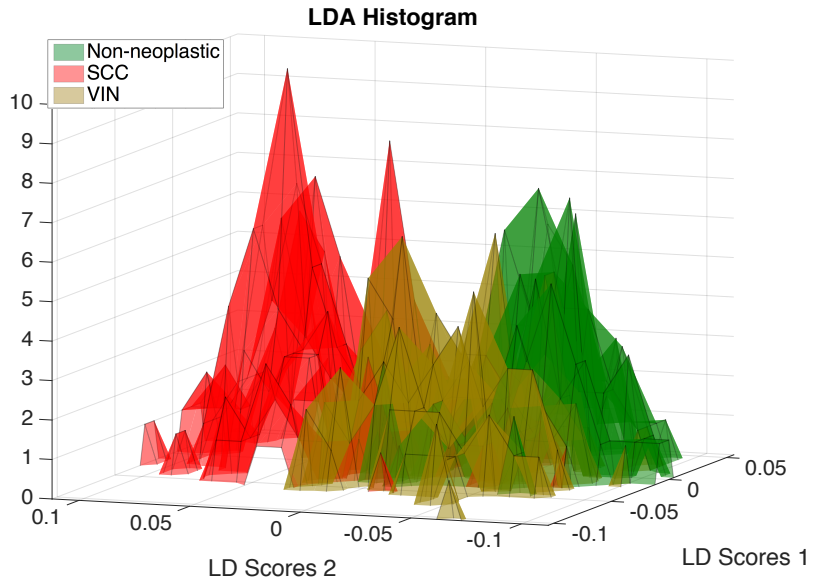


Figure 8-19 – Histogram of linear discriminant scores, non-neoplastic vulval tissue, SCC – squamous cell carcinoma and VIN – vulval intraepithelial neoplasia

The four principal components that accounted for the greatest variability between the pathology groups are shown in Figure 8-20. These components do not account for the greatest variability in the spectral data but they are all in the top six components by percentage variance suggesting that a significant proportion of the variation in the spectra is due to the differences between pathology groups.

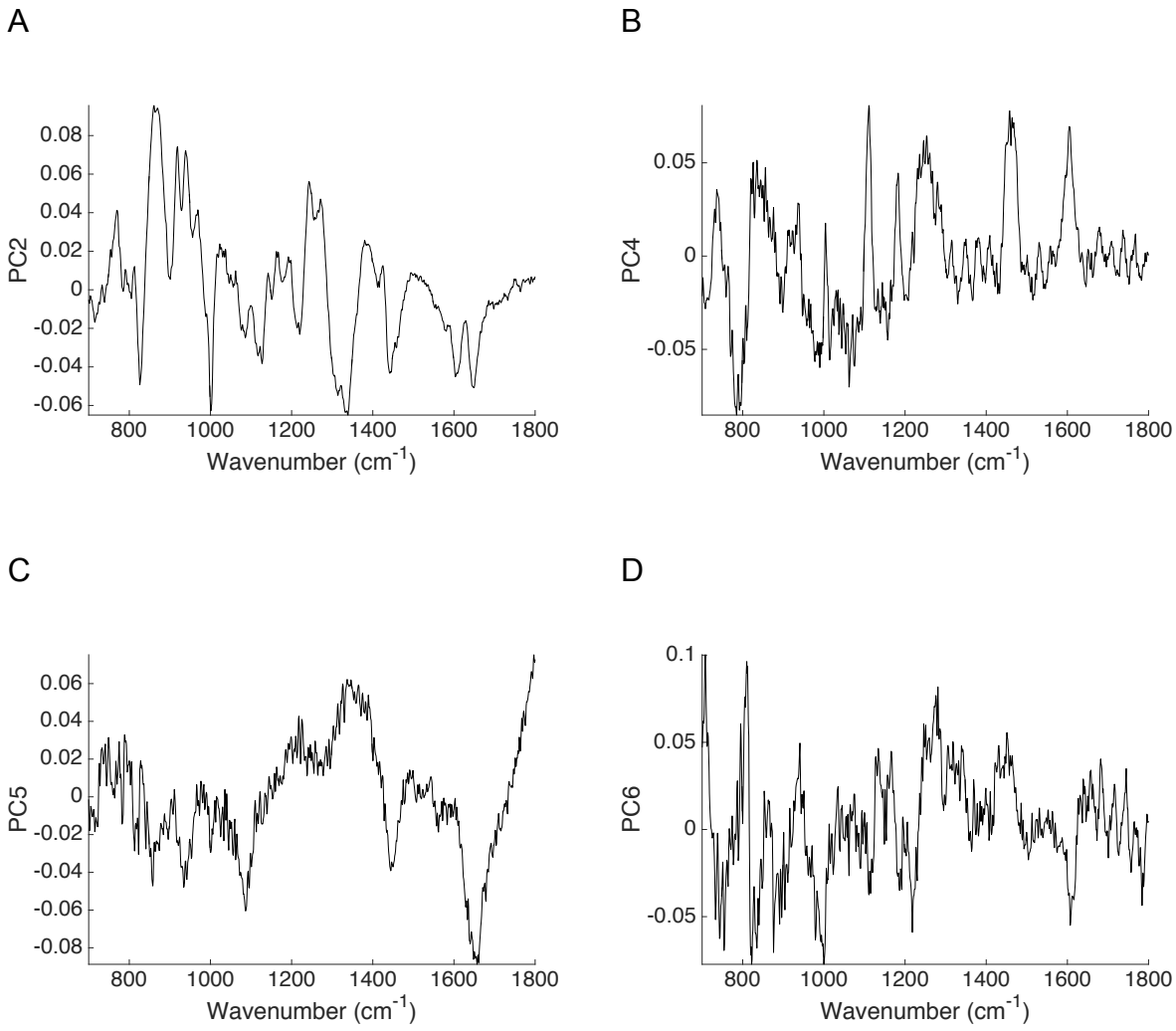


Figure 8-20 – Principal component loadings of four most significant scores as determined by ANOVA.

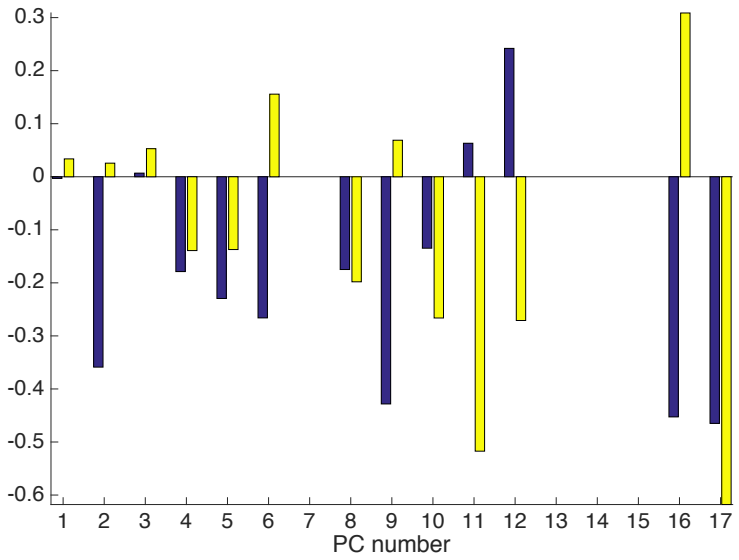


Figure 8-21 – Weight of the contribution of each principal component score to the linear discriminant model. In a three group model two linear discriminant functions are required to separate the groups and these are represented by function1 (blue) and function 2 (yellow).

Table 8-2 and Table 8-3 show the confusion matrices for the cross-validated classification model when analysed by individual spectra and as whole samples respectively. The classifications in these tables are a representation of the classification performance when the linear discriminant boundary is midway between the pathology groups. Given the relatively small number of spectra taken from each sample the results by spectra and by group are quite similar. The classification performance is similar to that obtained in the previous investigation with an overall sensitivity for the detection of VIN or SCC from non-neoplastic vulval skin of 81% with a specificity of 57% when the linear discriminant boundary is midway between the pathology groups. When analysing by sample the specificity is improved but the sensitivity is similar (77% sensitivity and 66% specificity).

Table 8-2 – Confusion matrix for spectra based on classification with linear discriminant boundary midway between the pathology groups, rows are true pathology group and columns are predictions.

Pathology group	Non-neoplastic	Vulval intraepithelial neoplasia	Squamous cell carcinoma
Non-neoplastic	99	64	11
Vulval intraepithelial neoplasia	56	59	46
Squamous cell carcinoma	19	15	190
Sensitivity	57	37	85
Specificity	81	80	83

Table 8-3 – Confusion matrix for samples based on modal classification of spectra with linear discriminant boundary midway between the pathology groups, rows are true pathology group and columns are predictions.

Pathology group	Non-neoplastic	Vulval intraepithelial neoplasia	Squamous cell carcinoma
Non-neoplastic	15	9	1
Vulval intraepithelial neoplasia	11	9	10
Squamous cell carcinoma	3	2	26
Sensitivity	60	30	84
Specificity	77	80	80

The ternary plot in Figure 8-22 shows the predicted probabilities of a sample belonging to each of the three pathology groups with the sample's actual pathology. There appears to be continuum between the neoplastic and non-neoplastic tissues in keeping with the pathogenesis of SCC rather than clearly separable groups by probability of group membership. No samples of SCC tissue have a high predicted

probability of being in the non-neoplastic group with all samples having a predicted probability of belonging this group of less than 50% and all samples having a probability of belonging to the cancer group of 41% or less. Likewise no samples of non-neoplastic tissue have a greater than 55% probability of belonging to the SCC group. This corroborates the findings in the previous chapter that the technique has significant promise in identifying SCC from non-neoplastic vulval skin. The distinction between the VIN and SCC groups and the VIN and non-neoplastic groups is less clear. Some samples of VIN were found to have a high predicted probability of being in the non-neoplastic or SCC group and some samples of non-neoplastic tissue were found to have a high predicted probability of being in the VIN group with some samples of SCC have high predicted probability of being in the VIN group.

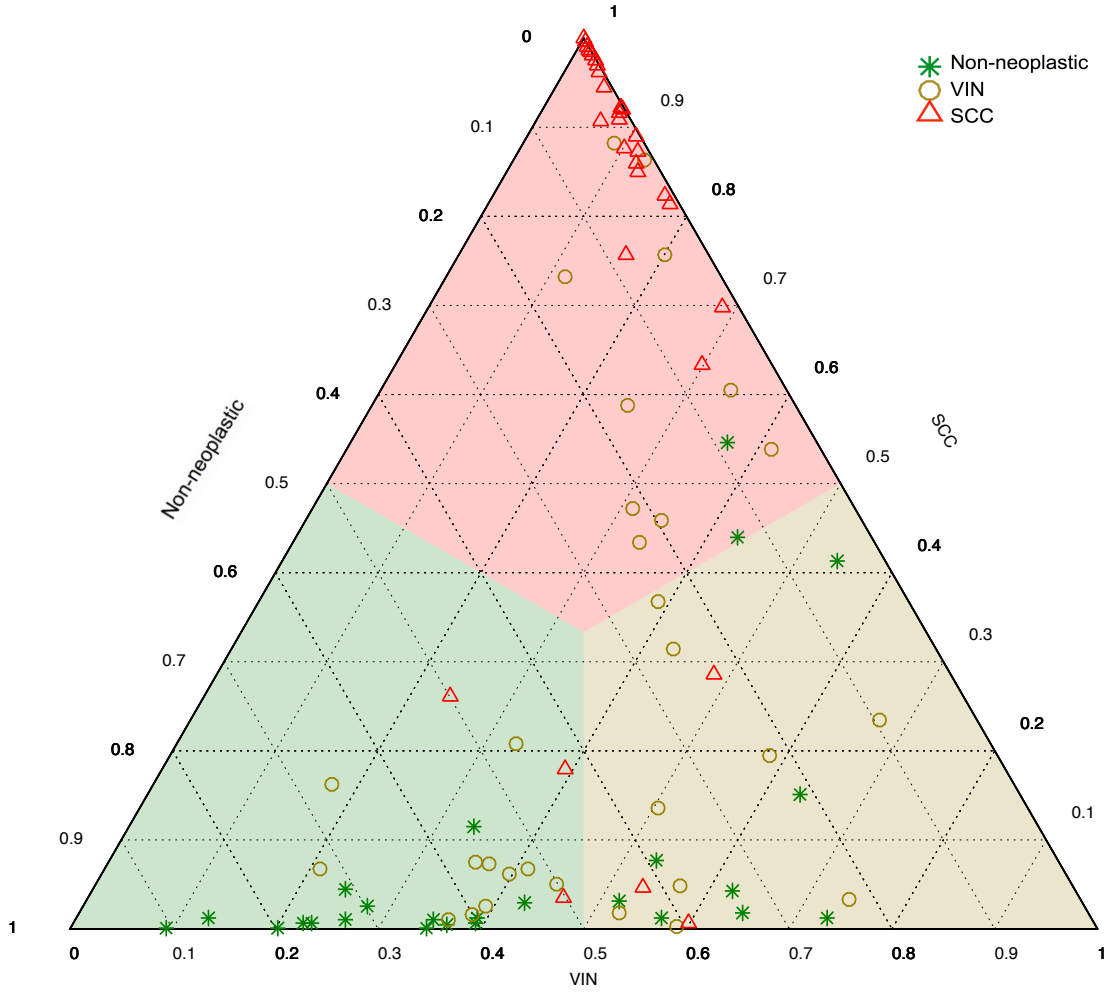


Figure 8-22 – Ternary plot of the predicted probabilities of a sample belonging to each of the three pathology groups under analysis. Non-neoplastic vulval tissue, VIN – vulval intraepithelial neoplasia, SCC – squamous cell carcinoma. Legend indicates actual pathology of samples, axes are probability of group membership based on linear discriminant distance.

The receiver operator curves demonstrating the diagnostic performance of the Raman probe with varying linear discriminate boundaries are shown in Figure 8-23, Figure 8-24 and Figure 8-25. These figures demonstrate that by altering the linear discriminate boundary between the groups sensitivities of 86% to 95% can be achieved with specificities of 72% to 57% respectively for the detection of non-neoplastic tissue from SCC or VIN (Figure 8-23). Likewise sensitivities of 81% to

89% can be achieved with specificities of 85% to 60% respectively for the detection of SCC from VIN (Figure 8-24) and sensitivities of 84% to 92% can be achieved with specificities of 84% to 64% respectively for the detection of SCC from non-neoplastic tissue and VIN (Figure 8-25).

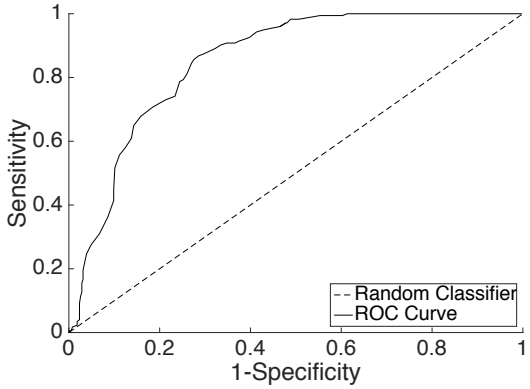


Figure 8-23 – Receiver operator curve for the detection of non-neoplastic tissue from SCC and VIN per spectrum. Area under curve, 0.85, Risk of falsely rejecting null hypothesis, <0.05

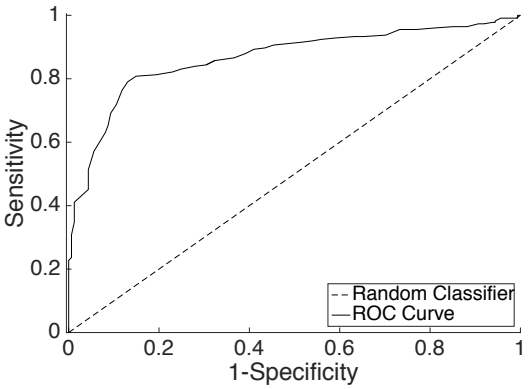


Figure 8-24 – Receiver operator curve for the detection of SCC from VIN per spectrum. Area under curve, 0.87, Risk of falsely rejecting null hypothesis, <0.05

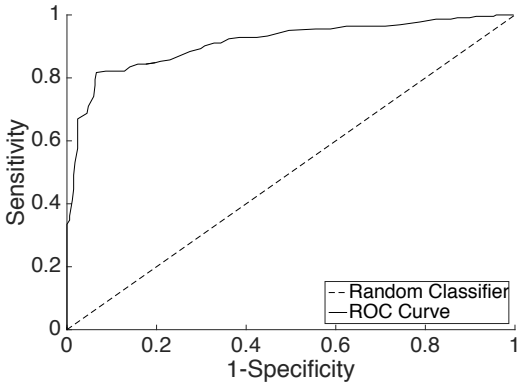


Figure 8-25 – Receiver operator curve for the detection of SCC from VIN and non-neoplastic tissue per spectrum. Area under curve, 0.91, Risk of falsely rejecting null hypothesis, <0.05

8.6 Spectral peak assignments

As previously discussed the exact nature of the biomolecular changes that leads to disease classification using Raman spectra are difficult to identify with confidence (chapter 6.7). In both the streamline investigation (Chapter 6) and the transepidermal investigation in Chapter 7, spectral differences between the pathology groups were identified that were consistent with the published literature investigating malignant processes at other skin sites. These included a right shift in the 1661 cm^{-1} Amide I band, a widening of the peak at 1451 cm^{-1} and an increase in peak intensity at 1080 cm^{-1} with malignant progression (Gniadecka *et al.*, 1997; De Gelder *et al.*, 2007). In the current data obtained using the fibre probe the right shift in the 1661 cm^{-1} band is less clearly seen although the changes at 1451 cm^{-1} and 1080 cm^{-1} are replicated (Figure 8-26).

A number of the novel spectral differences between the malignant and non-malignant groups seen in the streamline model and the transepidermal investigation in Chapter 7 are again replicated in the spectra obtained using the fibre optic probe. These include the increased prominence of the peaks at 723 cm^{-1}, 782 cm^{-1}, and 1101 cm^{-1} likely to represent DNA and RNA consistent with the increase rate of cell proliferation in neoplastic tissue. Increased prominence in the peak at 830 cm^{-1} (nucleic acids) is again seen in this data but it is partially obscured by the peak at 814 cm^{-1} likely from the presence of dermal collagens within the spectra. The previously seen reduction in the mono/disaccharide peak at 899 cm^{-1} in the streamline investigation is not clearly identifiable due to masking from adjacent

spectral peaks. Other amino acid peak changes previously seen including the increased prominence of the amino acid peak at 756 cm^{-1} and the reduction in the prominence in the peak at 1622 cm^{-1} are again seen but the changes are subtler in the data acquired through the probe system. In the spectra obtained with the probe system there is a striking difference in the Amide III region of the spectrum between 1224 cm^{-1} and 1269 cm^{-1}. In this region there is increased prominence of the spectral peaks in the VIN and non-neoplastic groups. This could be accounted for by the signal from dermal collagens or from changes in other structure proteins. The change in structural proteins can be explained by the desmoplastic/ fibroblastic response that can be seen in association with carcinoma formation (Holthoff *et al.*, 2017). In the adjoining region 1287 cm^{-1} to 1342 cm^{-1} there is a significant increase in prominence of the peaks in the SCC group and to a lesser extent the VIN group. This is likely due to amide II proteins (1283 cm^{-1} & 1311 cm^{-1}), lipids (1283 cm^{-1}), elastin (1273 cm^{-1}), collagen (1323 cm^{-1}) carbohydrates (1342 cm^{-1}) and RNA (1328 cm^{-1}). It is not clear why this feature was not seen in the trans-epidermal spectra collected by the microscope system but it is likely that there is a significant difference in sampling volume between the microscope and probe systems. The microscope system has a degree of confocality due to the use of microscope objectives and a relatively narrow spectrometer slit. This results in inclusion of signal from only a narrow volume in the collected spectra, as previously demonstrated. By contrast the probe system has no focusing objectives and no confocality, so signal from a wider tissue area will be included in the collected spectra. The changes seen in the spectra measured with the probe system between 1224 cm^{-1} and 1342 cm^{-1} were also noted

by Ferreira et al. in an ex vivo and in vivo study of Raman spectra collected from normal tissue, basal cell carcinoma, actinic keratosis and SCC (Ferreira Lima *et al.*, 2019). In this study the authors compared the in vivo and ex vivo Raman spectra of seven SCC lesions and ten areas of normal skin using a dispersive Raman spectrometer. The areas of greatest variability between the SCC and normal group were found in the regions between 1250 cm^{-1} and 1350 cm^{-1} and between 1440 cm^{-1} and 1670 cm^{-1}. These same variations in peak intensity are seen in this current study confirming these changes are likely due to changes in the pathology rather than as a result of interference from the experimental technique. The identified spectral differences are also reflected in the difference spectrum between overtly malignant and non-neoplastic tissue (Figure 8-27) as well as the composite PCA LDA loading (Figure 8-28) and appear to be significant in the predicted classification of the pathology groups.

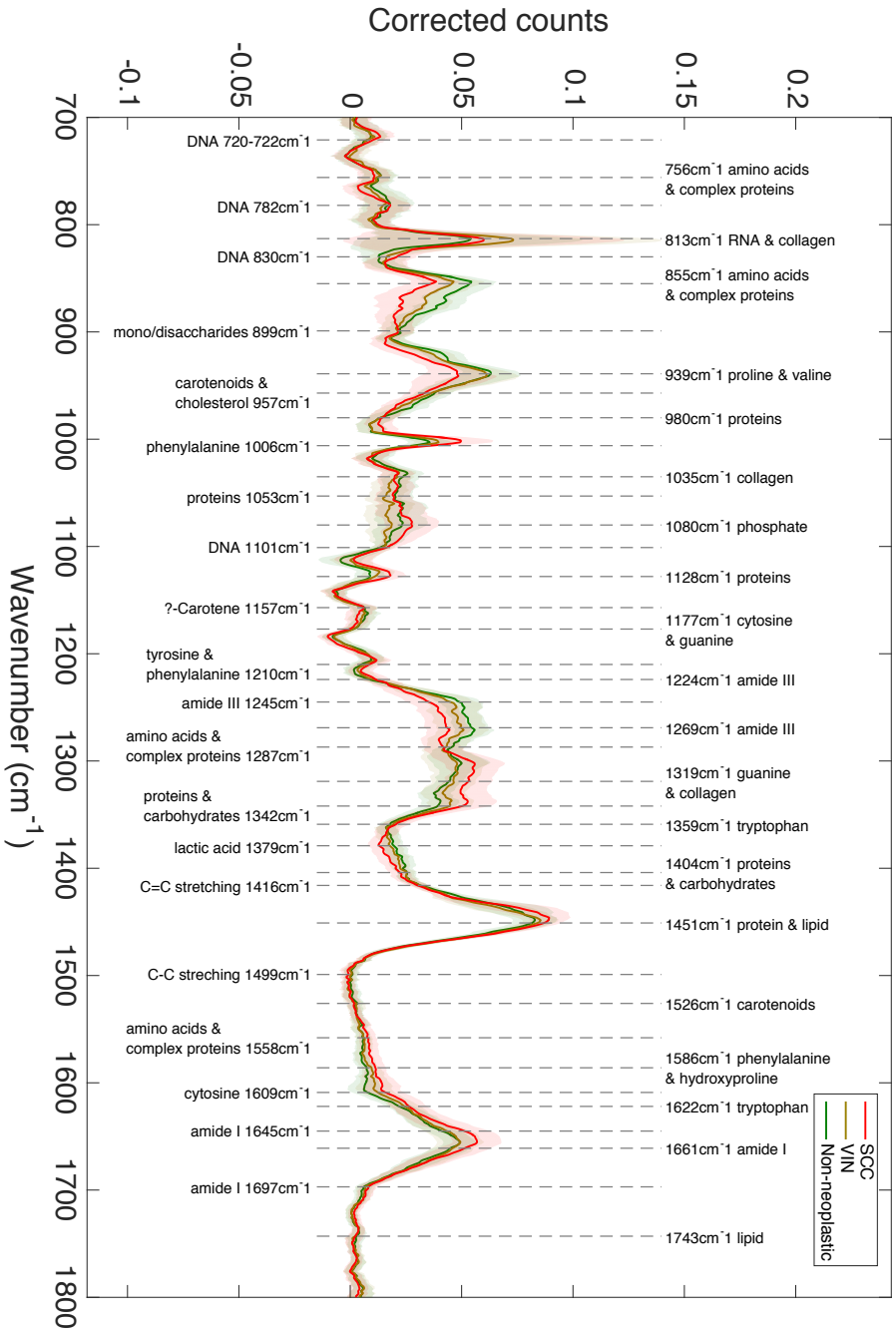


Figure 8-26 – Mean spectra of the pathology groups. Non-neoplastic vulval tissue, VIN – vulval intraepithelial neoplasia and SCC – squamous cell carcinoma. +/- one standard deviation marked by shading either side of line.

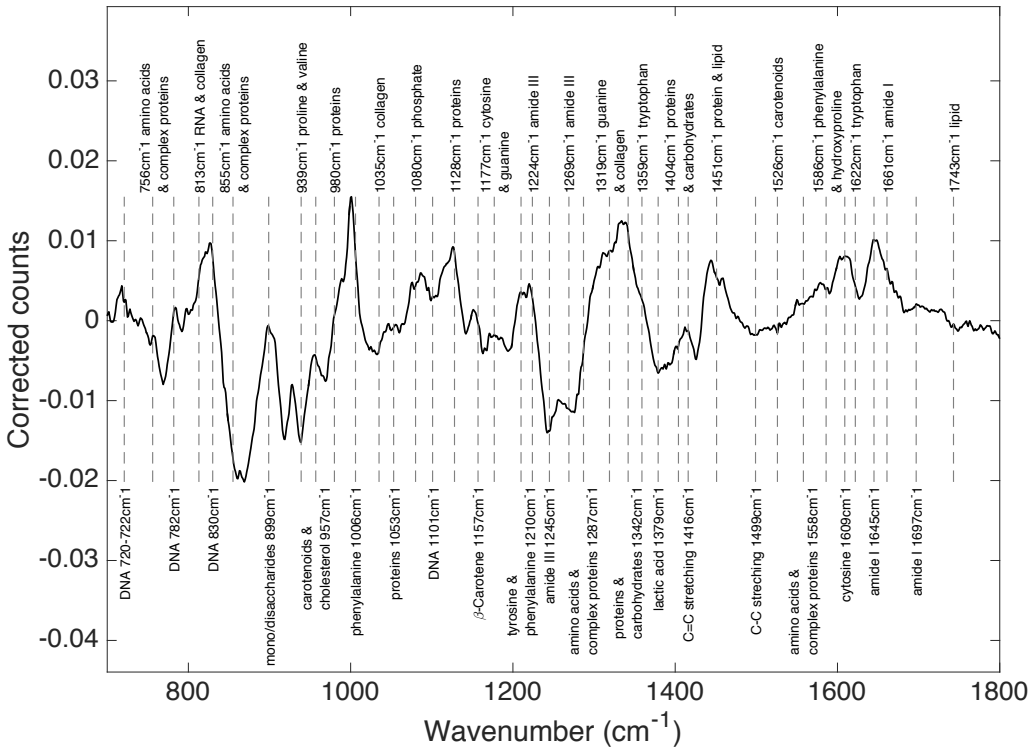


Figure 8-27 – Difference spectrum: mean of the non-neoplastic group spectrum subtracted from the mean of the SCC group spectrum

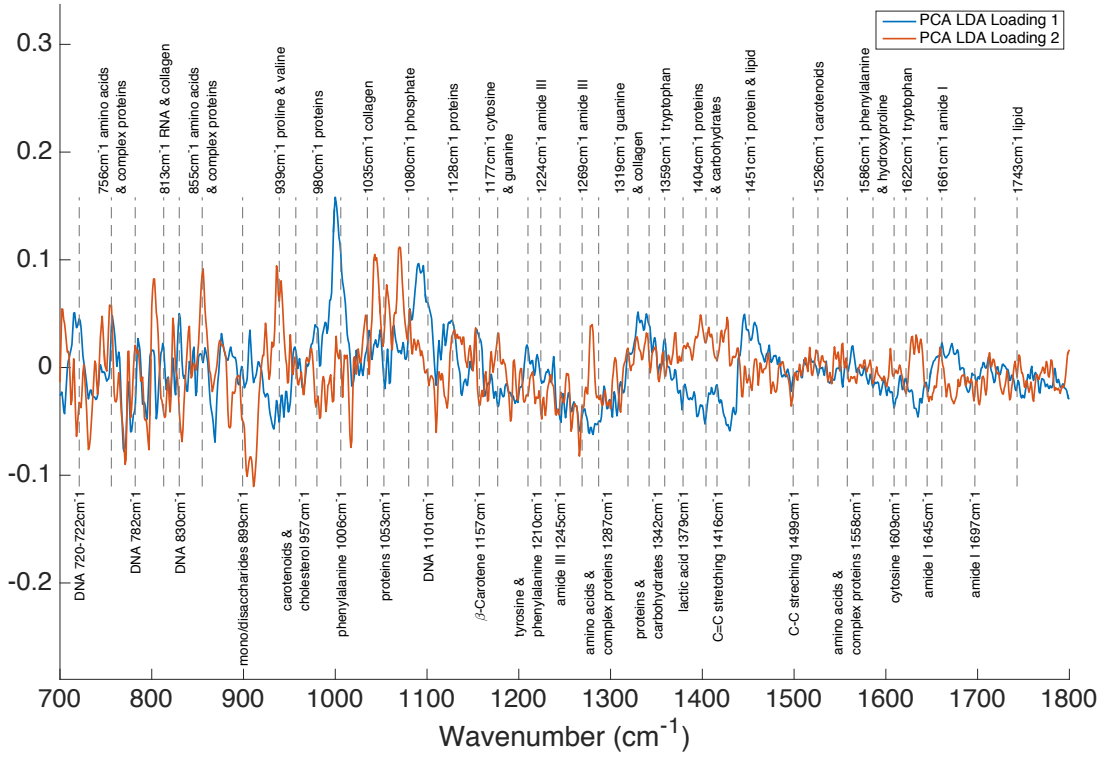


Figure 8-28 – Composite PCA LDA loading

8.7 Interim discussion

[Summative discussion and conclusions relating to the investigations outlined in Section C relating to Raman spectroscopy for the assessment of vulval disease can be found in Chapter 13]

The aim of this investigation was to evaluate the ability of a fibre optic probe and portable Raman spectrometer to classify non-neoplastic, premalignant and malignant disease of the vulva. This is the first study to demonstrate classification of vulval disease using Raman spectra collected using a probe based system. In this investigation it was possible to use individual Raman spectra to accurately classify

the pathology of the tissue being investigated with clinically relevant accuracy. This has replicated the results of the previous two analyses of this tissue set as outlined in Chapter 6 and Chapter 7. When comparing the probe technique with the mapping approach (Chapter 6) and the point measurement approach (Chapter 7), a significant fall in diagnostic accuracy would be expected as the probe system is associated with more noise and other sources of interference for example the optical etaloning outlined above. The study demonstrated that the ability of the probe technique in differentiating the pathology groups was found to be comparable to that of the point measurement approach (Chapter 7) as outlined in Table 13-2. The areas under the ROC curves were very similar for all three of the main diagnostic comparisons demonstrating a portable probe based system is capable of diagnostic performance similar to that of the bench based laboratory system. The relatively good performance of the probe in relation to the trans-epidermal spectra collected in Chapter 7 may be due to differences in sampling volume of the probe compared to microscope system as suggested by the additional spectral features identified as outlined in chapter 8.6. The same samples were analysed using both techniques.

Further discussion of future work and the steps necessary to translate these finding into clinical practice is located in Chapter 13.

Section D: Fourier Transform Infrared Spectroscopy for the Assessment of Vulval Disease

Chapter 9 Introduction, Aims and Objectives

9.1 Introduction

As previously discussed in chapter 3.4.3.1 there are currently no established methods for automating elements of the histopathological diagnostic pathway in vulval disease. Histopathologists' time is a valuable and finite resource and the increasing incidence of cancer and the growing complexity of referrals are leading to a potential crisis in this key diagnostic resource (Bainbridge *et al.*, 2016). The ability to automate elements of the histopathological pathway or triage tissue samples has the potential to speed up the diagnosis of patients with vulval cancer. The improvements in the diagnostic pathway could arise out of ability to rapidly triage samples at high risk of cancer to full pathology review or in providing an automated rapid double reporting of specimens (Hirschowitz, Wells and Lowe, 2013). Recently there has been a great deal of interest in spectroscopic methods for the identification of malignancies. The use of FTIR for differentiating between different tissue types has been gaining significant research interest because of its potential as a tool in histological and cytological diagnosis. FTIR is a non-destructive method that can detect changes in molecular functional groups in tissues under assessment, in the mid infrared region. The technique has been demonstrated to be an accurate

instrument for the discrimination of malignant and non-malignant tissue in the colon, breast, prostate, cervix, stomach, skin and oesophagus. To date there have been only a small number of studies exploring the role of FTIR in the assessment of skin disease. Tfayli et al. demonstrated that the areas of FTIR spectra relevant to melanin and DNA could be used to differentiate spectra of melanoma and pigmented naevi collected from paraffin fixed tissue sections (Tfayli *et al.*, 2005). Hammody et al. differentiated melanoma from surrounding epidermal tissue using markers found in the regions of the IR spectrum associated with nucleic acids (Hammody *et al.*, 2008). Ly et al demonstrated IR spectroscopic micro-imaging of paraffin fixed tissue sections could be used to differentiate BCC, SCC, Bowen's disease and normal or reactional epidermis (Ly *et al.*, 2009). More recently Kryriakidou et al showed IR spectral changes in samples of BCC, melanoma and benign naevi however the study was too small to demonstrate the classification ability of the technique (Kyriakidou *et al.*, 2017). These studies suggest FTIR has promise for the differentiation of skin pathologies and for the identification of malignancies. To date there are no published studies evaluating the use of FTIR for the assessment of the unique set of pathologies present in the human vulva.

In addition to simply distinguishing between different histopathological diagnoses FTIR like Raman spectroscopy has the potential to offer a unique biomolecular assessment of tissue in a way that can be easily integrated in the normal diagnostic pathway. FTIR can be used to assess the protein, carbohydrate, nucleic acid and lipid content of human tissues. The composition of these constituents will change during the process of carcinogenesis (chapter 1.4) and these changes have the

potential to be an early marker for the development of cancer (Bhargava, 2007; Kumar *et al.*, 2013). Detecting these changes could allow the identification of cancer before the typical histopathological morphological changes are seen, leading to earlier diagnosis and treatment. Kumar et al used FTIR to detect changes in key structural proteins within breast tissue that where found to be present in tumour tissue, in adjacent histologically normal tissue and with decreasing intensity with increasing distance from the primary tumour (Kumar *et al.*, 2013). Detectable changes in tissues adjacent to tumours raises the possibility of detectable biomolecular changes associated with cancer that could be used for early diagnosis or risk assessment.

As well as early diagnosis, biomolecular assessment of tissues using FTIR has the potential to provide useful information on disease course or prognosis not provided by traditional histopathology. Baker et al associated FTIR based spectral characteristics with clinically aggressive behaviour in prostate cancer (local and distant spread) (Baker *et al.*, 2008). This study importantly demonstrates that FTIR has the potential to enhance traditional histopathology with information not currently accessible, such as enhanced assessment of aggressive tumour behaviours. Mackanos et al demonstrated that histopathological grading score (Gleason grading) could be determined from FTIR spectra with 92.3% sensitivity and 99.4% specificity (Mackanos and Contag, 2009). The authors also demonstrated that tumour staging could be predicted from FTIR spectra of the primary tumour with 91.2% sensitivity and 93% specificity. It should be noted that this study was an exploratory analysis with only 39 patients included with no validation of the classification model. Minnes

et al showed that ATR-FTIR could be used to identify metastatic and non-metastatic melanoma cell lines, information not currently available using histopathology or cytology (Minnes *et al.*, 2017).

As previously discussed FTIR spectroscopy is particularly well suited to these applications as the technique can be used to assess biomolecular features of tissue prepared in a routine manner in a timely fashion and could be integrated into the standard sample processing pathway.

9.2 Aims

To investigate the ability of FTIR Spectroscopy to classify vulval pathology including lichen sclerosus, vulval intraepithelial neoplasia and vulval squamous cell carcinoma.

To investigate the potential of FTIR spectroscopy to work as an adjunct to histopathology to aid in the early diagnosis of cancer and offer additional information about the molecular risk profile of vulval tissue.

To achieve this aim, the following objectives were set to progress experimentation with the technique.

9.3 Objectives

1. Evaluate the ability of FTIR spectroscopy to classify significant vulval pathologies including vulval intraepithelial neoplasia, vulval squamous cell

carcinoma & non-neoplastic vulval tissue and explore underlying spectral differences.

   a. To investigate the ability of FTIR spectroscopy applied to routinely fixed and embedded tissue sections for the triage of vulval pathologies.

   b. To investigate the ability of FTIR spectroscopy to act as a second opinion to confirm routine histological diagnosis.

2. Evaluate the ability of FTIR spectroscopy and multivariate analysis to identify molecular changes in precancerous vulval conditions that are associated with malignant transformation.

Chapter 10 FTIR General Technical Considerations and Ethical Approval

10.1 Ethical approval and consent

Research ethics committee approval was gained for the use of archived fixed vulval tissue from the East of Scotland Research Ethics Service (14/ES/1066).

10.2 Instrumentation

The Fourier transform infrared spectroscopy presented in this thesis was performed using a FTIR spectrometer linked to an imaging system. The system used was a Perkin Elmer Spectrum One spectrometer together with a Perkin Elmer Spotlight 400 microscope system (Ohio, USA). This system has a liquid nitrogen cooled linear focal plane array detector (16x1) that can be used in imaging mode to produce spectral maps of tissue sections. This system can be used in a reflectance or transmission mode and for this study all samples were analysed in transmission.

10.3 Sample indentification

The Gloucestershire Hospitals NHS Foundation Trust pathology database was screened to locate relevant samples and review of the accompanying H&E stained tissue sections was used to locate the fixed tissue blocks of interest. Tissue sections from these blocks were cut for FTIR analysis. Tissue blocks with visible tissue ink were avoided to prevent contamination of the ink onto the tissue during cutting. All tissue analysed was embedded within paraffin and no chemical deparaffinisation was undertaken prior to spectral acquisition.

10.4 Sample preparation

Samples prepared for analysis by FTIR spectroscopy must meet the criteria outlined in Table 10-1 to ensure the biomolecular information contained within the infrared absorbance spectra is effectively recorded and available for analysis.

10.4.1 Sample substrate

Circular 20 mm diameter calcium fluoride slides were chosen as the most appropriate tissue substrate. The low absorption of calcium fluoride in the infrared range means the tissue signal is not adversely affected by the slide. In addition circular slides mounted for analysis in a custom made slide tray (Figure 10-1) allowed orientation of the slide to make best use of the rectangular mapping area of

the FTIR system. The fabrication workshop, department of physics and astronomy, University of Exeter, manufactured the custom slide tray.

Table 10-1 – Desirable properties of sample preparation for FTIR spectroscopy

Criteria	**Rational**
Should allow transmission of infrared radiation	Samples should not be completely opaque to infrared radiation as this will result in saturation of the absorption spectrum and loss of spectral information.
The infrared spectrum of the substrate should not obscure the bio-relevant spectral range	If the spectra peaks of the substrate overlap the tissue related spectral peaks the tissue signal may be obscured and relevant biomolecular information lost.
Should not contain other substances that will obscure the bio-relevant spectral range	Other substance within the tissue such and inks and paraffin may result in spectral signal that obscures relevant biomolecular information.
Should be readily correlated to pathological diagnosis	The ability to map the measured spectrum to histopathological classifiable stained sections allows correlation of spectral and histological information.
Should be orientatable to allow efficient use of the rectangular measurement window of the spectrometer	The FTIR system used was limited to acquisition of rectangular maps of samples and therefor the ability to securely orientate the samples is key to make most efficient use of the spectrometer.

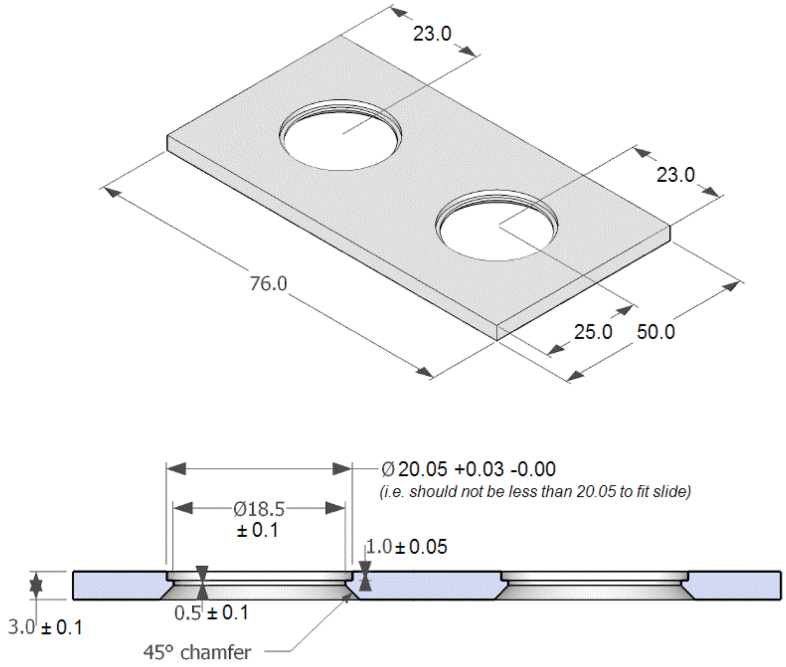


Figure 10-1 – Custom built slide tray for 20 mm calcium fluoride slides

10.4.2 Sample tissue section thickness

An initial experiment was undertaken to ascertain the optimal thickness for tissue sections to undergo FTIR mapping. As external vulval skin like skin elsewhere in the body forms a protecting barrier it is highly keratinised. This keratin layer is dense and relatively opaque to light. For this reason an assessment of the limit of the FTIR spectrometer's ability to detect photons penetrating through tissues of different thicknesses was undertaken. This was done to ensure that the spectra collected from mapping the tissue sections would not be adversely affected by high absorbance in the tissues and the limit of the FTIR detector.

The assessment was formed of two parts. The first was to establish the FTIR detector's response to decreased signal due to an increasing absorbance by the

sample under analysis. Spectral maps were obtained of semi opaque polymer windows that were added together to form and increasingly absorbent layer. Spectral maps were collected for an increasing number of windows up to a total of twelve. The absorbance at the wavenumber with the greatest absorbance in the spectra (1600 cm-1) (Figure 10-2) was plotted against the increasing number of semi opaque windows. This revealed that the response of the detector is linear to a limit after which no light is detected by the spectrometer, the saturation point (Figure 10-3).

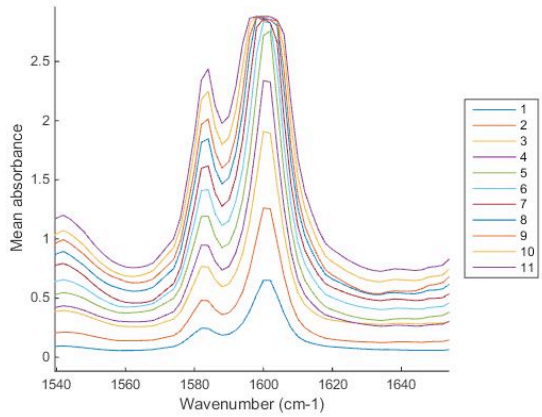


Figure 10-2 – FTIR absorbance spectra in the region 1540 to 1645 cm^{-1} with increasing numbers of semi opaque windows (Spectral resolution 4 cm-1; Scan per pixel 16; Interferometer Speed 2.2 cm/s)

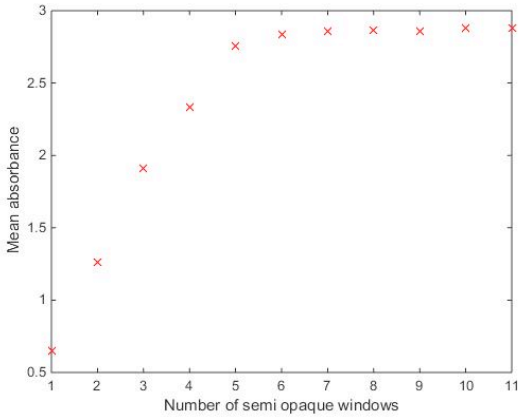


Figure 10-3 – Absorbance plotted against increasing numbers of semi opaque windows (Spectral resolution 4 cm-1; Scan per pixel 16; Interferometer Speed 2.2 cm/s)

This process was then repeated for increasing thicknesses of tissue sections taken from a block of normal vulval skin ranging from 1 to 8 microns. The greatest absorbance peak (Figure 10-4) (1650-1660 cm-1) was plotted against the thickness of the tissue sections and the same linear relationship was seen (Figure 10-5). The thickness of tissue sections chosen for the analysis was 4 µm as this thickness of

section is readily prepared and correlated with adjacent H&E stained sections and it lies firmly within the detection range of the FTIR system. If tissue section density varies between different samples and pathologies at 4 µm, denser tissues should still be within the range of the detector.

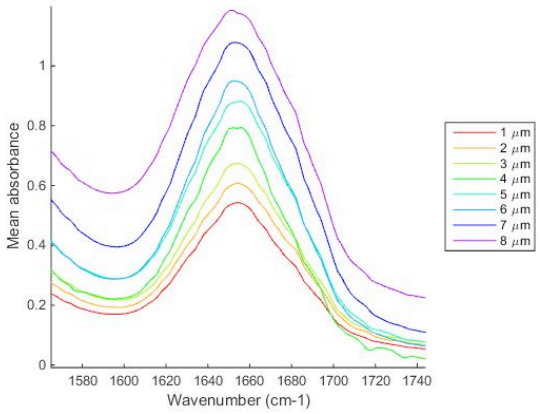


Figure 10-4 – FTIR absorbance spectra in the region 1560 to 1745 cm^{-1} with increasing thickness of tissue section (Spectral resolution 4 cm-1; Scan per pixel 16; Interferometer Speed 2.2 cm/s)

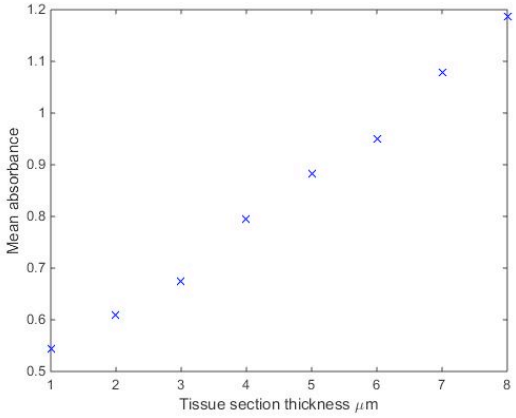


Figure 10-5 – Absorbance plotted against increasing numbers of semi opaque windows (Spectral resolution 4 cm^{-1}; Scan per pixel 16; Interferometer Speed 2.2 cm/s)

10.5 Optimisation of acquisition parameters for spectral measurement

Initial experiments were undertaken to determine the optimal parameters for spectral measurement. The aim of the optimisation was to maximise the signal to noise ratio in the spectra collected (chapter 4.1.2), preserve useful detail in the spectra and allow the measurements to be undertaken in a technically feasible timeframe.

For the purposes of determining the optimal acquisition parameters the signal to noise ratio was calculated from the signal height of the main tissue peak between 1600 and 1800 cm^{-1} over the adjacent spectrum at 1692 cm^{-1} divided by the noise. The noise was calculated by estimating the noiseless spectrum by applying a median filtering technique at a section of the spectrum where there is no significant tissue signal. The mean of the absolute difference between the actual spectrum and the estimated median filtered spectrum was then taken to be the noise (chapter 4.1.2).

10.5.1 Maximum length of aquisition

The length of spectral acquisition with the Perkin Elmer Spotlight 400 system is limited by the size of the liquid nitrogen Dewar flask within the instrument. An experiment was performed to evaluate the length of acquisition that was possible. The liquid nitrogen Dewar flask was filled from room temperature and a 7 hour map of homogeneous tissue commenced. The signal to noise ratio was calculated and plotted (Figure 10-6). After around five and a half hours acquisition, the detector starts to warm and the signal to noise ratio decreases and eventually at around six and a quarter hours, spectral acquisition stops altogether (Figure 10-6).

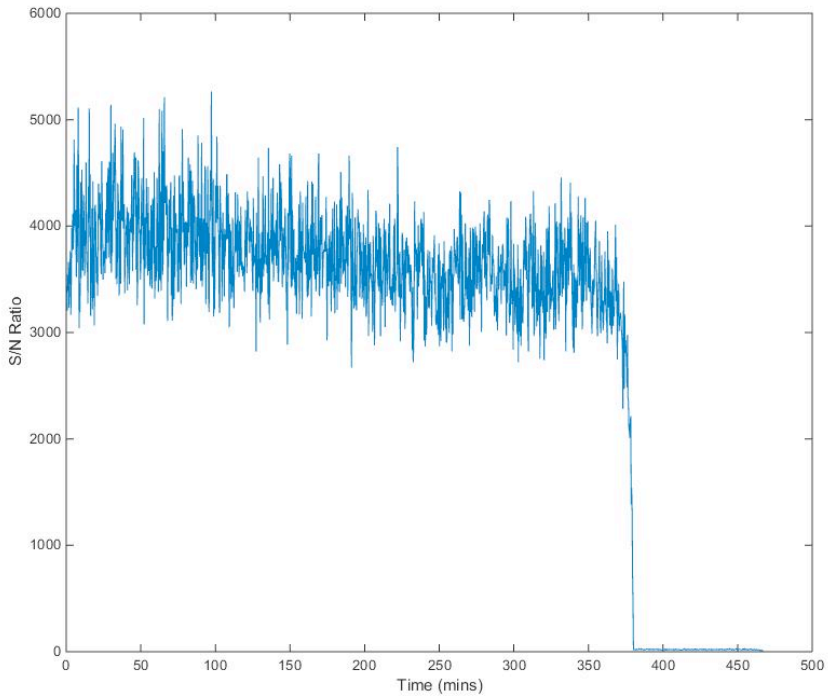


Figure 10-6 – Signal to noise ratio in mapping model with increasing time from filling of liquid nitrogen dewar in Perikin Elmer Spotlight 400 FTIR system

10.5.2 Acquisition parameters

Using a 4 μm thick paraffin fixed section of normal vulval epithelium a comparison was undertaken of the effect of the scans performed per pixel, spectral resolution and interferometer speed on signal to noise and acquisition time.

To assess the role of the spectral resolution, spectra were collected from normal vulval epithelium for resolutions of 2, 4, 8 and 16 cm^{-1}. At the higher resolutions, the spectra have increased spectral detail however a reduced signal to noise ratio (Figure 10-7). At 16 cm^{-1} the spectra is significantly smoothed and the noise in spectrum is difficult to assess accurately. At the lower resolutions the noise reduces

however the detail spectra information is lost. At resolutions below 4 cm^{-1} the ability to distinguish the paraffin peaks between 450 cm^{-1} and 480 cm^{-1} is lost (Figure 10-7). The resolution of the spectra collected significantly impacts on the time taken to collect the spectra. The greatest reduction in time taken for map acquisition is seen between a spectral resolution of 2 and 4 cm^{-1} (Figure 10-9). In order to record good quality spectra with a high spectral resolution and signal to noise ratio in an acceptable time frame a spectral resolution of 4 cm^{-1} was chosen for the analysis.

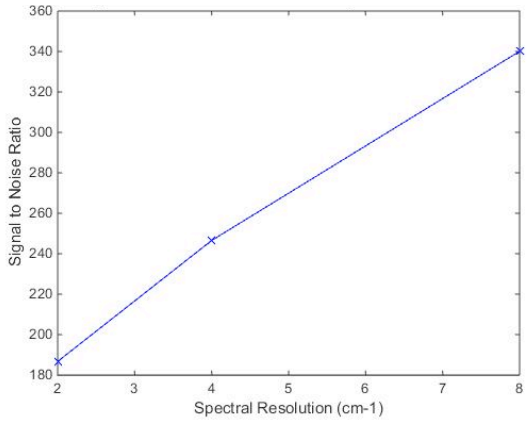


Figure 10-7 – Signal to noise ratio with decreasing spectral resolution

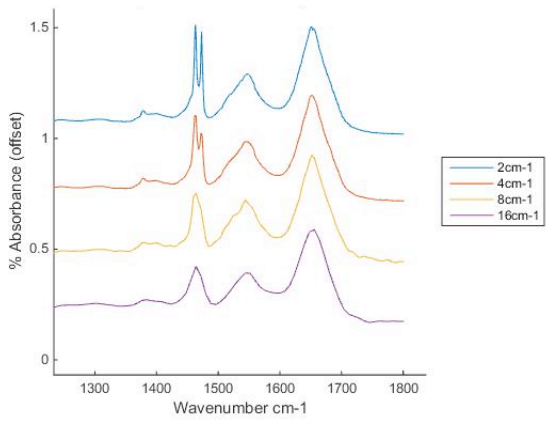


Figure 10-8 – Tissue absorbance spectra for different spectral resolutions

Spectra were collected for the two different interferometer speed settings 1 cm/s and 2.2 cm/s. There did not seem to be any difference in the resultant spectra or signal to noise ratio (Figure 10-10). The time taken for spectral acquisition is directly proportional to the speed of the interferometer with a typical tissue map of 750x750 microns area at resolution of 6.25microns, 16 scans per pixel and spectral resolution taking 301 minutes and 153 minutes at interferometer speeds of 1 cm/s and 2.2 cm/s respectively. Therefore as the absorbance of the light though 4 µm sections of tissue

does not approach the detection limit of the detector (Chapter 10.3) an interferometer speed of 2.2 cm/s was chosen for the analysis.

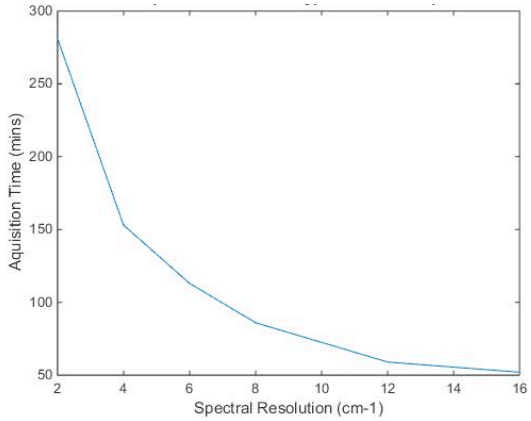


Figure 10-9 – Acquisition time for a typical map of a 750x750 microns area at spatial resolution of 6.25 µm,16 scans per pixel, interferometer speed 2.2 cm/s

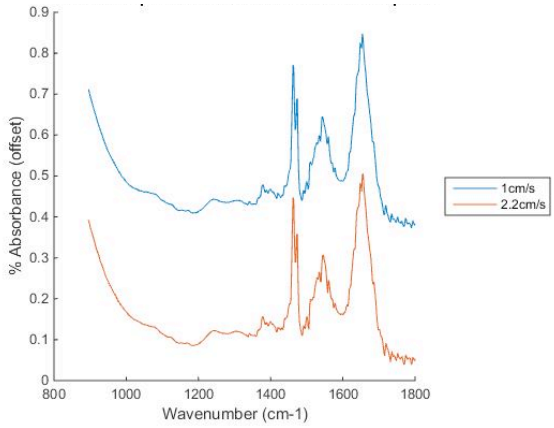


Figure 10-10 – Tissue spectra for different interferometer speeds, 16 scans per pixel, spectral resolution of 4 cm-1

The number of scans per pixel (SPP) was evaluated by the acquisition of spectral maps of normal vulval epithelium at a 2, 4, 6, 16, 32, 64 and 128 SPP using an interferometer speed of 2.2 cm/s, a spectral resolution of 4 cm^{-1} and spatial resolution of 6.25 microns. The signal to noise ratio of individual spectra was then calculated and is shown in Figure 10-11. Overall the signal to noise ratio for all the spectra collected across all numbers of SPP was high and visually there is little noticeable difference between the spectra. However the signal to noise increases significantly when the SPP increases from 2 to 32 with the majority of the increase happening between 2 and 16 SPP. Beyond 32 SPP there is not such a pronounced increase in the signal to noise ratio. The spectral acquisition time is directly proportional to the number of SPP and doubling of acquisition time between 16 and

32 SPP does not produce such a rise in the signal to noise ratio. Therefore for the purposes of this study, 16 SPP was used when analysing tissue samples.

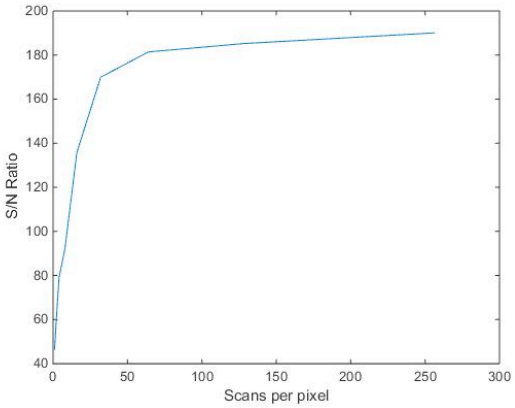


Figure 10-11 – Signal to noise ratio with increasing scans per pixel for spatial resolution of 6.25 µm, interferometer speed 2.2 cm/s and spectral resolution of 4 cm^{-1}

Parameter	Optimised Setting
Resolution	6.25 µm
Spectral resolution	4 cm^{-1}
Scan per pixel	16
Interferometer Speed	2.2 cm/s

Table 10-2 – Optimised parameters for acquisition of FTIR spectral maps

The optimised parameters selected (Table 10-2) allowed for the collection of high quality spectra of histological relevant regions of tissue in an acquisition time that is within the constraints of the FTIR system.

10.6 Management of non tissue FTIR absorption

FTIR absorption due to substances other than tissue has the potential to obscure tissue signal and confound spectral analysis. Two main sources of non-tissue absorption were considered in this study: paraffin absorption and absorption due to atmospheric water. In both cases steps were taken in the processing of the spectral data to minimise the impact of this absorption on the analysis.

10.6.1 Paraffin correction

Spectra obtained from FTIR of sections of paraffin embedded fixed tissue contain large variations due to the variation in paraffin content. In order that the classification model developed did not classify pathology groups based on the differences in paraffin content, a spectral correction was undertaken. This was done to ensure the inter- and intra-sample variation in the areas of the spectrum associated with paraffin was minimised.

To correct for the paraffin, representative areas of tissue free paraffin were scanned using an identical technique as the tissue samples. The resultant spectra underwent principal component analysis (chapter 4.2.1) to identify the key variation in the spectra. The principal component loads were then used with an extended multiplicative signal correction (EMSC) to correct the tissue spectral data for paraffin content. This method neutralised the variability in the signal from the paraffin by constraining the paraffin bands in the spectra to the same amplitude, regardless of the measured spectra (Tfayli *et al.*, 2009; Nallala, Lloyd and Stone, 2015).

10.6.2 Atomospheric water correction

Water vapour signal is typically seen in FTIR spectra between 1300 and 2000 cm^{-1} and again between 3500 and 4000 cm^{-1}. Several precautions are taken to reduce the variation in the impact of atmospheric water on spectral measurements. The FTIR spectrometer was situated in an air-conditioned room with low and stable relative humidity. In addition, the sample stage of the spectrometer is enclosed to reduce

variation in humidity due to the presence of the operator and the spectrometer undergoes continual active dehumidification with desiccant within the apparatus. Despite these measures, artefacts from IR absorption from water are still seen within the tissue spectra.

The control and data acquisition software of the Perkin Elmer spectrometer (Spectrum Image, Perkin Elmer, Ohio, USA) contains an algorithm for the automatic correction of background absorption based on the measurement of a background signal undertaken prior to the acquisition of a spectral map. This background was measured from a clean area of the calcium fluoride slide on which the tissue section was mounted. The intention is that this built in correction algorithm should correct for water vapour. The persistence of a water vapour artefact in the corrected data and principal component loadings suggested the correction algorithm was not adequate and further correction was required. Algorithms for the correction of water vapour artefact have been developed and a technique previously established for the correction of water vapour artefact from this instrument was applied to the collected data prior to analysis (Bruun *et al.*, 2006; Old, 2015). This correction technique uses background measurements taken while the spectrometer is purged with nitrogen to reduce the water vapour present within the spectrometer. The second derivative of the ratio of the background spectrum measured in normal atmosphere and the background spectrum measured during a nitrogen purge underwent a least squares fit to the second derivative of the measured spectra. The area of the spectrum associated with carbon dioxide (2280-2394 cm^{-1}), which is the other major contributor to the atmospheric spectrum was excluded from the fit. This prevented

the fit from under correcting for water vapour and preferentially correcting for atmospheric carbon dioxide. The resultant coefficient of fit for each individual spectrum was then used to correct for the water content in the spectrum (Figure 10-12 and Figure 10-13).

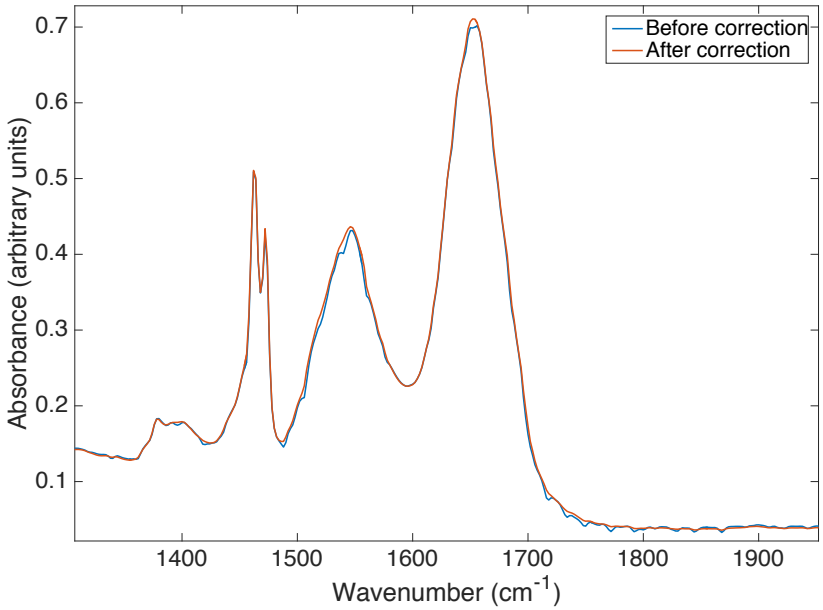


Figure 10-12 – Example mean tissue spectrum before and after water vapour correction

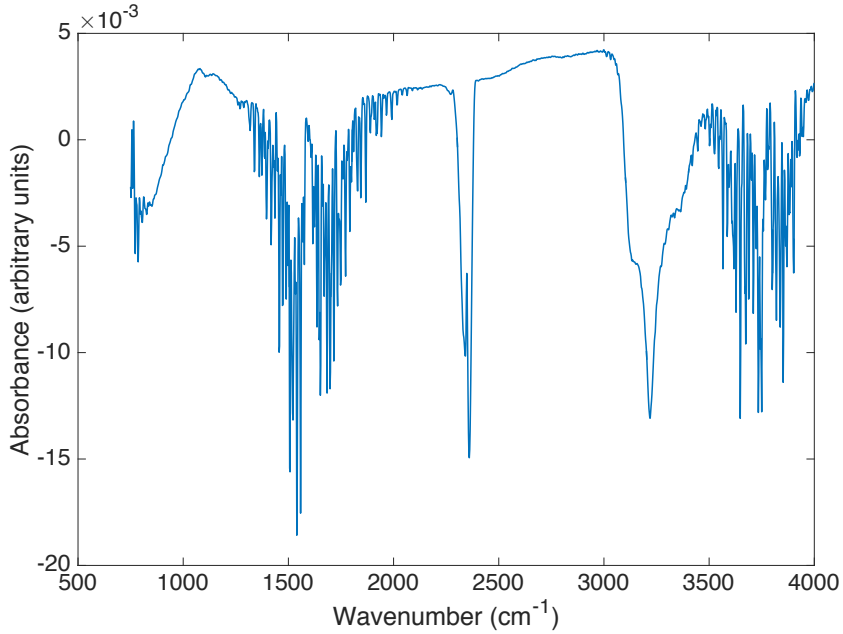


Figure 10-13 – Example mean atmospheric correction using nitrogen purge background ratio fit correction

Chapter 11 FTIR Spectroscopic Aided Detection of Disease

11.1 Introduction

As previously discussed (chapter 3.4.3.1 and chapter 9.1), there is the need for an objective technique for the classification vulval disease. Such a technique would enable the triage of cases to enable the histopathologist's workload to be prioritised and would offer a second opinion of the histopathological diagnosis to help reduce reporting errors and reduce the burden of double histopathological reporting. The following comprises a preliminary evaluation of FTIR spectroscopy as a suitable technique for this purpose.

11.2 Objectives

This investigation assesses objective 1a and 1b of the aims and objectives outlined in Chapter 9.

11.3 Sample selection, preparation and measurement

Three tissue groups were selected for the analysis based on their common pathological origins and development. The first tissue group selected was non-neoplastic tissue comprising normal vulva and lichen sclerosus affected vulva as these represent the benign pre-neoplastic state of the vulva. The second group chosen was vulva affected by vulval intraepithelial neoplasia (VIN) as this represents an important premalignant vulval condition that often requires treatment to avoid progression to vulval cancer. The third group chosen for the comparison was vulval squamous cell carcinoma (SCC) as this is an important diagnosis to be accurately made and a condition that can follow as a consequence of lichen sclerosus and VIN. The electronic pathology database of Gloucestershire Hospitals NHS Foundation Trust was interrogated to identify vulval tissue samples from each of the three selected pathology groups. The H&E stained slides from the identified cases were reviewed by an expert gynaecological histopathologist to select the areas in the tissue blocks containing the pathologies of interest. In cases where the presence of pathology of interest was confirmed the paraffin blocks containing the tissue were identified and four μm sections cut onto circular 20 mm calcium fluoride slides (Chapter 10.4). A 4 μm section was also cut on to a glass slide for H&E stain before and after a section was cut on to calcium fluoride. This was completed to allow accurate correlation of the unstained section and the pathology present. In total 32 sections were cut on to calcium fluoride from each of the pathology groups. This number of samples was selected to account for failures in sample processing aiming

to have 30 completed IR maps of each pathology group. As the significance of the spectral changes between the pathology groups was not known, it was not possible to perform a meaningful sample size calculation to estimate the number of samples required.

The H&E stained tissue sections cut from either side of the unstained sections were reviewed by a pathologist and the area containing the pathology of interest marked on the slide. The unstained sections on calcium fluoride then underwent FTIR spectroscopic mapping. When mapping the unstained sections a white light image of the samples was measured on the Perkin Elmer Spotlight system and the image compared to the area selected by the pathologist on the H&E stained section (Figure 11-1). The relevant area was then selected for spectroscopic analysis. The area selected for spectroscopic analysis included the full thickness of the epidermis and an area of paraffin for EMSC correction of the paraffin peaks (chapter 10.6.1) (Figure 11-1). Due to the irregular and heterogeneous nature of the samples and the need to collect spectra from the full thickness of the epidermis the shape and size of the spectroscopic maps varied according to the sample under analysis. Approximately four and a half hour long maps were taken from each sample using the acquisition parameters outlined in chapter 10.5. This resulted in varied number of spectra acquired due to the variable dimensions of the acquisition area. Prior to each map a background was measured as described in chapter 10.5.

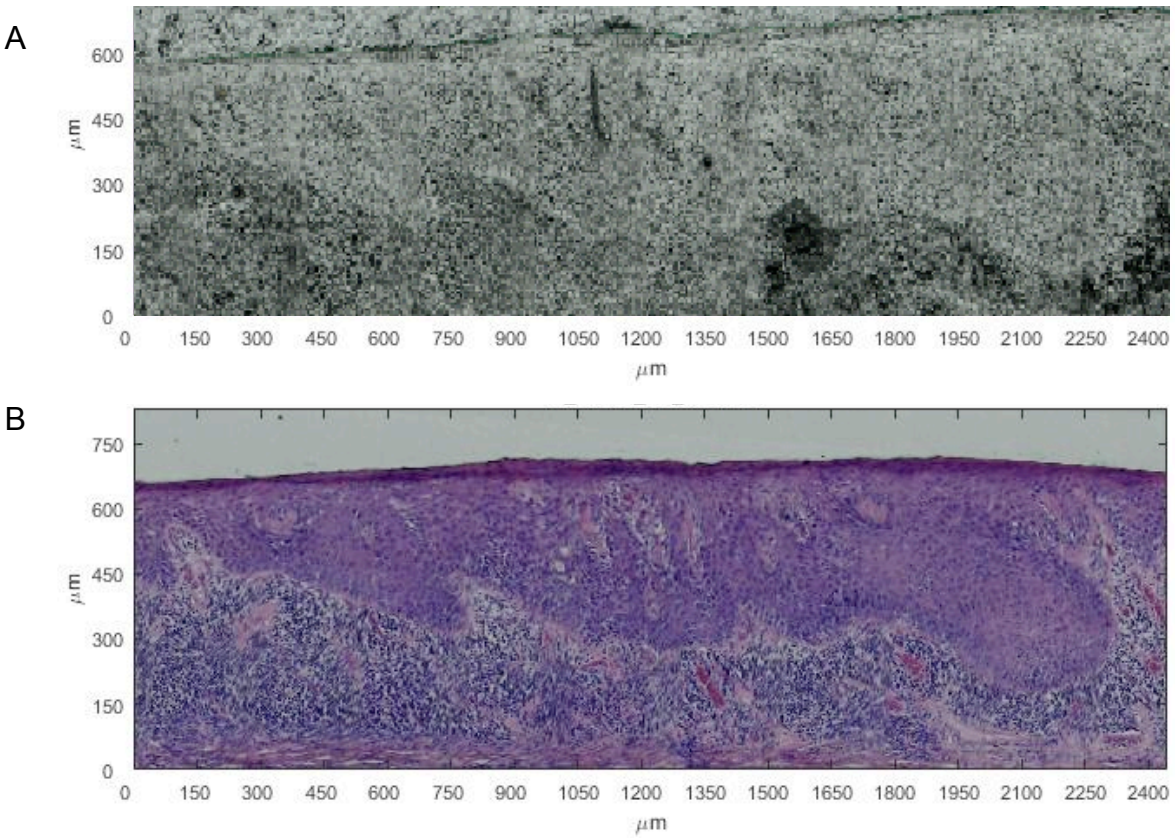


Figure 11-1 – White light image of unstained section of a sample of VIN (A) and corresponding H&E stained section (B)

11.4 Spectral preprocessing and analysis

Following acquisition of the spectral data map, the measured transmission was converted to absorbance by logarithmic transformation and corrected for atmospheric water vapour and paraffin as described in chapter 10.6.2. After correction for water, a two-stage k means clustering was applied isolating the spectra from the paraffin, epidermis, dermis and stratum corneum where present (Duda, Hart and Stork, 2001). The first stage was used to isolate the paraffin spectra in the map using a k means clustering with two centroids. The two centroids were

seeded with a manually selected tissue spectrum and a manually selected paraffin spectrum. This split the spectral map into paraffin and tissue spectra and spectra only containing paraffin signal. The paraffin spectra were saved for analysis and removed from the spectral map for the second stage of the k means clustering. In the second stage eight centroid clustering was performed with seed spectra from manually selected epidermis, dermis and stratum corneum and five other tissue spectra randomly selected from the spectral map. This enabled isolation of the spectra measured from the epidermis, dermis and stratum corneum as illustrated in Figure 11-2. The spectra from the resultant k means were then saved for analysis.

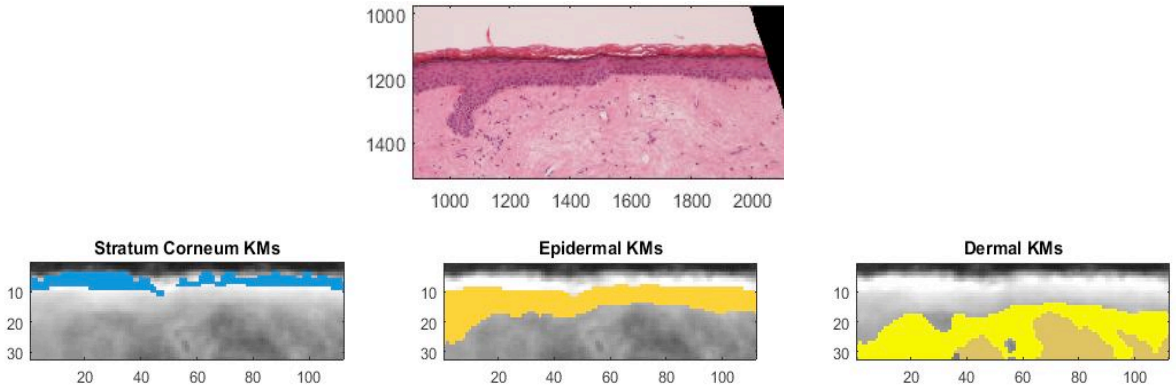


Figure 11-2 – Example of k means clustering for the isolution of differnent tissue layers in a sample of lichen sclerosus. H&E stained section for comparision (top) and k means clusters (bottom)

Earlier work on the differentiation of normal and pathological abnormal tissues have suggested that the wavenumber range 800 to 1800 cm^{-1} contains the most bio relevant area of the infrared spectrum (D Naumann, 2001; Eikje, Aizawa and Ozaki, 2005; Sahu and Mordechai, 2005; Hammody *et al.*, 2008; Movasaghi, Rehman and ur Rehman, 2008a). For this reason and to exclude the contribution of atmospheric water (1300 - 2000 cm^{-1} and 3500 - 4000 cm^{-1}) and carbon dioxide (2220 - 2390 cm^{-}

[1]) the wavenumber range was restricted to 800 – 1800 cm^{-1} for the analysis. It has been suggested that the high wavenumber range (2800 – 3000 cm^{-1}) where stretching vibration of CH_2 and CH_3 groups in some proteins and lipids are detected can be useful in the classification of disease however in paraffin embedded tissue this spectral range is masked by a strong paraffin signal. Initial inspection of the resultant significant loadings from the PCA and ANOVA revealed that some of the loadings described what appeared to be noise at the lower extreme of the wavenumber range (Figure 11-3). To prevent this from having a negative impact on the classification accuracy the wavenumber range used for the analysis was shortened to 850-1800 cm^{-1}. This excluded the noise between 800 and 850 cm^{-1} whilst retaining the bulk of the bio relevant spectral data.

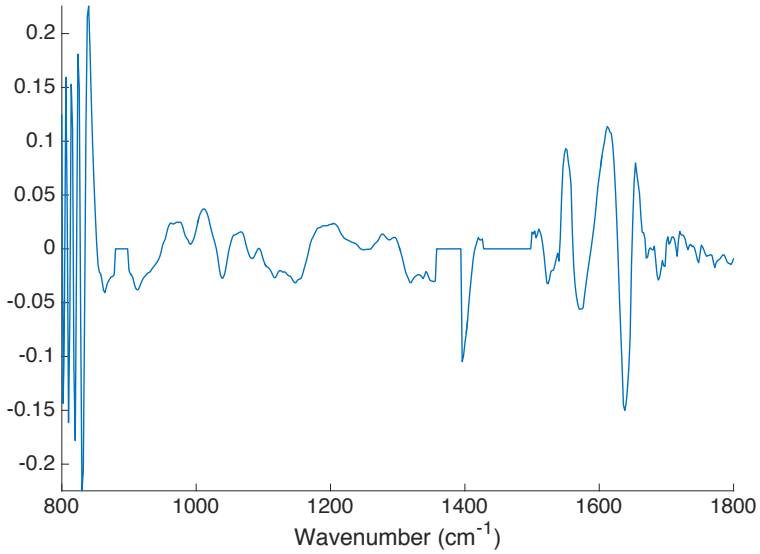


Figure 11-3 – Example principal component loading describing variance due to noise at the lower end of the wavenumber range (main paraffin spectral peaks at 880-898 cm^{-1}, 1358-1394 cm^{-1} and 1428-1498 cm^{-1} removed)

The paraffin spectra isolated from all the samples underwent PCA and the mean spectra combined with the PCA loads were used to correct for the variation in paraffin content of the combined tissue spectra as described in chapter 10.6.1. In addition to this, the areas of the spectrum associated with the main paraffin spectral peaks at 880-898 cm^{-1}, 1358-1394 cm^{-1} and 1428-1498 cm^{-1} (Figure 11-4) were excluded from the analysis to reduce the impact of non tissue signal contributing to the classification.

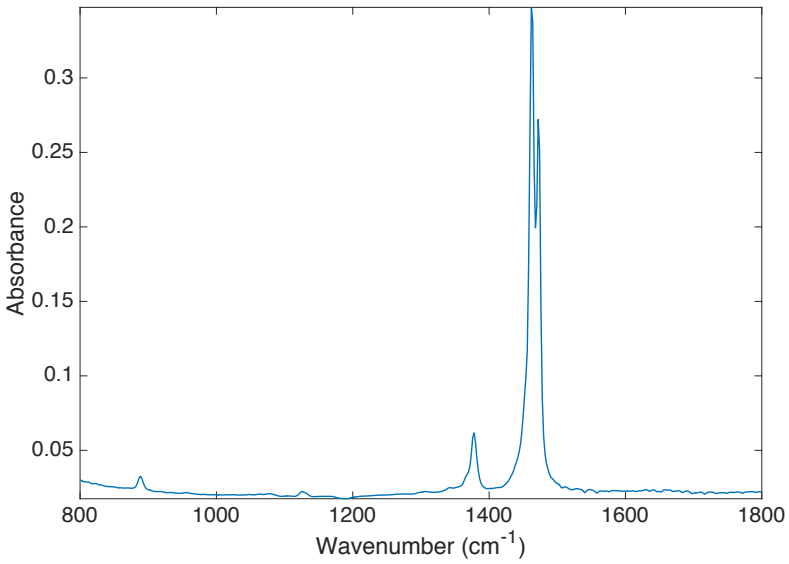


Figure 11-4 – Mean paraffin spectrum

The tissue spectra were then filtered for outlying spectra using the D and Q classifiers (chapter 4.2.6.1). Classification performance was assessed using PCA LDA in a leave one sample out cross validation loop (chapter 4.2.5) and the final prediction filtered by Mahalanobis distance to the mean of each of the LDA groups (chapter 4.2.6.2). The data analysis performed is summarised in Figure 11-5.

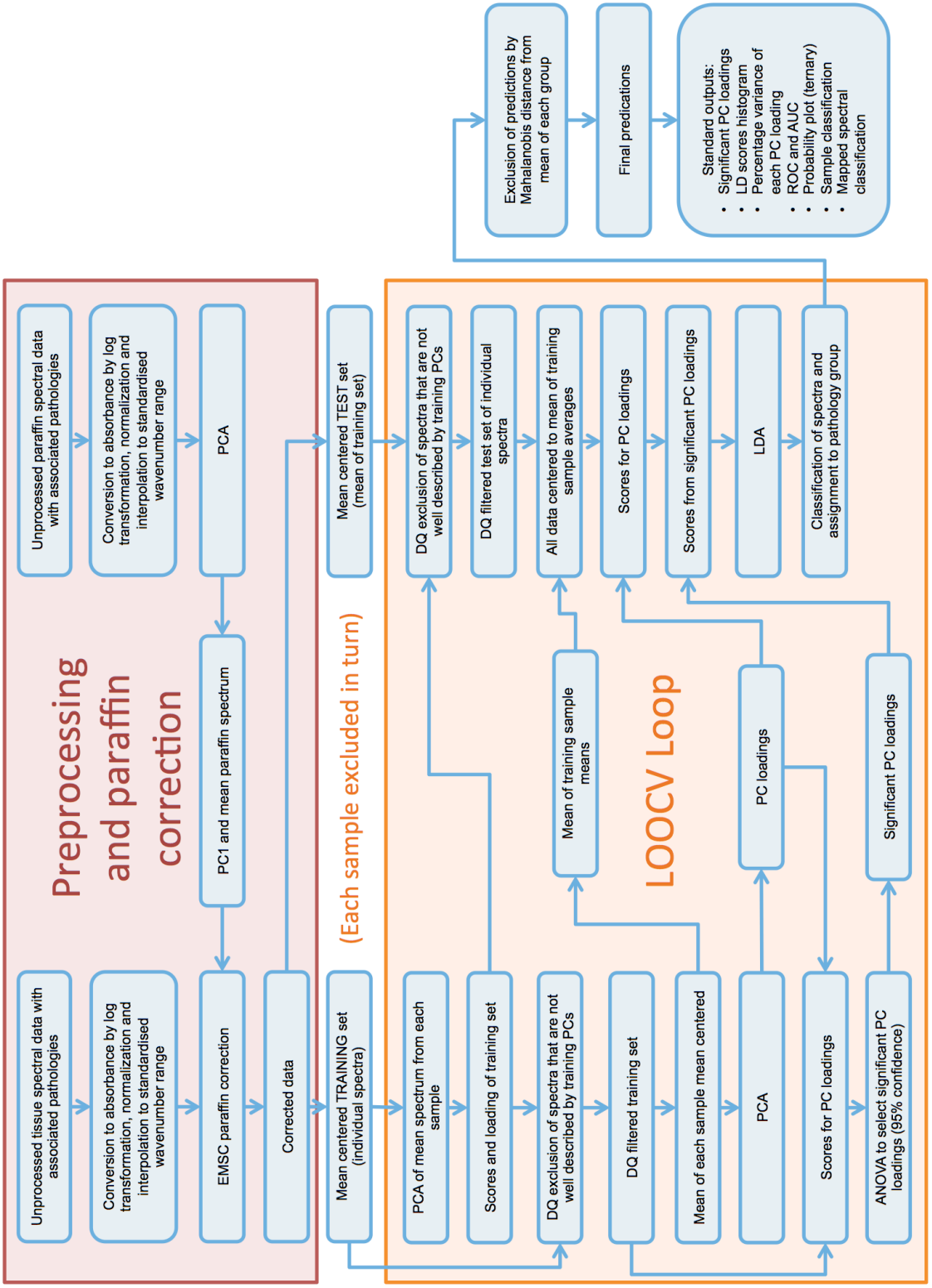


Figure 11-5 – Data analysis flow chart for the analysis of FTIR spectral maps for evaluation of spectroscopic assisted diagnosis.

11.5 Summary of measurements

A total of 93 FTIR spectral maps were collected from 93 tissue samples (Table 11-1). These tissue samples came from the specimens from 81 patients as some of the larger specimens originating from women undergoing surgery for VIN or SCC contained multiple pathologies (e.g. VIN or LS and concurrent SCC) and so contributed more than one tissue type for analysis (Figure 11-6).

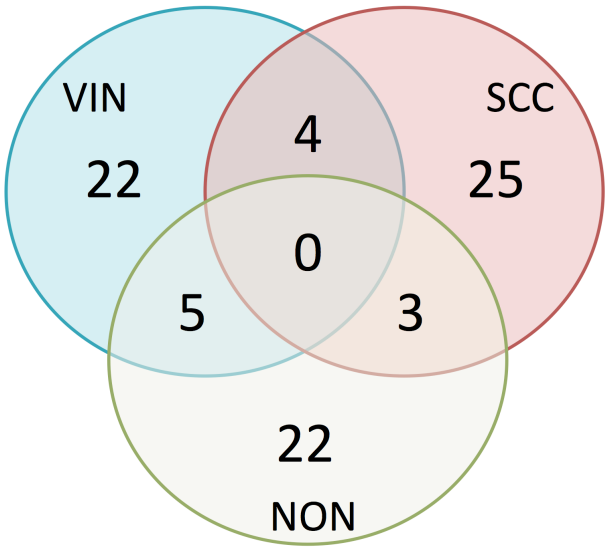


Figure 11-6 – Summary of the participants included in the investigation (VIN – vulval intraepithelial neoplasia, SCC – squamous cell carcinoma & NON – non-neoplastic vulval skin)

There were no instances where more than one sample of the same pathology originated from the same patient. The tissue maps were designed to completely sample the epidermis of the tissue sections. In order to achieve this map dimensions were individualised to each sample.

Table 11-1 – Number of samples and epidermal spectra measured.

Pathology group	Number of samples / maps measured	Total number of epidermal spectra measured
Non-neoplastic	30	245662
Vulval intraepithelial neoplasia	31	881406
Squamous cell carcinoma	32	1328825

After pre-processing (Figure 11-5), the mean spectra from each of the samples were visually reviewed to identity any technical errors in spectral acquisition prior to the inclusion of the data in the classification model. The mean spectra from all of the samples exhibited typical features of a tissue spectrum and therefore were all included in the analysis (Figure 11-7).

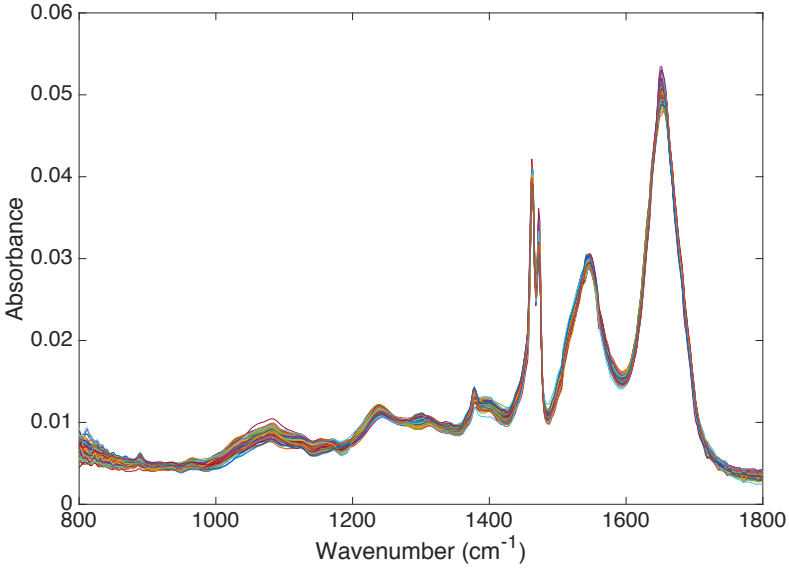


Figure 11-7 – Mean epidermal spectra from each of the 93 samples measured

11.6 Results

The initial analysis assessed the classification performance of model for discriminating between non-neoplastic vulval skin, VIN and vulval SCC. The reason for this approach is that biopsies from women with either non-neoplastic vulval skin or VIN will often be taken to investigate for the presence of vulval SCC. The ability to use FTIR to triage or offer a second opinion on biopsies taken in this situation would be advantageous as previously discussed (chapter 3.4.3).

As described (chapter 11.4), PCA was used to explore the variance in the data set with up to 24 PCs calculated. The PCs that reach the 95% confidence threshold set by ANOVA when comparing the PCA scores between the pathological groups as determined by the critical value of the F ratio were included in the LDA. The ability of

FTIR to differentiate between the tissue types was assessed in a leave one sample out cross validation loop (chapter 4.2.5).

The initial PCA of the spectral data revealed that 95% of the variance in the data was accounted for in the first ten PCs with 85% of the variance in the first five PCs (Figure 11-8). When ANOVA was applied to the PC scores and the known pathology groups of the whole data set the significance of each PC for the classification of the disease state was assessed. The ANOVA F ratios for each PC are shown in Figure 11-9. As much of the variance in the data set is due to variation within individual samples or between samples of the same pathology the most significant PCs as selected by ANOVA do not correspond to those that describe the most variance in the data (Figure 11-8 & Figure 11-9).

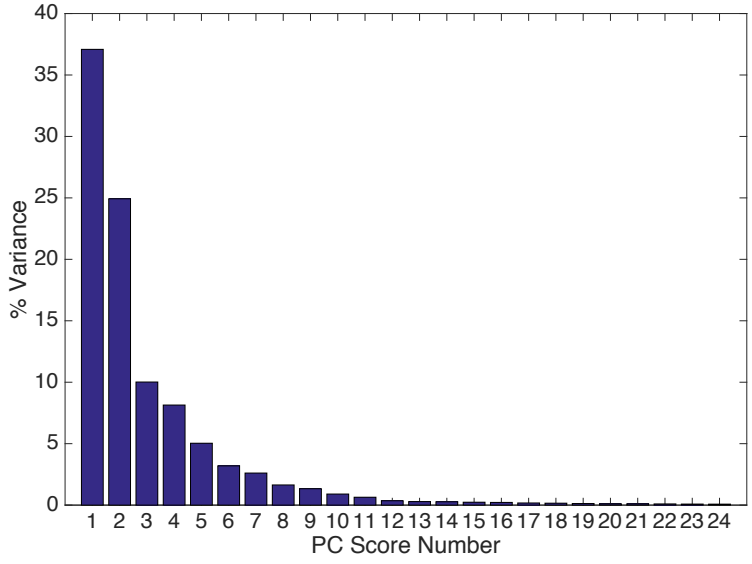


Figure 11-8 – Variance in data described by each principal component.

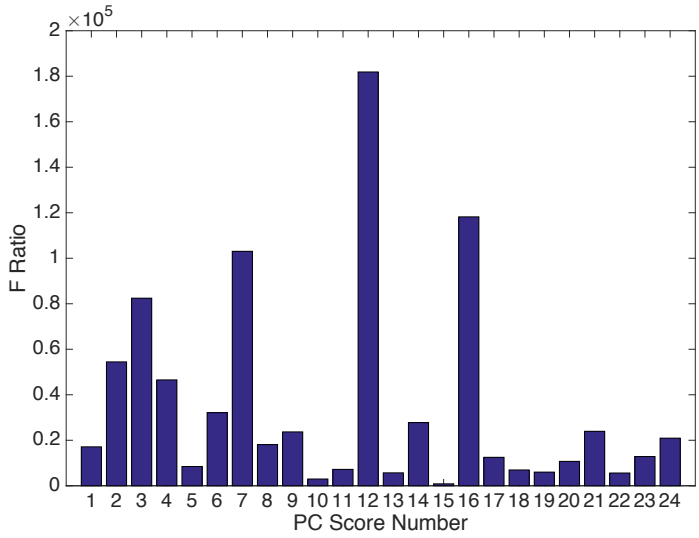


Figure 11-9 – ANOVA F ratios for principal component scores

LDA of the significant PC scores revealed the pathology groups could be separated by LD scores as shown in the histogram of the LD scores by pathology group (Figure 11-10). The LD scores plot showed some overlap between the groups that is most significant between the non-neoplastic group and the VIN group, the model performed less well in separating these pathologies.

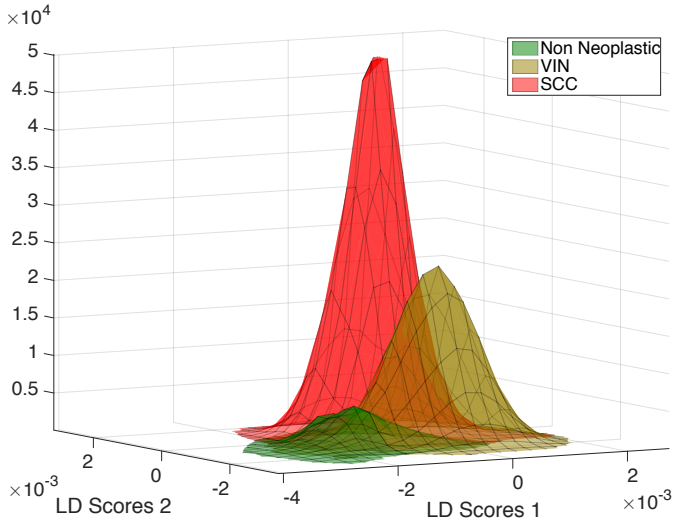


Figure 11-10 – Histogram of linear discriminant scores

In order to separate the three pathology groups the LDA requires two linear discriminant functions. Each of these functions uses the PC scores to separate the pathology groups. The weighting given to each PC loading by the linear discriminant functions are illustrated in Figure 11-11.

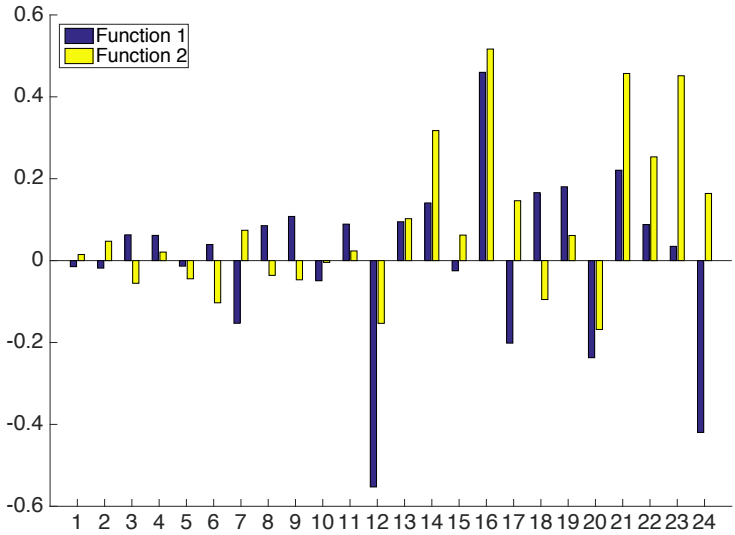


Figure 11-11 – Weight of each principal component score to the linear discriminant model.

The leave one out cross-validated PCA LDA classification model with DQ filtering of outlying spectra (chapter 4.2) showed good classification of the three pathology groups. The confusion matrix for the classification of the samples analysed with the linear discriminant boundary set midway between the pathology groups is shown in Table 11-2.

Table 11-2 – Confusion matrix for whole samples based on modal classification of spectra with linear discriminant boundary midway between the pathology groups, rows are true pathology group and columns are predictions.

Pathology group	Non-neoplastic	Vulval intraepithelial neoplasia	Squamous cell carcinoma
Non-neoplastic	27	2	1
Vulval intraepithelial neoplasia	5	23	3
Squamous cell carcinoma	2	4	26

The overall probabilities of group membership were calculated from the combined individual discriminant distances of the spectra in each group and plotted on a ternary plot to allow visualisation of the distribution of correctly and incorrectly classified samples (Figure 11-12). The ternary plot demonstrates that it is possible to define alternative discriminatory boundaries to tailor the classification to the clinical need. In Figure 11-12 an alternative classification boundary is shown that increases the sensitivity of detecting cancer, the result of which are shown in Table 11-3. With the modified boundary SSC is detected with a sensitivity of 100% and specificity of 85% with a positive predictive value of 0.78 and false positive rate of 0.15. This approach would be suitable for screening samples for further urgent histopathological assessment.

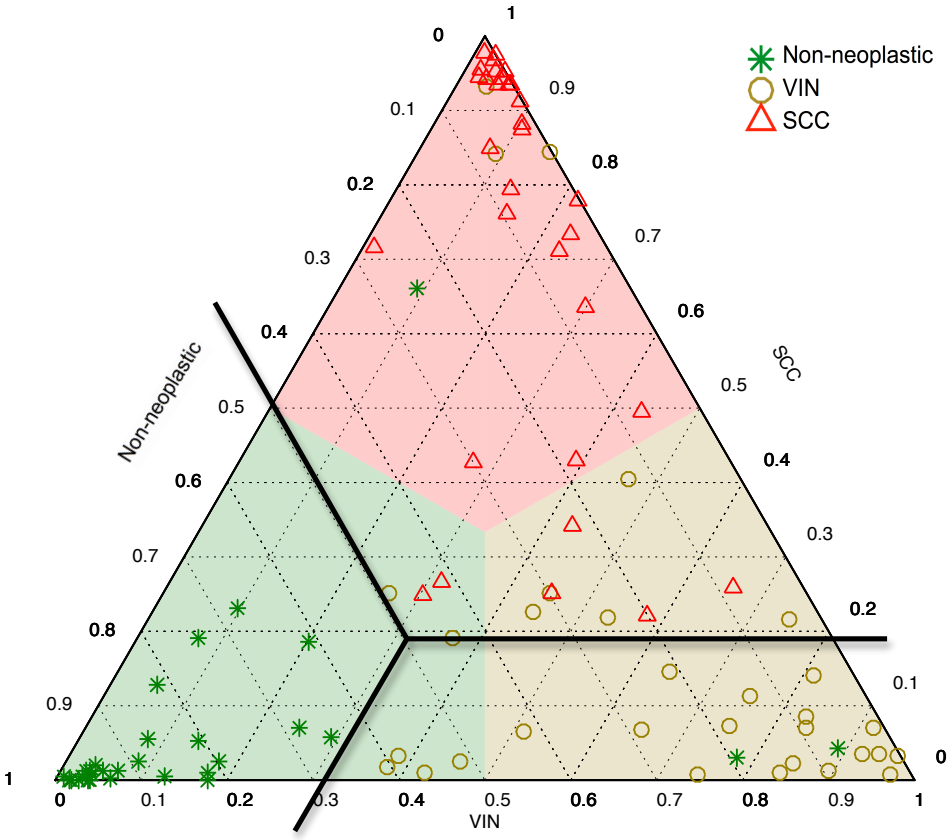


Figure 11-12 – Ternary plot of the predicted probabilities of a sample belonging to the three pathology groups under analysis. Non-neoplastic vulval skin, VIN - vulval intraepithelial neoplasia , SCC - squamous cell carcinoma. Legend indicates actual pathology of sample shaded areas show classification of samples based on discriminant boundary midway between groups. Solid lines indicates alternative discriminatory boundary.

Table 11-3 – Confusion matrix for whole samples based on combined probability of spectra with modified discriminatory boundary between the pathology groups, rows are true pathology group and columns are predictions.

Pathology group	Non-neoplastic	Vulval intraepithelial neoplasia	Squamous cell carcinoma
Non-neoplastic	27	2	1
Vulval intraepithelial neoplasia	1	22	8
Squamous cell carcinoma	0	0	32

To further assess the classification performance of the technique receiver operator curves were plotted whilst varying the linear discriminant boundary. Comparisons were undertaken of the most clinically useful classifications as outlined in chapter 3.4.3.1. The first comparison looked at the ability of the technique as a technique for triaging samples and highlighting those that are likely to contain SCC for urgent histopathological assessment. The performance of the technique for the correct identification of SCC from non-neoplastic skin and VIN on a per spectrum basis is shown by the receiver operator curve in Figure 11-13. For individual spectra the area under the receiver operator curve is 0.89 and sensitivities of 81-90% can be achieved with specificities of 82-66%. In the context of triaging whole samples it is more useful to consider the diagnostic performance in relation to all spectra measured from each sample and this more analogous to true histopathological assessment. Each sample was classified as SCC or not SCC based a threshold number of spectra that were classified as SCC. Using this approach and varying the threshold proportion of spectra to classify the sample as SCC, a receiver operator curve was produced (Figure 11-14). The area under the curve is 0.96 indicating good performance for the identification of samples of SCC from benign pathologies. Using this approach a sensitivity of approximately 100% can be achieved with a corresponding specificity of 79%.

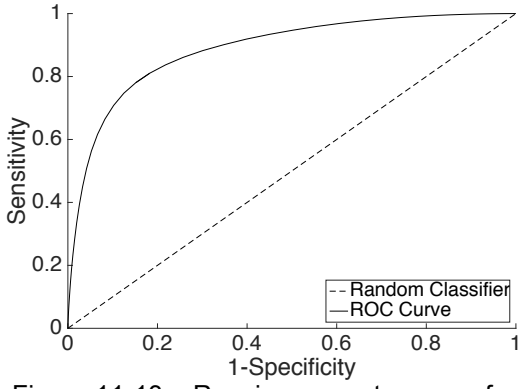


Figure 11-13 – Receiver operator curve for the detection of SCC from non-neoplastic skin and VIN per spectrum. Area under curve, 0.89, Risk of falsely rejecting null hypothesis, <0.05

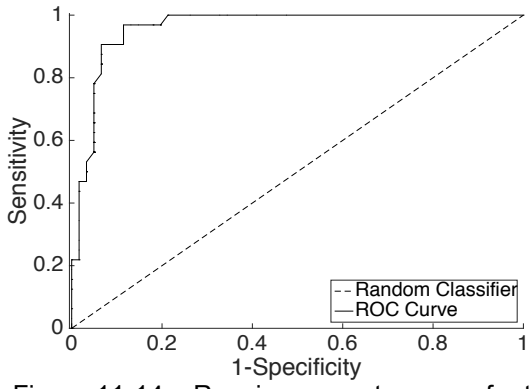


Figure 11-14 – Receiver operator curve for the detection of SCC from non-neoplastic skin and VIN per sample. Area under curve 0.96, Risk of falsely rejecting null hypothesis <0.05

The second comparison undertaken assessed the role of FTIR for correctly identifying SCC from VIN to confirm or assist the histopathologist in making the correct diagnosis using a spectroscopic second opinion. In this comparison the receiver operator curves for the technique is shown in Figure 11-15 and Figure 11-16. For individual spectra the area under the receiver operator curve is 0.89 and sensitivities of 84-90% can be achieved with specificities of 80-70%. Using the whole sample approach outlined above the area under the receiver operator curve is 0.96 and a sensitivity of approximately 100% can be achieved with a corresponding specificity of 77%. These results indicate that the technique shows promise as an adjunct to histopathology for the identification of SCC from VIN with high sensitivity and specificity.

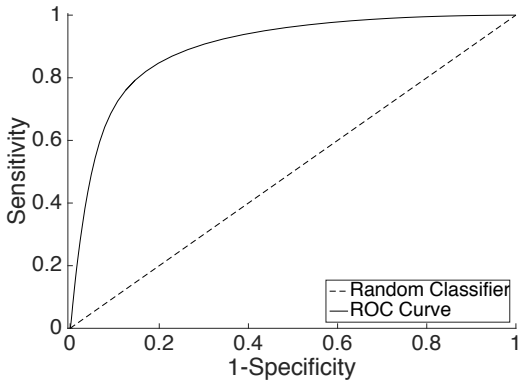


Figure 11-15 – Receiver operator curve for the detection of SCC from VIN per spectrum. Area under curve, 0.89, Risk of falsely rejecting null hypothesis, <0.05

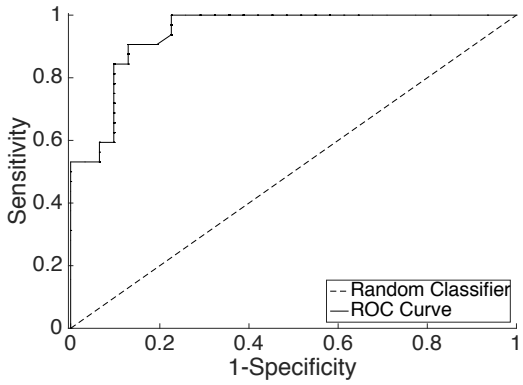


Figure 11-16 – Receiver operator curve for the detection of SCC from VIN per sample. Area under curve 0.94, Risk of falsely rejecting null hypothesis <0.05

11.7 Spectral analysis of biomolecular differences between pathologies

Analysis of the spectral differences between pathology groups was undertaken to understand the basis for disease classification from FTIR spectra. Understanding which biomolecules contribute to which peaks can be achieved though consulting the relevant literature and measurements of pure biomolecules made on the same FTIR spectrometer. Identification of specific biomolecular contributions to the spectra can be challenging, as at many wavelengths there is significant overlap between a number of possible bond vibrations and hence possible biomolecules present within the tissue. This overlap means that a given spectral appearance may have more than one possible molecular cause and hence assignment of spectral peaks to individual molecules is purely tentative.

11.7.1 Spectral peak assignment

The mean spectra from each pathology group were analysed to identify the key spectral peaks present. The largest peak seen in all pathology groups is the amide I peak at 1650 cm^{-1}. The amide I peak is primarily due to protein and the stretching of the C=O bond however there are also contributions from DNA, glycogen and glycoproteins demonstrating the complexity of spectral peak assignment (Barry, Edwards and Williams, 1992). The second largest peak excluding the paraffin peaks between 1358 and 1394 cm^{-1} is the Amide II peak at 1547 cm^{-1}. The Amide II peak is produced from the C-N and N-H stretching in tissue proteins (Movasaghi, Rehman and ur Rehman, 2008a). Peaks attributable to DNA are broadly spread across the spectrum and can be found at 965, 980, 1071, 1084, 1095 and 1245 cm^{-1} (Skrebova, Aizawa and Aras, 2003; Eikje, Aizawa and Ozaki, 2005). Other putative peak assignments are shown in Table 11-4.

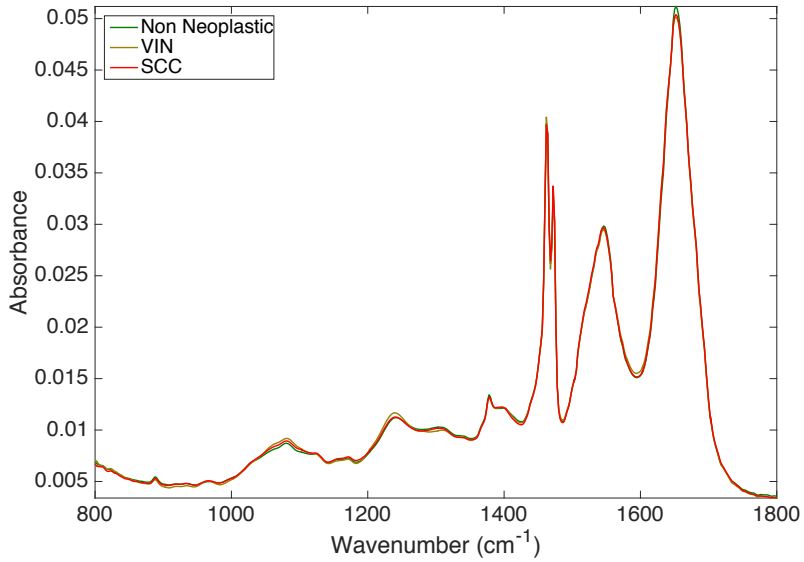


Figure 11-17 – Mean normalised spectra of the pathology groups. Non-neoplastic; VIN – vulval intraepithelial neoplasia and SCC – squamous cell carcinoma.

Table 11-4 – Putative FTIR spectral peak assignments for vulval epidermis

Wavenumber (cm^{-1})	**Molecular attribution**	**Reference**
965,1071,1084,1095,1245	DNA & Amide III	(Skrebova, Aizawa and Aras, 2003; Eikje, Aizawa and Ozaki, 2005)
980	Ribose group of nucleic acids	(McIntosh *et al.*, 1999; Eikje, Aizawa and Ozaki, 2005)
995, 1028, 1157	Glycogen	(Movasaghi, Rehman and ur Rehman, 2008a)
1080	Phosphate group of nucleic acids	(McIntosh *et al.*, 1999; Eikje, Aizawa and Ozaki, 2005)
1084	Glycogen	
1000-1150	DNA/protein (chromatin)	(Eikje, Aizawa and Ozaki, 2005)
1240	Phosphate group of nucleic acids	(McIntosh *et al.*, 1999; Eikje, Aizawa and Ozaki, 2005)
1278	Collagen and Amide III	
1480-1575	Amide II	(Eikje, Aizawa and Ozaki, 2005)
1600-1700	Amide I	(Eikje, Aizawa and Ozaki, 2005)
1634	Collagen	(McIntosh *et al.*, 1999)
1740	Lipid C=O	(Eikje, Aizawa and Ozaki, 2005)

11.7.2 Spectral differences between groups

As the differences between the mean spectra of the different groups were found to be subtle and not easily identifiable from the mean spectra (Figure 11-17) the spectral differences between the groups were assessed in more detail by subtracting the mean spectrum from each pathology group and creating a difference spectrum. The spectral differences used by the classification model were explored by examining the composite PCA LDA loadings and individual PCA loadings.

The difference spectra of the mean of the non-neoplastic spectra from the mean of the SCC (Figure 11-18) spectra indicates stronger peaks at 964, 1066, 1094, 1184 and 1220 cm^{-1} consistent with changes in the protein content between the two groups and a higher signal from nucleic acids as you would expect from an increase in nuclear material in SCC. Shifts in the Amide I and II bands are also seen in the higher wavenumber range consistent with changes in proteins and amino acids with the transition to SCC. Similar changes can be seen in the difference spectra of the mean of the non-neoplastic spectra from the mean of the VIN group indicating an analogous biomolecular transition from non neoplastic skin to SCC and VIN (Figure 11-19).

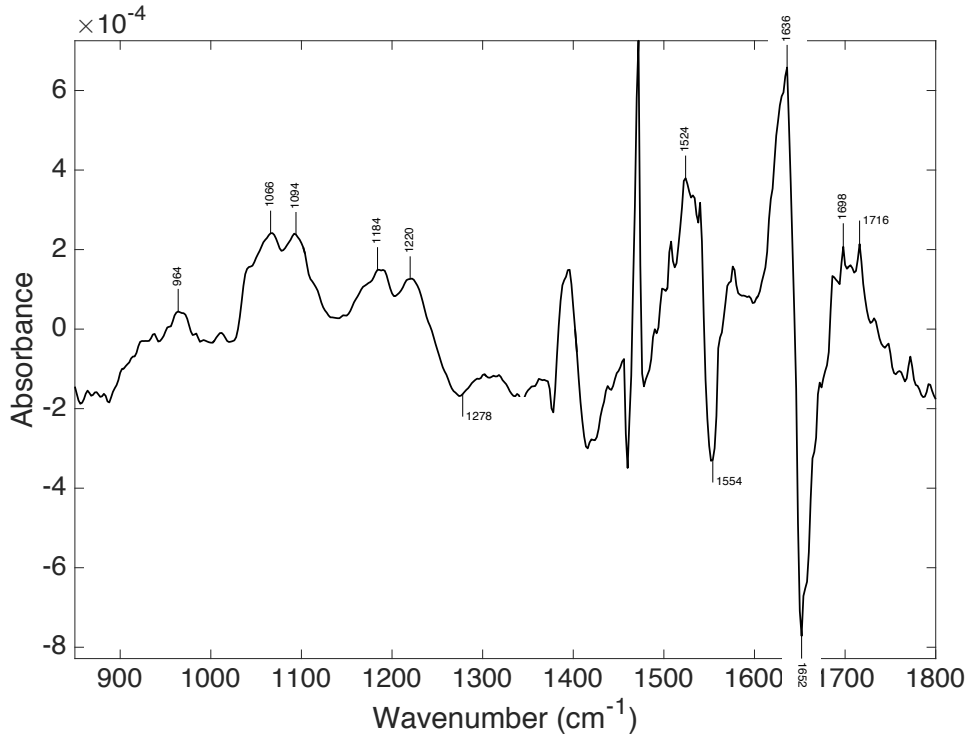


Figure 11-18 – Difference spectrum of the mean of the non-neoplastic group spectrum from the mean of the SCC group spectrum

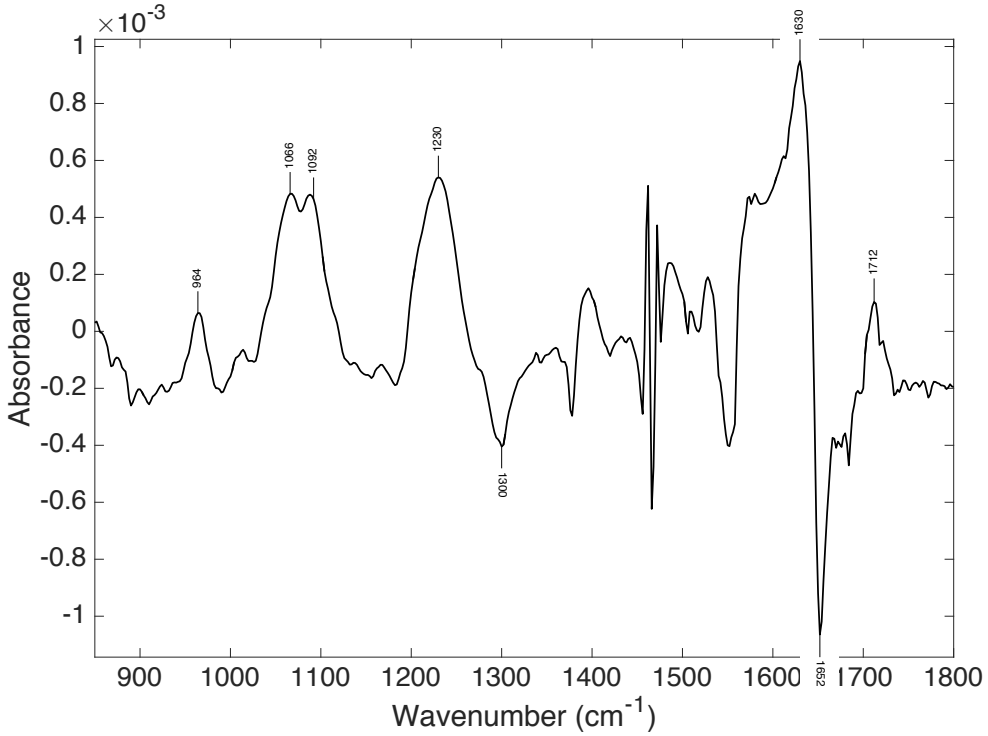


Figure 11-19 – Difference spectrum of the mean of the non neoplastic group spectrum from the mean of the VIN group spectrum

Composite loadings for each of the linear discriminant functions were created (Figure 11-20) using the weighting given to each PC by the linear discriminant functions (Figure 11-11) combined with all respective PC loadings. Examining these composite PCA LDA loadings together with the four most significant PC loadings identified by ANOVA (Figure 11-21) we can assess areas of the spectrum important in separating the pathology groups. These plots indicate that shifts in the Amide I and II peaks combined with changes in the absorption from regions of the spectrum associated with DNA, glycogen and other proteins are important in the classification of the different pathology groups.

The most significant PC loadings were examined in more detail by considering their contribution to the classification of the three pathology groups. The LD scores plot (Figure 11-10) shows a positive LD score from the first LD function is associated with classification of VIN from non-neoplastic vulva and a positive score from the second LD function is associated with classification of SCC from VIN and non-neoplastic tissue. Using the LD scores and the LD weighting (Figure 11-11) of the PC loading to this we can deduce the significance of a positive or negative shift in the PC loading in the LDA classification.

For PC loading 12 (Figure 11-21) the weighting for both the first and second LD functions is negative therefore a negative shift in the loading is associated with classifying a spectra as being from and area of SCC or VIN. As the weighting of LD function one is greater than the weight for function two for the loading it is likely it is more strongly associated with classifying VIN from non-neoplastic tissue. The

loading contains negative shifts at 925-9 cm^{-1} (left handed helix DNA), 1040 cm^{-1} (Symmetric phosphate stretching in RNA and DNA), 1086 cm^{-1} (phosphate stretching modes from the phosphodiester groups in nucleic acids), and 1121 cm^{-1} (RNA) which is suggestive of an increase in the nucleic acids in premalignant and malignant tissues. Positive peaks in the loading at 972 cm^{-1}, 1018 cm^{-1} and 1204 cm^{-1} known to be associated with polysaccharides may indicate a decrease in polysaccharides in premalignant and malignant tissue which is consistent with the higher metabolic activity of these tissues (Movasaghi, Rehman and ur Rehman, 2008a).

Both the LD weightings of PC 16 loading (Figure 11-21) are strongly positive suggesting a positive change in the loading is associated classifying a sample as containing SCC. Strong negative peaks on this loading at 1110 cm^{-1}, 1122 cm^{-1} and 1419 cm^{-1} associated with polysaccharides and carbohydrate may indicate decreased signal from these molecules in SCC. A positive peaks at 1200 cm^{-1} within the Amide III region, 1586 cm-1 within the Amide II region and 1652 cm^{-1} within the Amide I region with negative peaks at 1620 cm^{-1} and 1670 cm^{-1} suggests a change in protein composition within SCC. The positive peak at 1040 cm^{-1} (Symmetric phosphate stretching in RNA and DNA) is consistent with the changes seen in PC loading 12.

In PC three (Figure 11-21) the situation is more complicated with a positive weight for the first LD function and a negative weight for the second. Therefore a positive shift in the loading is associated with classifying a sample as VIN. In this loading we

see shifts in the Amide III band at 1184 cm^{-1}, Amide II band at 1558 cm^{-1} and Amide I band at 1670-1690 cm^{-1}. More specific changes are difficult to interpret.

In PC seven (Figure 11-21) a negative weighting in the first LD function and a positive weighting in the second LD function imply a positive shift in this loading is associated with classifying non-neoplastic vulva from VIN and SCC from VIN. This does not link well with the known pathophysiological understanding of the pathologies and biomolecular interpretation of the loading is unlikely to be useful in understanding the molecular basis for classification.

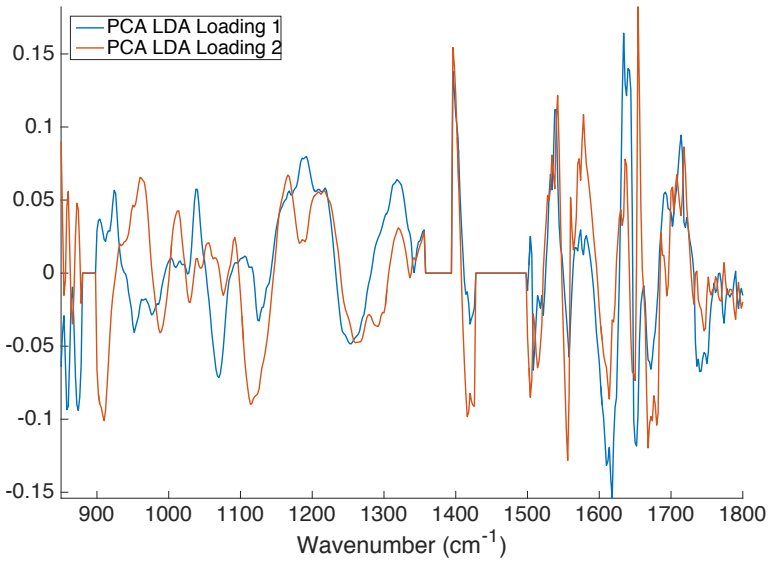


Figure 11-20 – Composite PCA LDA loading

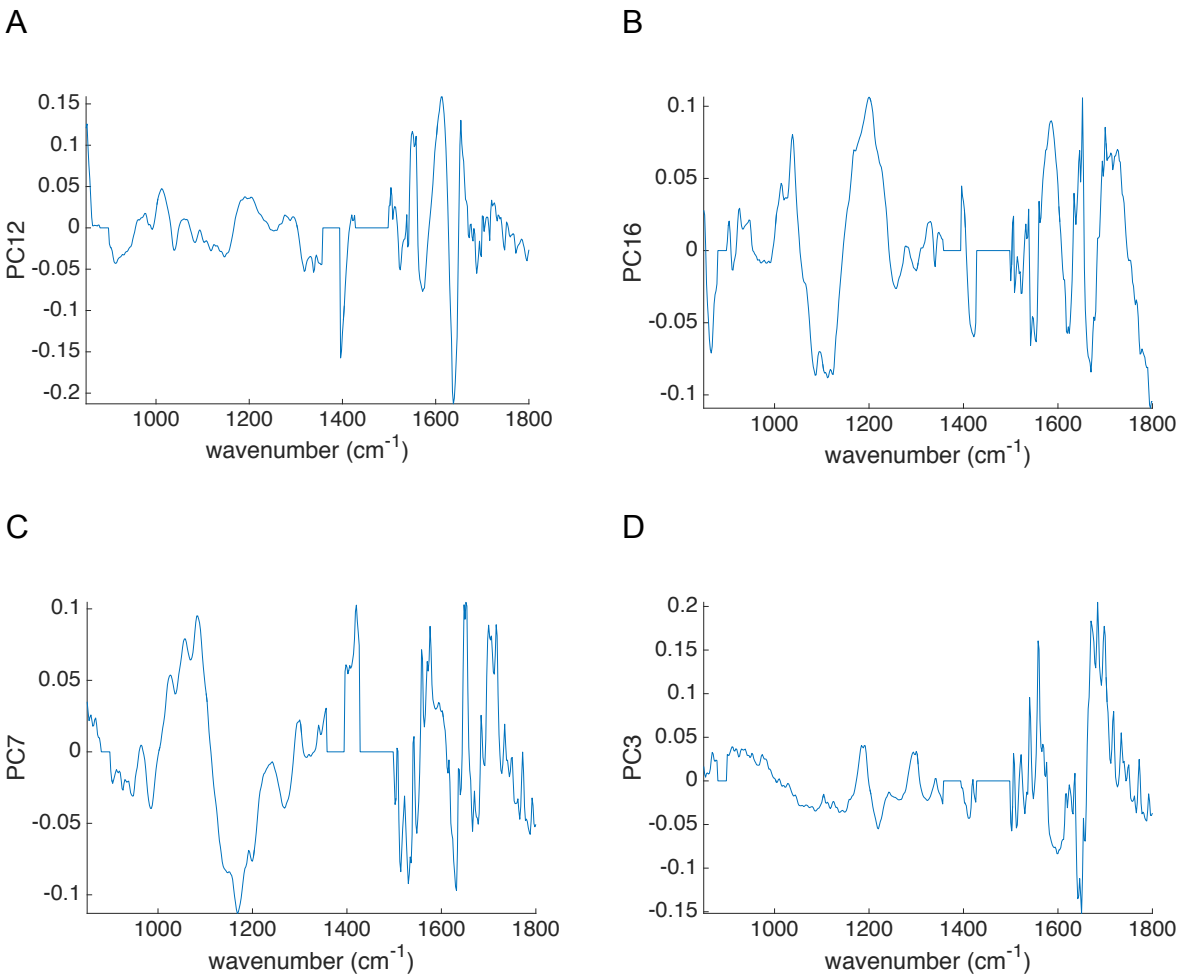


Figure 11-21 – Principal component loadings of four most significant scores as determined by ANOVA.

11.8 Interim discussion

[Summative discussion and conclusions relating to the investigations outlined in Section D can be found in Chapter 13]

In this investigation the aim was to evaluate FTIR spectroscopy for the purposes of classifying non-neoplastic vulval skin, premalignant and malignant disease of the vulva. This was accomplished with the mapping of sections of tissue that had

undergone the standard fixation and paraffin embedding required for routine histopathology. This is the first reported study to attempt classifying these vulval pathologies using FTIR spectra collected from fixed tissue sections. In this investigation it was possible to use the individual spectra and spectral maps to accurately classify the pathology present within the tissue with clinical relevant classification accuracy.

In the first comparison it was demonstrated that the technique could be used to differentiate SCC from VIN and non-neoplastic vulval tissue using individual epidermal spectra with sensitivities of 81-90% and specificities of 82-66%. On a per sample basis the sensitivity can be improved to approximately 100% with specificity of 79%. This demonstrates the technique may have a useful roll in screening samples for urgent histopathological review.

In the second comparison the ability of the technique to identify SCC from VIN was demonstrated. Using individual spectra the technique could achieve sensitivities of 84-90% with specificities of 80-70%. Using the whole sample map a sensitivity of approximately 100% can be achieved with a corresponding specificity of 77%. The differentiation of SCC from VIN is important in the accurate diagnosis of SCC as the neoplastic changes in VIN can be difficult to differentiate from SCC and a second opinion is often required. This investigation demonstrates FTIR has promise in offering the opportunity to provide an automated second opinion.

In addition, the analysis of the FTIR spectra also appears to demonstrate a continuum of biomolecular changes from non-neoplastic tissue through VIN to frank

SCC (Figure 11-12). This fits with our understanding of the carcinogenic processes leading to the development of vulval SCC (chapter 1.4).

Further discussion of future work and the steps necessary to translate these finding into clinical practice is located in Chapter 13.

Chapter 12 FTIR Augmented Pathology

12.1 Introduction and objectives

There are thought to be two distinct subtypes of vulval SCC each with unique identifiable precursor disorders (Ueda et al. 2011). The first is linked with oncogenic subtypes of the human papilloma virus (HPV) and HPV associated vulval intraepithelial neoplasia (VIN). The second is keratinising SCC which is associated with lichen sclerosus and non HPV associated (differentiated) VIN and typically occurs in older women (van de Nieuwenhof, van der Avoort and de Hullu, 2008). Each type of vulval SCC has a unique set of molecular changes that occur as vulval tissue is driven towards a malignant phenotype (Dakubo *et al.*, 2007). The ability to readily detect these molecular changes would be invaluable in the risk stratification of preneoplastic and malignant vulval disease and in the early diagnosis of vulval SCC.

The current gold standard for the diagnosis of precancerous and malignant vulval conditions is histopathological examination. Traditional histological examination gives little information on the molecular profile of the tissue being examined and a novel technique is required to probe the molecular composition of vulval tissue to

augment the histological assessment. The molecular information obtained through FTIR spectroscopy has the potential to work as an adjunct to histopathology to aid in the early diagnosis of cancer and may offer additional information about the molecular risk profile of the tissue.

Lichen sclerosus and vulval intraepithelial neoplasia are ideal subjects for investigating augmented pathological assessment with FTIR due to the established link between these conditions and vulval SCC and the need for indefinite follow up of patient with these conditions. The early identification of those women with these conditions who develop SCC or the molecular precursors to SCC has a clear clinical advantage over traditional histopathology. As previously discussed (1.4.1.1) in usual type VIN malignant transformation of vulval epithelial cells by high-risk HPV subtypes is mediated through integration of the HPV DNA with in the host genome leading to HPV induced expression of the E6 and E7 oncoproteins. These oncoproteins have numerous oncogenic effects including interference with control of the cell cycle, producing numeric and structural oncogenic chromosomal abnormalities (Vinokurova *et al.*, 2005; Zekan, Sirotkovic-Skerlev and Skerlev, 2011). In addition, methylation or mutations of the hosts' genome can occur that distorts transcriptional control, cell differentiation and viral gene expression. This results in altered molecular expression in tissues undergoing malignant transformation (Trietsch *et al.*, 2015).

LS is thought to develop into SCC through a HPV independent process which is not fully understood (1.4.1.2) (Ueda *et al.*, 2011). Numerous genetic and epigenetic changes have been noted to occur in areas of LS associated with SCC that are not

present in LS found in isolation. These include mutation of TP53 and hypermethylation of MGMT and RASSF2A, suggesting a potential role for these genes in non HPV associated carcinogenesis however the exact mechanism is not well understood (Guerrero *et al.*, 2011; Trietsch *et al.*, 2015).

Overall the pathogenesis of vulval SCC is not as well described as that for more common cancers, however it is clear that in both the HPV dependent and HPV independent oncogenesis, biomolecular changes occur in the vulva that lead to the development of invasive cancer. Detecting these changes may give valuable prognostic information for women with precursor conditions such as LS and VIN as well as for those who have already developed SCC.

This study explores the role of Fourier transform infrared spectroscopy (FTIR spectroscopy) and multivariate analysis in the evaluation of molecular changes in HPV dependent and HPV independent vulval carcinogenesis.

12.2 Objectives

This investigation assesses objective 2 of the aims and objectives outlined in Chapter 9.

Investigate the ability of FTIR spectroscopy and multivariate analysis to identify molecular changes in precancerous vulval conditions that are associated with malignant transformation.

This experiment is designed to establish if there are molecular changes detectable by FTIR spectroscopy in HPV related (uVIN) and HPV independent (LS) vulval disease that are associated with the development of vulval SCC.

12.3 Sample preparation and FTIR spectral acquisition

The pathology database at Gloucestershire Hospitals NHS Foundation Trust was interrogated and 24 cases of uVIN and 24 cases of LS with and without concurrent vulval SCC were selected. uVIN was chosen as the commonest identifiable HPV dependent precancerous vulval disease and LS was chosen as the commonest HPV independent preneoplastic vulval disease. The haematoxylin and eosin (H&E) stained slides from the cases identified were scrutinised by a gynaecology histopathologist and areas typical of uVIN and LS selected. Contiguous 4 μm tissue sections were cut from the corresponding paraffin embedded blocks onto a glass slide, a calcium fluoride substrate and a further glass slide. The sections cut onto glass were H&E stained and examined by a histopathologist blinded to the previous histological diagnosis to confirm the pathology present. The sections cut onto to the calcium fluoride substrates underwent FTIR spectroscopic mapping using a Perkin Elmer Spectrum One Spotlight 400 imaging system in transmission mode using the experimental conditions outlined in chapter 10.5. The number of spectra collected from each sample was determined by the size of the lesion identified on the tissue sections. In total tissue sections from 48 different patients were analysed (Figure 12-1).

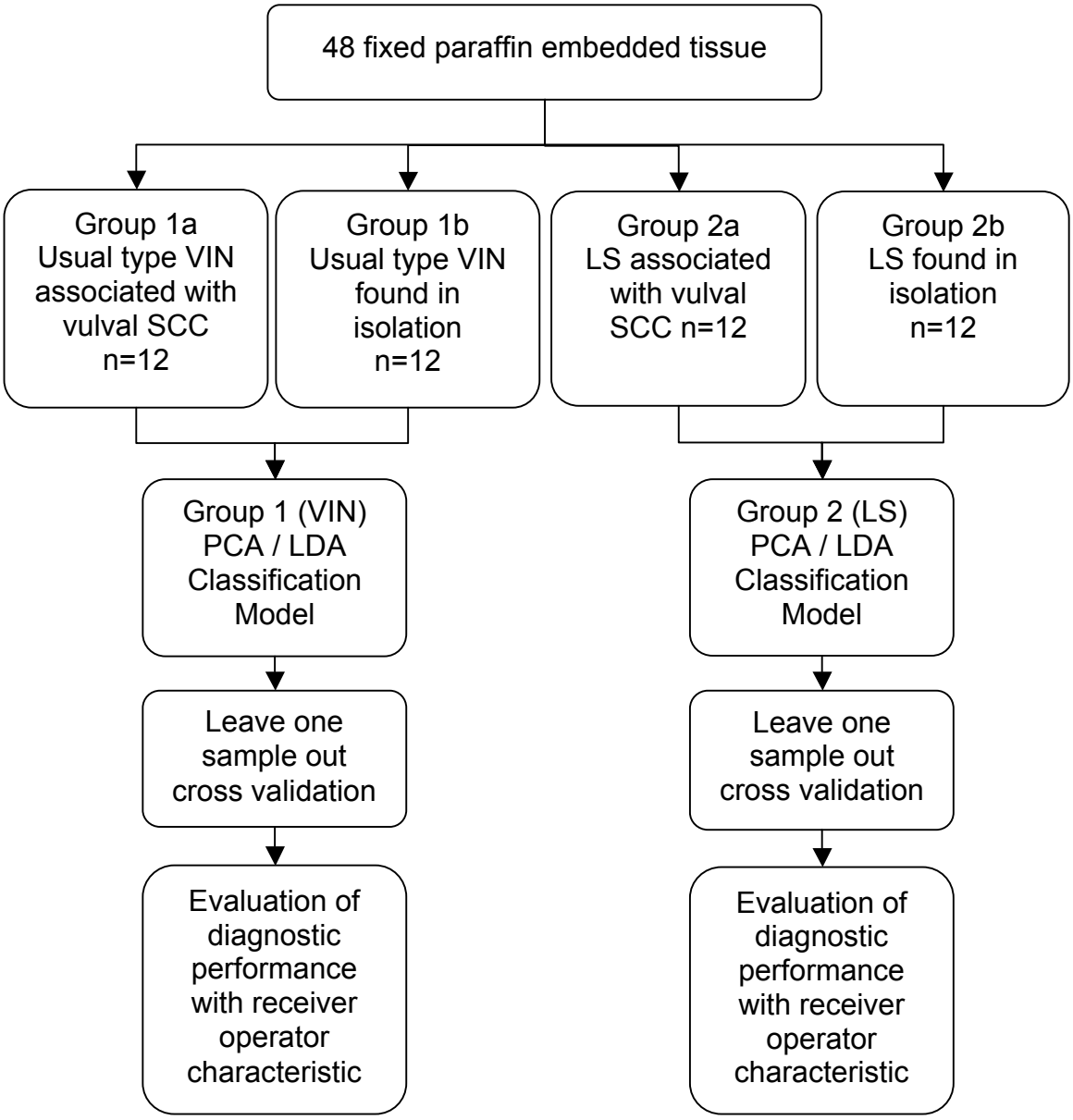


Figure 12-1 – Overview showing the four sample groups that underwent FTIR spectroscopic mapping and subsequent analysis (PCA - principal component analysis, LDA - Linear discriminant analysis)

A B

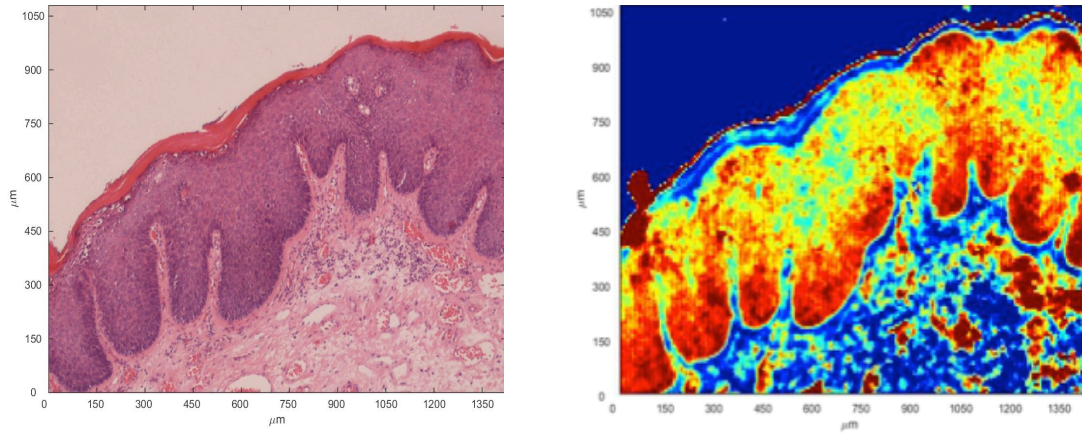


Figure 12-2 – H&E stained tissue section (A) and corresponding spectral principal component scores plot (first principal component) (B)

12.4 Data analysis

Data collected for the LS group (1) and the uVIN group (2) were analysed independently (Figure 12-1). The measured spectra were converted into absorbance and data points outside of the 800 to 1800 cm^{-1} biological fingerprint range were excluded. The remaining spectra were then corrected for the paraffin content of the tissue using an extended multiplicative scatter correction (EMSC) to remove the effect of the variation in the paraffin in the tissue samples (Ly *et al.*, 2008; Tfayli *et al.*, 2009; Nallala, Lloyd and Stone, 2015) (chapter 10.5). The spectra were then vector normalized.

In order to select the areas of epidermis (stratum germinosum and stratum spinosum) within the spectral maps (Figure 12-2) a two stage k-means cluster

analysis was performed as outlined in chapter 11.4. The areas of epidermis identified were then isolated for further analysis and the non-epidermal spectra excluded. These areas were selected as the bulk of molecular changes that result in malignant transformation occur within the cellular epidermis. Principal component analysis (PCA) was applied to reduce the dimensionality of the data into loadings and corresponding scores in preparation for linear discriminant analysis (LDA). ANOVA was performed on the principal component (PC) scores to determine which of the PC scores were significantly different between those tissue samples from women with concurrent SCC (groups 1a and 2a) and those without (groups 1b and 2b). The confidence level for the ANOVA was set at 95%. LDA was applied to the significant principal component scores and their associated sample groups (Figure 12-1). The resultant LDA model was then used to predict whether each spectrum was taken from a tissue sample from a woman with concurrent vulval SCC.

The classification ability of FTIR spectroscopy to detect whether a tissue sample of either LS or uVIN was from a patient with concurrent SCC was evaluated using leave one sample out cross validation loop (chapter 4.2.5). The predicted groups from each cycle of validation were then collated for analysis. The specificity and sensitivity of the technique for detecting tissue associated with a SCC was calculated across a range of diagnostic thresholds (i.e. the proportion of spectra classified as being associated with SCC for the whole sample to be classified as being associated with SCC).

12.5 Results

The diagnostic models for both group 1 (uVIN) and group 2 (LS) were each derived from 24 tissue samples. The total number of FTIR spectra included in each group was 509,000 for group 1 and 65,000 for group 2. The two stage k-means clustering analysis was able to correctly isolate the epidermal spectra in all but one case where manual selection of the epidermal spectra was successfully employed. The result of the initial PCA ANOVA LDA analysis demonstrated that the technique was able to classify the majority of spectra correctly using the linear discriminant distance (Figure 12-3).

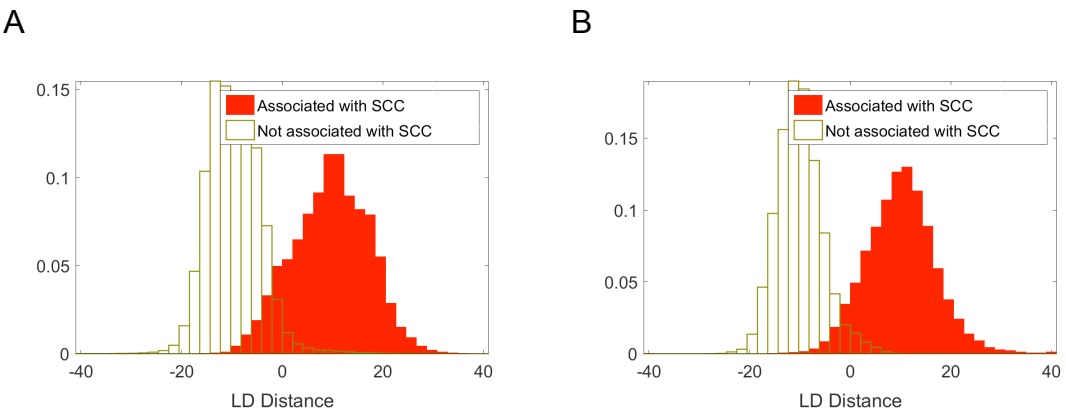


Figure 12-3 – Histogram of linear discriminant distance of individual spectra for the uVIN group (A) and the LS group (B), (LD Distance-linear discriminant distance)

Defining the biochemical constituent peaks responsible for discrimination within the LDA model is difficult. As previously discussed FTIR spectroscopy does not allow simple biochemical quantification of target tissues due to the compound effect on the absorption spectra from the numerous molecular constituents of the tissue. Examining the two most significant principal component loadings and the composite

linear discriminant loading we can speculate on the molecular basis for classification within the LDA model (Figure 12-4). The spectral differences between those with concurrent SCC and those without appear to occur primarily in the amide I, amide II and amide III regions of the absorbance spectra. In addition changes are seen in lipid, fatty acid, nucleic acid and carbohydrate regions. These findings are suggestive of a difference in protein expression as uVIN and LS progresses towards malignant transformation.

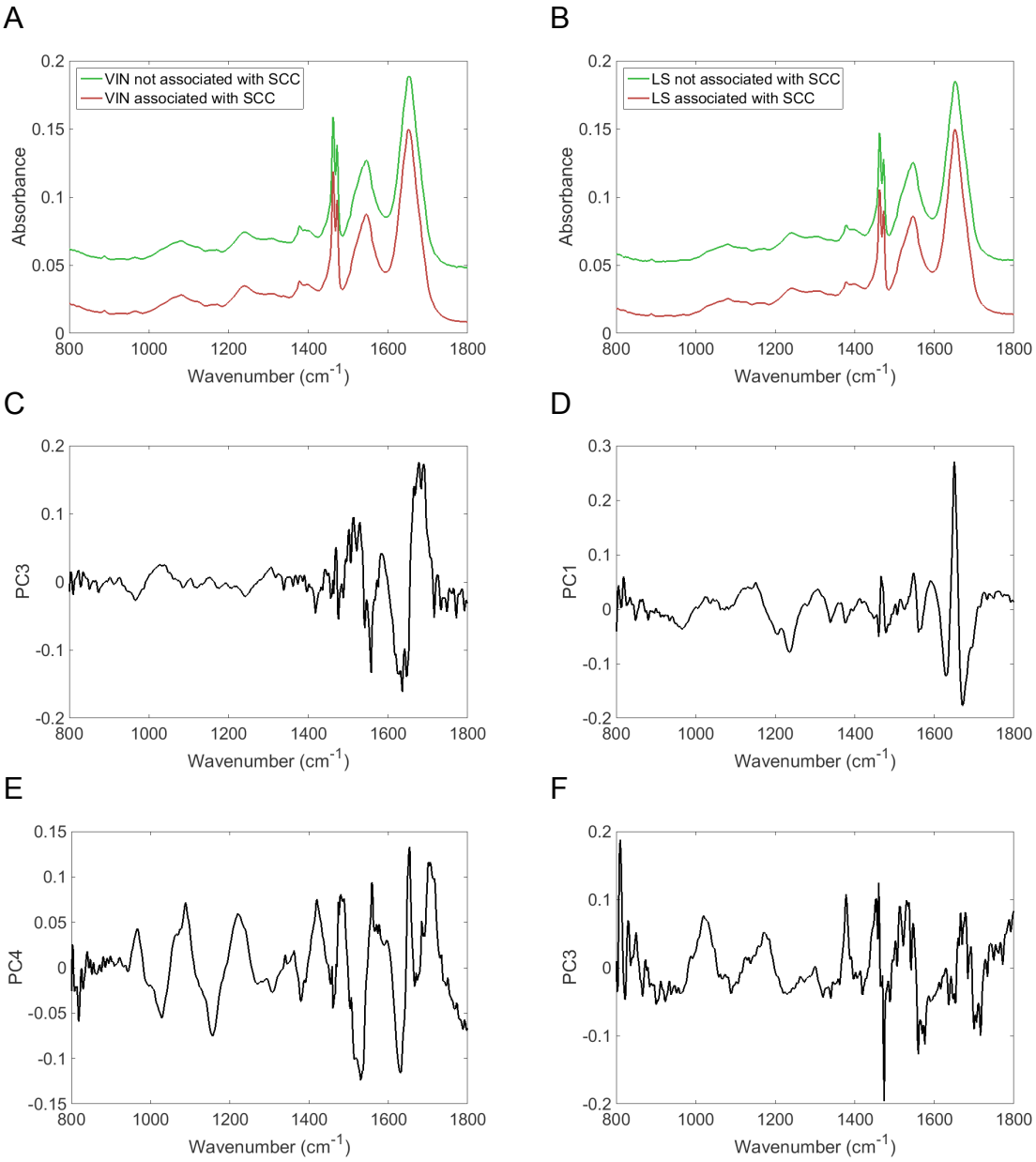
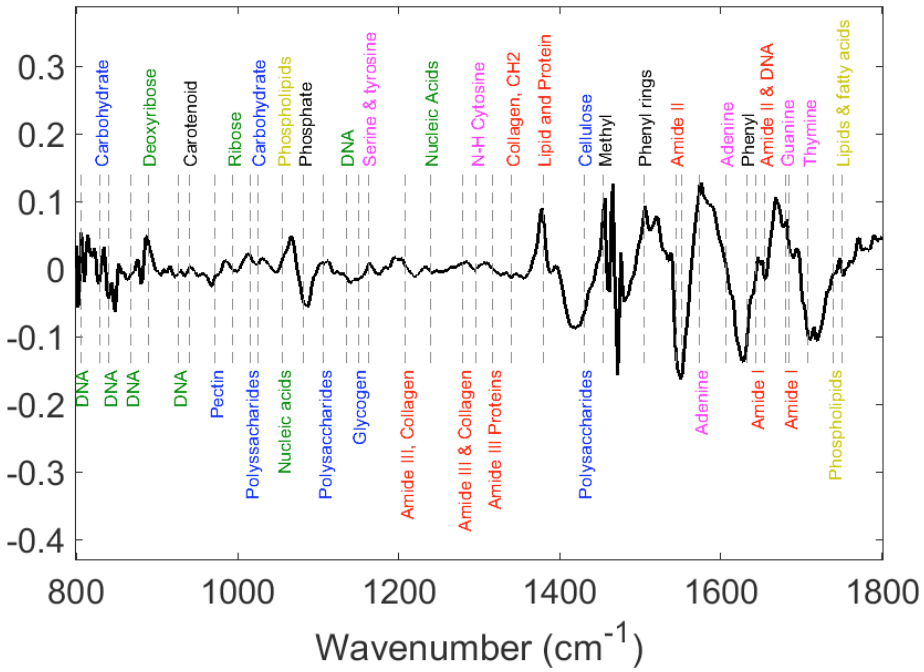


Figure caption on next page

G



H

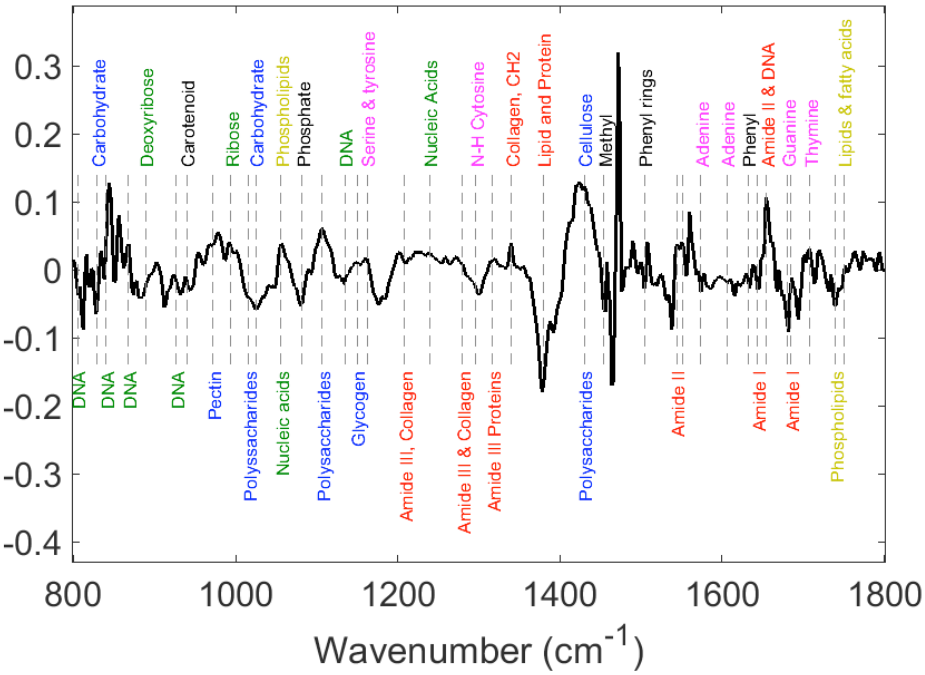


Figure 12-4 – Mean spectra for the uVIN group (A) and the LS group (B) (offset to separate spectra), the two most significant principal components for the uVIN model (C and E) and for the LS model (D and F), and composite LDA loading for the uVIN group (G) and for the LS group (H)

Table 12-1 – Proposed principal spectral changes responsible for discrimination between uVIN and LS found in association with SCC and that found in isolation with suggested biomolecular constituent assignments

Wavenumber (cm^{-1})	Suggested biomolecular constituent allocation
Group 1 (uVIN)	
966	C-O deoxyribose, C-C DNA
968	Left-handed helix DNA (Z form)
1000–1140	Protein amide I absorption
1022-4, 1028	Glycogen absorption
1040-100, 1025, 1107	Carbohydrates (including glucose, fructose, glycogen, etc.)
1056/7	Stretching C-O deoxyribose
1080	Collagen & phosphodiester groups of nucleic acids
1117	C-O stretching vibration of C-OH group of ribose (RNA)
1172	Serine, threonine, and tyrosine residues of cellular proteins
1204, 1339	Collagen, proteins-amide III and polysaccharides
1235	Composed of amide III as well as nucleic acids
1236	Amide III and nucleic acids
1236-42	Collagen and nucleic acids
1307-17	Amide III band components of proteins
1390	Carbon particle
1451, 1455-6	Methyl groups of proteins
1480-600, 1540, 1549	Amide II regions
1545	Protein band
1600-720, 1630–700	Amide I regions
1730, 1743	Fatty acid esters and lipids
Group 2 (LS)	
968	DNA
985	polysaccharides-cellulose
1018	polysaccharides, pectin
1000–140, 1600-720, 1630–700, 1656, 1670, 1717	Protein Amide I absorption
1030	Glycogen vibration, Collagen & phosphodiester groups of nucleic acids
1040-100	Carbohydrates (including glucose, fructose, glycogen, etc.)
1066, 1220, 1235	Nucleic acids
1084–6	Phosphate of nucleic acids
1164	Serine, threosine, & tyrosine of proteins
1200, 1284, 1339	Collagen
1235, 1284, 1307-17	Amide III
1396, 1455-6	Methyl groups of proteins
1419	Polysaccharides, pectin
1470	Methylene chains in lipids
1504	Phenyl rings
1545	Protein band
1480-600, 1517, 1540, 1549	Amide II regions
1670, 1750	lipids, fatty acids

Wavenumber assignments from (Dieter Naumann, 2001; Movasaghi, Rehman and ur Rehman, 2008b)

The ability of FTIR spectroscopy combined with multivariate analysis to identify tissue samples from women with SCC was evaluated using leave one sample out cross validation. After validation the discriminant model demonstrated FTIR spectroscopy was able to correctly differentiate spectra from uVIN associated with SCC from that not associated with SCC with a sensitivity of 75% and specificity of 94%. Similarly FTIR spectroscopy was able to correctly differentiate spectra from LS associated with SCC from that not associated with SCC with a sensitivity of 82% and specificity of 93%. Each tissue sample was then classified as being associated with SCC or found in isolation by the number of spectra the LDA model classified as being associated with SCC. The threshold number of spectra to classify the tissue sample was adjusted and the receiver operator characteristic was determined for both groups. This demonstrated the sensitivities and specificities that can be achieved using this technique per sample. LS associated with SCC can be differentiated from LS found in isolation with a sensitivity of 100% and a specificity of 84% with area under the curve (AUC) of 0.98. uVIN associated with SCC can be differentiated from uVIN found in isolation with a sensitivity of 100% and a specificity of 58% with AUC of 0.87. The specificity of the technique can be improved by conceding sensitivity (Figure 12-5). In the uVIN group, a higher threshold number of spectra to classify the samples can be used to give a sensitivity of 75% and a specificity of 83%.

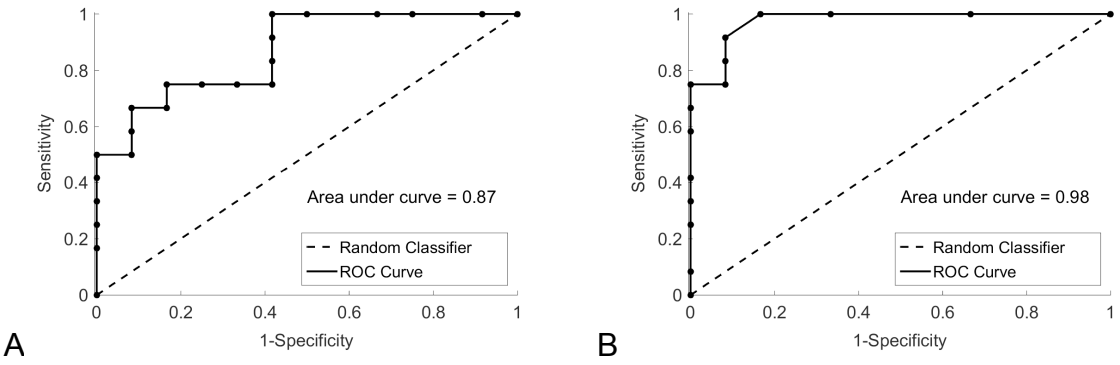


Figure 12-5 – Receiver operating characteristic for the identification of tissue associated with SCC in the uVIN group (A) and the LS group (B)

12.6 Interim discussion

[Summative discussion and conclusions relating to the investigations outlined in Section D can be found in Chapter 13]

This preliminary study has demonstrated FTIR spectroscopy is able distinguish between LS or uVIN found in isolation and LS or uVIN found with concurrent SCC at another anatomical location on the vulva. Discrimination between these groups is based on the molecular variations detected by FTIR spectroscopy and multivariate analysis. This is the first reported study to demonstrate the classification of these groups of histologically similar vulval tissue and separate them based on the concurrence of vulval SCC. The ability to separate these groups according to the presence of concurrent SCC is suggestive of cancer associated molecular changes occurring across the vulva. Demonstrating the ability of FTIR spectroscopy to differentiate between vulval disease found with concurrent cancer and vulval disease found in isolation highlights the potential role FTIR spectroscopy could have as a tool

for assessing preneoplastic molecular changes in the vulva. Potential applications include the early identification of vulval cancer and molecular risk stratification of patients with VIN or LS and the identification of those at high risk of developing vulval cancer.

Further discussion of future work and the steps necessary to translate these finding into clinical practice is located in Chapter 13.

Section E: Summative Discussion and Conclusions

Chapter 13 Summative Discussion and Conclusions.

13.1 Discussion: Raman spectroscopy for the assessment of vulval disease

The translational research presented in section C of this thesis aims to assess the potential of Raman spectroscopy for the discrimination of vulval SCC, VIN and non-neoplastic vulval in the clinical setting. The technique was evaluated in both laboratory experimental conditions and in conditions that would be technically suitable for in vivo application using a probe based system. To our knowledge this is the first study that has assessed the use of Raman spectroscopy for disease classification in the vulva. In consideration of the data reported in Chapter 6, Chapter 7 and Chapter 8 it is necessary to discuss how vulval Raman spectroscopy could be added to the clinician armamentarium for the benefit for patients with pre-malignant or malignant vulval disease. The potential role of Raman spectroscopy as a tool for assessing disease progression or monitoring response to treatment is also explored. The requirements for future refinements of the technique and for future research are also discussed and clinical and scientific conclusions drawn.

At the start of this investigation two areas of clinical need were identified. The first was the need for improved diagnostic tests to monitor and diagnose vulval conditions reducing the reliance on invasive tissue biopsy. The second was the need for a tool that is able to accurately determine disease margins at the time of surgical treatment of malignant or premalignant vulval disease. The translational research presented in Section C of this thesis aims to assess the potential role of Raman spectroscopy in addressing these needs.

The use of Raman spectroscopy for the discrimination of non-genital benign and malignant lesions has been widely investigated (Table 3-1). In the largest in vivo study to date 655 cases of benign and malignant lesions were measured from 560 patients incorporating the cohort from a previous study and an entirely new cohort of patients (Lui *et al.*, 2012; Zhao *et al.*, 2015). The study included 23 samples of in situ SCC and 41 samples of invasive SCC amongst a number of other benign and malignant skin lesions. The investigators found that after LOOCV and validation with an entirely separated training (518 lesions) and test group (127 lesions) that the technique was able to differentiate malignant from non-malignant lesions with high sensitivity and specificity, ROC AUC was 0.88. Interestingly the method of validation did not have a statistically significant impact on the performance of the technique suggesting in studies of this nature and scale LOOCV is a valid method for assessing the true accuracy of the technique. Other recent studies have also confirmed Raman spectroscopy combined with multivariate analysis can differentiate non-cancerous, pre-cancerous and cancerous skin conditions with high degrees of accuracy but none have demonstrated the technique in the vulva (Bodanese *et al.*,

2010; Silveira Jr *et al.*, 2012; Schleusener *et al.*, 2015; Silveira *et al.*, 2015; Wang *et al.*, 2015; Ferreira Lima *et al.*, 2019). Based on the studies assessing Raman spectroscopy for the diagnosis of skin cancers occurring in predominantly sun exposed areas we could be hopeful that the technique will be as successful in the vulva. There are however some important pathogenic differences between vulval cancers and cancers that occur elsewhere on the human body that necessitate experimental validation of the technique in the vulva and prohibit direct extrapolation of the results of studies performed on other skin types (Chapter 1.4.1). These include the different pathogenic pathways for carcinogenesis in the vulva compared with elsewhere in the body and the presence other conditions such as Lichen Sclerosus that occur predominantly in anogenital regions (Fleming *et al.*, 1995; Yap *et al.*, 2017; Hinten *et al.*, 2018). In this study we sought to explore if promising results achieved at other skin sites can be demonstrated in the vulva.

Three clinically important classifications were assessed in this study. These were the ability of the technique to identify non-neoplastic vulval tissue from vulval SCC and VIN; the ability of the technique to identify SCC from VIN and the ability of the technique to identify SCC from VIN and non-neoplastic vulval skin (Table 13-1).

Table 13-1 – Clinically relevant diagnostic comparisons

Diagnostic comparison	Comparator 1		Comparator 2
1	SCC and VIN	from	Non-neoplastic vulva
2	SCC	from	VIN
3	SCC	from	VIN and non-neoplastic vulva

The ability to identify vulval SCC and VIN from non-neoplastic vulval skin is important in the surgical setting where the ability to accurately determine the difference in these tissues has the potential to facilitate the minimisation of the volume of the surgical excision whilst being able to excise all neoplastic tissue. The ability of the technique to identify SCC from VIN or non-neoplastic vulval skin and SCC from VIN (comparisons 2 and 3, Table 13-1) would be valuable for both the diagnosis of vulval SCC and the monitoring of patients with VIN undergoing medical or conservative management of their disease. In addition, the technique may have utility as a research tool in providing an objective measure of response to non-surgical treatments of VIN or LS although this is beyond the scope of this thesis.

13.1.1 Raman spectroscopy for reducing clinicians reliance on invasive tissue biopsy in vulval disease.

Up to the present time clinicians have had to rely on their clinical impression of vulval lesions with recourse to invasive biopsy in order to diagnose vulval conditions. Although the established approach, the subjective evaluation of skin lesions lacks sensitivity and specificity (chapter 3.2.1). Furthermore, although a dependable approach, histological analysis necessitates invasive tissue biopsy combined with time consuming and expensive processing and reporting using subjective histological criteria. This diagnostic approach does not allow immediate diagnosis and therefore immediate treatment. In addition, invasive biopsies may not yield an accurate result if there are not taken from the most appropriate areas. Currently the site of biopsies is guided by the clinical appearances but this may miss areas of

more substantial abnormality and a non-invasive diagnostic technique could be used to guide the clinician to biopsy an area most likely to yield a diagnostically relevant result.

In those with established pre-malignant vulval disease the role for a non-invasive diagnostic technique is particularly compelling as this group undergo regular follow up to detect malignant progression. If suspicious lesions are clinically identified then invasive biopsies are taken to confirm or refute the presence of cancer. A non-invasive diagnostic technique could reduce the need for biopsies in this group.

Previous studies discussed in Chapter 3 have shown that Raman spectroscopy can be applied to the accurate diagnosis of skin conditions at other body sites but none to date have demonstrated the role of the technique in the vulva and the unique set of pathologies present in this area of the human body.

In this thesis the diagnostic capability of Raman spectroscopy was evaluated in two clinical scenarios. The first evaluated the role of the technique in diagnosing lichen sclerosus when the diagnosis was uncertain on clinical grounds alone. The second was an evaluation of the technique in the role of identifying VIN and SCC from non-neoplastic vulval skin. These studies constitute the first reported Raman spectroscopic evaluation of premalignant and malignant vulval skin and the first reported Raman spectroscopic analysis of lichen sclerosus at any body site.

The results from the first investigation demonstrate that in patients with clinically suspected lichen sclerosus, Raman spectroscopy was able to correctly identify lichen sclerosus from other diagnoses with clinically acceptable performance

(sensitivity of 91% and specificity of 80%, area under the ROC curve 0.95). This is comparable to diagnostic accuracy of commercially available diagnostic tools currently in clinical use such as the SIAscope and the SolarScan. The SIAscope is based on narrow-band spectral imaging and can identify melanoma from benign pigmented lesions with a sensitivity of 83% and specificity of 80% (Moncrieff *et al.*, 2002). The SolarScan which uses image analysis of digital surface microscopy is able to identify melanoma from other pigmented lesions with a sensitivity of 85% and specificity of 65% (Menzies *et al.*, 2005).

In our investigation, assessing the diagnosis of LS, by analysing spectra on a per patient basis and selecting a diagnostic threshold that optimizes for specificity, a near zero false positive rate can be achieved with a clinically relevant true negative rate. Using this approach we were able to correctly identify all 12 patients that did not have LS (test negative) with a negative predictive value of 0.8. In clinical practice only those with an uncertain diagnosis require invasive biopsy. These results suggest that by using Raman spectroscopy and avoiding biopsy in those who have a positive optical diagnosis of LS we could expect to reduce the number of invasive biopsies required. Using this approach in the study group only the 66% of patients that tested negative for LS would require a biopsy, reducing the number of invasive biopsies required by 44%. This has the potential to reduce the morbidity for the women; reduce the time taken to diagnose the condition and reduce the significant direct and indirect costs of invasive biopsy. In addition, this technique would allow non-invasive assessment of multiple regions of the vulva, as assessment would not be limited by the morbidity associated with invasive biopsy.

Due to the limitations of the study design it was not possible to evaluate the identification of lichen sclerosus using a probe based system. The concordance of results between the probe system and the microscope based point measurement system when comparing neoplastic and non-neoplastic lesions suggests accurate identification of lichen sclerosus should be possible using probe based spectra measurements (Chapter 8.7). Further research is however required to validate this technique for the diagnosis of lichen sclerosus using probe-based systems.

The second investigation evaluated the diagnostic role of the technique in two clinical scenarios. The first was the assessment of patients with vulval lesions suspicious for premalignant or malignant disease. In this clinical scenario it is important to be able to identify non-neoplastic lesions from neoplastic lesions and malignant disease from non-malignant disease. The second clinical scenario was the assessment of patients under long term follow up for VIN due to the risk of malignant progression. In this clinical scenario it is necessary to differentiate between SCC and underlying VIN so malignant progression can be identified appropriately and between neoplastic tissue and non-neoplastic tissue so new neoplastic lesions can be identified and treated.

This second investigation was comprised of three distinct phases. The first phase was the evaluation of the technique under optimised laboratory conditions on tissue sections where the precise area of abnormality could be identified (Chapter 6). The second phase was an evaluation of Raman spectroscopy in optimised laboratory conditions as applied to whole tissue samples with a transepidermal approach to

spectral collection (Chapter 7). The third and final phase was an evaluation of the technique using a fibre optic probe using an experimental configuration that could technically be applied to in vivo spectral collection (Chapter 8). In each phase the spectral characteristics and diagnostic performance were assessed.

In the first phase large numbers of spectra were collected from sections of vulval tissue and just those spectra from the epidermis were selected for analysis (Chapter 6). This achieved the most accurate discrimination between the pathology groups. The technique was able to identify non-neoplastic tissue from VIN or SCC with sensitivities of 82-98% with corresponding specificities of 92-72% (area under the receiver operator curve is 0.95). The technique was also able to identify SCC from VIN with sensitivities of 84-95% with corresponding specificities of 86-61% (area under the receiver operator curve is 0.92). This demonstrated the technique had promise in the chosen applications in optimised laboratory conditions and that the diagnostic performance is comparable to other diagnostic technologies currently in clinical use. This streamline approach may be useful to augment histopathological analysis however it is clearly not suitable as a clinical tool to reduce the reliance on invasive tissue biopsy.

To evaluate the technique in conditions closer to the clinical environment the second two phases of investigation were completed. In phase two the technique was able to identify non-neoplastic tissue from VIN or SCC with sensitivities of 80 to 90% with corresponding specificities of 66 to 54% (area under the receiver operator curve of 0.85). The technique was also able to identify SCC from VIN with sensitivities of 84

to 88% with corresponding specificities of 72 to 65% (area under the receiver operator curve of 0.88).

In the third phase probe measurements were taken from the tissue using a fibre optic probe with which it would be technically but not yet regulatorily feasible to take in vivo measurements. Using this approach the technique was able to identify non-neoplastic tissue from VIN or SCC with sensitivities of 86% to 95% with corresponding specificities of 72% to 57% (area under the receiver operator curve of 0.85). The technique was also able to identify SCC from VIN with sensitivities of 84% to 92% with corresponding specificities of 84% to 64% (area under the receiver operator curve of 0.87). The diagnostic performance obtained in each of these three phases is summarised in Table 13-2 where the area under the receiver operator curves for the different approaches are compared.

Table 13-2 – Comparison of the diagnostic accuracy of streamline spectroscopic mapping of epidermal cells, transepidermal Raman measurements and Raman probe measurements.

Diagnostic comparison	Comparator 1		Comparator 2	Area under ROC curve streamline mapping	Area under ROC curve point measurement	Area under ROC curve probe measurement
1	SCC and VIN	from	Non-neoplastic vulva	0.95	0.85	0.85
2	SCC	from	VIN	0.92	0.88	0.87
3	SCC	from	VIN and non-neoplastic vulva	0.95	0.90	0.91

It is challenging to assess whether a diagnostic test possesses an acceptable diagnostic accuracy for clinical application. When it comes to diagnosing or ruling out a cancerous or potentially cancerous condition a high sensitivity is expected

(Liang *et al.*, 2003). There are no published studies reporting clinician or patient views on the accuracy of diagnostic tests for vulval cancer or pre-cancerous conditions. There is however reported accuracy data on other established diagnostic interventions for comparison. The diagnostic sensitivity achieved in this study is comparable to the SIAscope and the SolarScan devices used for the detection of melanoma, see above. Cervical cytology is another example of a diagnostic test that is used to detect precancerous or cancerous changes and one that has widespread use. Cervical cytology combined with HPV virus detection appears to have 89% sensitivity and 94% specificity for the detection of high-grade precancerous changes or worse (Kitchener *et al.*, 2009). In the detection of cervical cancer, cytology is often combined with colposcopy, a commonly used technique for the detection of cervical pre-cancer or cancer. Colposcopy achieves a sensitivity of 85% and a specificity of 69% (AUC 0.80) for the detection of the normal cervix and low grade cervical intraepithelial neoplasia (CIN) from high-grade (CIN) and cancer (Mitchell *et al.*, 1998). Cytology / HPV detection and colposcopy are used in combination to increase the overall sensitivity and specificity for the identification of cervical cancer. By comparison to established and accepted elements of the cervical pre-cancer diagnostic pathway Raman spectroscopy appears to show promise in being a useful and sufficiently accurate test for vulval diagnostics. The sensitivities and specificities whilst not matching that of cervical cytology are close. It is feasible that with improved instrumentation; further optimisation of the technique for vulval tissue; tailored chemometrics and a larger training set, that improvements in diagnostic accuracy could be achieved to bring the technique in line with other commonly used

diagnostic tests used to identify cancer. It should be noted that the diagnostic accuracies described above relating to cervical assessment are determined from an asymptomatic screening population in contrast to the population in this study that are symptomatic and with identifiable vulval disease. The population in this study therefore have a higher pre-test probability of having detectable disease and this is not accounted for in the reported diagnostic performance.

It is important to note that, in general, diagnostic tests in symptomatic patients are not interpreted in isolation. In the clinical setting the diagnosis of vulval precancer or cancer is actually solidified by considering all the available evidence collectively. This includes the symptoms the patient reports; any previous history of related disease; the appearances on examination and specialised investigations such as histopathology. This is all significantly influenced by clinical expertise and experience. No single element of the diagnostic process is perfectly accurate and it is taken in the context of the other information obtained. Without any individual element of the clinical assessment the diagnostic accuracy will suffer. Likewise adding a novel test into the diagnostic pathway can be used to increase the diagnostic accuracy. If one element of the available information such as the histopathology does not fit with the remaining information this does not necessarily mean the diagnosis is abandoned. For example, if a clinician strongly suspects a vulval cancer but a biopsy and histopathology does not demonstrate cancer, then the clinician will repeat the biopsy or treat the patient to avoid missing an important diagnosis. Therefore the assessment of performance of an individual diagnostic test needs to be taken in the context of its clinical application. Raman spectroscopy

should be regarded as a tool for assisting the evaluation of vulval lesions rather than being a final and definitive arbiter of diagnosis.

In this study the diagnostic performance of the technique, as evaluated for the identification of lichen sclerosus, is to an extent clinically contextualised. Those selected for inclusion in the study were patients in whom a diagnosis of lichen sclerosus was suspected, but in whom there was insufficient clinical certainty to make the diagnosis without biopsy. In this context Raman spectroscopy performed well and correctly identify all 12 patients that did not have lichen sclerosus theoretically reducing the need for biopsy by 44% in this selected group.

The evaluation of the diagnostic performance of Raman spectroscopy for the identification of vulval SCC and VIN was not clinically contextualised in the same way as the lichen sclerosus investigation. In this arm of the investigation, the tissue samples analysed were taken from patients undergoing surgical treatment for VIN or SCC. These patients are a subset of the original patient group in whom diagnostic biopsies are undertaken and as such the data obtained is an approximation to the group in whom the technique would be applied in practice.

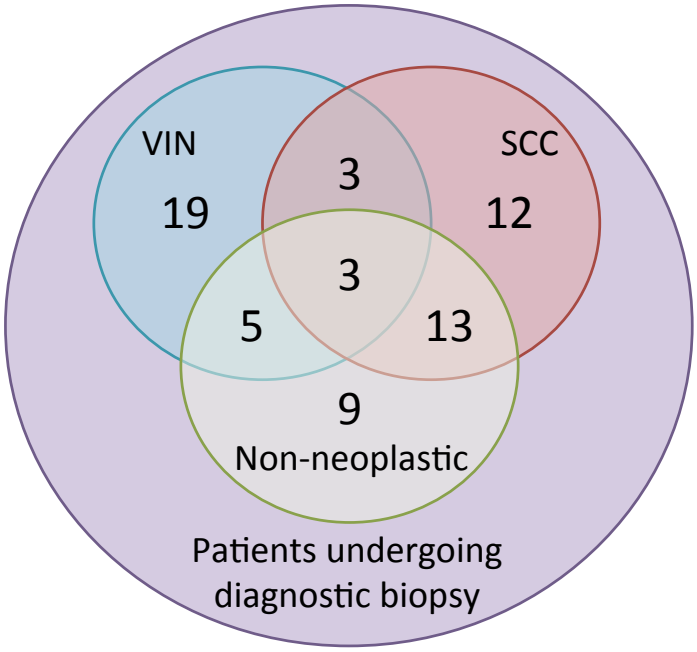


Figure 13-1 – Illustration of the patient cohort who have been recruited to this study (VIN, SCC, Non-neoplastic) in the context of all those undergoing diagnostic biopsy demonstrating the group of patients not evaluated in this study.

Figure 13-1 illustrates the group of patients not evaluated in this study i.e. those undergoing diagnostic biopsy but not undergoing excisional treatment. Figure 13-2 illustrates this within the diagnostic pathway for patients with suspected VIN or SCC. A and B are points at which Raman spectroscopy could be deployed as an optical diagnostic technique to reduce clinicians reliance on invasive tissue biopsy. Point C is however the point at which the tissue samples were collected in this study. Tissue collected at point C in the pathways is only an approximation to the types of tissue present at points A and B. It is probable that there are significant biomolecular differences in the tissue at points A and B when compared to that present at point C. In particular there is a subset of non-neoplastic tissue and VIN that is not subject to surgery (Figure 13-1) that is likely to possess a more benign biochemistry than the

non-neoplastic tissue and VIN found at point C. Further more all of the non-neoplastic tissue measured was obtained from diseased vulvas with either concurrent VIN, SCC, adenocarcinoma, BCC, melanoma or Paget's disease. Consequently given the field change found with vulval neoplastic disease (chapter 1.4.1.3) it is likely there are biomolecular abnormalities in the non-neoplastic group that would not be found in tissue from vulvas that do not contain concurrent disease. Assuming equivalence of the tissue analysed to the clinical context of disease diagnosis at points A and B in Figure 13-2 is likely to have resulted in an underestimation of the diagnostic accuracy of the technique.

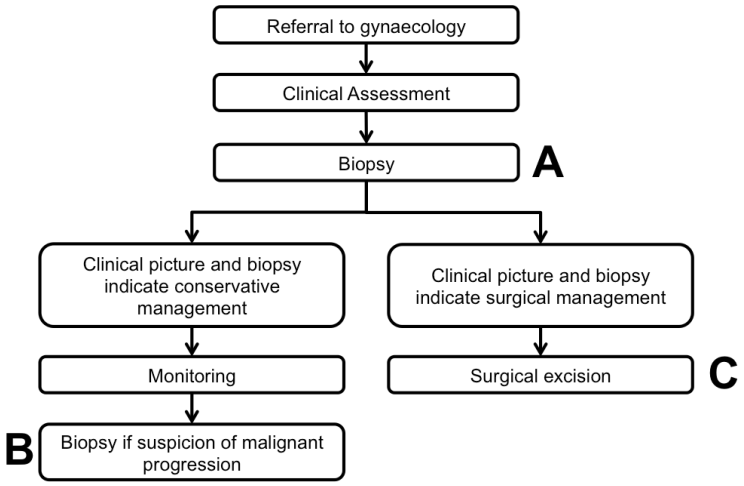


Figure 13-2 – Diagnostic pathway showing the points A and B at which a diagnositc optical biospy would be applied and C the point of tissue collection for this analysis.

In this study collection of tissue at point C (Figure 13-2) was necessary to permit the collection of sufficient samples of tissue of the three types included in the analysis. It was not possible to collect and store tissue from biopsies at point A and B for the reasons outlined in chapter 7.3.1.1.

The chemometric approach used in this study was tailored to allow identification of the biomolecular changes in the tissue that result in the classification of the different tissue types. This led to the choice of PCA combined with LDA as this combination allows for analysis of the spectral differences behind the classification and permits insight as to whether the classification is based on biomolecular changes or other noise within the data. The use of PCA also has the added advantage of producing orthogonal components so the contribution of the changes seen in any individual component should not be confused with any other component. This approach is well suited to an exploratory analysis such as that in this study as it avoids the analysis and classification becoming a black box from which the basis for classification cannot be not easily understood. This approach also enables the detection of incidences of over fitting to non-relevant spectral data.

Over fitting resulting in a type one error is a real danger in this type of classification analysis and some elements of the diagnostic approach used in this study risk increasing the chance of over fitting the data to the desired result. Where possible these have been countered by measures to validate the result of the analysis. The variant of PCA LDA used in the final analysis was selected by comparison to the diagnostic performance of the technique (Chapter 7.5.7). Likewise the selection of the thresholds for data exclusion by the D and Q statistic was selected based on diagnostic performance of the technique (Chapter 7.5.7). Selecting analytical techniques based on the performance of the technique introduces the risk of over fitting. Steps were taken to reduce the risk of over fitting including exclusion of data based solely on the fit to the training data set and ensuring strict separation of the

training and test sets within the internal cross validation. In selecting the samples to be analysed, care was taken to avoid selecting two samples of the same pathology from the same participant. In the investigations assessing SCC, VIN and non-neoplastic vulval skin, 21 participants contributed two samples each to different pathology groups and 3 participants contributed samples to all three pathology groups. This approach should not have increased the risk of a type one error and should have improved the validity of the resultant classification models by reducing any group separation based on differences in the biomolecular composition of individual patients rather than the pathology groups.

In the analysis of the classification accuracy of the technique it became apparent that there appeared to be a continuum of disease progressing from non-neoplasic vulval tissue through VIN to SCC. This was evident in the spectra and the ternary plots resulting from each of the three spectral collection methods used (Figure 6-25, Figure 7-29 and Figure 8-22). In each of the plots there was overlap between the SCC and VIN groups and between the VIN and non-neoplastic group. There was however minimal over lap between the non-neoplastic group and the SCC group. This suggests that rather than detecting three distinct pathology groups it may be possible to detect a continuum of disease from low risk of malignancy to overt malignancy. A recent retrospective study of 201 cases of vulval SCC demonstrated epithelial neoplastic disorders (e.g. VIN) are found adjacent to vulval SCC in 72% of patients (Yap *et al.*, 2016). The majority of vulval SCCs develop in these neoplastic lesions which precede the development of SCC by a variable period of time (Del Pino, Rodriguez-Carunchio and Ordi, 2013b). The findings of a continuum of spectral

changes from non-neoplastic to in situ neoplastic and malignant disease fit with this understanding of the carcinogenic process in the vulva. Longitudinal clinical studies would be required to assess if the spectral changes that lead to the classification of VIN as SCC or non-neoplastic tissue as VIN are associated with an increase risk of malignant progression. If this hypothesis was born out then Raman spectroscopy has the potential to offer clinicians an invaluable opportunity to perform real time in vivo biomolecular risk assessment of non-neoplastic vulval skin or VIN and target more intensive treatment or observation to this group. In addition these spectral changes could be used to non-invasively assess the extent of field change in the vulva (chapter 1.4.1.3). This in turn could be used to target medical or surgical management strategies to the wider vulva.

13.1.2 Raman spectroscopy as an intraoperative tool to identify disease margins

Obtaining adequate and complete excision during surgical management of vulval disease is important for both local disease control in patients with VIN and vulval SCC and for overall survival in patients with vulval SCC. Residual disease after excisional surgery is associated with the need for further surgery and adjuvant radiotherapy (Yap *et al.*, 2016). Current guidelines suggest that when excising an area of SCC a margin of non-neoplastic tissue at least 10-15 mm should be taken to ensure a clear histopathological margin of greater than 8 mm and complete excision of the tumour (RCOG, 2014; European Society of Gynaecological Oncology, 2016). This recommendation is based on a small retrospective cohort study in which none

of the patients with pathological margins of greater or equal to 8 mm had recurrent disease and the only recurrences were seen in those with pathological margins of less than 8 mm (Heaps *et al.*, 1990). This study did not account for stage of disease, a major confounding factor. Most of the patients with a margin of less than 8 mm in the study had advanced disease and there are likely other factors contributing to the higher recurrence rate seen in this group. Recent cohort studies have questioned the traditional importance placed on the pathological margin and have suggested the most important prognostic factor in early stage vulval SCC is the completeness of the excision rather than the size of the tumour free margin (Woelber *et al.*, 2011; Nooij *et al.*, 2016; Micheletti *et al.*, 2018; Te Grootenhuis *et al.*, 2018). The retrospective cohort study and meta analysis by Nooij et al. evaluated 148 patients undergoing surgical treatment for vulval SCC to assess the association between surgical margin and recurrence at two years. The study found that the only independent risk factor for recurrence was the presence of a tumour positive margin and not the size of the tumour free margin. Similarly Micheletti et al. did not find an association between an 8 mm pathological margin and recurrence in a retrospective cohort study of 114 patients with vulval SCC. A retrospective cohort study by Yap et al. assessed 201 patients with surgically treated vulval SCC and found no association between tumour margin of less than 8 mm and local recurrence or disease specific survival over a five year period (Yap *et al.*, 2016). Another contemporary series of 107 patients reported by De Simone et al. did not show an associated between disease recurrence and survival and margin status unless the pathological margins were less than 2 mm (DeSimone *et al.*, 2006). The extent of

the surgical margin is important in the excision of vulval disease, as small increases in the surgical margin are associated with significant changes in the area of vulval tissue that needs to be removed. For example, if we have a circular two centimetre tumour then the increase in tissue surface area that is removed with a 15 mm surgical margin is 36% larger than with a 10 mm margin. This is especially important adjacent to important structures that affect continence such as the urethra and anal sphincter as well as important functional structures such as the clitoris. Green et al. surveyed 41 women before and following vulvectomy and found the procedure was associated with a significant fall in sexual frequency alongside increases in arousal disorders and sexual aversion disorders (Green *et al.*, 2000). Hazewinkel et al. reported the results of a cross-sectional study of 76 women that had undergone surgical treatment for vulval cancer and found that of the 43% that were sexually active 42% reported significantly worse sexual function after surgery (Hazewinkel *et al.*, 2012). To date reported studies have not assessed the relationship between radically of excision and excision margin on post-operative psychosexual function. Although there is lack of evidence of the disutility of anxiety and psychosexual repercussions of vulval excisional surgery it is accepted that this type of surgery is associated with significant adverse effects (Cavanagh and Hoffman, 1996).

The reasons for requiring a larger surgical margin as compared to the pathological margin appear to be twofold. Firstly specimens undergo shrinkage during fixation. DeSimone et al. studied surgical specimens from 19 women undergoing vulval resections and reported mean specimen shrinkage of 15% after fixation (DeSimone *et al.*, 2006). Secondly the clinical assessment of disease margins at the time of

surgery through visual inspection and palpation can be imprecise and difficult. The ability to objectively assess disease margins at the time of surgery using Raman spectroscopy has the potential to reduce the size of surgical excisions, preserve important functional structures whilst ensuring all abnormal neoplastic tissue is removed preventing the need for further surgery or adjuvant radiotherapy. This is especially important when clinicians make decisions about the need to excise the clitoris, part of the urethra or part of the anal sphincter.

Currently the only available support for the intraoperative assessment of resection margins is frozen section. Even though this procedure is successful for assessment of resection margins it has a number of key limitations that prevent its use in vulval cancer surgery. Frozen section is very time consuming which extends the duration of anaesthesia and surgery; sampling errors can occur as only a small part of the margin can be assessed and the rapid freezing required can make the final histopathological analysis more difficult (Santos *et al.*, 2017). All these limitations can lead to false negative results. In recent years there has been a great deal of interest in the use of Raman spectroscopy for guiding oncological surgery. In the excision of neurological tumours an intraoperative Raman probe capable of direct measurement of brain tissues has be shown to distinguish cancer from benign tissue with sensitivity and specificity of 93% and 91% respectively (Jermyn *et al.*, 2015). In breast cancer excision a Raman system was shown to be able to distinguish cancer from normal breast tissue on freshly excised specimens with a sensitivity and specificity of 83% and 93% respectively (Haka *et al.*, 2014). Kong et al. used a two step auto-fluorescence and Raman system for the detection of basal cell carcinomas

(BCC) in skin at the time of surgery and demonstrated the automated detection of BCC with sensitivity and specificity of 95% and 94% respectively. Instigating Raman spectroscopy for margin assessment can have key benefits. It may facilitate representative sampling, avoiding sampling bias and objectively identify resection margins reducing incomplete excision and the need for further surgery or radiotherapy. In the surgical setting all neoplastic tissue needs to be excised during treatment so it is important to be able to reliably differentiate SCC and VIN from non-neoplastic skin. In the current investigation, sensitivities of 86% to 95% were achieved with specificities of 72% to 57% using the probe based system. If applied clinically the technique could be optimised for sensitivity to ensure that all areas of cancer are detected. In this scenario tissue within the traditional clear resection margin that was characterised as non-neoplastic according to Raman spectroscopy could be left in in situ with increased confidence given the 95% sensitivity. The sacrifice in specificity could be overcome with a pragmatic approach to resection. This is illustrated in the schematic of a vulval resection in Figure 13-3. At area A the visually abnormal lesion will be excised irrespective of the Raman characterisation; at point B the tissue would be resected if the Raman characterised the tissue as cancer and tissue at point C would not be resected provided it is clinically normal. Using this strategy it should be possible to safely reduce the size of the area excised whilst ensuring complete excision of the abnormal area. In practice it may also be desirable to define a fall back clinical safety margin around the tissue to define the minimum extent of the incision (Figure 13-3). This could be used to prevent too small a margin being removed.

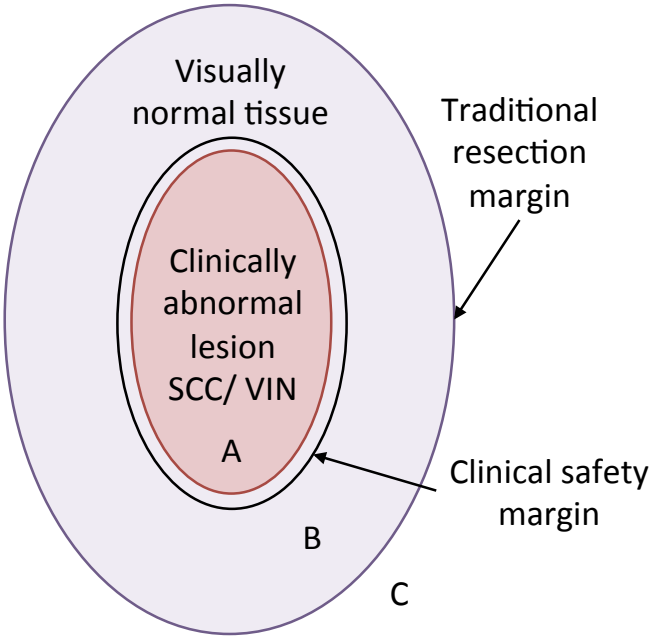


Figure 13-3 – Schematic of vulval resection

If subsequent studies improved the sensitivity and specificity then alternative strategies could be employed. A high specificity would enable a reverse of the approach where tissue in area B (Figure 13-3) could be left in situ provided Raman did not characterise the tissue as neoplastic.

As described in chapter 13.1.1 the tissue types compared are not exactly equivalent to those that would be under investigation in vivo. The tissue used in this investigation was collected at point C on Figure 13-2, which is the same point in the pathway as the proposed application. There are however some important differences in the tissue as compared to margin assessment in vivo. The most relevant are that the tissue samples were taken so as to avoid disrupting the assessment of the margins by the histopathologist. Tissue was not taken from the closest surgical margin or from the edge of the specimen. The result of this is that the tissue

analysed did not include the tissue furthest from the abnormal lesion. This may have led to an underestimation of the diagnostic performance in the whole of the area between the traditional resection margin and abnormal lesion (Figure 13-3). In addition the classification accuracy given in this investigation does not reflect the combined diagnostic approach adopted with any investigation used to aid diagnosis. In the real world use of Raman spectroscopy for the identification of disease margins at the time of surgery the underlying pathologies present and clinical assessment would both contribute to decisions surrounding the extent of the surgery and the surgical margins taken. Clinical trials of the technique would be needed to assess the multimodal nature of clinical decision making with advanced diagnostic techniques such as Raman spectroscopy.

13.1.3 Potiential role for Raman spectroscopy in detecting disease progression and monitoring response to treatment

Although not directly evaluated in this study, Raman spectroscopy may have an important role in measuring response to non-surgical treatment of vulval conditions. This is particularly important in the two commonest vulval conditions known to cause an increased risk of progression to vulval cancer, VIN and LS (Chapter 1).

Lichen sclerosus is associated with an increased lifetime risk of developing vulval SCC with the estimated risk varying between 3% and 7% in population based studies (Cooper *et al.*, 2004; Bleeker *et al.*, 2016). Non-randomised cohort studies suggest that good control of LS with ultra potent topical steroids may reduce the risk

of progression of LS to SCC (Cooper *et al.*, 2004; Renaud-Vilmer *et al.*, 2004; Andrew Lee, Bradford and Fischer, 2015). Lee et al. undertook the largest analysis assessing risk of disease progression in a prospective longitudinal cohort study of 507 women with vulval lichen sclerosus. This study demonstrated that treatment with ultra potent corticosteroids targeted towards maintenance of normal skin colour and texture was associated with a decreased rate of progression to vulval cancer (A Lee, Bradford and Fischer, 2015). This suggests a role for the treatment of LS in order to prevent progression to vulval cancer. Currently treatment of LS is targeted to symptoms or clinical response and many women with progressive LS will be asymptomatic. In this study we have demonstrated that Raman spectroscopy is capable of detecting patient with LS with high sensitivity and specificity 91% and 80% respectively. After further study Raman spectroscopy may provide a valuable technique for the objective assessment of response to treatment of women with asymptomatic vulval LS. Raman spectroscopy could be used to direct treatment towards biochemical and histological response rather than symptomatic control of disease or clinical response. This in turn may further prevent the development of a cellular environment conducive to malignant transformation and prevent the occurrence of vulval cancer. Treatment of LS with ultra potent topical steroids is considered safe and so this strategy is associated with minimal risk to patients.

Usual type VIN can be treated with topical treatments or surgical excision. Imiquimod, an immune modulator, and cidovir, an antiviral agent, are the primary topical agents used in the treatment or usual type VIN. Complete response to treatment occurs in 45% of those treated with imiquimod and 46% of those treated

with cidofovir (Ta *et al.*, 2016). Both these treatments require administration over a prolonged period of time and are associated with significant and distressing side effects (Ta *et al.*, 2016). Side effects include headaches, fatigue and vulval pain and are slightly more common with imiquimod than cidofovir. Currently response to treatment is assessed clinically with patients needing to complete a 16 to 24 week course of treatment to assess if it is efficacious. In the research setting phase II studies investigating novel agents often rely on histological response to treatment as a primary end point (Tristram *et al.*, 2014). This limits the longitudinal assessment of disease response due to the morbidity associated with invasive biopsy especially during treatment.

In this thesis we have demonstrated an apparent disease continuum between non-neoplastic vulval tissue through VIN to overt malignancy using Raman spectroscopy (chapter 13.1.1). This suggests Raman spectroscopy may be able to identify patients with VIN who have disease progressing towards malignancy and those responding to topical treatments. Further in vivo longitudinal studies would be required to assess the efficacy of Raman spectroscopy for this application. As a research tool Raman spectroscopy has the potential to be used to assess response to novel treatments without recourse to invasive biopsy. This would mean objective assessment of response to treatment would not be limited by the morbidity of invasive biopsy and patient acceptability and objective longitudinal response could be assessed with increased precision.

13.1.4 Raman spectroscopy for improving prognosis assessment in vulval cancer

Recurrent disease occurs commonly after the primary treatment of vulval SCC, affecting at least 25% of patients (Coulter and Gleeson, 2003; Te Grootenhuis *et al.*, 2018). Around half of recurrent cases recur locally within the vulva/ perineum and half at more distant sites (Salom and Penalver, 2002; Coulter and Gleeson, 2003). Currently our estimation of prognosis and risk of recurrence in vulval cancer is based on clinical and histopathological factors. Inadequate excision margins have been thought to be the main determinate of recurrent disease but this is increasingly being challenged as previously discussed. Other factors thought to have prognostic relevance include: presence of LS; groin lymph node metastasis and primary tumour characteristics. These tumour characteristics include: tumour size, tumour focality; depth of invasion; grade of differentiation; tumour localisation; presence of HPV and lymphovascular space invasion (Yap *et al.*, 2017; Te Grootenhuis *et al.*, 2018). These factors are considered when making decisions about adjuvant radiotherapy to reduce the risk of disease recurrence.

There has been increasing interest in Raman spectroscopy as tool for helping predict prognosis and response to treatment (Upchurch *et al.*, 2018). In the management of vulval cancer Raman biomolecular assessment of tumour or residual tissue may help guide follow up and the use of adjuvant radiotherapy. Studies assessing the role of Raman spectroscopy for prognostic assessment have already been undertaken at other cancer sites. Analysis of Raman spectra from prostate cancer specimens

demonstrated the technique can identify those patients who would not respond to adjuvant hormone therapy with high sensitivity and specificity 88% and 88% respectively (Wang, He, *et al.*, 2013). The same study also found the Raman spectra could be used to identify those patients who were initially hormone sensitive but who became hormone resistant in the following year with sensitivity and specificity of 86% and 89% respectively. Another study of patients with lung adenocarcinoma demonstrated Raman spectroscopy could identify patients sensitive to tyrosine kinase inhibitors with a sensitivity and specificity of 90% and 85% respectively (Wang, Zhang, *et al.*, 2013). These studies illustrate the potential of Raman spectroscopy to go beyond replicating the results of histopathology and immunohistochemistry and to analyse molecular markers of response to treatment or prognosis. It is helpful to think about the words of Donald Rumsfeld when considering the role of Raman spectroscopy in prediction of prognosis or response to treatment. Donald Rumsfeld said “there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns”. In this context the “known knowns” are specific biomarkers that we know have a link to disease course such as oestrogen receptor immunohistochemistry and sensitivity to anti-oestrogen treatment. There are also “unknown unknowns” which describes tumour characteristics that are completely unknown to us. In the middle sits the “known unknowns” these are those biomolecular changes that can be speculatively identified using techniques such as Raman spectroscopy and that can be correlated with disease outcomes. These are changes that we can identify but that we cannot

precisely define due to the limitations of the technique, hence “known unknowns”. Identification of these “known unknowns” using Raman spectroscopy has the potential to yield useful prognostic information but is beyond the scope of this non-longitudinal study. The ability to identify patients with high risk of recurrence could be used to identify those who would benefit from adjuvant treatment for example radiotherapy or possibly steroids, antivirals or immune modulators.

13.1.5 Raman spectroscopy study limitations

The investigation into the role of Raman spectroscopy in the management of vulval disease has methodological limitations that require highlighting. Many of the important limitations have been discussed at the relevant points in previous chapters and not all are repeated again in this chapter. The limitations in this study relate to the sample collection, spectral acquisition, data processing, multivariate analysis and core assumptions of the study.

The tissue available for analysis in this ex vivo study was limited by the limits of acceptability for the study participants. It was not acceptable to take biopsies or excise tissue solely for research purposes. This limited tissue collection to the end of the diagnostic pathway in the assessment of neoplastic disease and to the uncertain category of patients in the assessment of LS. This is likely to be reflected in under estimation of the performance of the technique as a tool for diagnosis. The effect of this tissue selection bias is likely to be less marked in the margin assessment comparisons as the tissue analysed is a closer approximation to the real life clinical scenario under investigation. It should also be noted that tissue measured in vitro is

by its nature ischaemic and tissue in vivo has active blood flow and this may adversely affect the Raman signal as the endogenous fluorophores in blood are know to induce fluorescence. In addition the tissue hypoxia at excision may significantly alter the tissue biochemistry. A study of spectra collected in vivo and ex vivo in the same tissue by Lima et al. demonstrated a reduction in intensity of Amide I and II peaks in ex vivo measurements compared to in vivo measurements (Ferreira Lima *et al.*, 2019).

The fibre optic probe used in the final stage of the study (Chapter 8) had a significant background signal and this was apparent in all the spectra collected and may have masked tissue signal. In addition the background combined with tissue fluorescence caused detector saturation in the probe measurements. This resulted in the need to exclude spectral data acquired at less than 700 cm^{-1} from the analysis. This may have negatively impacted the classification performance due to the loss of discriminatory features below this point.

The spectra taken using the probe were taken over an integration time of 20 seconds to optimise the signal to noise ratio in this proof of concept study. In the clinical environment the spectral acquisition time would need to be in the order of a few seconds to allow the assessment of multiple area of vulval skin. This should be achievable with improved collection optics in the probe and the application of multiple collection fibres around a central excitation fibre. The spectra acquired using the probe were also measured in a dark laboratory and the impact of ambient light was not assessed. The problem of ambient light interfering with the Raman signal

has been extensively investigated and a combination of selecting ambient light that does not interfere with the Raman signal and filtering the ambient light out of the collected spectra can be used to overcome this issue (Nyberg, Ramser and Lindahl, 2013; Zhao *et al.*, 2014). In addition this study did not assess inter-probe variability as all measurements were taken using one probe.

13.2 Discussion: FTIR spectrosocpy for the assessment of vulval diease

13.2.1 FTIR as an adjunct to aid routine histopathological assessment in the detection of vulval disease

The role of FTIR spectroscopy as an automated diagnostic tool was investigated in the study outline in Chapter 11. The acceptable accuracy of an automated tool for translation into clinical practice depends on the role the tool is to take. Given the significance of the diagnosis of vulval cancer or VIN it is not likely an automated FTIR based diagnostic assessment would be considered sufficient for diagnosis alone until sufficient validation of the approach versus current techniques over many years was performed. Given the increasing workload faced by histopathologist and the need for second reporting in cases of vulval cancer, the most appropriate application of this new technology would be in triaging samples for analysis by histopathologists and acting as a second diagnostic assessment in women with suspected vulval SCC.

The results obtained with FTIR mapping of selected samples of non-neoplastic vulval tissue, VIN and SCC were found to be promising. Two diagnostic comparisons were evaluated reflecting the two potential roles for the technique outlined above.

The first comparison assessed the ability of the technique to act as a triage tool to prioritise the histopathological workload. In this setting it is important to identify SCC with high sensitivity, as the consequence of delayed diagnosis is more severe than an inappropriately urgent histopathology review. Although the diagnostic accuracy was assessed on a per spectrum basis the only valid evaluation is that performed on a per sample basis. In practice the whole of a biopsy would need to be evaluated to be confident that it did not contain cancer. The use of individual spectra that are not necessarily representative of the whole specimen would not be acceptable. When analysing on a per sample map basis all of the cases of SCC were identified (100% sensitivity) and only 21% of samples were incorrectly classified as containing SCC (79% specificity). This indicates the technique has promise in the triage of diagnostic biopsies to enable histopathologists to prioritise those cases that are more likely to contain SCC. If FTIR mapping were to be translated into clinical practice for the purpose of triaging vulval pathology specimens several major limitations would need to be overcome. In this study only a small part of each tissue sample was assessed by FTIR. For the technique to be introduced into routine clinical practice it is likely that whole biopsy samples would need to be mapped to avoid missing a diagnosis of cancer due to a sampling error. This would be possible using the technique outlined in this study but it would take a prohibitively long time. The time of spectral map acquisitions could however be reduced by adjusting the collection parameters (e.g.

spatial resolution or wavenumber range) to allow faster spectral acquisition. Using different spectrometers with larger detector arrays or supercontinuum sources and detectors could also be used to increase the speed of acquisition (Naranjo *et al.*, 2015). Another approach to reduce the time taken to process samples would be a sampling only part of the epithelium. This could be done either using a random sampling technique or by using another technique to screen the sample for areas likely to contain cancer. Care would need to be taken with a random sampling technique for the reasons previously stated. Fluorescence spectroscopy would be a suitable technique for identifying areas for more detailed analysis by IR spectroscopy as it is rapid and highly sensitive, but lacks specificity (Kong *et al.*, 2013). Another approach would be to limit the analysis to a smaller number of relevant wavenumbers rather than collecting the whole spectrum. This would lead to faster spectral acquisition and potentially reduce the high cost of the instrument needed to analyse the tissue.

The second comparison evaluated the ability of the technique to act as a second reporter for samples of cancer. This was assessed by looking at the ability of the technique to differentiate between SCC and VIN as diagnostic uncertainty can arise between the classification of VIN and SCC when invasion is equivocal. As non-neoplastic tissue and SCC should have more biomolecular differences than SCC and VIN it was expected that the technique would perform less well when differentiating SCC from VIN. When analysing on a per sample map basis all of the cases of SCC were identified (100% sensitivity) from those of VIN and only 23% of samples of VIN were incorrectly classified as containing SCC (77% specificity).

These results suggest the technique has promise for confirming case of SCC with accuracy when histopathology suggests SCC is present and therefore it has the potential to provide an automated second opinion in cases of SCC. As confirming a diagnosis of cancer only requires spectral mapping of the area thought to contain the cancer there is no need to map the entire sample. As a result of this the previously discussed limitations in relation to completeness of sample mapping and time taken to scan the sample are less significant in this application. A significant limitation in this comparison is that the tissue selected for analysis does not exactly match tissue that would be analysed if the technique were integrated into the clinical work flow. If the technique was to be implemented as a tool for confirming the presence of cancer then the tissue sections under analysis would be those where the presence of SCC was not clear and those with severe VIN where the presence of SCC was suspected. In this comparison all stages and grades of SCC were compared together with unrefined samples of VIN. It may therefore be the case that the classification performance seen in this comparison is an overestimation of the performance that would be seen in routine practice.

Both these comparisons are only indicative of the potential of the technique to aid the histopathological workflow. Further work is required as outlined in chapter 13.4 to evaluate the technique closer to conditions seen in routine histopathological workflows.

13.2.2 FTIR spectroscopy for augmenting histopathology

This preliminary study has demonstrated FTIR spectroscopy is able distinguish between LS or uVIN found in isolation and LS or uVIN found alongside concurrent SCC at another anatomical location on the vulva. The ability to separate these groups according to the presence of concurrent SCC is suggestive of cancer associated molecular changes occurring across the vulva. The detection of these molecular changes supports the concept of preneoplastic field cancerisation in the vulva (Rosenthal *et al.*, 2002; Braakhuis *et al.*, 2003; Dakubo *et al.*, 2007). In the development of a preneoplastic field of cancerisation, a stem cell acquires a genetic or epigenetic alteration that gives it a growth advantage over its neighboring cells. Replication of the stem cell results in an expanding patch of altered daughter cells. With additional genetic or epigenetic alterations this patch replaces the normal epithelium and develops into a field of cancerisation. Within the abnormal preneoplastic field the epithelium may appear histologically benign however molecular alterations within cells result in an increased risk of malignant transformation. The development of preneoplastic field change can be used to explain the multiple primary tumours and distant tumour recurrences that are often observed in the vulva. The molecular basis for field cancerisation in the vulva has previously been explored by examining the distribution of X-chromosome inactivation, TP53 mutations and viral integration sites (Rosenthal *et al.*, 2002; Rolfe *et al.*, 2003; Vinokurova *et al.*, 2005). The data published demonstrated molecular monoclonality of anatomically distant vulval lesions supporting the concept of

preneoplastic field change in the vulva. These studies of specific molecular changes in the vulva only show us part of the picture as the transformation of normal cells to cancer is a complex multistep process involving more that just a few molecular variations. In this study we have used FTIR spectroscopy to demonstrate the molecular field changes associated with the development of vulval cancer. FTIR spectroscopy allows the analysis of a broad range of molecular changes within the vulva. This technique not only examines molecular changes resulting from known provocations (e.g. TP53 mutation and HPV integration) but also assesses other undescribed molecular changes (Ostrowska *et al.*, 2011). This gives a broader picture of the molecular changes present, but at the expense of detailed biomolecular causation, as FTIR spectroscopy does not allow us to determine the exact molecular differences underlying spectral differences identified.

Demonstrating the ability of FTIR spectroscopy to differentiate between vulval disease found with concurrent cancer and vulval disease found in isolation highlights the role FTIR spectroscopy could have as a tool for assessing preneoplastic molecular changes in the vulva.

Traditional histopathological analysis gives limited information about the biomolecular changes associated with an increased risk of malignant transformation. FTIR spectroscopy has the potential to be used to augment traditional histopathology giving pathologists additional molecular information from the tissue under examination. The FTIR spectroscopic analysis of preneoplastic fields and molecular risk stratification of patients with VIN or LS may facilitate the identification of those at

high risk of developing vulval cancer. In addition, FTIR spectroscopy has the potential to be used as a histopathological adjunct to improve the early diagnosis of SCC through the analysis of molecular changes that identify those in whom further investigation is likely to find an occult malignancy. This early diagnosis would be an advantage as when vulval SCC is diagnosed early, when the depth of invasion is less than 1.0 mm (i.e. FIGO stage 1a), the risk of lymph node metastasis is very low (Hacker and Van der Velden, 1993). The absence of lymph node metastasis is an important prognostic factor in vulval SCC, with reported five-year survival in those who are node negative ranges from 70-93% compared with 25-42% in node positive women (Schilder and Stehman, 2012).

13.3 FTIR spectroscopy study limitations

The investigation into the role of FTIR spectroscopy in the assessment of vulval disease has methodological limitations that require highlighting. Many of the important limitations (e.g. time taken to perform the technique and cost of the technique) have been discussed at the relevant points in previous chapters and are not repeated again here. A significant limitation in this proof of concept study was the bias inherent in the sample selection. In this study tissue samples were retrospectively selected that displayed the typical characteristics of the pathologies of interest and were not representative of all the possible pathologies present in the vulva. Introducing a broader range of pathologies into the classification model may result in a poorer diagnostic performance and needs to be evaluated further. In

addition, the process for measuring and analysing spectra from the samples in this study was not fully automated and required skilled technical input to select the correct areas for mapping and then to select the epidermal layers after k means analysis. If the technique is to be implemented in to the routine diagnostic pathway then either all elements of the process will need to be automated or training provided to those delivering the service introducing the possibility of human error.

13.4 Future work

13.4.1 Development of an in vivo Raman probe

The laboratory based ex vivo studies presented in this thesis demonstrate the Raman technique has real promise for delineating disease margins at the time of surgery and for the diagnosis of vulval conditions. The next proposed phase of research outlined below is the development and evaluation of a probe suitable for the in vivo assessment of vulval tissue within the clinical environment to allow the progression to large scale clinical trials. The aim would be to provide sufficient evidence to lead to large scale trials to assess if the technique can be used to minimise the mutilating and disfiguring surgery suffered by women with vulval cancer and VIN by accurately detecting disease margins at the time of surgery. In addition this phase of research would be needed prior to evaluating the ability of the technique to reduce the need for invasive tissue biopsies when diagnosing vulval conditions in the clinical environment. The main objectives of this next research phase are listed in Table 13-3.

Table 13-3 – Objectives for the next phase of research in the develpoment of Raman spectrscopy for the assessment of vulval disease.

Objectives
1. To develop a safe and reproducible non-invasive Raman spectroscopic probe for the assessment of vulval disease in vivo.
2. To demonstrate the prototype probe can identify disease specific Raman spectroscopic signal characteristics (spectral fingerprints) ex vivo across a range of known vulval skin conditions within the clinical environment.
3. Obtain regulatory and ethical approval for an in vivo feasibility study and demonstrate safety of the device in the first human pilot.
4. To demonstrate the probe can be used in vivo to detect the molecular changes that occur at the margins of pre-cancerous and cancerous vulval disease.
5. Develop the understanding of how the Raman probe will improve the care for women with vulval cancer and VIN.

This research would build on the initial laboratory based proof of concept studies presented in this thesis and would lay the foundation for a full clinical trial to evaluate the efficacy of the technique in clinical practice.

It is proposed that future research is split into five packages of work:

13.4.1.1 Work package one - development of Raman probe prototype suitable for in vivo assessment of vulval disease and probe optimisation for manufacture.

In this work package the design for a Raman probe suitable for in vivo evaluation should be developed. Data from analysis of ex vivo Raman spectral measurements of vulval tissue presented in this thesis would be used to inform the design of the

probe. The design should include features necessary for in vivo use, including the development of disposable sterile probe tips suitable for intraoperative use. A novel prototype probe device would then need to be evaluated with respect to interprobe and intersystem variability in preparation for human tissue studies.

In collaboration with a commercial medical device design, development and manufacturing partner, the probe would need to be optimised to suit commercial production facilities.

Production records would need to be kept and a technical file constructed to gain the necessary Medicines and Healthcare Products Regulatory Agency (MHRA) approval. Testing and quality criteria to be fulfilled at each stage of component procurement and production would need to be agreed. Sterilisation/ cleansing process to be used should be confirmed and dummy probes produced for bioburden tests to validate that process.

13.4.1.2 Work package two - probe manufacture and testing / regulatory (MHRA) and ethical approvals / initial ex vivo diagnostic study

A commercial partner would manufacture the finalised probe design. The quality of probe spectral signal and reproducibility should be evaluated by testing on existing samples of excised vulval tissue, ensuring signals are of similar quality to those obtained using the tested prototype design.

An initial ex vivo diagnostic study of the existing human vulval tissue samples used in this study could then be completed under existing ethical approvals. This would ensure the diagnostic performance demonstrated in early laboratory studies could be replicated using the new probe device. A multivariate discriminant classification model would be developed to allow disease classification using the novel probe. The classification model should be evaluated with 'leave one out cross-validation' to optimise the use of the available spectral data. MHRA and ethical approvals would need to be obtained to allow evaluation of the probe in vivo.

13.4.1.3 Work package three - in vivo pilot, safety evaluation and in vivo feasibility study

This work package would be used to identify and overcome the technical challenges of moving the technique from the laboratory and into the clinical setting (e.g. variations in lighting, temperature and humidity).

An in vivo feasibility study should be undertaken. This would assess the technique for safety, practical implementation and diagnostic accuracy on volunteer patients. This would include basic safety for the patient and operator; logistics and ergonomics; tissue biochemical signal to noise versus time and anatomical location as well as a basic signal reproducibility evaluation. Clinician training materials for the safe and effective use of the device in the theatre setting should also be developed and evaluated.

This in vivo study should be conducted during conventional surgical excision of vulval tissue for treatment of VIN or vulval cancer. Before removal of the tissue from the patient Raman spectra should be measured radially at intervals from the centre of the disease lesion to the margin of the resection specimen, typically 10-15 millimetres from the visible disease margin. The measurement sites should be marked and the spectra obtained correlated with the subsequent gold standard histopathological analysis. The aim would be to replicate the biochemical signals we have measured ex vivo within the clinical setting, thus providing evidence that the approach would provide in vivo real-time discrimination of disease.

13.4.1.4 Work package four - pathway to implementation and applications in clinical practice

In parallel with the other work packages clinical viability and preference evaluations will be undertaken. This should include an assessment of clinicians' preferences for probe design to ensure the probe is technically suited to the proposed application. Prototype probe hand pieces should be generated using rapid prototyping techniques and usability tests performed to ensure smooth ergonomic action of the probes within the clinical environment. Clinicians should also be engaged to co-create instrument readout methodologies to ensure the spectral analysis can be presented to clinicians in a safe, accurate and clear format.

Alongside this the barriers to implementation should be assessed through clinician engagement exercises with national and international clinical organisations. This

work programme would pave the way for stakeholder engagement with the wider clinical trials needed prior to implementation in clinical practice.

13.4.1.5 Work package five - PPI Involvement

This work package involves engagement with gynaecological patients and would create a focus group of interested individuals from the patient population to collaborate on the experimental design of this and future studies. This will help develop the next phase of clinical studies to follow on from the research and create a drive for adoption of the technique from the patient base. The focus group would also be used to help understand the potential impact of novel diagnostic technology on patients and patients' attitudes towards minimising vulval surgery. Patient representatives would need to be involved in the development of the recruitment packs and information leaflets key for any future studies.

13.4.1.6 Summary

The study suggested as the next phase of this research would translate the Raman technique from the laboratory to the clinical environment by creating and validating a novel diagnostic instrument ready for future clinical trials. This would be necessary for a large-scale multicentre in vivo clinical trial to evaluate the efficacy of the technique. The design of the large scale in vivo clinical trial is beyond the scope of this thesis and would be part of any future research.

13.4.2 Further FTIR spectroscopic diagnostic studies

To further evaluate FTIR spectroscopy as a tool for aiding the histopathological diagnostic pathway studies are required to evaluate the performance of the technique in conditions closer to the standard pathological workflow. Additional types of vulval pathology and unclassifiable vulval tissue need to be added to the classification model to allow the true 'real life' performance of the technique to be assessed. In addition the role of the technique as a second opinion in the diagnosis of vulval SCC should also be assessed within the clinical pathway. This could be achieved by the FTIR spectroscopic evaluation of difficult to diagnose cases of possible SCC. This approach could more accurately assess the potential for the technique to act as a spectroscopic second opinion for the confirmation of the presence of SCC.

13.4.3 Longitudinal studies

In addition to the future research proposed in chapter 13.4.1 longitudinal studies should be undertaken to explore the role of Raman and FTIR spectroscopy as a biomarker of malignant progression from LS and VIN to SCC and as a biomarker for recurrence of VIN or SCC. This could be undertaken using high quality ex vivo spectra measurements using the methodologies described in Chapter 6 and Chapter 12. To assess a correlation between Raman and FTIR spectra and malignant progression samples of tissue containing LS or VIN from patients receiving topical treatment rather than excisional surgery could be assessed with Raman and FTIR

spectroscopy at baseline and the patients followed up to monitor for recurrence. As the rate of progression from LS to vulval SCC is much lower than for VIN this is likely to prove difficult, as a large trial would be required. In addition the acquisition of tissue from all patients with LS would be challenging, as many do not undergo diagnostic biopsy. It is likely that longitudinal studies of the prognostic value of Raman spectroscopy in LS may need to be postponed until an in vivo probe is available. Assessing the correlation between Raman and FTIR spectral characteristics and recurrence of vulval cancer is less challenging given the availability of tissue from excisional surgery and the higher rate of recurrence compared with malignant progression in LS and VIN. The main challenge would be how to assess the biomolecular changes present in the residual tissue left in the vulva as this is the site of disease recurrence. The non-neoplastic tissue margin excised at the time of surgery could however be used as a proxy for the residual tissue elsewhere on the vulva.

13.5 Conclusions

Raman spectroscopy is a sophisticated technique well suited as an adjunct to clinical assessment and histopathology. The technique has a well-established ability to accurately discriminate a variety of different skin pathologies in non-genital sites (Feng *et al.*, 2018; Zhang *et al.*, 2018; Ferreira Lima *et al.*, 2019; Khristoforova *et al.*, 2019). The focus of these studies at non genital skin sites has been on common cancers at these sites, including basal cell carcinomas and melanomas which have different Raman spectral profiles to SCC (Feng *et al.*, 2018). No other studies to date have demonstrated the ability of Raman spectroscopy in differentiating the unique set of pathologies present in the vulva. The research presented in the previous chapters of this thesis constitutes an exploratory analysis of Raman spectroscopy in vulval disease. This study has demonstrated the technique is effective at differentiating a core set of vulval diseases not previously evaluated by Raman spectroscopy.

Considering the potential sources for type one and type two errors outlined above, on balance it appears that Raman spectroscopy has real potential as a tool for reducing clinicians reliance on invasive tissue biopsy and for targeting biopsies to greatest effect. The diagnostic accuracies achieved in this study do not match that of histopathology and may not be sufficient to replace invasive tissue biopsy entirely but they show the technique has the potential to reduce reliance on invasive biopsy.

Furthermore the technique has real potential for the determination of disease margins at the time of excisional surgery for neoplastic vulval disease and for increasing clinicians' confidence in minimising surgical margins. In this context Raman spectroscopy would be used alongside clinical assessment in a multimodal approach to decide whether tissue needs to be excised.

The data suggests a biomolecular continuum of disease is evident in the Raman spectra, from frank malignancy at one end to benign vulval tissue at the other. This suggests that the technique may be useful for monitoring disease progression and response to non-surgical treatments.

Development of an in vivo probe and an in vivo trial is warranted and feasible. This would allow assessment of the potential impact of the technique on the care of patients with vulval disease. Further longitudinal studies are also warranted to assess if the technique has potential in monitoring risk of disease progression and response to treatment.

FTIR spectroscopy offers a technique for molecular assessment that is straightforward to apply to routinely prepared tissue; that has a relatively low cost; that only requires a small amount of tissue and that can be performed in a relatively short amount of time (Old et al. 2014). These characteristics make FTIR spectroscopy suitable as an adjunct to routine histopathology where additional analysis of tissue samples is required. Previous research by Baker et al. demonstrated the potential of FTIR spectroscopy in evaluating the biopotential of prostate cancer, however to our knowledge this is the first time FTIR spectroscopy

has been used to assess the molecular changes associated with vulval cancer (Baker et al. 2008).

FITR-S has the potential to provide a technique that detects the genesis of malignant disease rather than just the presence of cancer. Such a technique would be useful for early cancer detection and informing prognosis. This study highlights the potential of FTIR spectroscopy for the molecular risk stratification of women with LS or uVIN. Further work in the form of longitudinal studies is required to determine if this preliminary work can be transformed into a technique that will improve outcomes for women.

In summary advanced diagnostic techniques show significant promise in improving the care of women with vulval diseases by reducing the reliance on traditional biopsy and histopathology and augmenting clinicians ability to differentiate different types of vulval disease. The work outlined in this thesis evaluates the role of emerging techniques in vibrational spectroscopy to address this need with three key themes underpinning the various aspects of this thesis:

1. Development of a diagnostic technique to reduce the reliance on traditional biopsy and histopathological diagnosis
2. Development of a diagnostic technique for improving the delineation of disease margins at the time of surgery for pre-malignant and malignant vulval conditions.

3. Evaluation of a tool for augmenting and automating aspects of histopathology to reduce the workload on busy histopathology departments.

This thesis demonstrates the considerable potential of vibrational spectroscopy in this clinical setting. The research has made significant progress in each of the three themes outlined above and indicates that further work is warranted to develop the techniques towards routine clinical application.

Section F: Appendices

Appendix A Regulatory Issues

A.1 Ethical approval

Ethical approval for the analysis of fresh and fresh frozen vulval and lymph node tissue was granted by the National Research Ethics Service Committee South West – Exeter (REC reference - 14/SW/1077).

Ethical approval for the analysis of archived vulval and lymph node tissue was granted by the East of Scotland Research Ethics Service (REC reference - 14/ES/1066).

A.2 Patient consent and patient information

Consent was obtained for the use of fresh or fresh frozen vulval and lymph node tissue. Consent was sought from each participant only after a full explanation had been given, an information sheet offered (Appendix 402) and time allowed for consideration. Signed participant consent (Appendix C.2) was obtained prior to analysis of tissue samples. The right of the participant to refuse to participate without giving reasons was respected.

A.3 Confidentiality

The confidentiality of participants taking part in this study was preserved in accordance with the Data Protection Act 1998.

A.4 Indemnity

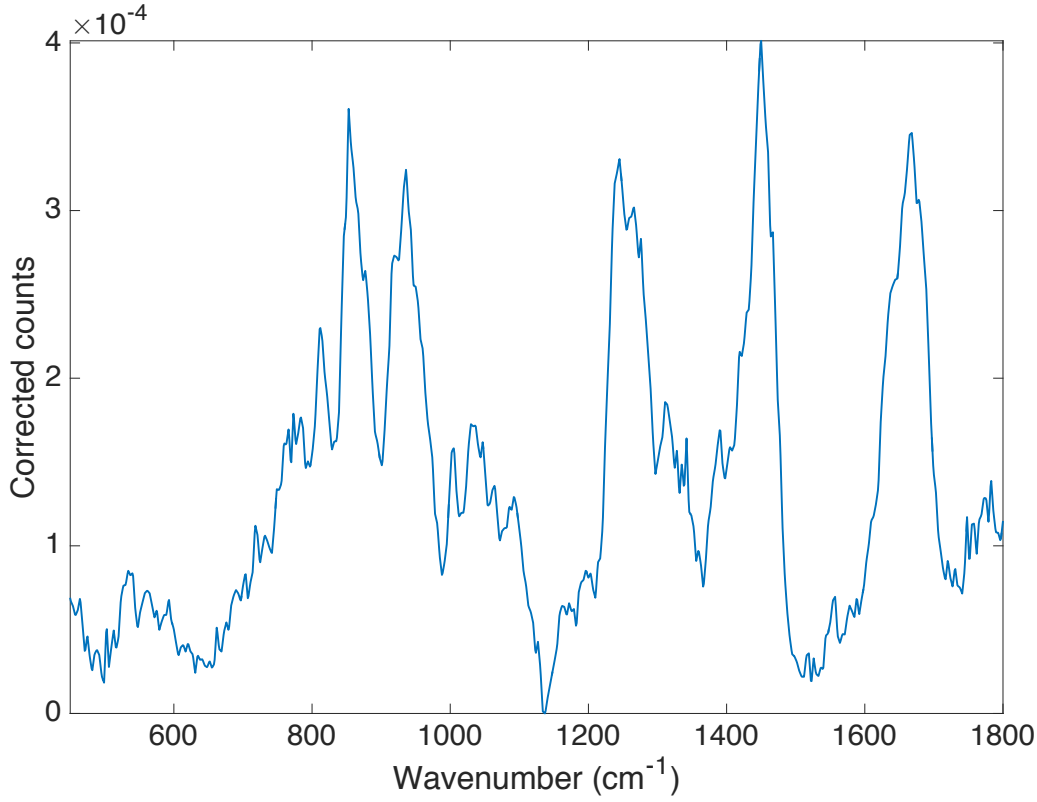
Gloucestershire Hospitals NHS Foundation Trust held standard NHS hospital indemnity and insurance cover with NHS Litigation Authority for NHS Trusts in England, which applied to this study.

A.5 Adverse events

Errors in processing fresh tissue samples required for histopathological diagnosis as part of patients routine care could have led to failure of diagnosis. The failure to diagnose a patient's condition has a significant impact on the care they receive. In light of this sample processing errors made during the study and were reportable. The impact of any processing errors was assessed and if necessary the clinical team and patient informed.

Appendix B Reference Spectra

B.1 Raman Reference Spectra



Raman spectrum of human collagen type I Sigma-Aldrich 068k3775

Appendix C Participant Information and Consent

C.1 Participant information sheet

Gloucestershire Hospitals NHS
NHS Foundation Trust

Project Title: Vulval Raman Spectroscopy

Project Title (Full): Evaluation of the role of Raman spectroscopy in the diagnosis and management of disease of the vulva.

Biophotonics Research Unit
Gloucestershire Royal Hospital
Great Western Road
Gloucester
GL1 3NN
Switchboard: 0300 422 2222

Participant Information Sheet: Vulval Raman Spectroscopy

Study Title: Evaluation of the role of Raman spectroscopy in the diagnosis and management of disease of the vulva.

(Developing a new technique for the assessment of vulval skin conditions)

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

Thank you for reading this.

What is the purpose of the study?

This study is being undertaken to develop a new non invasive method for use in the assessment of vulval skin conditions. It is hoped that this technique will improve the care of women with these conditions and reduce the need for invasive tissue biopsies.

Why have I been chosen?

You have been chosen to take part in this study because your gynaecologist would like you to have a biopsy or surgery as part of the investigation or treatment of a vulval skin condition. You may have already had the biopsy and this does not necessarily stop you from participating in the study.

What is the technique that is being tested?

The diagnostic technique being tested is called Raman spectroscopy. This technique allows us to determine the chemical composition of tissues by measuring the changes to specific colours of light shone onto the tissue. This technique is currently being researched as a way of diagnosing medical conditions in other parts of the body such as the cervix, breast and oesophagus.

This technique has not been used in this part of the body before and this is an initial study to try and establish if the type of vulval disease women have can be identified using this technique. The study will also look at the ability of the new test to identify women at risk of developing different types of vulval disease. If the study were successful then we would aim to make a device to assess different types of vulval disease using a painless technique without the need for tissue biopsies.

Do I have to take part?

It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason.

A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you receive.

What will happen to me if I take part?

If you decide to take part in the study your normal care will not be affected. The study involves answering a few questions about your health and an additional test of your biopsy specimen or tissue removed at the time of surgery. Participation in the study will not significantly affect or delay the results of your routine tests. The research team will then monitor your health records to compare the results of the new test with the routine tests, treatments and your long term vulval health. After answering the initial questions you will not be required to contribute any more of your time to the study. The research team will gather information about you and will follow your progress by referring to your medical records without contacting you further. These medical records may be followed up for some years.

If you take part in the study then your tissue sample will immediately be transported to the research laboratory where it will undergo the new diagnostic test. The sample will then be transported direct to the hospital laboratory for routine testing. After the new test has been performed on samples from many different women analysis will be carried out to see if the test can be used to diagnose different vulval diseases without the need for biopsies.

If you are undergoing surgical treatment of a vulval condition the research team may ask for your consent to retain for analysis some of the surplus tissue that is normally discarded without testing.

Who else will be taking part?

The research team anticipate that a total of 100 women will participate in the study.

What are the possible disadvantages and risks of taking part?

We do not anticipate that taking part in this study will affect the care you receive from your doctor. As the start of normal sample processing occurs daily (first thing in the morning) all samples collected during the day are processed at the same time the subsequent day regardless of when they arrive. Samples from women participating in the study will be undergo routine testing at the same time as samples from women not participating in the study. It is therefore not anticipated that participation in this study will cause a delay in the issuing of histology results and should not affect the timing of any further treatment you may require. Normally it takes around a week to ten days for the laboratory to produce the results from tissue samples sent for routine analysis.

As we are comparing the new diagnostic technique to the standard analysis the research findings will not affect your care or give additional information, but could contribute to improved tests for the future.

What are the possible benefits of taking part?

There is no benefit to your care if you participate in the study, however you will be contributing to the development of a new diagnostic technique. If successful this will give patients access to instant diagnosis of vulval conditions without the need for tissue biopsies.

Will my taking part in this study be kept confidential?

All information which is collected about you during the course of the research will be kept strictly confidential. Any information about you which leaves the hospital will have your name and address removed so that you cannot be recognised from it.

What will happen to the results of the research study?

You will not be identified in any report or publication produced as part of this study. Results will be submitted for publication in scientific journals and may be presented at scientific meetings. Initial results will be submitted for publication in 2016. If you would like a copy of the results of the study please contact the principal investigator or project supervisor listed at the end of this information sheet.

Who is organising and funding the research?

The study is being organised by the Biophotonics Research Unit and is funded by the Research and Innovation Forum, Gloucestershire Hospitals NHS Foundation Trust.

Who has reviewed the study?

This study has been reviewed by the Exeter Research Ethics Committee.

What if something goes wrong?

Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. Please raise your concerns in the first instance with the Principal Investigator (that is the lead researcher), Dr Jonathan Frost, his contact details are at the end of this form. If you wish to make a more formal complaint, please contact the hospital's Patient Advice and Liaison Service (PALS) (PALS Office, Gloucestershire Royal Hospital, Great Western Road, Gloucester, GL1 3NN, telephone: 0800 019 3282, e-mail: pals.gloucestershirehospitals@glos.nhs.uk).

What could I do to help researchers use my own experience more when they are planning their next project?

Do you want to get more involved and help researchers improve future project ideas and research information leaflets? Please contact "People in Research – Opportunities for public involvement in research" http://www.peopleinresearch.org/. If you would like to help, you can also contact INVOLVE which is a national advisory group, funded by the National Institute for Health Research (NIHR). Its role is to support and promote active public involvement in NHS, public health and social care research. http://www.invo.org.uk/ or Wessex House, Upper Market Street, Eastleigh, Hampshire, SO50 9FD, Telephone: 02380 651088 Email admin@invo.org.uk

Which insurance provisions are in place?

In the unlikely event that something does go wrong and you are harmed during the research and this is due to someone's negligence, then you may have grounds for a legal action for compensation against the Gloucestershire Hospitals NHS Foundation Trust but you may have to pay your legal costs. The normal National Health Service complaints mechanism will still be available to you.

This research study is being completed as part of a MD (Doctor of Medicine) qualification being undertaken by the principal investigator.

Thank you for taking the time to read this information and considering participation in this research study.

Contact for Further Information

Principal Investigator
Dr Jonathan Frost
Clinical Research Fellow & MD Student
Biophotonics Research Unit
Gloucestershire Royal Hospital
Great Western Road,
Gloucester,
GL1 3NN
Tel: +44 (0)300 422 5473
email: jonathan.a.frost@glos.nhs.uk

Project Supervisor
Dr Catherine Kendall
Scientific Lead, Biophotonics Research Unit
Gloucestershire Royal Hospital
Great Western Road,
Gloucester,
GL1 3NN
Tel: +44 (0)300 422 5470
email: Catherine.kendall@glos.nhs.uk

Version 1.2 03/11/2014

C.2 Participant consent form

Gloucestershire Hospitals NHS

NHS Foundation Trust

Project Title: Vulval Raman Spectroscopy

Project Title (Full): Evaluation of the role of Raman spectroscopy in the diagnosis and management of disease of the vulva.

Biophotonics Research Unit
Gloucestershire Royal Hospital
Great Western Road
Gloucester
GL1 3NN
Switchboard: 0300 422 2222

CONSENT FORM

Study Number: 14-SW-1077 / 14-072-GHT

Patient Identification Number for this trial:

Please initial all boxes

1. I confirm that I have read and understand the information sheet (version 1.1, September 2014) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected. ☐
3. I understand that sections of my medical notes will be looked at by the research team, now and in the future, where it is relevant to the research. I give permission for these individuals to have access to my records. ☐
4. I understand that relevant sections of my medical notes and data collected during the study may be looked at by individuals from Gloucestershire Hospitals NHS Foundation Trust or from regulatory authorities, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. ☐
5. I agree to surplus tissue not required for the care or diagnosis of my condition to be retained by the research team. ☐
6. I agree to take part in the above study. ☐

Name of participant	Date	Signature
Name of person taking consent.	Date	Signature

Please note that one copy of this form will be kept in your clinical notes, one copy will be retained by the research team and one will be available for your personal use.

Version 1.2 03/11/2014

BETTER FOR YOU

Appendix D Prizes, Awards, Presentations and Publications during PhD

Awards and prizes

1st prize in the poster competition at SPEC 2016 (clinical spectroscopy) conference, Montreal, Canada	2016
Awarded CLIRSPEC bursary to present research work at SPEC 2016 (clinical spectroscopy) conference, Montreal, Canada	2016
Award for best oral presentation of the oncology and colposcopy session, RCOG World Congress	2016
Awarded Royal Society of Chemistry bursary to present research work at the prestigious Advanced Vibrational Spectroscopy for Biomedical Applications: Faraday Discussion	2016
Royal Society of Chemistry Skinner Prize for work presented at the Advanced Vibrational Spectroscopy for Biomedical Applications: Faraday Discussion	2016
High scoring abstract, RCOG Annual Academic Meeting	2016
Prize winning paper – Oral Presentation at British Gynaecological Cancer Society Annual Scientific Meeting	2015
British Society for the Study of Vulval Disease Research Award	2014

Peer reviewed papers

Frost J, Ludeman L, Hillaby K, Gornall R, Kendall C, Shore A, Stone N. Raman spectroscopy and multivariate analysis for the non invasive diagnosis of clinically inconclusive vulval lichen sclerosus. The Analyst. 2017;142(8):1200-1206

Frost J, Ludeman L, Hillaby K, Gornall R, Kendall C, Shore A, Stone N. Identification of cancer associated molecular changes in histologically benign vulval disease found in association with vulval squamous cell carcinoma using Fourier transform infrared spectroscopy. Anal. Methods. 2016;8(48);8452-8460

Peer reviewed published abstracts and commentaries

Frost J, Ludeman L, Hillaby K, Gornall R, Kendall C, Shore A, Stone N. Raman spectroscopy and multivariate analysis for the non invasive diagnosis of Vulval lichen sclerosus. BJOG An Int J Obstet Gynaecol. 2016;123:78-82

Frost J, Ludeman L, Hillaby K, Gornall R, Kendall C, Shore A, Stone N. Infrared spectroscopy for the biomolecular assessment of vulval intraepithelial neoplasia and lichen sclerosus found in association with squamous cell carcinoma. BJOG An Int J Obstet Gynaecol. 2016;123:14-18

International presentations

Oral Presentation at RCOG World Congress 2016 – Raman spectroscopy and multivariate analysis for the non invasive diagnosis of Vulval lichen sclerosus. Frost J, Ludeman L, Hillaby K, Gornall R, Kendall C, Shore A, Stone N

Oral Presentation at RCOG World Congress 2016 – Infrared spectroscopy for the biomolecular assessment of vulval intraepithelial neoplasia and lichen sclerosus found in association with squamous cell carcinoma. Frost J, Ludeman L, Hillaby K, Gornall R, Kendall C, Shore A, Stone N

Invited Speaker SPEC 2016 (clinical spectroscopy) Conference Montreal, Canada – Raman spectroscopy for the assessment of vulval disease.

Poster Presentation at SPEC 2016 (clinical spectroscopy) Conference Montreal, Canada – Raman Spectroscopy and Multivariate Analysis for the Non Invasive Diagnosis of Benign Vulval Disease. Frost J, Ludeman L, Hillaby K, Gornall R, Kendall C, Shore A, Stone N

Poster Presentation at SPEC 2016 (clinical spectroscopy) Conference Montreal, Canada – Infrared spectroscopy for the biomolecular assessment of vulval disease associated with squamous cell carcinoma. Frost J, Ludeman L, Hillaby K, Gornall R, Kendall C, Shore A, Stone N

Poster Presentation at Advanced Vibrational Spectroscopy for Biomedical Applications: Faraday Discussion 2016 – Vibrational spectroscopy for the biomolecular assessment of vulval intraepithelial neoplasia and lichen sclerosus. Frost J, Ludeman L, Hillaby K, Gornall R, Kendall C, Shore A, Stone N

National presentations

Poster Presentation at British Gynaecological Cancer Society Annual Scientific Meeting 2017 - Raman spectroscopy for the identification of vulval squamous cell carcinoma and vulval intraepithelial neoplasia. Frost J, Ludeman L, Hillaby K, Kendall C, Shore A, Stone N

Invited Speaker CLIRSPEC 2017 (clinical spectroscopy) Conference Manchester, UK – Raman spectroscopy for the assessment of vulval disease.

Oral Presentation at RCOG Annual Academic Meeting 2016 - Infrared Spectroscopy for the biomolecular assessment of molecular field change in the vulva. Frost J, Ludeman L, Hillaby K, Gornall R, Kendall C, Shore A, Stone N

Invited Speaker at British Society of the Study of Vulval Disease Annual Scientific Meeting 2016 - Vibrational Spectroscopy for the Assessment of Vulval Disease

Poster Presentation at British Gynaecological Cancer Society Annual Scientific Meeting 2016 - Infrared spectroscopy for the biomolecular risk assessment of early stage vulval squamous cell carcinoma. Frost J, Ludeman L, Hillaby K, Kendall C, Shore A, Stone N

Poster Presentation at Infrared and Raman Discussion Group Conference 2016 - Fourier Transform Infrared Spectroscopy in the biomolecular assessment of VIN and lichen sclerosus found in association with vulval SCC**.** Frost J, Ludeman L, Hillaby K, Kendall C, Shore A, Stone N

Oral Presentation at British Medical Laser Association, Annual Scientific Meeting 2015 - Vibrational spectroscopy for the assessment of vulval disease
Frost J, Ludeman L, Hillaby K, Kendall C, Shore A, Stone N

Oral Presentation at British Gynaecological Cancer Society Annual Scientific Meeting 2015 - Fourier transform infrared spectroscopy in the assessment of vulval intraepithelial neoplasia and lichen sclerosus
Frost J, Ludeman L, Hillaby K, Kendall C, Shore A, Stone N

Invited Speaker Cancer Research UK Vulval Neoplasia Study Day, University of Birmingham 2015 - Mutational and spectroscopic profiling of the vulval field of cancerisation

Oral Presentation at University of Exeter Annual Research Meeting 2015 - Vibrational Spectroscopy in Premalignant and Malignant Vulval Disease
Frost J, Ludeman L, Hillaby K, Kendall C, Shore A, Stone N

Poster Presentation at CLIRSPEC Scientific Meeting 2015 - Vibrational Spectroscopy for the Assessment of Vulval Disease - Frost J, Ludeman L, Hillaby K, Kendall C, Shore A, Stone N

References

ACOG (2008) ‘ACOG Practice Bulletin No. 93: diagnosis and management of vulvar skin disorders’, *Obstet Gynecol*, 111(93), pp. 1243–1253.

Aernouts, B. *et al.* (2014) ‘Dependent scattering in Intralipid® phantoms in the 600-1850 nm range’, *Optics Express*, 22(5), p. 6086. doi: 10.1364/OE.22.006086.

Aerts, L. *et al.* (2012) ‘Sexual, psychological, and relational functioning in women after surgical treatment for vulvar malignancy: a literature review.’, *The journal of sexual medicine*, 9(2), pp. 361–371.

Ali, S. M. *et al.* (2012) ‘Raman spectroscopic analysis of human skin tissue sections ex-vivo : evaluation of the effects of tissue processing and dewaxing’, *Journal of Biomedical Optics*, 18(6), p. 061202. doi: 10.1117/1.JBO.18.6.061202.

Ali, S. M. *et al.* (2013) ‘A comparison of Raman, FTIR and ATR-FTIR micro spectroscopy for imaging human skin tissue sections’, *Analytical Methods*, 5(9), p. 2281. doi: 10.1039/c3ay40185e.

Ali, S M *et al.* (2013) ‘Raman spectroscopic mapping for the analysis of solar radiation induced skin damage.’, *Analyst*, 138(14), pp. 3946–3956.

Almond, L. M. (2012) *Towards objective endoscopic diagnosis of Barrett's associated early neoplasia using fibre-oprtic Raman spectroscopy*.

Almond, L. M. *et al.* (2014) 'Endoscopic Raman spectroscopy enables objective diagnosis of dysplasia in Barrett's esophagus', *Gastrointestinal Endoscopy*. Elsevier Ltd, 79(1), pp. 37–45. doi: 10.1016/j.gie.2013.05.028.

Andersen, B. L. *et al.* (1988) 'Sexual functioning after treatment of in situ vulvar cancer: preliminary report.', *Obstetrics and gynecology*, 71(1), pp. 15–19.

Andersen, B. L. (1993) 'Predicting sexual and psychologic morbidity and improving the quality of life for women with gynecologic cancer.', *Cancer*, 71(4 Suppl), pp. 1678–1690.

Astner, S. *et al.* (2008) 'Clinical applicability of in vivo fluorescence confocal microscopy for noninvasive diagnosis and therapeutic monitoring of nonmelanoma skin cancer.', *Journal of biomedical optics*, 13(1), p. 014003. doi: 10.1117/1.2837411.

Van der Avoort, I. a. M. *et al.* (2006) 'Vulvar Squamous Cell Carcinoma is a Multifactorial Disease Following Two Separate and Independent Pathways', *International Journal of Gynecological Pathology*, 25(1), pp. 22–29. doi: 10.1097/01.pgp.0000177646.38266.6a.

Bae-Jump, V. L., Bauer, M. and Van Le, L. (2007) 'Cytological evaluation correlates poorly with histological diagnosis of vulvar neoplasias.', *Journal of lower genital tract disease*, 11(1), pp. 8–11. doi: 10.1097/01.lgt.0000229566.57482.f3.

Baggish, M. S. and Karram, M. M. (2016) *Atlas of Pelvic Anatomy and Gynecologic Surgery*. Fourth Edi. London: Elsevier Inc. doi: 10.1016/B978-0-323-22552-6.00065-7.

Bahreini, M. (2015) 'Role of optical spectroscopic methods in neuro-oncological sciences', *Journal of Lasers in Medical Sciences*, 6(2), pp. 51–61.

Bainbridge, S. *et al.* (2016) *Testing Times To Come? An Evaluation of Pathology Capacity Across the UK*.

Baker, M. J. *et al.* (2008) 'FTIR-based spectroscopic analysis in the identification of clinically aggressive prostate cancer.', *British Journal of Cancer*, 99(11), pp. 1859–66. doi: 10.1038/sj.bjc.6604753.

Baker, M. J. *et al.* (2014) 'Using Fourier transform IR spectroscopy to analyze biological materials.', *Nature protocols*, 9(8), pp. 1771–91. doi: 10.1038/nprot.2014.110.

Barry, B. W., Edwards, H. G. M. and Williams, A. C. (1992) 'Fourier-Transform Raman and Infrared Vibrational Study of Human Skin - Assignment of Spectral Bands', *Journal of Raman Spectroscopy*, 23(11), pp. 641–645. doi: 10.1002/jrs.1250231113.

Barth, A. (2007) 'Infrared spectroscopy of proteins', *Biochimica et Biophysica Acta - Bioenergetics*, 1767(9), pp. 1073–1101. doi: 10.1016/j.bbabio.2007.06.004.

Bazant-Hegemark, F. and Stone, N. (2008) 'Near real-time classification of optical coherence tomography data using principal components fed linear discriminant

analysis', *Journal of Biomedical Optics*, 13(3), p. 034002. doi: 10.1117/1.2931079.

Begley, R. F., Harvey, A. B. and Byer, R. L. (1974) 'Coherent anti-Stokes Raman spectroscopy', *Applied Physics Letters*, 25(7), pp. 387–390. doi: 10.1063/1.1655519.

Berek, J. S. (2012) *Berek and Novak's Gynecology*. Fifteenth. Philadelphia: Lippincott Williams and Wilkins. Available at: https://books.google.co.uk/books?id=P3erl0J8tEQC&pg=PA75&lpg=PA75&dq=anderson+anatomy+and+embryology&source=bl&ots=l4wrRGcGE7&sig=bc-Od5vPfw-4EoPEn5TphONUbUo&hl=en&sa=X&ved=0ahUKEwjJvpaN_YDLAhXLVRQKHR6lBZAQ6AEIJDAC#v=onepage&q=anderson anatomy and embryol.

Bergholt, M. S. *et al.* (2011) 'Characterizing variability in in vivo Raman spectra of different anatomical locations in the upper gastrointestinal tract toward cancer detection.', *Journal of biomedical optics*, 16(3), p. 037003. doi: 10.1117/1.3556723.

Bergholt, M. S. *et al.* (2013) 'Fiber-optic Raman spectroscopy probes gastric carcinogenesis in vivo at endoscopy', *Journal of Biophotonics*, 6(1), pp. 49–59. doi: 10.1002/jbio.201200138.

van Beurden, M. *et al.* (1995) 'Multifocal vulvar intraepithelial neoplasia grade III and multicentric lower genital tract neoplasia is associated with transcriptionally active human papillomavirus.', *Cancer*, 75(12), pp. 2879–84.

Bhargava, R. (2007) 'Towards a practical Fourier transform infrared chemical imaging protocol for cancer histopathology', *Analytical and Bioanalytical Chemistry*, 389(4), pp. 1155–1169. doi: 10.1007/s00216-007-1511-9.

Bigio, I. J. and Mourant, J. R. (1997) 'Ultraviolet and visible spectroscopies for tissue diagnostics: fluorescence spectroscopy and elastic-scattering spectroscopy.', *Physics in medicine and biology*, 42(5), pp. 803–814. doi: 10.1088/0031-9155/42/5/005.

Bleeker, M. C. G. *et al.* (2016) 'Lichen sclerosus: Incidence and risk of vulvar squamous cell carcinoma', *Cancer Epidemiology Biomarkers and Prevention*, 25(8), pp. 1224–1230. doi: 10.1158/1055-9965.EPI-16-0019.

Bodanese, B. *et al.* (2010) 'Differentiating Normal and Basal Cell Carcinoma Human Skin Tissues In Vitro Using Dispersive Raman Spectroscopy: A Comparison Between Principal Components Analysis and Simplified Biochemical Models', *Photomedicine and Laser Surgery*, 28(S1), p. S-119-S-127. doi: 10.1089/pho.2009.2565.

Bodanese, B. *et al.* (2012) 'Discrimination of basal cell carcinoma and melanoma from normal skin biopsies in vitro through Raman spectroscopy and principal component analysis.', *Photomedicine and Laser Surgery*, 30(7), pp. 381–387.

Borisova, E. *et al.* (2010) 'Qualitative optical evaluation of malignancies related to cutaneous phototype', in Tuchin, V. V., Duncan, D. D., and Larin, K. V. (eds), p. 75630X. doi: 10.1117/12.852791.

Borisova, E. *et al.* (2012) 'Optical biopsy of human skin - A tool for cutaneous tumours' diagnosis', *International Journal Bioautomotion*, 16(1), pp. 53–72.

Bourke, J., Coulson, I. and English, J. (2009) 'Guidelines for the management of

contact dermatitis: An update', *British Journal of Dermatology*, 160(5), pp. 946–954. doi: 10.1111/j.1365-2133.2009.09106.x.

Braakhuis, B. J. M. *et al.* (2003) 'A Genetic Explanation of Slaughter ' s Concept of Field Cancerization : Evidence and Clinical Implications', *Cancer Res*, 63, pp. 1727–1730.

Brancaleon, L. *et al.* (2001) 'Attenuated Total Reflection–Fourier Transform Infrared Spectroscopy as a Possible Method to Investigate Biophysical Parameters of Stratum Corneum In Vivo', *Journal of Investigative Dermatology*, 116(3), pp. 380–386. doi: 10.1046/j.1523-1747.2001.01262.x.

Brereton, R. (2003) *Chemometrics*. 1st edn. Wiley.

Brereton, R. (2009) *Chemometrics for pattern recognition*. 1st edn. Chichester: Wiley.

Brereton, R. G. and Lloyd, G. R. (2016) 'Re-evaluating the role of the Mahalanobis distance measure', (January). doi: 10.1002/cem.2779.

Bruun, S. W. *et al.* (2006) 'Correcting Attenuated Total Reflection – Fourier Transform Infrared Spectra for Water Vapor and Carbon Dioxide', *Applied Spectroscopy*, 60(9), pp. 1029–1039.

Burke, T. W. *et al.* (1995) 'Surgical therapy of T1 and T2 vulvar carcinoma: further experience with radical wide excision and selective inguinal lymphadenectomy.', *Gynecologic oncology*, pp. 215–220. doi: 10.1006/gyno.1995.1128.

Burkitt, H., Young, B. and Heath, J. (1995) *Wheater's Functional Histology*. Third

Edit. Churchill Livingstone.

Butler, H. J. *et al.* (2016) 'Using Raman spectroscopy to characterise biological materials', *Nature Protocols*. Nature Publishing Group, 11(4), pp. 1–47. doi: 10.1038/nprot.2016.036.

Calin, M. A. *et al.* (2013) 'Optical techniques for the noninvasive diagnosis of skin cancer', *Journal of Cancer Research and Clinical Oncology*, 139(7), pp. 1083–1104. doi: 10.1007/s00432-013-1423-3.

Cancer Research UK (2016) *Cancer Research UK. UK Vulva Cancer statistics.* Available at: http://info.cancerresearchuk.org/cancerstats/ (Accessed: 20 February 2016).

Cappel, U. B., Bell, I. M. and Pickard, L. K. (2010) 'Removing cosmic ray features from Raman map data by a refined nearest neighbor comparison method as a precursor for chemometric analysis.', *Applied spectroscopy*, 64(2), pp. 195–200. doi: 10.1366/000370210790619528.

Carli, P. *et al.* (1995) 'Squamous cell carcinoma arising in vulval lichen sclerosus: a longitudinal cohort study.', *European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation (ECP)*, 4(6), pp. 491–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8580785.

Carlson, J. a *et al.* (1998) 'Vulvar lichen sclerosus and squamous cell carcinoma: a cohort, case control, and investigational study with historical perspective; implications for chronic inflammation and sclerosis in the development of neoplasia.',

Human pathology, 29(9), pp. 932–48. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9744309.

Cartaxo, S. B. *et al.* (2010) 'FT-Raman spectroscopy for the differentiation between cutaneous melanoma and pigmented nevus.', *Acta Cirurgica Brasileira*, 25(4), pp. 351–356.

Caspers, P. *et al.* (1998) 'In vitro and in vivo Raman spectroscopy of human skin.', *Biospectroscopy*, 4(5 Suppl), pp. S31-9.

Caspers, P. (2003) *In vivo skin characterization by confocal Raman microspectroscopy*. Erasmus Universiteit Rotterdam.

Caspers, P. J. *et al.* (2001) 'In vivo confocal Raman microspectroscopy of the skin: noninvasive determination of molecular concentration profiles.', *Journal of Investigative Dermatology*, 116(3), pp. 434–442.

Caspers, P., Lucassen, G. W. and Puppels, G. J. (2003) 'Combined in vivo confocal Raman spectroscopy and confocal microscopy of human skin.', *Biophysical Journal*, 85(1), pp. 572–580.

Cavanagh, D. and Hoffman, M. S. (1996) 'Controversies in the management of vulvar carcinoma', *BJOG: An International Journal of Obstetrics and Gynaecology*, 103(4), pp. 293–300. doi: 10.1111/j.1471-0528.1996.tb09731.x.

Centre for Workforce Intelligence (2010) *Medical Specialty Workforce Factsheet - Histopathology*. Available at: http://www.cfwi.org.uk.

Chan, J. K. *et al.* (2007) 'Margin distance and other clinico-pathologic prognostic

factors in vulvar carcinoma: A multivariate analysis', *Gynecologic Oncology*, 104(3), pp. 636–641. doi: 10.1016/j.ygyno.2006.10.004.

Chi, C.-C. *et al.* (2012) 'Systematic review and meta-analysis of randomized controlled trials on topical interventions for genital lichen sclerosus.', *Journal of the American Academy of Dermatology*. Elsevier Inc, 67(2), pp. 305–12. doi: 10.1016/j.jaad.2012.02.044.

Choi, J. *et al.* (2005) 'Direct observation of spectral differences between normal and basal cell carcinoma (BCC) tissues using confocal Raman microscopy.', *Biopolymers*. Department of Applied Chemistry, Hanyang University, Ansan 426-791, South Korea., 77(5), pp. 264–272.

Cinotti, E. *et al.* (2013) 'Characterization of cutaneous foreign bodies by Raman spectroscopy.', *Skin Research & Technology*, 19(4), pp. 508–509. doi: 10.1111/srt.12065.

Clark, R. (2005) *Scientific Examination of Art*, *Scientific Examination of Art: Modern Techniques in Conservation and Analysis*. Washington, D.C.: National Academies Press. doi: 10.17226/11413.

Cooper, S. M. *et al.* (2004) 'Does Treatment of Vulvar Lichen Sclerosus Influence Its Prognosis?', *Archives of Dermatology*, 140(6), pp. 702–6. doi: 10.1001/archderm.140.6.702.

Cooper, S. M. and Wojnarowska, F. (2006) 'Influence of treatment of erosive lichen planus of the vulva on its prognosis.', *Archives of dermatology*, 142(3), pp. 289–94.

doi: 10.1001/archderm.142.3.289.

Coulter, J. and Gleeson, N. (2003) 'Local and regional recurrence of vulval cancer: Management dilemmas', *Best Practice and Research: Clinical Obstetrics and Gynaecology*, 17(4), pp. 663–681. doi: 10.1016/S1521-6934(03)00050-6.

Dakubo, G. D. *et al.* (2007) 'Clinical implications and utility of field cancerization.', *Cancer cell international*, 7, p. 2. doi: 10.1186/1475-2867-7-2.

Das, K. *et al.* (2006) 'Raman spectroscopy of parathyroid tissue pathology', *Lasers in Medical Science*, 21(4), pp. 192–197. doi: 10.1007/s10103-006-0397-7.

Das, K. *et al.* (2008) 'FTIR of touch imprint cytology: A novel tissue diagnostic technique', *Journal of Photochemistry and Photobiology B: Biology*, 92(3), pp. 160–164. doi: 10.1016/j.jphotobiol.2008.05.012.

Day, J. C. C. *et al.* (2009) 'A miniature confocal Raman probe for endoscopic use', *Physics in Medicine and Biology*, 54(23), pp. 7077–7087. doi: 10.1088/0031-9155/54/23/003.

Day, J. C. C. and Stone, N. (2013) 'A subcutaneous Raman needle probe', *Applied Spectroscopy*, 67(3). doi: 10.1366/12-06651.

Decesare, S. L. *et al.* (1997) 'A pilot study utilizing intraoperative lymphoscintigraphy for identification of the sentinel lymph nodes in vulvar cancer.', *Gynecologic oncology*, 66(3), pp. 425–8. doi: 10.1006/gyno.1997.4798.

Deckert, V. (2009) 'Tip-Enhanced Raman Spectroscopy', *Journal of Raman Spectroscopy*, 40(10), pp. 1336–1337. doi: 10.1002/jrs.2452.

Defernez, M. and Kemsley, E. K. (1997) ‘The use and misuse of chemometrics for treating classification problems’, *TrAC - Trends in Analytical Chemistry*, 16(4), pp. 216–221. doi: 10.1016/S0165-9936(97)00015-0.

DeSimone, C. P. *et al.* (2006) ‘Concordance of gross surgical and final fixed margins in vulvar intraepithelial neoplasia 3 and vulvar cancer.’, *The Journal of reproductive medicine*, 51(8), pp. 617–20. doi: 10.1097/IGC.0000000000001195.

Dominiak-Felden, G. *et al.* (2013) ‘Impact of human papillomavirus-related genital diseases on quality of life and psychosocial wellbeing: results of an observational, health-related quality of life study in the UK.’, *BMC public health*, 13, p. 1065. doi: 10.1186/1471-2458-13-1065.

Drakaki, E. *et al.* (2009) ‘Laser-induced fluorescence and reflectance spectroscopy for the discrimination of basal cell carcinoma from the surrounding normal skin tissue’, *Skin Pharmacology and Physiology*, 22(3), pp. 158–165. doi: 10.1159/000211912.

Drakaki, E. *et al.* (2013) ‘Spectroscopic methods for the photodiagnosis of nonmelanoma skin cancer.’, *Journal of biomedical optics*, 18(6), p. 61221. doi: 10.1117/1.JBO.18.6.061221.

Duda, R., Hart, P. and Stork, D. (2001) *Pattern Classification*. 2nd edn. Wiley-Interscience.

Edwards, S. K. *et al.* (2015) ‘UK national guideline on the management of vulval conditions’, *International Journal of STD & AIDS*, 26(9), pp. 611–624. doi:

10.1177/0956462414554271.

Eerens, A. (2010) *GROningen INternational Study on Sentinel nodes in Vulvar cancer (GROINSS-V) II Protocol*.

Egodage, K. D. *et al.* (2015) 'The combination of optical coherence tomography and Raman spectroscopy for tissue characterization', 72(June), pp. 169–177.

Eikje, N. S., Aizawa, K. and Ozaki, Y. (2005) 'Vibrational spectroscopy for molecular characterisation and diagnosis of benign, premalignant and malignant skin tumours.', *Biotechnology Annual Review*, 11, pp. 191–225.

Van den Einden, L. *et al.* (2012) 'Cytology of the vulva: feasibility and preliminary results of a new brush.', *British Journal of Cancer*, 106(2), pp. 269–273. Available at: Available.

Epstein, E. J. and Munderloh, N. (1978) 'Human skin collagen. Presence of type I and type III at all levels of the dermis.', *J Biol Chem*, 253(5), pp. 1336–7.

European Society of Gynaecological Oncology (2016) 'Vulvar Cancer Guidelines', pp. 1–75.

Eva, L. J. *et al.* (2009) 'Differentiated-Type Vulval Intraepithelial Neoplasia Has a High-Risk Association With Vulval Squamous Cell Carcinoma', *International Journal of Gynecological Cancer*, 19(4), pp. 741–744. doi: 10.1111/IGC.0b013e3181a12fa2.

Farage, M. and Maibach, H. (eds) (2016) *The Vulva: Anatomy, Physiology, and Pathology*.

Farries, M. *et al.* (2015) ‘Mid infra-red hyper-spectral imaging with bright super continuum source and fast acousto-optic tuneable filter for cytological applications.’, *Journal of Physics: Conference Series*, 619(1), p. 012032. doi: 10.1088/1742-6596/619/1/012032.

Fehr, M. K. *et al.* (2013) ‘Disease progression and recurrence in women treated for vulvovaginal intraepithelial neoplasia’, *J Gynecol Oncol*, 24(3), pp. 236–241. doi: 10.3802/jgo.2013.24.3.236.

Feng, X. *et al.* (2018) ‘Raman biophysical markers in skin cancer diagnosis’, *Journal of Biomedical Optics*, 23(05), p. 1. doi: 10.1117/1.jbo.23.5.057002.

Ferreira Lima, A. M. *et al.* (2019) ‘Discrimination of non-melanoma skin cancer and keratosis from normal skin tissue in vivo and ex vivo by Raman spectroscopy’, *Vibrational Spectroscopy*. Elsevier, 100(October 2018), pp. 131–141. doi: 10.1016/j.vibspec.2018.11.009.

Fistarol, S. K. and Itin, P. H. (2013) ‘Diagnosis and treatment of lichen sclerosus: An update’, *American Journal of Clinical Dermatology*, 14(1), pp. 27–47. doi: 10.1007/s40257-012-0006-4.

Fleming, I. D. *et al.* (1995) ‘Principles of management of basal and squamous cell carcinoma of the skin’, *Cancer*, 75(2 S), pp. 699–704. doi: 10.1002/1097-0142(19950115)75:2+<699::AID-CNCR2820751413>3.0.CO;2-Q.

Fox, H. and Wells, M. (2003) ‘Recent advances in the pathology of the vulva’, *Histopathology*, 42(3), pp. 209–216. doi: 10.1046/j.1365-2559.2003.01578.x.

Frable, W. J. (2006) ‘Surgical pathology--second reviews, institutional reviews, audits, and correlations: what’s out there? Error or diagnostic variation?’, *Archives of pathology & laboratory medicine*, 130(5), pp. 620–5. doi: 10.1043/1543-2165(2006)130[620:SPRIRA]2.0.CO;2.

Frienkel, R. and Woodley, D. (2001) *The biology of the skin*. 1st edn. New York: Parthenon.

Gajjar, K. *et al.* (2012) ‘Diagnostic segregation of human brain tumours using Fourier-transform infrared and/or Raman spectroscopy coupled with discriminant analysis’, *Analytical Methods*, 5, pp. 89–102. doi: 10.1039/c2ay25544h.

Gartner, L. P. (ed.) (2017) *Textbook of Histology*. Fourth Edi. Philadelphia: Elsevier Inc.

De Gelder, J. *et al.* (2007) ‘Reference database of Raman spectra of biological molecules’, *Journal of Raman Spectroscopy*, 38(9), pp. 1133–1147. doi: 10.1002/jrs.1734.

Gerger, A. *et al.* (2005) ‘Confocal examination of untreated fresh specimens from basal cell carcinoma: implications for microscopically guided surgery’, *Arch Dermatol*, 141(10), pp. 1269–1274. doi: 10.1001/archderm.141.10.1269.

Gniadecka, M. *et al.* (1997) ‘Distinctive molecular abnormalities in benign and malignant skin lesions: studies by Raman spectroscopy.’, *Photochemistry & Photobiology*, 66(4), pp. 418–423.

Gniadecka, M. *et al.* (2004) ‘Melanoma diagnosis by Raman spectroscopy and

neural networks: structure alterations in proteins and lipids in intact cancer tissue.', *Journal of Investigative Dermatology*, 122(2), pp. 443–449.

Gniadecka, M., Wulf, H. C. and Mortensen, N. N. (1997) 'Diagnosis of Basal Cell Carcinoma by Raman Spectroscopy', *J Raman Spectrosc*, 28(September 1996), pp. 125–129. doi: 10.1002/(SICI)1097.

Godoy, C. *et al.* (2015) 'Unusual remodeling of the hyalinization band in vulval lichen sclerosus by type V collagen and ECM 1 protein', *Clinics*, 70(5), pp. 356–362. doi: 10.6061/clinics/2015(05)09.

Green, M. S. *et al.* (2000) 'Sexual dysfunction following vulvectomy', *Gynecologic Oncology*, 77(1), pp. 73–77. doi: 10.1006/gyno.2000.5745.

Grimbergen, M. C. M. *et al.* (2010) 'Signal-to-noise contribution of principal component loads in reconstructed near-infrared Raman tissue spectra', *Applied Spectroscopy*, 64(1), pp. 8–14. doi: 10.1366/000370210790572052.

Groenen, S. M. A., Timmers, P. J. and Burger, C. W. (2010) 'Recurrence rate in vulvar carcinoma in relation to pathological margin distance', *International Journal of Gynecological Cancer*, 20(5), pp. 869–873. doi: 10.1111/IGC.0b013e3181df7423.

Te Grootenhuis, N. C. *et al.* (2018) 'Prognostic factors for local recurrence of squamous cell carcinoma of the vulva: A systematic review.', *Gynecologic oncology*. Elsevier Inc., 148(3), pp. 622–631. doi: 10.1016/j.ygyno.2017.11.006.

Guerrero, D. *et al.* (2011) 'Differential hypermethylation of genes in vulvar cancer and lichen sclerosus coexisting or not with vulvar cancer.', *International journal of*

cancer. Journal international du cancer, 128(12), pp. 2853–64. doi: 10.1002/ijc.25629.

Guitera, P. *et al.* (2012) 'In Vivo Confocal Microscopy for Diagnosis of Melanoma and Basal Cell Carcinoma Using a Two-Step Method: Analysis of 710 Consecutive Clinically Equivocal Cases', *Journal of Investigative Dermatology*. Nature Publishing Group, 132(10), pp. 2386–2394. doi: 10.1038/jid.2012.172.

Hacker, N. F. *et al.* (1981) 'Radical vulvectomy and bilateral inguinal lymphadenectomy through separate groin incisions.', *Obstetrics and gynecology*, 58(5), pp. 574–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7301232.

Hacker, N. F. and Van der Velden, J. (1993) 'Conservative management of early vulvar cancer.', *Cancer*, 71(4 Suppl), pp. 1673–1677.

Haefner, H. K. *et al.* (2014) 'The impact of vulvar lichen sclerosus on sexual dysfunction.', *Journal of women's health (2002)*, 23(9), pp. 765–70. doi: 10.1089/jwh.2014.4805.

Haka, A. S. *et al.* (2014) 'Diagnosing breast cancer using Raman spectroscopy: prospective analysis.', *Journal of biomedical optics*, 14(5), p. 054023. doi: 10.1117/1.3247154.

Hammody, Z. *et al.* (2008) 'Distinction of malignant melanoma and epidermis using IR micro-spectroscopy and statistical methods', *Analyst*, 133(3), pp. 372–378. doi: 10.1039/B712040K.

Hampl, M. and Sarajuuri, H. (2006) 'Effect of Human Papillomavirus Vaccines on

Vulvar , Vaginal , and Anal Intraepithelial Lesions and Vulvar Cancer', *OBSTETRICS & GYNECOLOGY*, 108(6), pp. 1361–1368.

Han, X. *et al.* (2010) 'Near-infrared autofluorescence imaging of cutaneous melanins and human skin in vivo.', *Journal of biomedical optics*, 14(2), p. 024017. doi: 10.1117/1.3103310.

Haris, P. I. and Chapman, D. (1992) 'Does Fourier-transform infrared spectroscopy provide useful information on protein structures?', *Trends in Biochemical Sciences*, 17(9), pp. 328–333. doi: 10.1016/0968-0004(92)90305-S.

Harrington, C. I. and Dunsmore, I. R. (1981) 'An investigation into the incidence of auto-immune disorders in patients with lichen sclerosus and atrophicus.', *The British journal of dermatology*, 104(5), pp. 563–6.

Harris, A. T. *et al.* (2010) 'Raman spectroscopy in head and neck cancer.', *Head & neck oncology*, 2(1), p. 26. doi: 10.1186/1758-3284-2-26.

Hartman, K. A., Clayton, N. and Thomas, G. J. (1973) 'Studies of virus structure by Raman spectroscopy I. R17 virus and R17 RNA', *Biochemical and Biophysical Research Communications*, 50(3), pp. 942–949. doi: 10.1016/0006-291X(73)91336-3.

Hata, T. R. *et al.* (2000) 'Non-invasive Raman spectroscopic detection of carotenoids in human skin', *Journal of Investigative Dermatology*, 115(3), pp. 441–448. doi: 10.1046/j.1523-1747.2000.00060.x.

Hauspy, J. *et al.* (2007) 'Sentinel lymph node in vulvar cancer.', *Cancer*, 110(5), pp.

1015–1023. doi: 10.1002/cncr.22874.

Hazewinkel, M. H. *et al.* (2012) ‘Long-term sexual function in survivors of vulvar cancer: A cross-sectional study’, *Gynecologic Oncology*. Elsevier Inc., 126(1), pp. 87–92. doi: 10.1016/j.ygyno.2012.04.015.

Heal, C. F. *et al.* (2008) ‘Accuracy of clinical diagnosis of skin lesions.’, *The British journal of dermatology*, 159, pp. 661–668. doi: 10.1111/j.1365-2133.2008.08715.x.

Heaps, J. M. *et al.* (1990) ‘Surgical-pathologic variables predictive of local recurrence in squamous cell carcinoma of the vulva’, *Gynecologic Oncology*, 38(3), pp. 309–314. doi: 10.1016/0090-8258(90)90064-R.

Hillemanns, P. *et al.* (2000) ‘Photodynamic therapy of vulvar intraepithelial neoplasia using 5-aminolevulinic acid.’, *International journal of cancer*, 85(5), pp. 649–53. doi: 10.1006/gyno.2000.6028.

Hinten, F. *et al.* (2018) ‘Vulvar cancer: Two pathways with different localization and prognosis’, *Gynecologic Oncology*. Elsevier Inc., 149(2), pp. 310–317. doi: 10.1016/j.ygyno.2018.03.003.

Hirschowitz, L., Wells, M. and Lowe, J. (2013) *Double-reporting in histopathology*, *Royal college of pathologists*.

Höckel, M. *et al.* (2010) ‘Vulvar field resection: Novel approach to the surgical treatment of vulvar cancer based on ontogenetic anatomy’, *Gynecologic Oncology*. Elsevier Inc., 119(1), pp. 106–113. doi: 10.1016/j.ygyno.2010.06.019.

Holthoff, E. R. *et al.* (2017) ‘Vulvar squamous cell carcinoma aggressiveness is

associated with differential expression of collagen and STAT1', *Clinical Proteomics*. BioMed Central, 14(1), pp. 1–13. doi: 10.1186/s12014-017-9175-8.

Hørding, U. *et al.* (1995) 'Vulvar intraepithelial neoplasia III: a viral disease of undetermined progressive potential.', *Gynecologic oncology*, 56(2), pp. 276–9. doi: 10.1006/gyno.1995.1046.

Horn, M. *et al.* (2007) 'The use of confocal laser-scanning microscopy in microsurgery for invasive squamous cell carcinoma', *British Journal of Dermatology*, 156(1), pp. 81–84. doi: 10.1111/j.1365-2133.2006.07574.x.

Horsnell, J. *et al.* (2010) 'Raman spectroscopy--a new method for the intra-operative assessment of axillary lymph nodes.', *Analyst*, 135(12), pp. 3042–3047.

Horsnell, J. D. *et al.* (2012) 'Raman spectroscopy -- A potential new method for the intra-operative assessment of axillary lymph nodes', *The Surgeon*, 10(3), pp. 123–127. doi: 10.1016/j.surge.2011.02.004.

Horsnell, J. D. (2012) *The use of Raman Spectroscopy for the Intra-Operative Assessment of Axillary Lymph Nodes in Breast Cancer*.

Huang, Z. *et al.* (2014) 'In vivo early diagnosis of gastric dysplasia using narrow-band image-guided Raman endoscopy.', *Journal of biomedical optics*, 15(3), p. 037017. doi: 10.1117/1.3420115.

De Hullu, J. a. *et al.* (2002) 'Vulvar carcinoma. The price of less radical surgery.', *Cancer*, 95(11), pp. 2331–8. doi: 10.1002/cncr.10969.

Hutchings, J. *et al.* (2009) ‘The potential for histological screening using a combination of rapid Raman mapping and principal component analysis’, 103(1), pp. 91–103. doi: 10.1002/jbio.200810070.

Iping Petterson, I. E. *et al.* (2015) ‘Characterisation of a fibre optic Raman probe within a hypodermic needle Raman4Clinics’, *Analytical and Bioanalytical Chemistry*, 407(27), pp. 8311–8320. doi: 10.1007/s00216-015-9021-7.

Isaac, B. H. and Young, L. (2013) *Cancer of the Vulva*. Thrid Edit, *Women’s Cancers*. Thrid Edit. Elsevier Inc. doi: 10.1002/9781118786635.ch13.

Isabelle, M. (2009) *Evaluation of metastatic status of lymph nodes using vibrational spectroscopy*. Cranfield University.

Isabelle, M. *et al.* (2015) ‘FDVIBSPC16: Multi-Centre Raman Spectral Mapping of Oesophageal Cancer Tissues: a study to assess system transferability’, *Faraday Discuss.*, 00, pp. 1–17. doi: 10.1039/C5FD00183H.

ISSVD (1976) ‘New nomenclature for vulvar disease.’, *Obstetrics and gynecology*, 47(1), pp. 122–4. Available at: http://www.ncbi.nlm.nih.gov/pubmed/174037 (Accessed: 30 December 2014).

ISSVD (2015) *The 2015 International Society for the Study of Vulvovaginal Disease Terminology of Vulvar Squamous Intraepithelial Lesions*. Available at: http://issvd.org/wp-content/uploads/2015/09/2015-ISSVD-VIN-terminology-for-the-website-v5.pdf.

Ito, H. *et al.* (2014) ‘Use of surface-enhanced Raman scattering for detection of

cancer-related serum-constituents in gastrointestinal cancer patients', *Nanomedicine: Nanotechnology, Biology, and Medicine*. Elsevier Inc., 10(3), pp. 599–608. doi: 10.1016/j.nano.2013.09.006.

Iversen, T. and Tretli, S. (1998) 'Intraepithelial and invasive squamous cell neoplasia of the vulva: trends in incidence, recurrence, and survival rate in Norway.', *Obstetrics and gynecology*, 91(6), pp. 969–72. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9611006.

Jacques, S. L. (2009) 'Spectral Imaging and Analysis To Yield Tissue Optical Properties', *Journal of Innovative Optical Health Sciences*, 02(02), pp. 123–129. doi: 10.1142/S1793545809000528.

Jacques, S. L. (2013) 'Optical properties of biological tissues: a review.', *Physics in medicine and biology*, 58(11), pp. R37-61. doi: 10.1088/0031-9155/58/11/R37.

Jacques, S. L. and Pogue, B. W. (2008) 'Tutorial on diffuse light transport.', *Journal of biomedical optics*, 13(4), p. 041302. doi: 10.1117/1.2967535.

Jermyn, M. *et al.* (2015) 'Intraoperative brain cancer detection with Raman spectroscopy in humans', 7(274), pp. 1–10.

Ji, M. *et al.* (2013) 'Rapid, label-free detection of brain tumors with stimulated Raman scattering microscopy.', *Science Translational Medicine*, 5(201), pp. 201ra119-201ra119. doi: 10.1126/scitranslmed.3005954.

Jones, R. W., Baranyai, J. and Stables, S. (1997) 'Trends in squamous cell carcinoma of the vulva: the influence of vulvar intraepithelial neoplasia.', *Obstetrics*

and gynecology, 90(3), pp. 448–52. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9277660 (Accessed: 30 December 2014).

Jones, R. W., Rowan, D. M. and Stewart, A. W. (2005) 'Vulvar intraepithelial neoplasia: aspects of the natural history and outcome in 405 women.', *Obstet Gynecol*, 106(6), pp. 1319–1326.

Jorgensen, T. M. *et al.* (2008) 'Machine-learning classification of non-melanoma skin cancers from image features obtained by optical coherence tomography', *Skin Research and Technology*, 14(3), pp. 364–369. doi: 10.1111/j.1600-0846.2008.00304.x.

Joura, E. A. *et al.* (1999) 'Differentiating Vulvar Intraepithelial Neoplasia From Nonneoplastic Epithelial Disorders', *Obstetrical & Gynecological Survey*, 54(1), pp. 28–29. doi: 10.1097/00006254-199901000-00015.

Joura, E. A. *et al.* (2000) 'Trends in vulvar neoplasia. Increasing incidence of vulvar intraepithelial neoplasia and squamous cell carcinoma of the vulva in young women.', *The Journal of reproductive medicine*, 45(8), pp. 613–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10986677.

Judson, P. L. *et al.* (2006) 'Trends in the incidence of invasive and in situ vulvar carcinoma.', *Obstetrics and gynecology*, 107(5), pp. 1018–22. doi: 10.1097/01.AOG.0000210268.57527.a1.

Kallaway, C. *et al.* (2013) 'Advances in the clinical application of Raman spectroscopy for cancer diagnostics', *Photodiagnosis and Photodynamic Therapy*,

10, pp. 207–219. doi: http://dx.doi.org/10.1016/j.pdpdt.2013.01.008.

Kast, R. E. *et al.* (2008) ‘Raman spectroscopy can differentiate malignant tumors from normal breast tissue and detect early neoplastic changes in a mouse model.’, *Biopolymers*, 89(3), pp. 235–241.

Kaushik, P. L. *et al.* (2015) ‘Medical interventions for high-grade vulval intraepithelial neoplasia (Review)’, *Cochrane Database of Systematic Reviews*, (8). Available at: http://www.thecochranelibrary.com.

Kazarian, S. G. and Chan, K. L. A. (2013) ‘ATR-FTIR spectroscopic imaging: recent advances and applications to biological systems’, *The Analyst*, 138(7), p. 1940. doi: 10.1039/c3an36865c.

Kendall, C. (2002) *A study of Raman spectroscopy for the early detection and classification of malignancy in oesophageal tissue*.

Kendall, C. *et al.* (2009) ‘Vibrational spectroscopy: a clinical tool for cancer diagnostics’, *The Analyst*, 134(6), p. 1029. doi: 10.1039/b822130h.

Kenter, G. G. *et al.* (2009) ‘Vaccination against HPV-16 oncoproteins for vulvar intraepithelial neoplasia.’, *The New England journal of medicine*, 361(19), pp. 1838–1847. doi: 10.1056/NEJMoa0810097.

Khristoforova, Y. A. *et al.* (2019) ‘Portable spectroscopic system for in vivo skin neoplasms diagnostics by Raman and autofluorescence analysis’, *Journal of Biophotonics*, 12(4). doi: 10.1002/jbio.201800400.

Kitchener, H. C. *et al.* (2009) 'HPV testing in combination with liquid-based cytology in primary cervical screening (ARTISTIC): a randomised controlled trial', *The Lancet Oncology*. Elsevier Ltd, 10(7), pp. 672–682. doi: 10.1016/S1470-2045(09)70156-1.

Knudsen, L. *et al.* (2002) 'Natural variations and reproducibility of in vivo near-infrared Fourier transform Raman spectroscopy of normal human skin', *Journal of Raman Spectroscopy*, 33(7), pp. 574–579. doi: 10.1002/jrs.888.

Kollias, N., Zonios, G. and Stamatas, G. N. (2002) 'Fluorescence spectroscopy of skin', *Vibrational Spectroscopy*, 28(1), pp. 17–23. doi: 10.1016/S0924-2031(01)00142-4.

Komachi, Y. *et al.* (2009) 'Improvement and analysis of a micro Raman probe', *Applied Optics*, 48(9), p. 1683. doi: 10.1364/ao.48.001683.

Kong, K. *et al.* (2013) 'Diagnosis of tumors during tissue-conserving surgery with integrated autofluorescence and Raman scattering microscopy', *Proceedings of the National Academy of Sciences*, 110(38), pp. 15189–15194. doi: 10.1073/pnas.1311289110.

Kong, K. *et al.* (2014) 'Towards intra-operative diagnosis of tumours during breast conserving surgery by selective-sampling Raman micro-spectroscopy.', *Physics in medicine and biology*, 59(20), pp. 6141–52. doi: 10.1088/0031-9155/59/20/6141.

Kronz, J. D., Westra, W. H. and Epstein, J. I. (1999) 'Mandatory second opinion surgical pathology at a large referral hospital', *Cancer*, 86(11), pp. 2426–2435. doi: 10.1002/(SICI)1097-0142(19991201)86:11<2426::AID-CNCR34>3.0.CO;2-3.

Kumar, S. *et al.* (2013) ‘Change in the microenvironment of breast cancer studied by FTIR imaging’, *Analyst*, 138(14), pp. 4058–4065. doi: 10.1039/c3an00241a.

Kwak, J. T. *et al.* (2015) ‘Improving prediction of prostate cancer recurrence using chemical imaging.’, *Scientific reports*, 5, p. 8758. doi: 10.1038/srep08758.

Kyriakidou, M. *et al.* (2017) ‘FT-IR spectroscopy study in early diagnosis of skin cancer’, *In Vivo*, 31(6), pp. 1131–1137. doi: 10.21873/invivo.11179.

Lasch, P. and Kneipp, J. (2008) *Biomedical Vibrational Spectroscopy*. First. Hoboken: Wiley.

Lawrie, T. a *et al.* (2014) ‘Sentinel node assessment for diagnosis of groin lymph node involvement in vulval cancer.’, *Cochrane Database of Systematic Reviews*, 6(6). doi: 10.1002/14651858.CD010409.pub2.

Lee, A, Bradford, J. and Fischer, G. (2015) ‘Long-term Management of Adult Vulvar Lichen Sclerosus: A Prospective Cohort Study of 507 Women’, *JAMA Dermatol*, 151(10), pp. 1061–1067. doi: 10.1001/jamadermatol.2015.0643.

Lee, Andrew, Bradford, J. and Fischer, G. (2015) ‘Long-term Management of Adult Vulvar Lichen Sclerosus’, *JAMA Dermatology*, 151(10), p. 1061. doi: 10.1001/jamadermatol.2015.0643.

Lee, C. Y., Kim, K. H. and Kim, Y. H. (2010) ‘The efficacy of photodynamic diagnosis in defining the lateral border between a tumor and a tumor-free area during mohs micrographic surgery’, *Dermatologic Surgery*, 36(11), pp. 1704–1710. doi: 10.1111/j.1524-4725.2010.01722.x.

Leemans, C. R., Braakhuis, B. J. M. and Brakenhoff, R. H. (2011) 'The molecular biology of head and neck cancer.', *Nature reviews. Cancer*. Nature Publishing Group, 11(1), pp. 9–22. doi: 10.1038/nrc2982.

Leibovitz, A. *et al.* (2000) 'Vulvovaginal examinations in elderly nursing home women residents', *Archives of Gerontology and Geriatrics*, 31(1), pp. 1–4. doi: 10.1016/S0167-4943(00)00059-5.

Levenback, C. F. *et al.* (2009) 'Sentinel lymph node biopsy in patients with gynecologic cancers Expert panel statement from the International Sentinel Node Society Meeting, February 21, 2008.', *Gynecologic oncology*. Elsevier B.V., 114(2), pp. 151–156. doi: 10.1016/j.ygyno.2009.03.035.

Levine, T. S. *et al.* (2001) 'The use of cytospin monolayer technique in the cytological diagnosis of vulval and anal disease', *Cytopathology*, 12(5), pp. 297–305. doi: 10.1046/j.1365-2303.2001.00348.x.

Lewis, I. R. and Edwards, H. (2001) *Handbook of Raman Spectroscopy: From the Research Laboratory to the Process Line*. 1st edn. Marcel Dekker.

Li, Q. *et al.* (2013) 'Detection of Gastric Cancer with Fourier Transform Infrared Spectroscopy and Support Vector Machine Classification', 2013.

Liang, W. *et al.* (2003) 'Acceptability of diagnostic tests for breast cancer', *Breast Cancer Research and Treatment*, 79(2), pp. 199–206. doi: 10.1023/A:1023914612152.

Lieber, C. A. *et al.* (2008) 'In vivo nonmelanoma skin cancer diagnosis using Raman

microspectroscopy', *Lasers in Surgery and Medicine*, 40(7), pp. 461–467. doi: 10.1002/lsm.20653.

Lieber, C. a *et al.* (2008) 'In vivo nonmelanoma skin cancer diagnosis using Raman microspectroscopy.', *Lasers in surgery and medicine*, 40(7), pp. 461–7. doi: 10.1002/lsm.20653.

Lieber, C. and Mahadevan-Jansen, A. (2003) 'Automated Method for Subtraction of Fluorescence from Biological Raman Spectra', *Applied Spectroscopy*, 57(11), pp. 1363–1367. doi: 10.1366/000370203322554518.

Lind, A. C. *et al.* (1995) 'Prospective peer review in surgical pathology', *Am J Clin Pathol*, 104(5), pp. 560–6.

Lloyd, G. R. (2009) *Chemometrics and Pattern Recognition for the Analysis of Multivariate Datasets*. University of Bristol.

Lloyd, G. R. *et al.* (2012) 'Assessing the performance of spectroscopic models for cancer diagnostics using cross-validation and permutation testing', *SPIE BiOS*, 8219, pp. 82190C--82190C. doi: 10.1117/12.919864.

Lloyd, G. R. *et al.* (2013a) 'Discrimination between benign, primary and secondary malignancies in lymph nodes from the head and neck utilising Raman spectroscopy and multivariate analysis.', *The Analyst*, 138(14), pp. 3900–8. doi: 10.1039/c2an36579k.

Lloyd, G. R. *et al.* (2013b) 'Discrimination between benign, primary and secondary malignancies in lymph nodes from the head and neck utilising Raman spectroscopy

and multivariate analysis', *The Analyst*, 138(14), p. 3900. doi: 10.1039/c2an36579k.

Lord, R. C. and Yu, N. (1970) 'Laser-excited Raman spectroscopy of biomolecules', *Journal of Molecular Biology*, 50(2), pp. 509–524. doi: 10.1016/0022-2836(70)90208-1.

Lui, H. *et al.* (2012) 'Real-time Raman spectroscopy for in vivo skin cancer diagnosis.', *Cancer Research*, 72(10), pp. 2491–2500.

Ly, E. *et al.* (2008) 'Combination of FTIR spectral imaging and chemometrics for tumour detection from paraffin-embedded biopsies.', *The Analyst*, 133(2), pp. 197–205. doi: 10.1039/b715924b.

Ly, E. *et al.* (2009) 'Differential diagnosis of cutaneous carcinomas by infrared spectral micro-imaging combined with pattern recognition', *Analyst*, 134(6), pp. 1208–1214. doi: 10.1039/b820998g.

Ly, E. *et al.* (2010) 'Probing tumor and peritumoral tissues in superficial and nodular basal cell carcinoma using polarized Raman microspectroscopy.', *Experimental dermatology*, 19(1), pp. 68–73. doi: 10.1111/j.1600-0625.2009.00992.x.

Mackanos, M. A. and Contag, C. H. (2009) 'FTIR microspectroscopy for improved prostate cancer diagnosis', *Trends in Biotechnology*, 27(12), pp. 661–663. doi: 10.1016/j.tibtech.2009.09.001.

Mackanos, M. A. and Contag, C. H. (2010) 'Fiber-optic probes enable cancer detection with FTIR spectroscopy', *Trends in Biotechnology*. Elsevier Ltd, 28(6), pp. 317–323. doi: 10.1016/j.tibtech.2010.04.001.

MacKenzie-Wood, a R., Milton, G. W. and de Launey, J. W. (1998) 'Melanoma: accuracy of clinical diagnosis.', *The Australasian journal of dermatology*, 39(1), pp. 31–3. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9602875.

Maclean, A. and Reid, W. (2011) *Benign disease of the vulva and the vagina*. Fourth Edi, *Gynaecology*. Fourth Edi. Elsevier Ltd. doi: 10.1016/B978-0-7020-3120-5.00040-0.

Madsen, S. J., Patterson, M. S. and Wilson, B. C. (1992) 'The use of India ink as an optical absorber in tissue-simulating phantoms.', *Physics in medicine and biology*, 37(4), pp. 985–93. Available at: http://www.ncbi.nlm.nih.gov/pubmed/1589459.

Manion, E., Cohen, M. B. and Weydert, J. (2008) 'Mandatory second opinion in surgical pathology referral material: clinical consequences of major disagreements.', *The American journal of surgical pathology*, 32(5), pp. 732–7. doi: 10.1097/PAS.0b013e31815a04f5.

Martin-Hirsch, P. L. *et al.* (1998) 'Photodynamic treatment for lower genital tract intraepithelial neoplasia', *Lancet*, 351(9113), p. 1403. doi: 10.1016/S0140-6736(98)24019-0.

Matousek, P. *et al.* (2005) 'Subsurface probing in diffusely scattering media using spatially offset Raman spectroscopy.', *Applied spectroscopy*, 59(4), pp. 393–400. doi: 10.1366/0003702053641450.

McIntosh, L. M. *et al.* (1999) 'Infrared spectra of basal cell carcinomas are distinct from non-tumor-bearing skin components.', *The Journal of investigative dermatology*,

112(6), pp. 951–6. doi: 10.1046/j.1523-1747.1999.00612.x.

Meffert, J. J., Davis, B. M. and Grimwood, R. E. (1995) 'Lichen sclerosus', *Journal of the American Academy of Dermatology*, 32(3), pp. 393–416; quiz 417–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7868709.

Menzies, G. E. *et al.* (2014) 'Fourier transform infrared for noninvasive optical diagnosis of oral, oropharyngeal, and laryngeal cancer', *Translational Research*. Mosby, Inc, 163(1), pp. 19–26. doi: 10.1016/j.trsl.2013.09.006.

Menzies, S. W. *et al.* (2005) 'The performance of SolarScan: an automated dermoscopy image analysis instrument for the diagnosis of primary melanoma.', *Archives of dermatology*, 141(11), pp. 1388–96. doi: 10.1001/archderm.141.11.1388.

Meyrick Thomas, R. H. *et al.* (1988) 'Lichen sclerosus et atrophicus and autoimmunity - a study of 350 women', *Br J Dermatol*, 118(1), pp. 41–46.

Micheletti, L. *et al.* (2018) 'Prognostic impact of reduced tumor-free margin distance on long-term survival in FIGO stage IB/II vulvar squamous cell carcinoma', *Journal of Gynecologic Oncology*, 29(5), p. e61. doi: 10.3802/jgo.2018.29.e61.

Middleton, L. P. *et al.* (2014) 'Second-opinion pathologic review is a patient safety mechanism that helps reduce error and decrease waste.', *American Society of Clinical Oncology*, 10(4), pp. 275–281. doi: 10.1200/JOP.2013.001204.

Minnes, R. *et al.* (2017) 'Using Attenuated Total Reflection–Fourier Transform Infra-Red (ATR-FTIR) spectroscopy to distinguish between melanoma cells with a different metastatic potential', *Scientific Reports*, 7(1), p. 4381. doi: 10.1038/s41598-

017-04678-6.

Mitchell, M. F. *et al.* (1998) ‘Colposcopy for the diagnosis of squamous intraepithelial lesions: A meta-analysis’, *Obstetrics and Gynecology*, pp. 626–631. doi: 10.1016/S0029-7844(98)00006-4.

Modesitt, S. C. *et al.* (1998) ‘Vulvar intraepithelial neoplasia III: Occult cancer and the impact of margin status on recurrence’, *Obstetrics and Gynecology*, 92(6), pp. 962–966. doi: 10.1016/S0029-7844(98)00350-0.

Mogensen, M. *et al.* (2009) ‘Assessment of optical coherence tomography imaging in the diagnosis of non-melanoma skin cancer and benign lesions versus normal skin: Observer-blinded evaluation by dermatologists and pathologists’, *Dermatologic Surgery*, 35(6), pp. 965–972. doi: 10.1111/j.1524-4725.2009.01164.x.

Moll, R. (1991) ‘Molecular diversity of cytokeratins: significance for cell and tumor differentiation.’, *Acta histochemica. Supplement*, 41, pp. 117–127.

Moncrieff, M. *et al.* (2002) ‘Spectrophotometric intracutaneous analysis: A new technique for imaging pigmented skin lesions’, *British Journal of Dermatology*, 146(3), pp. 448–457. doi: 10.1046/j.1365-2133.2002.04569.x.

Mordechai, S. *et al.* (2004) ‘Possible common biomarkers from FTIR microspectroscopy of cervical cancer and melanoma.’, *Journal of microscopy*, 215(Pt 1), pp. 86–91. doi: 10.1111/j.0022-2720.2004.01356.x.

Morton, C. A. and Mackie, R. M. (1998) ‘Clinical accuracy of the diagnosis of cutaneous malignant melanoma’, *British Journal of Dermatology*, 138(2), pp. 283–

287. doi: 10.1046/j.1365-2133.1998.02075.x.

Moskovits, M. (2005) ‘Surface-enhanced Raman spectroscopy: a brief retrospective’, *Journal of Raman Spectroscopy*, 36(6–7), pp. 485–496. doi: 10.1002/jrs.1362.

Movasaghi, Z., Rehman, S. and ur Rehman, D. I. (2008a) ‘Fourier Transform Infrared (FTIR) Spectroscopy of Biological Tissues’, *Applied Spectroscopy Reviews*, 43(2), pp. 134–179. doi: 10.1080/05704920701829043.

Movasaghi, Z., Rehman, S. and ur Rehman, D. I. (2008b) ‘Fourier Transform Infrared (FTIR) Spectroscopy of Biological Tissues’, *Applied Spectroscopy Reviews*, 43(2), pp. 134–179. doi: 10.1080/05704920701829043.

Nallala, J., Lloyd, G. R. and Stone, N. (2015) ‘Evaluation of different tissue de-paraffinization procedures for infrared spectral imaging’, *The Analyst*, pp. 2369–2375. doi: 10.1039/C4AN02122C.

Naranjo, V. *et al.* (2015) ‘MINERVA Project, mid- To near Infrared Spectroscopy for Improved Medical Diagnostics’, in *European Project Space on Intelligent Systems, Pattern Recognition and Biomedical Systems*. SCITEPRESS - Science and and Technology Publications, pp. 53–69. doi: 10.5220/0006162400530069.

Nascimento, A. F. *et al.* (2004) ‘Vulvar acanthosis with altered differentiation: a precursor to verrucous carcinoma?’, *The American journal of surgical pathology*, 28(5), pp. 638–43. doi: 00000478-200405000-00012.

Naumann, Dieter (2001) ‘Ft-Infrared and Ft-Raman Spectroscopy in Biomedical Research’, *Applied Spectroscopy Reviews*, 36(2–3), pp. 239–298. doi:

10.1081/ASR-100106157.

Naumann, D (2001) ‘FT-infrared and Raman spectroscopy in biomedical research’, *Applied Spectroscopy Reviews*, 36(2&3), pp. 239–298.

Nauth, H. F. and Schilke, E. (1982) ‘Cytology of the exfoliative layer in normal and diseased vulvar skin: correlation with histology.’, *Acta cytologica*, 26(3), pp. 269–83. Available at: http://www.ncbi.nlm.nih.gov/pubmed/6954808.

Neill, S. and Lewis, F. (2010) ‘British Association of Dermatologists’ guidelines for the management of lichen sclerosus 2010’, *Br J Dermatol*, (May), pp. 672–682. doi: 10.1111/j.1365-2133.2010.09997.x.

Nguyen, T. T. *et al.* (2013) ‘Characterization of type I and IV collagens by Raman microspectroscopy: Identification of spectral markers of the dermo-epidermal junction’, *Advances in Biomedical Spectroscopy*, 7(5), pp. 105–110. doi: 10.3233/978-1-61499-184-7-105.

van de Nieuwenhof, H. P., van der Avoort, I. a M. and de Hullu, J. a. (2008) ‘Review of squamous premalignant vulvar lesions’, *Critical Reviews in Oncology/Hematology*, 68(2), pp. 131–156. doi: 10.1016/j.critrevonc.2008.02.012.

Nijssen, A. *et al.* (2002) ‘Discriminating basal cell carcinoma from its surrounding tissue by Raman spectroscopy.’, *Journal of Investigative Dermatology*, 119(1), pp. 64–69.

Nijssen, A. *et al.* (2014) ‘Discriminating basal cell carcinoma from perilesional skin using high wave-number Raman spectroscopy.’, *Journal of biomedical optics*, 12(3),

p. 034004. doi: 10.1117/1.2750287.

Di Ninni, P., Martelli, F. and Zaccanti, G. (2010) 'The use of India ink in tissue-simulating phantoms.', *Optics express*, 18(26), pp. 26854–26865. doi: 10.1364/OE.18.026854.

Ninni, P. Di, Martelli, F. and Zaccanti, G. (2011) 'Intralipid: towards a diffusive reference standard for optical tissue phantoms.', *Physics in medicine and biology*, 56(2), pp. N21–N28. doi: 10.1088/0031-9155/56/2/N01.

Nooij, L. S. *et al.* (2016) 'Tumour-free margins in vulvar squamous cell carcinoma: Does distance really matter?', *European Journal of Cancer*. Elsevier Ltd, 65, pp. 139–149. doi: 10.1016/j.ejca.2016.07.006.

Nucci, M. (2013) *Diagnostic Histopathology of Tumors*. Fourth Edi. Edited by C. D. M. Fletcher. Philadelphia: Saunders. Available at: https://www.clinicalkey.com/#!/content/book/3-s2.0-B9781437715347000459.

Nyberg, M., Ramser, K. and Lindahl, O. A. (2013) 'Optical fibre probe NIR raman measurements in ambient light and in combination with a tactile resonance sensor for possible cancer detection', *Analyst*, 138(14), pp. 4029–4034. doi: 10.1039/c3an00243h.

Obuchowski, N. A., Lieber, M. L. and Wians, F. H. (2004) 'ROC curves in Clinical Chemistry: Uses, misuses, and possible solutions', *Clinical Chemistry*, 50(7), pp. 1118–1125. doi: 10.1373/clinchem.2004.031823.

Obuchowski, N. A. and McClish, D. K. (1997) 'Sample size determination for

diagnostic accuracy studies involving binormal ROC curve indices.', *Statistics in medicine*, 16(13), pp. 1529–42. doi: 10.1002/(SICI)1097-0258(19970715)16:13<1529::AID-SIM565>3.0.CO;2-H.

Office for National Statistics (2013) *Cancer Registration Statistics, England 2011*.

Old, O. (2015) *Detection of Barrett's neoplasia with vibrational spectroscopy*, *MD Thesis*.

Old, O. J. *et al.* (2014) 'Vibrational spectroscopy for cancer diagnostics', *Analytical Methods*, 6(12), p. 3901. doi: 10.1039/c3ay42235f.

Oliveira, A. F. *et al.* (2010) 'Differential diagnosis in primary and metastatic cutaneous melanoma by FT-Raman spectroscopy.', *Acta Cirurgica Brasileira*, 25(5), pp. 434–439.

Olmedo, J. M. *et al.* (2006) 'Optical coherence tomography for the characterization of basal cell carcinoma in vivo: A pilot study', *Journal of the American Academy of Dermatology*, 55(3), pp. 408–412. doi: 10.1016/j.jaad.2006.03.013.

Ostrowska, K. M. *et al.* (2011) 'Correlation of p16(INK4A) expression and HPV copy number with cellular FTIR spectroscopic signatures of cervical cancer cells.', *Analyst*, 136(7), pp. 1365–1373.

Oyama, N. *et al.* (2003) 'Autoantibodies to extracellular matrix protein 1 in lichen sclerosus', *The Lancet*, 362(9378), pp. 118–123.

Panjehpour, M. *et al.* (2002) 'Laser-induced fluorescence spectroscopy for in vivo

diagnosis of non-melanoma skin cancers', *Lasers in Surgery and Medicine*, 31(5), pp. 367–373. doi: 10.1002/lsm.10125.

Patterson, J. W. (2016) '11 – Disorders of collagen', in *Weedon's Skin Pathology*, pp. 347-379.e18. doi: 10.1016/B978-0-7020-5183-8.00011-4.

Philipsen, P. A. *et al.* (2013) 'Diagnosis of malignant melanoma and basal cell carcinoma by in vivo NIR-FT Raman spectroscopy is independent of skin pigmentation.', *Photochemical & Photobiological Sciences*, 12(5), pp. 770–776.

Phillips, G. R. and Harris, J. M. (1990) 'Polynomial filters for data sets with outlying or missing observations: application to charge-coupled-device-detected Raman spectra contaminated by cosmic rays', *Analytical Chemistry*, 62(21), pp. 2351–2357. doi: 10.1021/ac00220a017.

Del Pino, M., Rodriguez-Carunchio, L. and Ordi, J. (2013a) 'Pathways of vulvar intraepithelial neoplasia and squamous cell carcinoma.', *Histopathology*, 62(1), pp. 161–75. doi: 10.1111/his.12034.

Del Pino, M., Rodriguez-Carunchio, L. and Ordi, J. (2013b) 'Pathways of vulvar intraepithelial neoplasia and squamous cell carcinoma.', *Histopathology*, 62(1), pp. 161–75. doi: 10.1111/his.12034.

Platz, C. E. and Benda, J. A. (1995) 'Female genital tract cancer.', *Cancer*, 75(1 Suppl), pp. 270–294.

Pogue, B. W. and Patterson, M. S. (2006) 'Review of tissue simulating phantoms for optical spectroscopy, imaging and dosimetry', *Journal of Biomedical Optics*, 11(4),

pp. 041102-041102–16. doi: 10.1117/1.2335429.

Powell, J. and Wojnarowska, F. (1999) ‘Lichen sclerosus’, *The Lancet*, 353(9166), pp. 1777–1783. doi: 10.1016/S0140-6736(98)08228-2.

Rajaram, N. *et al.* (2010) ‘Pilot clinical study for quantitative spectral diagnosis of non-melanoma skin cancer’, *Lasers in Surgery and Medicine*, 42(10), pp. 716–727. doi: 10.1002/lsm.21009.

RCOG (2014) ‘Guidelines for the Diagnosis and Management of Vulval Carcinoma’, *RCOG*, pp. 1–35.

Regauer, S. (2011) ‘Residual anogenital lichen sclerosus after cancer surgery has a high risk for recurrence: A clinicopathological study of 75 women’, *Gynecologic Oncology*, 123(2), pp. 289–294. doi: 10.1016/j.ygyno.2011.07.010.

Renaud-Vilmer, C. *et al.* (2004) ‘Vulvar lichen sclerosus: effect of long-term topical application of a potent steroid on the course of the disease.’, *Archives of dermatology*, 140(6), pp. 709–12. doi: 10.1001/archderm.140.6.709.

Reyes, M. C. and Cooper, K. (2014) ‘An update on vulvar intraepithelial neoplasia: terminology and a practical approach to diagnosis’, *Journal of Clinical Pathology*, 67(4), pp. 290–294. doi: 10.1136/jclinpath-2013-202117.

Rhodes, C. A., Registrar, S. and Lecturer, C. C. (1998) ‘The management of squamous cell vulval cancer : a population based retrospective study of 411 cases’, *British Journal of Obstetrics and Gynaecology*, 105(February), pp. 200–205.

Robles, F. E. *et al.* (2011) ‘Molecular imaging true-colour spectroscopic optical coherence tomography.’, *Nature photonics*, 5(12), pp. 744–747. doi: 10.1038/nphoton.2011.257.

Rolfe, K. J. *et al.* (2003) ‘TP53 mutations in vulval lichen sclerosus adjacent to squamous cell carcinoma of the vulva.’, *British journal of cancer*, 89(12), pp. 2249–53. doi: 10.1038/sj.bjc.6601444.

Rosenthal, A. N. *et al.* (2002) ‘Molecular evidence of a common clonal origin and subsequent divergent clonal evolution in vulval intraepithelial neoplasia, vulval squamous cell carcinoma and lymph node metastases’, *International Journal of Cancer*, 99, pp. 549–554. doi: 10.1002/ijc.10362.

Ross, M. J. and Ehrmann, R. L. (1987) ‘Histologic prognosticators in stage I squamous cell carcinoma of the vulva.’, *Obstetrics and gynecology*, 70(5), pp. 774–84.

Rouzier, R. *et al.* (2002) ‘Local relapse in patients treated for squamous cell vulvar carcinoma: Incidence and prognostic value’, *Obstetrics and Gynecology*, 100(6), pp. 1159–1167. doi: 10.1016/S0029-7844(02)02501-2.

Sahu, R. and Mordechai, S. (2005) ‘Fourier transform infrared spectroscopy in cancer detection’, *Future Oncology*, 1(5), pp. 635–647. doi: 10.2217/14796694.1.5.635.

Salom, E. M. and Penalver, M. (2002) ‘Recurrent vulvar cancer.’, *Current treatment options in oncology*, 3(2), pp. 143–153. doi: 10.1097/01.grf.0000179671.98939.fe.

Santos, I. P. *et al.* (2017) ‘Raman spectroscopy for cancer detection and cancer surgery guidance: Translation to the clinics’, *Analyst*, 142(17), pp. 3025–3047. doi: 10.1039/c7an00957g.

Schilder, J. M. and Stehman, F. B. (2012) ‘Invasive Cancer of the Vulva’, in *Clinical Gynecologic Oncology: Eighth Edition*, pp. 219–244. doi: 10.1016/B978-0-323-07419-3.00008-4.

Schleusener, J. *et al.* (2015) ‘In vivo study for the discrimination of cancerous and normal skin using fibre probe-based Raman spectroscopy’, *Experimental Dermatology*, 24(10), pp. 767–772. doi: 10.1111/exd.12768.

Van Seters, M., Van Beurden, M. and de Craen, A. J. M. (2005) ‘Is the assumed natural history of vulvar intraepithelial neoplasia III based on enough evidence? A systematic review of 3322 published patients.’, *Gynecologic oncology*, 97(2), pp. 645–51. doi: 10.1016/j.ygyno.2005.02.012.

Shafer-Peltier, K. E. *et al.* (2002) ‘Raman microspectroscopic model of human breast tissue: Implications for breast cancer diagnosis in vivo’, *Journal of Raman Spectroscopy*, 33(7), pp. 552–563. doi: 10.1002/jrs.877.

Shetty, G. *et al.* (2006) ‘Raman spectroscopy: elucidation of biochemical changes in carcinogenesis of oesophagus.’, *British journal of cancer*, 94(10), pp. 1460–1464. doi: 10.1038/sj.bjc.6603102.

Shim, M. G. and Wilson, B. C. (1996) ‘The effects of ex vivo handling procedures on the near-infrared Raman spectra of normal mammalian tissues.’, *Photochemistry*

and photobiology, 63(5), pp. 662–671. doi: 10.1111/j.1751-1097.1996.tb05671.x.

Shylasree, T. S. *et al.* (2008) 'Contribution of demographic, psychological and disease-related factors to quality of life in women with high-grade vulval intraepithelial neoplasia', *Gynecologic Oncology*, 110(2), pp. 185–189.

Sigurdsson, S. *et al.* (2004) 'Detection of skin cancer by classification of Raman spectra.', *IEEE Transactions on Biomedical Engineering*, 51(10), pp. 1784–1793.

Silveira, F. L. *et al.* (2015) 'Discrimination of non-melanoma skin lesions from non-tumor human skin tissues in vivo using raman spectroscopy and multivariate statistics', *Lasers in Surgery and Medicine*, 47(1). doi: 10.1002/lsm.22318.

Silveira Jr, L. *et al.* (2012) 'Discriminating model for diagnosis of basal cell carcinoma and melanoma in vitro based on the Raman spectra of selected biochemicals.', *Journal of Biomedical Optics*, 17(7).

Singh, S. P. *et al.* (2012) 'Raman spectroscopy in head and neck cancers: toward oncological applications.', *Journal of cancer research and therapeutics*, 8 Suppl 1, pp. S126-32. doi: 10.4103/0973-1482.92227.

Skin Optics (no date). Available at: http://omlc.org/news/jan98/skinoptics.html.

Skrebova, N., Aizawa, K. and Aras, S. (2003) 'Data processing and analysis of benign , premalignant and malignant changes in skin tissue samples using FT-IR microspectroscopy', *Proceedings of SPIE*, 5047(47), pp. 378–385.

Slominski, A. *et al.* (2004) 'Melanin pigmentation in mammalian skin and its hormonal regulation.', *Physiological reviews*, 84(4), pp. 1155–1228. doi:

10.1152/physrev.00044.2003.

Smith, J. a (2005) *RAMAN SPECTROSCOPY IN THE ASSESSMENT OF AXILLARY LYMPH NODES IN BREAST CANCER INSTITUTE OF BIOSCIENCE AND TECHNOLOGY*. CRANFIELD UNIVERSITY.

Smith, J. H. F. (2010) *Other tumours and lesions of cervix, vulva and vagina*. Third Edit, *Diagnostic Cytopathology*. Third Edit. Elsevier Limited. doi: 10.1016/B978-0-7020-3154-0.00025-9.

Snow, J. B., Qian, S.-X. and Chang, R. K. (1985) 'Stimulated Raman scattering from individual water and ethanol droplets at morphology-dependent resonances', *Optics Letters*, 10(1), p. 37. doi: 10.1364/OL.10.000037.

Standring, S. (ed.) (2016) *Gray's Anatomy: The Anatomical Basis of Clinical Practice*. 41st edn. London: Elsevier Ltd. Available at: https://books.google.co.uk/books?id=b7FVCgAAQBAJ&printsec=frontcover#v=onepage&q&f=false.

Stone, N. *et al.* (2000) 'Raman spectroscopy for early detection of laryngeal malignancy: preliminary results.', *The Laryngoscope*, 110(10 Pt 1), pp. 1756–1763. doi: 10.1097/00005537-200010000-00037.

Stone, N. (2001) *Raman Spectroscopy of Biological Tissue for Application in Optical Diagnosis of Malignancy*. doi: 10.1016/j.phrs.2010.10.003.

Stone, N. *et al.* (2007) 'The use of Raman spectroscopy to provide an estimation of the gross biochemistry associated with urological pathologies', *Analytical and*

Bioanalytical Chemistry, 387(5), pp. 1657–1668. doi: 10.1007/s00216-006-0937-9.

Ta, L. *et al.* (2016) 'Medical and surgical interventions for the treatment of usual-type vulval intraepithelial neoplasia (Review)', *Cochrane Database of Systematic Reviews*, (1).

Talari, A. C. S. *et al.* (2015) 'Raman Spectroscopy of Biological Tissues', *Applied Spectroscopy Reviews*, 50(May 2016), pp. 46–111. doi: 10.1080/05704928.2014.923902.

Tfayli, A. *et al.* (2005) 'Discriminating nevus and melanoma on paraffin-embedded skin biopsies using FTIR microspectroscopy', *Biochimica et Biophysica Acta - General Subjects*, 1724(3), pp. 262–269. doi: 10.1016/j.bbagen.2005.04.020.

Tfayli, A. *et al.* (2009) 'Digital dewaxing of Raman signals: discrimination between nevi and melanoma spectra obtained from paraffin-embedded skin biopsies.', *Applied Spectroscopy*. MEDyC Unit, CNRS UMR 6237, Faculty of Pharmacy, University of Reims Champagne Ardenne (URCA), Reims, France., 63(5), pp. 564–570.

Tobin, D. J. (2006) 'Biochemistry of human skin - our brain on the outside.', *Chemical Society reviews*, 35(1), pp. 52–67. doi: 10.1039/b505793k.

Toms, S. A. *et al.* (2006) 'Neuro-oncological applications of optical spectroscopy', *Technol Cancer Res Treat*, 5(3), pp. 231–238. doi: d=3025&c=4205&p=14224&do=detail [pii].

Trietsch, M. D. *et al.* (2015) 'Genetic and epigenetic changes in vulvar squamous cell

carcinoma and its precursor lesions: A review of the current literature.', *Gynecologic oncology*. Elsevier Inc., 136(1), pp. 143–157. doi: 10.1016/j.ygyno.2014.11.002.

Tristram, A. *et al.* (2014) 'Activity, safety, and feasibility of cidofovir and imiquimod for treatment of vulval intraepithelial neoplasia (RT3VIN): A multicentre, open-label, randomised, phase 2 trial', *The Lancet Oncology*, 15(12), pp. 1361–1368. doi: 10.1016/S1470-2045(14)70456-5.

Ueda, Y. *et al.* (2011) 'Two distinct pathways to development of squamous cell carcinoma of the vulva.', *Journal of skin cancer*, 2011, p. 951250. doi: 10.1155/2011/951250.

Upchurch, E. *et al.* (2018) 'An update on the use of Raman spectroscopy in molecular cancer diagnostics: current challenges and further prospects', *Expert Review of Molecular Diagnostics*. Taylor & Francis, 18(3), pp. 245–258. doi: 10.1080/14737159.2018.1439739.

Vinokurova, S. *et al.* (2005) 'Clonal history of papillomavirus-induced dysplasia in the female lower genital tract', *Journal of the National Cancer Institute*, 97(24), pp. 1816–1821. doi: 10.1093/jnci/dji428.

Volmar, K. E. *et al.* (2014) 'Surgical pathology report defects: a College of American Pathologists Q-Probes study of 73 institutions.', *Archives of pathology & laboratory medicine*, 138(5), pp. 602–12. doi: 10.5858/arpa.2013-0099-CP.

De Vuyst, H. *et al.* (2009) 'Prevalence and type distribution of human papillomavirus in carcinoma and intraepithelial neoplasia of the vulva, vagina and anus: a meta-

analysis.', *International journal of cancer. Journal international du cancer*, 124(7), pp. 1626–36. doi: 10.1002/ijc.24116.

Walker, J. L. and Mathews, C. A. (2012) *Preinvasive Disease of the Vagina and Vulva and Related Disorders*. Eighth Edi, *Clinical Gynecologic Oncology: Eighth Edition*. Eighth Edi. Elsevier Inc. doi: 10.1016/B978-0-323-07419-3.00002-3.

Wallbillich, J. J. *et al.* (2012) 'Vulvar intraepithelial neoplasia (VIN 2/3): comparing clinical outcomes and evaluating risk factors for recurrence.', *Gynecologic oncology*. Elsevier Inc., 127(2), pp. 312–5. doi: 10.1016/j.ygyno.2012.07.118.

Wang, H. *et al.* (2012) 'Improving skin Raman spectral quality by fluorescence photobleaching.', *Photodiagnosis & Photodynamic Therapy*. Integrative Oncology Department, BC Cancer Research Centre, Vancouver, BC, Canada V5Z 1L3., 9(4), pp. 299–302.

Wang, L., Zhang, Z., *et al.* (2013) 'Evaluation of Raman spectroscopy for diagnosing EGFR mutation status in lung adenocarcinoma', *Analyst*, 139(2), pp. 455–463. doi: 10.1039/c3an01381b.

Wang, L., He, D., *et al.* (2013) 'Raman spectroscopy, a potential tool in diagnosis and prognosis of castration-resistant prostate cancer', *Journal of Biomedical Optics*, 18(8), p. 087001. doi: 10.1117/1.jbo.18.8.087001.

Wang, W. *et al.* (2015) 'Real-time in vivo cancer diagnosis using raman spectroscopy', *Journal of Biophotonics*, 8(7), pp. 527–545. doi: 10.1002/jbio.201400026.

Weir, M. M., Jan, E. and Colgan, T. J. (2003) 'Interinstitutional Pathology Consultations: A Reassessment', *American Journal of Clinical Pathology*, 120(3), pp. 405–412. doi: 10.1309/Q2HX-BQ17-4A1R-8H5Y.

Wessels, R. *et al.* (2012) 'Optical coherence tomography in vulvar intraepithelial neoplasia.', *Journal of Biomedical Optics*, 17(11 SP-EP-).

Winters, U. *et al.* (2008) 'Clinical and Immunologic Results of a Phase II Trial of Sequential Imiquimod and Photodynamic Therapy for Vulval Intraepithelial Neoplasia', *Clinical Cancer Research*, 14(16), pp. 5292–5299. doi: 10.1158/1078-0432.CCR-07-4760.

Woelber, L. *et al.* (2011) 'Prognostic Value of Pathological Resection Margin Distance in Squamous Cell Cancer of the Vulva', *Annals of Surgical Oncology*, 18(13), pp. 3811–3818. doi: 10.1245/s10434-011-1778-0.

Wong, P. T. *et al.* (1993) 'Distinct infrared spectroscopic patterns of human basal cell carcinoma of the skin.', *Cancer research*, 53(4), pp. 762–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8428355.

Woolderink, J. M. *et al.* (2006) 'Patterns and frequency of recurrences of squamous cell carcinoma of the vulva', *Gynecologic Oncology*, 103(1), pp. 293–299. doi: 10.1016/j.ygyno.2006.03.010.

Yap, J. *et al.* (2017) 'Current insights into the aetiology, pathobiology, and management of local disease recurrence in squamous cell carcinoma of the vulva', *BJOG: An International Journal of Obstetrics & Gynaecology*, 124(6), pp. 946–954.

doi: 10.1111/1471-0528.14560.

Yap, J. K. W. *et al.* (2016) ‘Adjacent Lichen Sclerosis predicts local recurrence and second fi eld tumour in women with vulvar squamous cell carcinoma’, *Gynecologic Oncology*. Elsevier Inc., pp. 4–10. doi: 10.1016/j.ygyno.2016.06.019.

Van Der Zee, A. G. J. *et al.* (2008) ‘Sentinel node dissection is safe in the treatment of early-stage vulvar cancer’, *Journal of Clinical Oncology*, 26(6), pp. 884–889. doi: 10.1200/JCO.2007.14.0566.

Zekan, J., Sirotkovic-Skerlev, M. and Skerlev, M. (2011) ‘Oncogenic Aspects of HPV Infections of the Female Genital Tract’, in *DNA Replication-Current Advances*. InTech. doi: 10.5772/19165.

Zhang, J. *et al.* (2018) ‘Accuracy of Raman spectroscopy for differentiating skin cancer from normal tissue’, *Medicine (United States)*, 97(34). doi: 10.1097/MD.0000000000012022.

Zhao, Jianhua *et al.* (2008) ‘Integrated real-time Raman system for clinical in vivo skin analysis.’, *Skin Research & Technology*. The Laboratory for Advanced Medical Photonics, Department of Dermatology and Skin Science, University of British Columbia, Vancouver, BC, Canada., 14(4), pp. 484–492. doi: 10.1111/j.1600-0846.2008.00321.x.

Zhao, J *et al.* (2008) ‘Real-time Raman spectroscopy for non-invasive skin cancer detection - preliminary results.’, *Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine & Biology Society*. Laboratory for

Advanced Medical Photonics (LAMP), Photomedicine Institute, Department of Dermatology and Skin Science, University of British Columbia & Vancouver Coastal Health Research Institute, B.C., Canada., 2008, pp. 3107–3109.

Zhao, J. *et al.* (2010) 'Real-time Raman Spectroscopy for Noninvasive in vivo Skin Analysis and Diagnosis', *Recent Advances in Biomedical Engineering*, (January), pp. 455–474. doi: 10.5772/7603.

Zhao, J. *et al.* (2014) 'Clinical Raman measurements under special ambient lighting illumination', *Journal of Biomedical Optics*, 19(11), p. 111609. doi: 10.1117/1.jbo.19.11.111609.

Zhao, J. *et al.* (2015) 'Real-time Raman spectroscopy for automatic in vivo skin cancer detection: an independent validation', *Analytical and bioanalytical chemistry*, 407(27), pp. 8373–8379. doi: 10.1007/s00216-015-8914-9.

Zhou, X.-H., Obuchowski, N. A. and McClish, D. K. (2011) *Statistical methods in diagnostic medicine*. 2nd Editio. Hoboken: Wiley.

Ziegler, L. D. (1990) 'Hyper-Raman spectroscopy', *Journal of Raman Spectroscopy*, 21(12), pp. 769–779. doi: 10.1002/jrs.1250211203.