# Prognostic Radiological Variables Derived From Oesophageal Cancer Staging Investigations

**Submitted for the degree of Doctor of Philosophy**

**at Cardiff University**

**Kieran G. Foley**

**2017**

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed …… (candidate) Date ……..…

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD.

Signed …… (candidate) Date ……

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated, and the thesis has not been edited by a third party beyond what is permitted by Cardiff University's Policy on the Use of Third Party Editors by Research Degree Students. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed …… (candidate) Date ……..

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed …… (candidate) Date ……..

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loans after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.

Signed …… (candidate) Date ……..

For my wife, Emma

**Acknowledgements**

Firstly, I wish to thank Dr Ashley Roberts. Without his supervision, hard work and commitment over the past 5 years, I would not have been able to complete this thesis. I'm extremely grateful for everything he's taught me, both academically and clinically. Thank you.

I would also like to thank Professor Howard Kynaston, my university supervisor, for all his advice and support. I'm extremely grateful to Drs Emiliano Spezi and Beatrice Berthon, who helped me pursue my interest in texture analysis and provided expert technical guidance. Thank you also to Craig Parkinson and Phil Whybra for their technical support.

Dr Adam Christian provided expert pathological evaluation and guidance in Chapter 3. I'm very grateful to Drs Tom Crosby and Carys Morgan for their input in Chapter 6 and to Professor Robert Hills for providing his statistical expertise. I'm also extremely grateful to Professor Wyn Lewis for his continued guidance and support.

I thank the staff in PETIC especially Dr Patrick Fielding and Professor Chris Marshall for their help. I acknowledge the support of the South Wales Radiology training scheme for the opportunity to undertake a period of research. I also acknowledge members of the surgical team and the Regional Upper GI cancer MDT.

Part of this research was funded by Tenovus Cancer Care and I'm very grateful for the support of the Wales Clinical Academic Training (WCAT) fellowship.

My dog, Pepper, took me on long walks which gave me the time to think things over. Thank you to Dr Lewis Connolly for lending his laptop at the start of this project.

Finally, thank you to Emma and my family for supporting me during this PhD.

**Summary**

Accurate radiological staging is vital in oesophageal cancer (OC). Radiological staging largely informs risk-stratification, treatment decisions and planning. However, the prognosis of OC remains poor, suggesting that radiological staging must improve. Therefore, the additional value of novel prognostic variables compared to current staging methods was assessed in a large cohort of OC patients managed by a regional upper gastrointestinal cancer network. Radiological-pathological correlation of resected lymph nodes assessed the accuracy of CT, EUS and PET/CT N-stage. The added value of PET-defined variables to predict circumferential resection margin (CRM) involvement was investigated. With EUS use declining, differences in PET and EUS measurements were assessed to understand potential implications for treatment planning should staging PET/CT be performed alone. Validation of two prognostic models; one in patients staged N0 on PET/CT and one incorporating novel PET features, was performed. The accuracy of CT, EUS and PET/CT N-stage was poor (54.5%, 55.4% and 57.1%, respectively) which greatly impacts on patient selection and treatment decisions. EUS continues to play an important role in OC staging, being significantly and independently associated with overall survival (OS; $p=0.012$) and CRM involvement ($p=0.022$). PET-defined variables had no additional value for predicting CRM status. The difference between PET and EUS length of disease was statistically significant ($p<0.001$), increasing the risk of geographical miss (38.1%) had PET/CT been used alone. Three novel PET image features (log (TLG), log(Histogram Energy) and Histogram Kurtosis) were independently associated with OS in the prognostic model. There was a significant OS difference between patient

quartiles ($p<0.001$) in the development and validation cohorts. Incorporation of these image features added prognostic value and improved model performance compared to current staging methods. These significant data demonstrate radiological prognostic variables that add value in OC management and highlight the importance of improved radiological staging.

## Table of Contents

## List of Tables

# List of Figures

**Abbreviations List**

| | |
|---|---|
| 3D | 3-dimensional |
| 5-FU | 5-fluorouracil |
| ACD | Annihilation coincidence detection |
| AIC | Akaike information criterion |
| AJCC | American Joint Committee on Cancer |
| APC | Argon plasma coagulation |
| AS | Automatic segmentation |
| AT | Adaptive threshold |
| ATLAAS | Advanced decision Tree-Learning Algorithm for Automatic Segmentation |
| BIC | Bayes information criterion |
| BMI | Body mass index |
| BSG | British Society of Gastroenterology |
| CaNISC | Cancer Network Information System Cymru |
| CAP | College of American Pathologists |
| CAVUHB | Cardiff and Vale University Health Board |
| CERR | Computational Environment for Radiotherapy Research |
| CI | Confidence interval |
| CR | Complete response |
| CRM | Circumferential resection margin |
| CT | Computed tomography |
| dCRT | Definitive chemo-radiotherapy |
| DFS | Disease-free survival |
| DSC | Dice similarity coefficient |
| ECX | Epirubicin, cisplatin and capecitabine |
| EDTV | Endoscopic derived tumour volume |
| ELoD | Endoscopic length of disease |
| EMR | Endoscopic mucosal resection |
| EPV | Events per variable |
| EUS | Endoscopic ultrasound |
| FCM | Fuzzy c-means |
| FDG | Fluorodeoxyglucose |
| FISH | Fluorescence in-situ hybridisation |
| G | Grade of differentiation |
| GC | Gradient contour |
| GCM | Gaussian c-means |
| GI | Gastrointestinal |
| GOJ | Gastro-oesophageal junction |
| GORD | Gastro-oesophageal reflux disease |
| GTV | Gross tumour volume |

| | |
|---|---|
| GUI | Graphical user interface |
| HGD | High-grade dysplasia |
| HR | Hazard ratio |
| IHC | Immunohistochemical |
| IQR | Inter-quartile range |
| IRB | Institutional review board |
| IV | Intra-venous |
| KM | K-means |
| LA | Level of agreement |
| LNM | Lymph node metastasis |
| LoD | Length of disease |
| LOR | Line of response |
| M | Metastasis |
| MAGIC | Medical Research Council Adjuvant Gastric Infusional Chemotherapy |
| MDT | Multi-disciplinary team |
| MRC | Medical Research Council |
| MRI | Magnetic resonance imaging |
| MTL | Metabolic tumour length |
| MTV | Metabolic tumour volume |
| MTW | Metabolic tumour width |
| N | Node |
| NACRT | Neo-adjuvant chemo-radiotherapy |
| NACT | Neo-adjuvant chemotherapy |
| NI | Number of intensity levels |
| NICE | National Institute for Health and Clinical Excellence |
| NOGCA | National Oesophago-Gastric Cancer Audit |
| NPV | Negative predictive value |
| NR | Non-response |
| OAC | Oesophageal adenocarcinoma |
| OC | Oesophageal cancer |
| OS | Overall survival |
| PET | Positron-emission tomography |
| PETIC | Positron Emission Tomography Imaging Centre |
| pM | Pathological metastasis |
| PMT | Photo-multiplier tube |
| pN | Pathological node |
| PPI | Proton pump inhibitor |
| PPV | Positive predictive value |
| PR | Partial response |
| pT | Pathological tumour |
| RCPath | Royal College of Pathologists |
| RCR | Royal College of Radiologists |
| RCT | Randomised control trial |

| | |
|---|---|
| RFA | Radio-frequency ablation |
| RG | Region growing |
| ROC | Receiver operator characteristic |
| ROI | Region of interest |
| SCC | Squamous cell carcinoma |
| SEER | Surveillance Epidemiology and End Results Program |
| SPSS | Statistical Package for Social Sciences |
| SUV | Standardised uptake value |
| T | Tumour |
| $TBR_{peak}$ | Peak target background ratio |
| Tis | Carcinoma in situ |
| TL | Tumour length |
| TLG | Tumour lesion glycolysis |
| TOF | Time of flight |
| TRG | Tumour regression grade |
| UICC | Union for International Cancer Control |
| UK | United Kingdom |
| USPIO | Ultrasmall superparamagnetic iron oxide |
| WECC | Worldwide Esophageal Cancer Collaboration |
| WT | Watershed transform |

**Contributions to Thesis**

Chapter 3

Dr Adam Christian (AC, Consultant Pathologist, University Hospital of Wales) performed measurements of lymph node and metastasis size in Chapter 3. Kieran Foley (KF) collected all radiological and clinical data, and performed statistical analyses.

Chapter 4

KF collected all data in Chapter 4 and performed statistical analyses, with support from Professor Wyn Lewis (Consultant Surgeon, University Hospital of Wales) in the original study. (Foley et al. 2014b)

Chapter 5

KF performed measurement of the PET-defined tumour variables following guidance from Dr Patrick Fielding (Consultant Radiologist, PETIC, Cardiff University). AC reviewed all cases for CRM involvement, confirming their status. KF performed statistical analyses with support from Professor Robert Hills (RH, Reader in Translational Statistics, Haematology Clinical Trials Unit, Cardiff University).

Chapter 6

PET LoD was retrospectively measured by KF. Dr Ashley Roberts (AR, Consultant Radiologist, University Hospital of Wales) performed all EUS examinations in this chapter, and formally recorded LoD in the final radiology report. KF performed all data collection and statistical analyses. Dr Carys Morgan and Dr Tom Crosby (Consultant Oncologists, Velindre Cancer Centre) reviewed this chapter and provided advice regarding radiotherapy decision-making and planning.

Chapter 7

KF collected all clinical data, segmented the tumours and performed texture analysis. Drs Emiliano Spezi (ES) and Beatrice Berthon (BB) (School of Engineering, Cardiff University) provided technical support and guidance regarding texture analysis. The prognostic model was developed and validated by RH, who has considerable expertise in this field.

Ethical and scientific committee applications were prepared and submitted by KF, with support from AR.

# Chapter 1. Introduction

Radiological staging of oesophageal cancer (OC) serves to define the extent of disease and influences management decisions made by the multi-disciplinary team (MDT). To ensure the best chance of survival for each patient, the most appropriate treatment should be decided upon and initiated promptly.

The purpose of this chapter is to provide the necessary background information for this thesis. Basic science of the oesophagus, including anatomy and histology, the diagnosis and management of the disease, advanced image analysis and the statistical background of prognostic modelling is discussed below.

## 1.1 Anatomy of Oesophagus

The oesophagus is a hollow muscular tube approximately 25 cm long, originating at the level of the cricoid cartilage in the hypopharynx. Its main purpose it to transport undigested food to the stomach. The thoracic oesophagus is located within the posterior mediastinal compartment and passes through the oesophageal hiatus of the diaphragm at the level of T10 before terminating at the gastro-oesophageal junction (GOJ). (Stevens and Lowe 1997)

Endoscopically, the oesophagus begins approximately 15-18 cm from the incisors to the GOJ at approximately 40 cm. (DeNardi and Riddell 1991) The oesophagus is conventionally classified into upper, middle and lower thirds. The Union for

International Cancer Control (UICC) define anatomical sub-sites of the oesophagus. (Table 1.1.1)

Table 1.1.1. Anatomical Description of Oesophageal Sub-sites

| Oesophageal sub-site | Anatomical Description |
|---|---|
| Cervical oesophagus | Commences at the lower border of the cricoid cartilage and ends at the thoracic inlet (suprasternal notch), approximately 18 cm from the upper incisor teeth. |
| Upper oesophagus | Extends from the thoracic inlet to the level of the tracheal bifurcation, approximately 24 cm from the upper incisor teeth. |
| Mid oesophagus | The proximal portion of the oesophagus between the tracheal bifurcation and the GOJ with the lower level approximately 32 cm from the upper incisor teeth. |
| Lower oesophagus | The distal portion of the oesophagus approximately 8 cm in length (including abdominal oesophagus), between the tracheal bifurcation and the GOJ with the lower level approximately 40 cm from the upper incisor teeth. |

The cervical oesophagus is supplied by the inferior thyroid artery. Oesophageal and bronchial branches from the anterior aspect of the thoracic aorta provide arterial blood supply to the thoracic oesophagus. Arterial blood supply to the abdominal oesophagus is from the left gastric and left inferior phrenic arteries. Venous drainage

of the thoracic oesophagus is via oesophageal veins, which enter the systemic azygous vein. The abdominal oesophagus drains via the left gastric vein into the porto-venous system. These two venous systems subsequently form a porto-systemic anastomosis. (Moore and Agur 2002)

Lymphatic drainage of the oesophagus is arranged in a bi-directional, free-flowing anastomosis of lymph vessels. The upper third of the oesophagus drains to the para-tracheal and internal jugular lymph nodes, the middle third of the oesophagus drains to the mediastinal nodes and the lower third of the oesophagus drains primarily to the left gastric and coeliac lymph nodes. (Riquet et al. 1993; Howard and Johnston 2013) This free-flowing lymphatic plexus is important in OC and allows for uninterrupted spread of lymph node metastases (LNMs), which usually originate in the peri-tumoural location. (Kayani et al. 2011)

## 1.2 Histology of Normal Oesophagus

The thoracic oesophagus is lined by stratified squamous epithelium, but the intra-abdominal oesophagus is lined by columnar epithelium. The transition between columnar and squamous epithelium is the GOJ, also called the z-line. This important region is a common site of pathology due to reflux of gastric contents into the distal oesophagus. (Stevens and Lowe 1997)

The oesophageal wall comprises 4 distinct histological layers; the mucosa, submucosa, muscularis propria and adventitia. (Fig. 1.2.1) The mucosa is subdivided

into 3 layers; the mucous membrane epithelium, lamina propria and muscularis mucosae. The mucosa comprises non-keratinised stratified squamous epithelium. The epithelium is further subdivided to a basal zone and a superficial zone, containing melanocytes, endocrine cells, T-lymphocytes and Langerhans cells. (Goldblum and Lee 2004)

The lamina propria is deep to the mucosa and consists of fibrovascular tissue containing capillaries, lymphatics, mononuclear cells and lymphoid aggregates. The fibrovascular tissue is folded into papillae and extends into the epithelium. Mucus secreting glands resembling those in the stomach are found within the lamina propria. (Goldblum and Lee 2004) The thickness of the muscularis mucosa varies and is maximally thick at the distal oesophagus where it approaches the GOJ. The muscularis mucosa comprises sheets of longitudinal and circular smooth muscle. (Stevens and Lowe 1997)

The submucosa contains mucous glands and has a rich supply of lymphoid tissue, blood vessels and nerves. (Crawford 2003) Each mucous gland comprises 2-5 lobes which drain into a duct lined by stratified columnar epithelium. These ducts traverse the mucosa and drain into the oesophageal lumen. (Stevens and Lowe 1997)

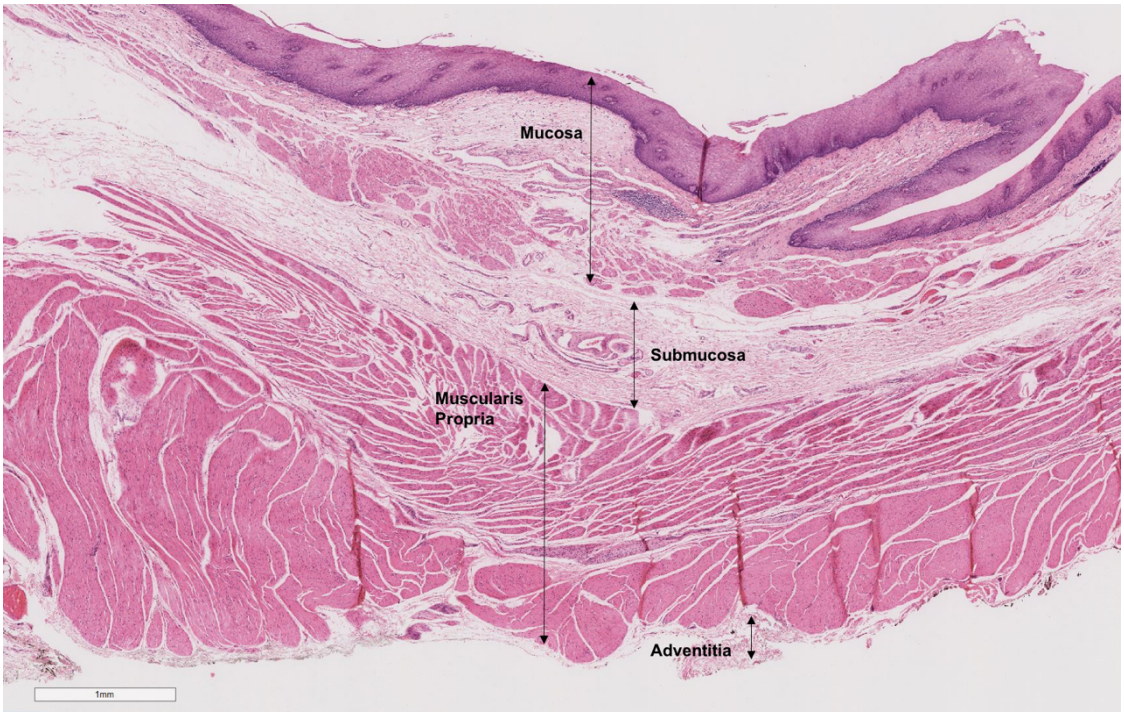



Figure 1.2.1. Longitudinal cross-sectional histological image of a normal oesophagus. The oesophagus has been stained with Haematoxylin & Eosin and the image was acquired at a magnification of 25x. The 4 distinct layers of the oesophagus are seen. The mucosa is divided into 3 layers; the epithelium, lamina propria and muscularis mucosae. The submucosa is deep to the mucosa. The inner circular and outer longitudinal layers of the muscularis propria are demonstrated. The adventitia is a thin covering layer.

The muscularis propria is generally arranged into discrete inner circular and outer longitudinal layers but can vary along the length of the oesophagus. The upper third consists mainly of striated muscle, gradually changing to smooth muscle in the middle third, where a combination of smooth and striated muscle is found. The muscularis propria of the lower third of the oesophagus is comprised mostly of smooth muscle. (Stevens and Lowe 1997) A developed network of nerves called Auerbach's plexus is located between the longitudinal and circular layers.

Unlike the remainder of the gastrointestinal (GI) tract, the thoracic oesophagus does not have a covering serosal layer, but an adventitia composed of loose connective tissue. Only a short segment of abdominal oesophagus has a serosa. Oesophageal tumours invade the mediastinum more readily due to the lack of serosa. (Crawford 2003) The histopathology of OC is discussed in section 1.7.4.2.

## 1.3 Epidemiology and Prognosis

This section discusses important gender, age-related and geographical differences in OC, whilst highlighting the poor prognosis.

### 1.3.1 Worldwide

OC is the eighth most common cancer, causing approximately 400,000 deaths per year. (Ferlay et al. 2015) There is significant variation in epidemiology worldwide. In 2012, there were an estimated 450,000 new cases of OC. Overall, squamous cell carcinoma (SCC) is more common (5.2 per 100,000) than adenocarcinoma (0.7 per 100,000). (Arnold et al. 2015)

Men have a higher incidence of OC, especially adenocarcinoma, with a male-to-female ratio of 4.4. SCC is most common in South-East and Central Asian countries such as Turkey, Iran and Northern China (79% of total cases worldwide). Adenocarcinoma is more common in developed countries in Northern and Western Europe, North America and Australasia (46% of total cases). There is a higher incidence of adenocarcinoma in high-income countries. (Edgren et al. 2013; Arnold et al. 2015) Similarly, ethnicity has a strong association with histological cell type. Adenocarcinoma is significantly more common in white populations, with SCC more common in black men. (Cooper et al. 2009)

### 1.3.2 United Kingdom

OC is the 14th most common malignancy in the United Kingdom (UK), representing 2% of all new cases of cancer. There were more than 8,900 new cases in 2014. The incidence of OC is significantly higher in the elderly. Fifty-six percent of new cases were diagnosed in patients over 70 years of age. (Cancer Research UK 2016a) OC is more common in men (3:1) but the incidence has increased in both men and women since the mid-1970s. (Adams and Jaunoo 2014)

Scotland has a significantly higher incidence of OC compared to England, Northern Ireland and Wales. There appears to be a North-South divide in the UK, with cancer networks across Scotland and Northern England reporting a higher incidence than South-East England. Wales has the 2nd highest crude rate of OC in the UK. (National Cancer Intelligence Network 2008; Cancer Research UK 2016a) Socio-economic status could account for some of these differences, as there are increased rates of SCC in deprived areas. (Cooper et al. 2009)

### 1.3.3 Prognosis

The prognosis of early OC is good. The 5-year disease-free survival (DFS) of early mucosal tumours (T1) is 95%. (Takeshita et al. 1997)

However, overall 5-year survival rates are highly dependent on disease stage. (Surveillance Epidemiology and End Results Program (SEER) 2006-2012). Table

1.3.1 shows 5-year overall survival (OS) rates for all new cases of OC in the USA between 2006 and 2012. In the UK, overall 5-year survival is 15%, with 40% of patients surviving for 1 year, 15% for five years and 10% for ten years. (Cancer Research UK 2016b)

Table 1.3.1. Overall 5-year Survival Rates for Localised, Regional and Distant Disease in USA Between 2006 and 2012.

| | Definition | 5-year OS (%) |
|---|---|---|
| Localised (Stage I/II) | T1-3, N0, M0 | 41.3 |
| Regional (Stage III) | T4, N1-3, M0 | 22.8 |
| Distant (Stage IV) | Any T, any N, M1 | 4.5 |
| Unknown | | 12.4 |

## 1.4 Aetiology and Risk Factors

There are several risk factors associated with OC. However, different risk factors are associated with the development of the two main histological types; adenocarcinoma and SCC.

### 1.4.1 Diet and Lifestyle

Tobacco smoking and alcohol intake are the main risk factors for SCC. (De Stefani et al. 1993; Lee et al. 2007) The risk of smokers developing adenocarcinoma is less conclusive. A multicentre, case-control study including over 500 patients showed an increased risk of adenocarcinoma, which persisted for 30 years after smoking cessation. (Gammon et al. 1997) Excessive alcohol consumption is associated with OC and the risk increases in different ethnic groups. Genetic mutations of alcohol and aldehyde dehydrogenases increase the risk in East Asian heavy drinkers. (Yokoyama and Omori 2003)

Diet is associated with the development of SCC and adenocarcinoma. The excessive consumption of processed meat, pickled foods, high-fat dairy products and hot drinks are associated with increased rates of OC. (Cheng et al. 1992; Bahmanyar and Ye 2006; Islami et al. 2009)

Some medications may reduce the risk of OC. In patients with known Barrett's oesophagus, the risk of progression to high-grade dysplasia (HGD) or

adenocarcinoma is reduced in up to 71% of patients taking proton pump inhibitors (PPI), although this risk reduction may only be seen following long-term use. (Singh et al. 2014) Regular use of cyclo-oxygenase inhibitors, non-steroidal anti-inflammatories such as aspirin, and statins may also reduce the risk. (Kantor et al. 2012; Zhang et al. 2014) Long term use of low-dose aspirin (for more than 5 years) reduces the risk of colorectal cancer by 27%, but further research investigating its utility in OC is required. (Friis et al. 2015)

### 1.4.2 Gastro-Oesophageal Reflux Disease and Obesity

The worldwide increase in obesity coupled with gastro-oesophageal reflux disease (GORD) has seen a rise in the "pre-malignant" condition Barrett's oesophagus, which is defined as any portion of the normal distal squamous epithelium replaced by metaplastic columnar epithelium, endoscopically visible 1 cm or more above the GOJ. (Fitzgerald et al. 2014)

GORD is the reflux of gastric contents into the distal oesophagus caused by inappropriate relaxation of the lower oesophageal sphincter. GORD may be asymptomatic but can cause retrosternal chest pain (dyspepsia), difficulty swallowing (dysphagia) or cough. GORD reduces the intraluminal pH, resulting in epithelial and mucosal damage.

Obesity itself is a cause of GORD, and obese people tend to have a poor diet, which further increases the risk of developing OC. There is an association between body

mass index (BMI) and adenocarcinoma. Patients in the highest BMI quartile have an adjusted odds ratio of 7.6 (95% confidence interval (CI), 3.8 to 15.2) compared to those in the lowest BMI quartile for developing adenocarcinoma. (Lagergren et al. 1999)

### 1.4.3 Barrett's Oesophagus and Adenocarcinoma

A large, population-based cohort study in Denmark surveilled over 11,000 patients diagnosed with Barrett's oesophagus between 1992 and 2009. The overall incidence of adenocarcinoma was 2.9 cases per 1,000 person-years, representing 7.6% of the total new cases of adenocarcinoma in the country. This equates to an annual risk of 0.12%, or 1 case of adenocarcinoma per 860 person-years. The authors concluded that routine surveillance of these patients is of doubtful value. (Hvid-Jensen et al. 2011) However, given that surveillance correlates with earlier cancer detection and consequently improved survival, the British Society of Gastroenterology (BSG) recommends surveillance in these patients. This is the standard of care in the UK. (Fitzgerald et al. 2014)

## 1.5 Diagnosis

OC is usually diagnosed following endoscopy and biopsy. OC can be diagnosed radiologically with a Barium swallow examination or computed tomography (CT).

### 1.5.1 Clinical Presentation

Dysphagia is the most common symptom of OC. Dysphagia is most often caused by tumour mass effect and often explains the frequently late presentation of patients with advanced disease. (Cancer Research UK 2016a)

### 1.5.2 Investigation

The National Institute for Health and Clinical Excellence (NICE) guidance states that urgent direct access to upper GI endoscopy should be performed within 2 weeks, in patients with dysphagia or aged 55 and over, with weight loss and either upper abdominal pain, reflux or dyspepsia. (National Institute for Health and Clinical Excellence 2015) Upper GI endoscopy is performed by a variety of health-care professionals and is the first-choice investigation for OC. Information regarding location, length of tumour and the degree of circumferential abnormality can also be evaluated.

Patients not wishing to undergo upper GI endoscopy, or that are felt to be unsuitable due to the risks of the procedure, can have a double-contrast barium swallow

examination performed as an alternative investigation. The patient ingests barium sulphate solution and effervescent granules whilst fluoroscopic images of the hypopharynx, thoracic oesophagus, GOJ, stomach and proximal duodenum are acquired. The degree of stenosis in a stricturing tumour can be assessed.

## 1.6 Management of Oesophageal Cancer

The management of OC is dependent upon a combination of factors including clinical MDT decisions, patient wishes and co-morbidities, but is heavily influenced by radiological staging. Randomised clinical trials (RCTs) have attempted to demonstrate the best treatment, but the optimum management is still unknown. Several studies have failed to show a significant difference in outcome between different treatment regimes.

### 1.6.1 Therapeutic Endoscopy Techniques for Early Cancer

Endoscopic treatment plays an important role in early OC. Such techniques include endoscopic mucosal resection (EMR), argon plasma coagulation (APC) and radio-frequency ablation (RFA). The aim is to remove areas of high-grade dysplasia and T1a cancers which, for adenocarcinoma, are normally present in areas of Barrett's oesophagus. There is significantly less morbidity following these techniques than

surgery. (Allum et al. 2011) These endoscopic techniques are restricted to T1a N0 tumours. (Griffin et al. 2011)

### 1.6.2 Treatment Options in Advanced Cancer

Patients with incurable disease receive palliative therapy, including placement of metal stents to improve dysphagia. For those with potentially curable disease, treatment options include surgery alone, neo-adjuvant chemotherapy (NACT), neo-adjuvant chemo-radiotherapy (NACRT) and definitive chemo-radiotherapy (dCRT).

#### *1.6.2.1 Surgery*

Surgical management aims to provide a definitive cure for the patient. Due to the commonly late presentation of OC, surgical management is an option for relatively few patients (20-30%). (Crosby and Evans 2009) The 2- and 5-year survival rates of patients treated with surgery alone is 34% and 17%, respectively. (Medical Research Council Oesophageal Cancer Working Group 2002; Allum et al. 2009) Quality of life post-oesophagectomy is an important consideration. Surgical patients only regain their quality of life if they live for at least 2 years following resection, therefore careful patient selection is critical. (Blazeby et al. 2000)

Two types of oesophagectomy are commonly performed in the UK. (National Oesophago-Gastric Cancer Audit 2016) The Ivor-Lewis oesophagectomy (trans-thoracic approach) involves a thoracotomy and en-bloc resection of oesophagus and

lymph nodes. The trans-hiatal approach involves a gastric pull-through and cervical anastomosis. A formal lymphadenectomy is not performed. Comparable survival statistics have been reported, however morbidity rates are significantly lower following trans-hiatal oesophagectomy than an Ivor-Lewis procedure. (Hulscher et al. 2002; Davies et al. 2014b)

NACT is given in attempt to reduce the tumour volume and disease stage prior to surgery and is currently the first-line treatment. (Medical Research Council Oesophageal Cancer Working Group 2002) Patients with less advanced disease may have surgery alone.

#### *1.6.2.2 Neo-Adjuvant Chemotherapy*

Patients receive 3 cycles of epirubicin, cisplatin and capecitabine (ECX) prior to surgery. (Cunningham et al. 2008) Survival rates from the Medical Research Council (MRC) OE02 and USA Intergroup 113 RCTs comparing NACT with surgery alone are shown in Table 1.6.1.

The USA Intergroup 113 RCT found no significant difference in survival between surgery with or without NACT, but the larger MRC OE02 trial (802 patients) found that two cycles of NACT improved OS ($p=0.03$). A subsequent Cochrane review demonstrated the likely benefit of NACT, but evidence was inconclusive due to associated toxicity. (Vogt et al. 2006)

Table 1.6.1. Five-year Overall Survival Rates from MRC OE02 and USA Intergroup Trials Comparing NACT with Surgery Alone

| Randomised Control Trial | 5-year Overall Survival (%) | |
|---|---|---|
| | Surgery Alone | Neo-Adjuvant Chemotherapy |
| MRC OE02 (Medical Research Council Oesophageal Cancer Working Group 2002; Allum et al. 2009) | 17.1 | 23.0 |
| USA Intergroup 113 (Kelsen et al. 1998; Kelsen et al. 2007) | 19.8 | 19.4 |

Both the MRC Adjuvant Gastric Infusional Chemotherapy (MAGIC) trial, a primarily gastric cancer trial that included GOJ tumours, and the MRC OE05 trial increased the number of NACT cycles to 4, but found no significant difference in survival. (Cunningham et al. 2006; Alderson et al. 2015)

#### *1.6.2.3 Neo-Adjuvant Chemo-radiotherapy*

The CROSS trial showed improved survival, pathological response and resection margin involvement following NACRT compared to surgery alone. (van Hagen et al. 2012) NACRT improved OS (median OS 49.4 vs 24.0 months, p=0.003). Long-term

survival data confirmed the improved OS (NACRT 48.6 months vs surgery alone 24.0 months, HR 0.68, 95% CI 0.53-0.88, p=0.003). (Shapiro et al. 2015) There was also an improvement in the number of involved resection margins (NACRT 8% vs surgery alone 31%, p<0.001). A potential disadvantage of NACRT is an increase in operative morbidity and mortality. (Urschel and Vasan 2003; Fiorica et al. 2004) RCTs comparing NACT and NACRT are lacking. At the time of writing, the Neo-AEGIS trial was randomising patients between NACT plus post-operative chemotherapy (modified MAGIC regime) and NACRT using the CROSS protocol. (Keegan et al. 2014)

#### *1.6.2.4 Definitive Chemo-radiotherapy*

dCRT is available for patients with potentially curable disease who have relatively poor performance status precluding them from a major operation, or those that do not wish to have surgery. Overall 3- to 5-year survival following dCRT ranges between 20-30%. (Minsky et al. 2002; Crosby et al. 2004; Stahl et al. 2005; Bedenne et al. 2007) The SCOPE1 phase II/III RCT compared dCRT regimes with and without cetuximab but found no significant difference in survival and was closed early on grounds of futility. (Crosby et al. 2017) A small Swedish study including 91 patients with SCC and adenocarcinoma found no significant difference in survival between dCRT and surgery alone. Four-year OS for surgery and dCRT was 23% and 29.6%, respectively. (Carstens et al. 2007) Similarly, no OS difference was found between dCRT and surgery in a propensity score analysis of 521 OC patients. (Karran et al. 2014)

The decision to use radiotherapy partly relies on the total length of disease (LoD), which can be measured from positron-emission tomography (PET) and endoscopic ultrasound (EUS) staging investigations. LoD is defined as the maximum cranio-caudal length of the primary tumour plus any regional LNMs. The prognostic significance of LoD is described in section 1.9.3 with further discussion and investigation of the importance of LoD for treatment planning in Chapter 6.

#### *1.6.2.5 Summary*

Although evidence regarding the best treatment of OC is lacking, this section highlights the importance of OC staging. Differentiation of early tumours from advanced but potentially curable disease allows alternative management techniques. Defining non-curable disease prevents unnecessary over-treatment, which affects patient's quality of life in their final stages.

## 1.7 Staging of Oesophageal Cancer

OC is staged according to the pathological Tumour, Node and Metastasis (TNM) classification. (Sobin et al. 2009) The TNM classification aims to separate patients into groups based on their prognosis.

### 1.7.1 Tumour Node Metastasis Classification

The UICC first introduced the TNM staging system in 1968. The American Joint Committee on Cancer (AJCC) released the 1st edition of TNM classification for OC in 1977. (Tangoku et al. 2012)

Several editions have subsequently been published. The TNM 5th edition for OC divided M-stage into M1a and M1b. The M1a category included cervical and coeliac LNMs, and M1b included all other distant metastases. (Sobin and Wittekind 1997) The TNM 6th edition did not differ from the 5th edition. (Sobin and Wittekind 2002) The TNM 7th edition (Table 1.7.1) was used to classify stage throughout this thesis period and has therefore been used in all chapters. (Sobin et al. 2009) The TNM 8th edition was published in December 2016 and is discussed in Chapter 8. (Rice et al. 2017) Radiological TNM staging is discussed in Section 1.8.

The 7th edition differed from the 6th edition with the inclusion of HGD and Tis (carcinoma in situ). T1 was separated into T1a and T1b, tumour invasion into mucosa and submucosa, respectively. T4 was also separated to differentiate resectable (T4a)

disease (tumour invading pleura, pericardium or diaphragmatic crus) and non-resectable (T4b) disease (tumour invasion into adjacent organs such as aorta, vertebral body or trachea). (Sobin et al. 2009)

The Siewert classification defines the location of GOJ tumours. (Siewert et al. 2000) Type I tumours are located in the distal oesophagus with the epicentre 5 cm or less from the GOJ. Type II tumours are 'true' tumours of the GOJ with the epicentre at the junction. Type III tumours are proximal gastric cancers that extend into the distal oesophagus. Tumours that do not extend into the oesophagus and have an epicentre more than 5 cm from the GOJ are staged as gastric cancers.

GOJ tumours have different lymphatic drainage pathways. Generally, type I tumours disseminate to the chest and type III tumours to the upper abdominal lymph nodes. The most common locations of LNMs are the left paracardial and lesser curvature nodes (67.8%), right paracardial nodes (56.9%) and left gastric artery and coeliac axis nodes (26.8%). (Siewert et al. 2000)

Regional lymph nodes are defined as those that drain the oesophagus, irrespective of the tumour location. They include the para-oesophageal lymph nodes in the neck cranially and the coeliac axis lymph nodes caudally. All other lymph nodes, including supraclavicular, are classified as non-regional (M1). Distant metastases are classified as dissemination to any organ including liver, lung and bone.

Table 1.7.1. UICC TNM 7th Edition Classification of Oesophageal Cancer

| Primary Tumour (T) | |
|---|---|
| TX | Primary tumour cannot be assessed |
| T0 | No evidence of primary tumour |
| Tis | Carcinoma in situ / high-grade dysplasia |
| T1a | Tumour invades lamina propria or muscularis mucosae |
| T1b | Tumour invades submucosa |
| T2 | Tumour invades muscularis propria |
| T3 | Tumour invades adventitia |
| T4a | Tumour invades pleura, pericardium or diaphragm |
| T4b | Tumour invades other adjacent structures such as aorta, vertebral body or trachea |
| **Regional Lymph Nodes (N)** | |
| NX | Regional lymph nodes cannot be assessed |
| N0 | No regional lymph node metastases |
| N1 | Metastases in 1-2 regional lymph nodes |
| N2 | Metastases in 3-6 regional lymph nodes |
| N3 | Metastases in 7 or more lymph nodes |
| **Distant Metastases (M)** | |
| M0 | No distant metastases |
| M1 | Distant metastases |

### 1.7.2 Staging Groups in Oesophageal Cancer

Patients can also be classified into stage groups related to prognosis (Table 1.7.2). (Sobin et al. 2009) The presence of lymph node and distant metastatic disease increases stage group and indicates a poorer prognosis.

Table 1.7.2. Stage Groups in Oesophageal Cancer

| Stage Group | T | N | M |
|---|---|---|---|
| Stage 0 | Tis | N0 | M0 |
| Stage IA | T1 | N0 | M0 |
| Stage IB | T2 | N0 | M0 |
| Stage IIA | T3 | N0 | M0 |
| Stage IIB | T1, T2 | N1 | M0 |
| Stage IIIA | T4a | N0 | M0 |
| | T3 | N1 | M0 |
| | T1, T2 | N2 | M0 |
| Stage IIIB | T3 | N2 | M0 |
| Stage IIIC | T4a | N1, N2 | M0 |
| | T4b | Any N | M0 |
| | Any T | N3 | M0 |
| Stage IV | Any T | Any N | M1 |

### 1.7.3 Prognostic Groups in Oesophageal Cancer

Prognostic groups are also described in the TNM 7th edition. (Sobin et al. 2009) The TNM stage, tumour location and grade of differentiation are used to classify patients accordingly.

The importance of non-anatomical data for staging purposes was identified by the AJCC and incorporated into the TNM 7th edition. These data included histopathological cell type, grade of differentiation and tumour location. For example, stages I and II were separated into adenocarcinoma and SCC groups to reflect differences in survival. In SCC, lower oesophageal tumours are grouped differently from upper and mid T2/3 N0 M0 tumours because prognosis is worse in the latter. (Table 1.7.3) In stage I and II adenocarcinoma (Table 1.7.4), the differentiation of G1 and G2 (well and moderately differentiated) from G3 (poorly differentiated) has prognostic significance. (Rice 2010)

Table 1.7.3. Prognostic Groups for Squamous Cell Carcinoma

| | T | N | M | Grade (G) | Location |
|---|---|---|---|---|---|
| Group 0 | Tis | 0 | 0 | 1 | Any |
| Group IA | 1 | 0 | 0 | 1, X | Any |
| Group IB | 1 | 0 | 0 | 2, 3 | Any |
| | 2, 3 | 0 | 0 | 1, X | Lower, X |
| Group IIA | 2, 3 | 0 | 0 | 1, X | Upper, middle |
| | 2, 3 | 0 | 0 | 2, 3 | Lower, X |
| Group IIB | 2, 3 | 0 | 0 | 2, 3 | Upper, middle |
| | 1, 2 | 1 | 0 | Any | Any |
| Group IIIA | 1, 2 | 2 | 0 | Any | Any |
| | 3 | 1 | 0 | Any | Any |
| | 4a | 0 | 0 | Any | Any |
| Group IIIB | 3 | 2 | 0 | Any | Any |
| Group IIIC | 4a | 1, 2 | 0 | Any | Any |
| | 4b | Any | 0 | Any | Any |
| | Any | 3 | 0 | Any | Any |
| Group IV | Any | Any | 1 | Any | Any |

Table 1.7.4. Prognostic Groups for Adenocarcinoma

| | T | N | M | Grade (G) |
|---|---|---|---|---|
| Group 0 | Tis | 0 | 0 | 1 |
| Group IA | 1 | 0 | 0 | 1, 2, X |
| Group IB | 1 | 0 | 0 | 3 |
| | 2 | 0 | 0 | 1, 2, X |
| Group IIA | 2 | 0 | 0 | 3 |
| Group IIB | 3 | 0 | 0 | Any |
| | 1, 2 | 1 | 0 | Any |
| Group IIIA | 1, 2 | 2 | 0 | Any |
| | 3 | 1 | 0 | Any |
| | 4a | 0 | 0 | Any |
| Group IIIB | 3 | 2 | 0 | Any |
| Group IIIC | 4a | 1, 2 | 0 | Any |
| | 4b | Any | 0 | Any |
| | Any | 3 | 0 | Any |
| Group IV | Any | Any | 1 | Any |

### 1.7.4 Pathological Staging

The following sections introduce important prognostic pathological features associated with the TNM stage classification. In patients undergoing surgical resection, core data should be reported as per the Royal College of Pathologists (RCPath) guidelines (Appendix A). (Mapstone 2007)

#### *1.7.4.1 T-stage*

T-stage classifies the depth of tumour invasion by anatomic landmarks. The depth of tumour invasion is an important predictor of survival. (Ide et al. 1994; Lieberman et al. 1995; Paraf et al. 1995; Khan et al. 2003) Five-year OS rates approach 82% when the tumour is limited to the mucosa or submucosa. (Paraf et al. 1995)

Depth of tumour invasion is associated with the presence of regional LNMs. The risk of regional LNMs is highly unlikely (0%) in T1a tumours, but rises to 12% in T1b tumours invading the submucosa. (Griffin et al. 2011) Approximately 74% of pT3 tumours and 83% of pT4 tumours have regional LNMs. (Ide et al. 1994)

An interesting phenomenon that occurs in patients with a tumour situated in a segment of Barrett's oesophagus is duplication of the muscularis mucosae. This has potentially significant implications when staging early, superficial tumours. (Abraham et al. 2007)

*1.7.4.2 Histopathology*

Adenocarcinoma and SCC are the two most common histological cell types. Less common histological cell types include neuro-endocrine, small cell carcinoma and leiomyosarcoma, comprising less than 1% of total cases. (Arnold et al. 2015) Adenocarcinoma is the more favourable histology in early cancers because local recurrence rates are lower than SCC. (Holscher et al. 1995) Histological cell type becomes less important in larger, more advanced tumours. (Lieberman et al. 1995)

Grade of tumour differentiation should be assessed by a Pathologist according to the WHO International Histological Classification of Tumours guidance. (Table 1.7.5) (Watanabe et al. 1990) The evidence regarding the significance of grade differentiation on OS is conflicting, with studies unable to demonstrate a significant difference between groups. (Robey-Cafferty et al. 1991; Ide et al. 1994; Paraf et al. 1995)

Table 1.7.5. Histological Grade of Differentiation in Oesophageal Cancer

| Grade | Degree of Differentiation |
|---|---|
| G1 | Well differentiated |
| G2 | Moderately differentiated |
| G3 | Poorly differentiated |
| G4 | Undifferentiated |
| GX | Grade of differentiation cannot be assessed |

#### *1.7.4.3 Resection Margins*

Assessment of proximal, distal and circumferential resection margins (CRM) should always be performed by a Pathologist. (Mapstone 2007) An R0 grading is defined as the absence of resection margin involvement following resection. R1 is evidence of microscopic involvement and R2 is clear, macroscopic evidence of resection margin involvement. (Mapstone 1998)

Involvement of the proximal and distal resection margins is associated with an increased likelihood of disease recurrence. The evidence for recurrence with proximal resection margin involvement is more compelling than for distal resection margin involvement. (Paraf et al. 1995; Mariette et al. 2003) Mariette et al found a significant difference in median survival between patients with positive and negative proximal resection margins (median survival 11.1 vs 36.3 months, respectively).

Different definitions of CRM involvement exist. The College of American Pathologists (CAP) defines CRM involvement as tumour at the cut resection margin. The RCPath define CRM involvement as tumour within 1 mm of the resection margin. (Mapstone 1998; College of American Pathologists 2005) Likewise, different terminology regarding pre-operative CRM assessment are used. A threatened CRM describes the radiological appearance of a CRM at risk (defined as tumour within 1 mm of the CRM on imaging), and is commonly used in rectal cancer staging. (Brown et al. 2003) A CRM can be described as involved or not involved following pathological examination.

Sagar et al produced the first major paper describing the prognostic significance of CRM involvement, but conflicting evidence exists. (Sagar et al. 1993) Khan et al found no such association in 329 patients. (Khan et al. 2003) Both studies used the RCPath definition of CRM involvement. A systematic review and meta-analysis including 2,433 patients found a significant difference in 5-year survival between CRM status using both RCPath and CAP definitions of a positive CRM. (Chan et al. 2013b)

Reid et al found a positive CRM to be significantly and independently associated with OS for all pT-stages. (Reid et al. 2012) Advanced EUS T-stage was independently associated with a positive CRM, with an almost 25-fold increased risk of CRM involvement, once the tumour was classified T3 or greater.

#### *1.7.4.4 Tumour Length*

Tumour length (TL) is an important prognostic factor and should be measured by the Pathologist following resection. TL is included in the core dataset for reporting. (Eloubeidi et al. 2002; Mapstone 2007) The resection specimen should be pinned immediately during preparation as the specimen can shrink by up to a third if not fixed adequately. (Siu et al. 1986)

#### *1.7.4.5 Tumour Regression Grade*

Mandard et al described a classification of pathological tumour response to NACRT. (Mandard et al. 1994) This classificiation is also routinely used following NACT. The

Mandard score represents a 5-point scale quantifying tumour regression grade (TRG). TRG 1 represents complete regression, defined by the absence of residual cancer with fibrosis extending through the layers of the oesophageal wall. TRG 2 is defined by scattered residual cancer cells in predominant fibrosis. In TRG 3, fibrosis predominates but there are increased numbers of cancer cells compared to TRG 2. TRG 4 represents largely residual cancer with some fibrosis and TRG 5 is absence of any treatment effect. TRG 1-3 is considered to represent treatment response whereas TRG 4-5 is regarded as minimal or no treatment response. Treatment response has significant implications for patient outcome, as pathological stage post neo-adjuvant therapy and pathological TRG are independent prognostic factors. (Mandard et al. 1994; Davies et al. 2014a)

#### *1.7.4.6 Lymph Node Metastases*

The presence of LNMs has been described as the single most important prognostic factor in patients with OC. (Kayani et al. 2011) Patients with LNMs have an overall 5-year survival of 18-47% following surgical resection compared to 70-92% without LNMs.

The number of regional LNMs and the ratio of pathological to normal lymph nodes have prognostic significance. (Wilson et al. 2008; Twine et al. 2009b; Liu et al. 2010) This finding was reflected in the TNM 7$^{th}$ edition, with formation of N0-N3 groups described in Table 1.7.1. The previous 6$^{th}$ edition only acknowledged the presence of regional LNMs (N0 or N1). Various studies have suggested different category

thresholds for the number of involved LNMs, but all studies agree that LNMs predict a poorer outcome. (Rizk et al. 2006; O'Riordan et al. 2007; Mariette et al. 2008; Hu et al. 2010) Five-year survival rates of patients with no LNMs range between 49.1%-57.0%, 1-2 LNMs between 19.5%-33% and more than 3 LNMs approximately 11.0%. (Hu et al. 2010; Liu et al. 2010)

The lymph node ratio is the total number of LNMs compared to the total number of resected nodes. Lymph node ratio is an independent predictor of OS (median survival 27 months with lymph node ratio <11% vs 13 months with lymph node ratio >33%). (Bogoevski et al. 2008) At least 10 lymph nodes should be examined before confidently staging as pN0. (Twine et al. 2009b) The current recommendation from the National Oesophago-Gastric Cancer Audit (NOGCA) is to resect a minimum of 15. (National Oesophago-Gastric Cancer Audit 2016)

#### *1.7.4.7 Micro-metastases*

Micro-metastases are small metastatic lesions within lymph nodes and can comprise a single or tiny cluster of cancer cells. Traditional histopathological methods involve bisecting a lymph node and evaluating the exposed tissue. (Jiao et al. 2003) Up to 50% of lymph nodes histologically diagnosed as N0, have further evidence of micro-metastases. (Luketich et al. 1998)

Doubt remains whether micro-metastases have malignant potential, simply represent non-proliferating cells or are tumour cells that have been cleared by the immune

system. (Izbicki et al. 1997; Grotenhuis et al. 2010) The prognostic significance of micro-metastases in OC is unclear and conflicting evidence has been published. Some studies failed to show an association between lymph node micro-metastases and outcome (Glickman et al. 1999), whilst others have found an association with relapse-free survival and OS. (Izbicki et al. 1997)

Importantly, micro-metastases cannot be directly imaged at present. Studies attempting to predict the presence of lymph node micro-metastases have made associations with tumour length, lymphatic infiltration and vascular invasion. (Eloubeidi et al. 2002; Wayman et al. 2002; Tanabe et al. 2003) A study in patients with SCC investigated metabolic parameters on PET and found an association between the combined maximum standardised uptake value of the primary tumour ($SUV_{max}$) and clinical T-stage for predicting micro-metastases. (Moon et al. 2014)

#### *1.7.4.8 Extra-capsular lymph node involvement*

Extra-capsular lymph node involvement is defined as extension of cancer cells through the capsule into adjacent fat. Lagarde et al studied 1,562 positive lymph nodes and found extra-capsular lymph node involvement in 456 nodes (29%). (Lagarde et al. 2006) Extra-capsular involvement was associated with advanced T-stage, number of LNMs and LNM ratio. Although there was no significant difference in disease recurrence, patients with extra-capsular involvement had poorer median survival (15 months vs 41 months, $p < 0.001$).

## 1.8 Radiological TNM staging of oesophageal cancer

In 2006, the Worldwide Esophageal Cancer Collaboration (WECC) was assembled to improve OC staging. A database of more than 4,600 surgical patients was collated following recruitment from numerous international centres. These data were used to define the TNM 7$^{th}$ edition classification. (Rice 2010)

In the UK, the Royal College of Radiologists (RCR) published guidance regarding the radiological staging pathway of upper GI cancer. After a diagnosis of OC is made, patients are initially staged with CT of the chest and abdomen. This is predominantly used to exclude distant metastases and irresectable locally advanced disease. If patients are deemed potentially curable, either with surgery, chemotherapy, radiotherapy or a combination, they undergo 18-Fluorine ($^{18}$F) fluorodeoxyglucose (FDG) PET/CT followed by EUS for more detailed staging. (Roberts and Kay 2014) Staging with MRI has been investigated in a research setting but is not routinely performed.

### 1.8.1 Normal Oesophageal and GOJ Appearance on Imaging

The normal oesophagus and GOJ can vary in appearance depending on the radiological modality and scan preparation. On CT, normal oesophageal wall thickness changes depending on luminal distension. When adequately distended, the wall is thin and well-defined.

EUS differentiates the individual layers of the oesophageal wall clearly, allowing superior assessment of the depth of tumour invasion. (Fig. 1.8.1) PET provides little information regarding the normal oesophagus. Benign, low-grade uptake can indicate oesophagitis and could mimic malignancy if reviewed independently and blinded to other investigations.

On T2-weighted MRI sequences, it is possible to differentiate the layers of the normal oesophageal wall, provided the sequence is optimised. Normal mucosa, submucosa and muscularis propria demonstrate intermediate, high signal and low signal, respectively. (Riddell et al. 2006)

### 1.8.2 T-stage

Radiological T-stage follows the same classification as pT-stage detailed above. (Table 1.7.1)

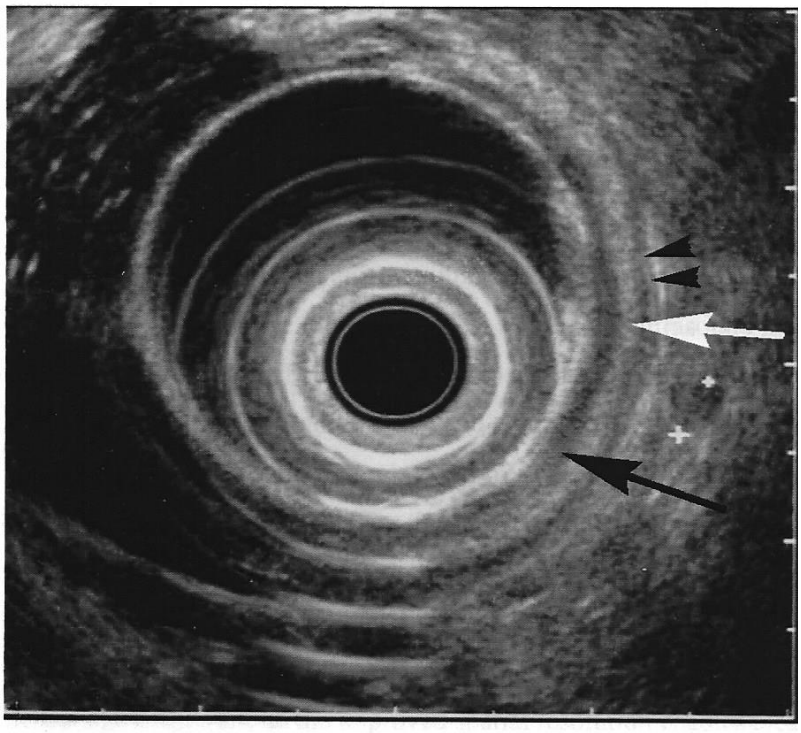


Figure 1.8.1. Selected radial endoscopic ultrasound image demonstrating different layers of the oesophageal wall. The inner hypoechoic layer (black arrow) represents the muscularis mucosa and the middle hyperechoic layer (white arrow) represents the submucosa. The outer hypoechoic layer (black arrowheads) represents the muscularis propria. A small, hypoechoic lymph node is marked with calipers. The image was acquired in the distal oesophagus at 36 cm using a UM-2000 Olympus Video EUS endoscope (Key Med, Southend, UK).

#### *1.8.2.1 Tumour Characteristics by Imaging Modality*

Imaging characteristics of oesophageal and junctional tumours on each modality (CT, EUS, PET/CT and MRI) are described in further detail below.

#### *1.8.2.2 CT*

OC can appear as eccentric or circumferential wall thickening on CT, with abnormal density and enhancement compared to adjacent normal tissue. For preparation, the patient drinks approximately 500 ml of water prior to CT to distend the stomach. However, this can be difficult for patients with dysphagia, which is a limitation of the technique. CT can demonstrate invasion into adjacent organs, an important factor for determining resectable disease, but cannot confidently differentiate early T-stages. (Takashima et al. 1991) Circumferential thickening on CT suggests T3 disease or greater. (Li et al. 2013)

#### *1.8.2.3 EUS*

EUS is considered the best modality for T-stage assessment because the individual layers of the oesophageal wall are well visualised. EUS provides excellent assessment of superficial tumours and is useful for distinguishing early T1 tumours from more advanced disease. A meta-analysis demonstrated sensitivity and specificity for T1a tumours as 85% and 87%, respectively and for T1b tumours, 86% and 86%, respectively. (Thosani et al. 2012)

EUS T-stage accuracy is approximately 70%. (Choi et al. 2010; O'Farrell et al. 2013) EUS has limited capability for Siewert type III GOJ tumours. Type III GOJ tumours are more technically challenging than type II due to anatomical constraints. T-stage agreement (defined by the weighted kappa statistic) for Siewert type II and type III tumours is 0.851 and 0.173, respectively. (Blackshaw et al. 2008)

#### *1.8.2.4 PET/CT*

Currently, the inherent poor spatial resolution of PET (approximately 5mm) limits T-stage assessment. (Levin 2012) However, PET provides useful data describing the metabolic activity of the tumour. Adenocarcinoma and SCC both have a high affinity for FDG, making PET particularly useful in OC. Quantification of FDG-uptake and metabolic tumour dimensions have prognostic significance. This is discussed further in section 1.9.3.

#### *1.8.2.5 MRI*

The staging accuracy of MRI has been investigated in research studies. MRI may be a useful, non-invasive technique, particularly in stenotic tumours that do not allow passage of the endoscope. MRI is technically difficult and prone to motion artefact given the proximity of the oesophagus to the heart, lungs and diaphragm.

Oesophageal tumours tend to be intermediate signal on T2 weighted images, with thickening and alteration of the anatomical layers. Mucinous tumours may be high signal due to their proteinaceous content.

Riddell et al investigated MRI T-staging with high-resolution 1.5 Tesla T2 sequences. (Riddell et al. 2007) Eighty-one percent of patients (28/37) were correctly T-staged when compared to the histopathological stage. Under-staging and over-staging were 16.2% (n=6) and 8.1% (n=3), respectively. A small number of studies have reported comparable MRI T-staging accuracy with EUS (EUS accuracy 82-94%). The ability of MRI to determine resectability is comparable to CT. (Takashima et al. 1991; van Rossum et al. 2013) T-staging with endoscopic MRI has been investigated but remains an invasive procedure with risk of complications. (Inui et al. 1995; Kulling et al. 1998; Wu et al. 2003; Dave et al. 2004) Wu et al found the accuracy of 1.5 Tesla endoscopic MRI was 60%. (Wu et al. 2003) The sensitivity and specificity of differentiating T1 & T2 from T3 & T4 tumours were 40% and 63%, respectively.

Higher strength MRI scanners have demonstrated excellent T-staging accuracy using ex-vivo specimens. Accuracy of a 4.7 Tesla MRI scanner was 94% compared to the resection specimens and was replicated with high resolution 1.5 Tesla MRI. (Yamada et al. 1997; Yamada et al. 2001) One hundred percent accuracy has been obtained using a 7 Tesla MRI in oesophageal specimens. (Yamada et al. 2014) Further in-vivo validation is required.

### 1.8.3 N-stage

Normal lymph nodes are flat or triangular, with preservation of the fatty hilum on imaging. (Richards et al. 2000) Size criteria are applied by measuring the short axis diameter of the lymph node. In general, lymph nodes measuring more than 10 mm are considered pathological. Radiological N-staging (TNM 7$^{th}$ edition) follows the same classification as pN-stage (Table 1.7.1). (Sobin et al. 2009)

#### *1.8.3.1 Lymph Node Metastases by Imaging Modality*

The imaging characteristics of LNMs and diagnostic ability of each modality are described below. Each modality has limitations for N-staging. CT provides anatomical information only, relies on size criteria and involves radiation. PET/CT also involves radiation but provides additional metabolic data which improves the positive predictive value (PPV) of LNMs. (Okada et al. 2009) The differentiation of peri-tumoural LNMs from adjacent avid tumour can be challenging due to the limited spatial resolution of PET. (Kapoor et al. 2004) This may increase 'false-negative' rates therefore under-staging the extent of nodal disease. EUS has better sensitivity compared to CT and PET/CT due to its superior contrast resolution.

##### *1.8.3.1.1 CT*

The main criterion for diagnosis of LNMs on CT is size. Peri-oesophageal lymph nodes are considered metastatic if the short axis diameter measures more than 10

mm, although it is recognised that the sensitivity of identifying LNMs on CT is suboptimal (as low as 18%). (Choi et al. 2000) CT cannot differentiate between benign and metastatic lymph nodes that are normal size, which may explain such poor sensitivity.

There is significant variation in results of CT N-staging. Benign lymph nodes adjacent to the trachea and main bronchi can appear large and round. One study found the sensitivity, specificity, accuracy, PPV and negative predictive value (NPV) were 56.0%, 97.3%, 92.4%, 73.7% and 94.2% respectively, using definitions of 10 mm for para-tracheal lymph nodes and 7 mm for all others. (Okada et al. 2009)

*1.8.3.1.2 EUS*

Malignant lymph nodes appear round, hypoechoic, homogenous and enlarged with loss of the normal fatty hilum on EUS. EUS is regarded as the best investigation for N-staging. Accuracy of EUS N-staging is approximately 65-70% (Choi et al. 2010) but is operator dependent. Van Vliet et al found that endoscopists performing less than 50 EUS procedures per year had inferior results compared to those performing more than 50 per year. (van Vliet et al. 2006)

A meta-analysis found the sensitivity of CT, EUS and PET/CT for the detection of regional LNMs was 50%, 80% and 57%, respectively. (van Vliet et al. 2008) The specificity was 83%, 70% and 85%, respectively suggesting EUS is better than CT or

PET for correctly excluding regional LNMs. The study also found the sensitivity and specificity of diagnosing coeliac LNMs with EUS was 85 and 96%, respectively.

A main limitation of EUS are stenotic tumours that are non-traversable with the endoscope. It has been reported that around 30% of tumours are non-traversable, however a published failure rate over a nine-year period by an experienced EUS operator was 2.9%. (Morgan et al. 2008)

Some centres also perform ultrasound of the neck as part of the routine diagnostic work-up (Omloo et al. 2009) but Blom et al found no additional value as a routine investigation. (Blom et al. 2012) In Blom et al's study, 140 of 170 patients (82.4%) had no suspicion of cervical disease, although 84% had tumours at or below the diaphragm. Neck ultrasound is justified in patients with clinically palpable nodes, or SCC in the cervical, upper or mid oesophagus. (Griffith et al. 2000)

*1.8.3.1.3 PET/CT*

Lymph nodes are considered involved on PET if identified on the CT component and show FDG-uptake that is appreciably higher than background values. Early studies found sensitivity, specificity and accuracy of PET/CT was 93.9%, 92.1% and 92.4% compared to PET alone (81.71%, 87.3% and 86.15%, respectively). (Yuan et al. 2006)

There is conflicting evidence regarding sensitivity of PET/CT for N-staging. Kato et al (Kato et al. 2008) found less additional value of PET to CT than Yuan et al. (Yuan et al. 2006) Sensitivity of PET/CT and PET was 46.0% vs 32.9%, specificity was 95.1% vs 93.9%, and accuracy was 99.4% vs 98.9%, respectively.

Another meta-analysis including 6 studies and 245 patients demonstrated that the pooled sensitivity and specificity of PET/CT was 55% and 76%, respectively. (Shi et al. 2013) These results are similar to the meta-analysis by van Vliet et al. (sensitivity 55%, specificity 85%) (van Vliet et al. 2008)

*1.8.3.1.4 MRI*

Regional LNMs tend to be round with intermediate T2 signal. A defined size threshold of a LNM has not been agreed upon as research is on-going. Recent studies have shown sensitivity, specificity and accuracy of 38-62%, 68-85% and 64-77%, respectively. (Wu et al. 2003; Nishimura et al. 2006) These results are comparable to CT, EUS and PET/CT but require further validation. Promising results were obtained using ferumoxtran-10, an ultra-small intra-venous (IV) ultrasmall superparamagnetic iron oxide (USPIO) compound. The sensitivity, specificity and accuracy was 100%, 95% and 96%, respectively. (Nishimura et al. 2006) However, ferumoxtran-10 and other USPIO agents were discontinued due to safety concerns. (Atri et al. 2015)

### 1.8.4 M-stage

The purpose of the initial CT in the routine staging pathway is to identify distant metastases or irresectable local disease that would preclude radical treatment. PET/CT is superior to CT for the detection of distant metastases. The sensitivity and specificity of CT and PET/CT is 52% and 91%, and 71% and 93%, respectively. (van Vliet et al. 2008) PET/CT can change management in up to 38% of cases. (Gillies et al. 2011; Blencowe et al. 2013) PET/CT has been considered as the initial staging investigation in OC, but is expensive and resource heavy. (van Vliet et al. 2008)

## 1.9 PET/CT

The integration of PET with CT images provides anatomical and metabolic data. The first clinical PET/CT scanner was installed at the University of Pittsburgh Medical Centre in 1998.

### 1.9.1 Physics of PET/CT

The most commonly used positron emitter used is $^{18}F$, which is labelled with FDG. $^{18}F$ has a half-life ($T_{1/2}$) of 110 minutes. A positron is released from $^{18}F$ and travels approximately 2 mm within the body before colliding with an electron. Two photons of equal energy (511 keV) are produced during this interaction, a process called annihilation. The photons then travel at 180° to each other and enter opposite detectors within the PET camera. (Allisy-Roberts and Williams 2008)

The PET camera is a ring of thousands of scintillation detectors surrounding the patient. The detectors are commonly made of bismuth germinate or lutetium oxyorthosilicate crystals, and are arranged into blocks, coupled to 4 photo-multiplier tubes (PMTs). These amplify the signal measured from each photon count. There are typically 20-30 rings of detectors around the body. (Allisy-Roberts and Williams 2008)

The ideal PET detector is a crystal material that is highly efficient at detecting the maximum number of photons, with a rapid processing time to enable further subsequent detections. A lead collimator ensures that a photon is channelled into a detector. There must be a balance between collimator size and photon count signal so diagnostic images are obtained. This compromise reduces spatial resolution.

### 1.9.2 PET/CT Image Production

The PET scanner analyses the annihilation coincidence detection (ACD) and assumes the interaction must have occurred along a path between detectors, called a line of response (LOR). Time of flight (TOF) software localises the annihilation along the LOR by measuring the time difference between photon detections, which can be as small as 500 picoseconds (ps).

The photon count and location is then processed by the systems computer, which builds a matrix (a grid consisting of pixels) using the co-ordinates of each photon

count in its memory locations. A PET image is constructed. (Allisy-Roberts and Williams 2008)

The PET and CT images must be co-registered. (Fig. 1.9.1) This is vitally important when assessing the anatomical location of increased FDG-uptake. Spatial matching is improved using an integrated PET/CT scanner so the patient does not move between scans. (Allisy-Roberts and Williams 2008) SUVs are calculated following a process called attenuation correction. This process adjusts the SUV depending on the CT tissue density.

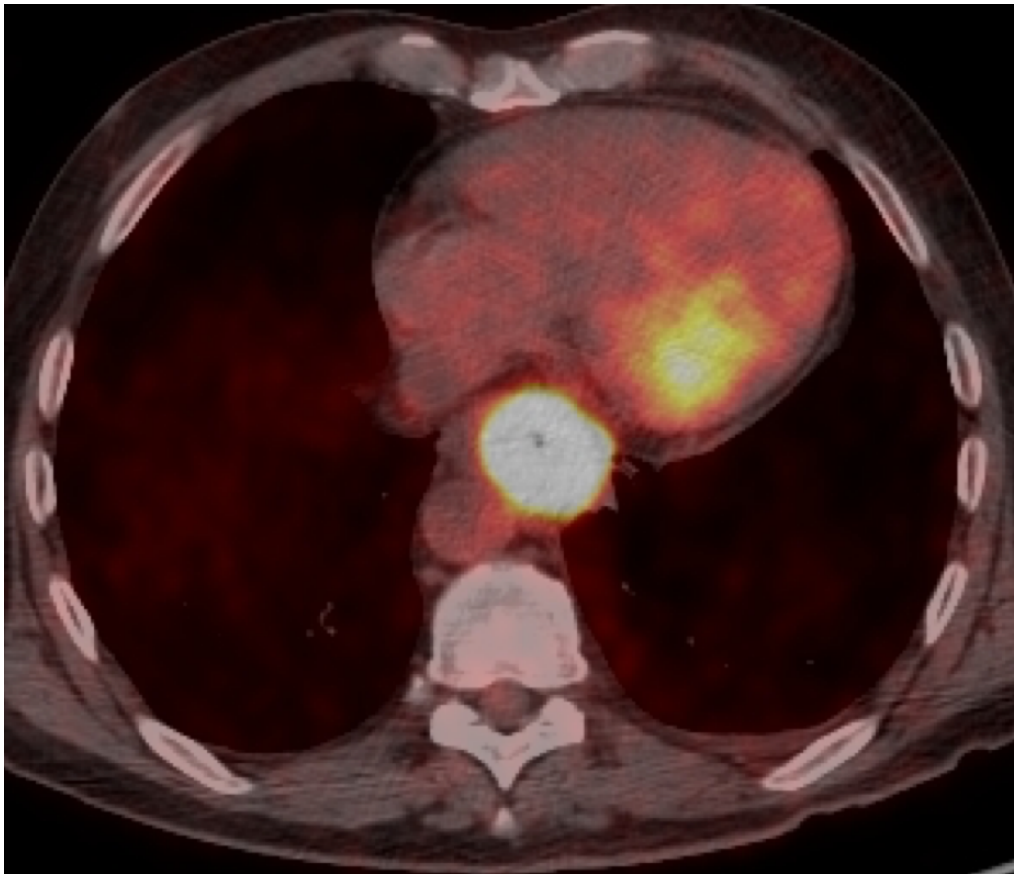


Figure 1.9.1. Example of a fused PET/CT axial image showing a distal oesophageal adenocarcinoma. PET images have been co-registered with the CT images. The oesophagus is circumferentially thickened at the level of the ventricles and the primary tumour is intensely FDG-avid. The image also shows FDG-uptake in the myocardium.

### 1.9.3 PET/CT Variables in Oesophageal Cancer Staging

PET/CT research in OC has increased considerably over the past decade. (Foley et al. 2016) Additional prognostic and predictive data can be recorded, which have the potential to guide management decisions, predict treatment response and improve risk stratification. (Schmidt et al. 2015)

A commonly used variable is $SUV_{max}$. This represents the voxel with the highest uptake value within a region of interest (ROI). $SUV_{max}$ is prone to random noise artefact and can provide falsely high values. (Kinahan and Fletcher 2010) $SUV_{max}$ has been extensively investigated as a prognostic indicator, although positive findings are not always reproduced. (Pan et al. 2009; Hatt et al. 2011b; Omloo et al. 2011) The mean standardised uptake value of the primary tumour ($SUV_{mean}$) is an alternative to SUVmax. (Higgins et al. 2012) SUVs can vary depending on the histological cell type of the primary tumour. SCC tends to have a higher $SUV_{max}$ than adenocarcinoma. (de Geus-Oei et al. 2007)

Metabolic tumour volume (MTV) represents the volume of metabolically active tumour. MTV has prognostic significance because larger tumours tend to be more advanced at presentation. (Li et al. 2014) A small study demonstrated an association between MTV and the number of LNMs. (I et al. 2012) Other studies have shown that $SUV_{max}$ and MTV correlate with T- and N-stage in patients with SCC. (Zhu et al. 2012) Importantly, MTV is highly dependent on the segmentation method used. (Gillies et al. 2016) This is discussed in more detail in sections 1.10.2.1 and 1.10.2.2.

Metabolic tumour length (MTL), defined as the maximum length of FDG-avid tumour, and length of disease (LoD), defined as the cranio-caudal length of primary tumour plus any regional LNMs, have been investigated for prognostic significance but further research is required to validate their use in risk-stratification and staging models. (Hatt et al. 2011b) PET and EUS LoD have implications for treatment planning (section 1.6.2.4).

Prediction of treatment response using PET/CT is currently of great interest. The MUNICON-1 trial investigated PET uptake pre- and post-NACT and found an $SUV_{max}$ reduction of 35% after 2 weeks of chemotherapy predicted response at the end of the 12-week cycle. (Javeri et al. 2009) The trial showed that discontinuation of chemotherapy in metabolic non-responders did not affect prognosis and prevented further exposure to treatment with potential side-effects. The MUNICON-2 trial investigated PET-guided management, adding radiotherapy to conventional NACT in non-responders. (zum Buschenfelde et al. 2011) PET-responders had longer 2-year survival than non-responders (71% vs 42%) but this difference did not reach statistical significance (p=0.10).

## 1.10 Introduction to texture analysis

Texture analysis is an advanced imaging technique that provides non-invasive, high-output feature extraction from a ROI. It is commonly referred to as 'Radiomics'. (Lambin et al. 2012) Texture analysis can be performed on CT, MRI and PET images. It is a post-processing technique that requires complex mathematical and statistical computation. Texture analysis is an evolving field of research that aims to improve prediction of patient outcome and treatment response. (Goh et al. 2011)

The potential value of texture analysis is substantial given that it can be performed retrospectively or prospectively from routinely-collected images, at little extra cost. There is no need for additional radiological investigations. Texture analysis in combination with traditional staging methods may improve radiological staging, clinical-decision tools and treatment pathways. (Aerts et al. 2014)

### 1.10.1 Background

Multiple sub-clonal populations of cells are known to co-exist within solid cancers, establishing the concept of tumour heterogeneity. (Gerlinger et al. 2012) Data from texture analysis could be used as surrogate markers of underlying tumour heterogeneity and may correlate with biological, genomic or tumour marker expression. (Chong et al. 2014) Variation in FDG-uptake is associated with underlying pathophysiological features such as cellular proliferation, vascularity, perfusion, hypoxia and necrosis. (Rajendran et al. 2006; Henriksson et al. 2007; Basu

et al. 2011; Humbert et al. 2016) Henriksson et al studied FDG-PET images of head and neck cancers in mice. There was good correlation between SUV and the number of tumour cells in different quadrants of the tumour. (Henriksson et al. 2007)

An image is composed of a matrix comprising discrete units called pixels. A pixel is a two-dimensional area that displays a corresponding representative value of density (CT), signal (MRI) or uptake (PET). A medical image is formed from multiple pixels spatially arranged in a matrix.

A voxel is a 3-dimensional (3D) pixel that represents a volume of space. Modern cross-sectional imaging techniques acquire scan data with isotropic voxels (those with equal dimensions) allowing 3D texture analysis to be calculated. (Aerts et al. 2014) Feature extraction algorithms provide first-, second- and higher-order statistics that quantify the spatial distribution and intensity values within selected voxels considered to represent the tumour volume. (Lambin et al. 2012; Gillies et al. 2016) 3D texture is more representative of tumour heterogeneity than 2D analysis. (Ng et al. 2013)

The concept of texture analysis originated outside of the medical field. A seminal article by Haralick et al described a series of texture features obtained from geological satellite images. These features could differentiate between forest, swamp, marsh, coast and urban areas. (Haralick et al. 1973) Visual texture perception is an area of research pioneered by a psychologist called Bela Julesz. His work investigated the human performance of second-order texture metric discrimination. He demonstrated

that computational methods of digital texture analysis are superior to human visual perception. (Julesz 1975)

### 1.10.2 Texture Analysis in Clinical Research

Texture analysis has been investigated in several solid tumours including oesophageal, nasopharyngeal, lung, cervical cancer. (Kidd and Grigsby 2008; Huang et al. 2012; Hatt et al. 2013; Win et al. 2013; Yip et al. 2014) Use of the image features as prognostic and predictive biomarkers is promising but its role in management decision pathways has not been validated. The texture metrics used in this thesis are discussed in more detail in section 2.4.5.

#### *1.10.2.1 Segmentation*

One of the most importance aspects of texture analysis is segmentation, which defines the contour or outline of the tumour. (Gillies et al. 2016) Manual, semi-automatic and automatic methods are available. Segmentation is critical in texture analysis, as this defines the voxels analysed in the selected ROI. Results are affected if adjacent tissue is included within the contour, for example, normal FDG-uptake in the myocardium.

Manual segmentation is very time-consuming, especially in a busy clinical context. It is operator dependent and has the potential for high inter-observer variability. (Ford et al. 2006; Egger et al. 2013) Automatic segmentation is an alternative to manual

delineation in PET imaging. Early automatic segmentation techniques use a fixed threshold, which rely on a pre-defined level set by the operator. The software identifies the voxel with the highest SUV then includes adjacent voxels within that threshold to be included in the segmentation.

Several limitations exist with this technique. Tumours with discontinuous uptake may not fully segment and would require manual adjustment. In addition, optimal threshold levels have not been agreed. Studies using fixed SUV thresholds of 2.5 or thresholds of 20%, 30%, 40%, 42% and 50% have been conducted, presenting differing evidence. (Biehl et al. 2006; Prieto et al. 2012)

*1.10.2.2 Automatic Segmentation Methods*

Several automatic thresholding methods have been proposed. (Hatt et al. 2017) These include adaptive thresholding (AT), gradient contour (GC), region growing (RG), watershed transform (WT) and clustering methods.

AT is an iterative method, developed by Drever et al, based initially on a fixed threshold method of 50%. (Drever et al. 2007) However, the contour is modified iteratively following comparison of the original and generated backgrounds. Segmentation terminates once there is no successive difference between backgrounds.

The GC algorithm assesses PET images slice-by-slice for the highest SUV. The next successive highest SUV is determined following systematic evaluation of voxels in clockwise direction. (Ford et al. 2006) The WT method identifies 'crests' of high intensity and the algorithm includes these in the defined ROI. (Geets et al. 2007)

The RG method is an iterative process that randomly selects voxels and propagates a contour from that voxel, depending on the intensity of the voxel adjacent to it. The contour terminates once the proportion of added voxels becomes small compared to the total volume, often set at 5%. (Day et al. 2009).

Clustering methods implement iterative steps to group clusters of voxels after comparison of their uptake values. The algorithm repeatedly tests first-order statistics to produce homogenous volumes within the cluster. Named methods of clustering include K-means (Zaidi and El Naqa 2010), Fuzzy C-means (Belhassen and Zaidi 2010), and Gaussian Fuzzy C-means (Hatt et al. 2011a).

An advanced decision tree-learning algorithm for automatic segmentation (ATLAAS) has been developed using machine-learning methods in attempt to standardise PET segmentation. (Berthon et al. 2016) Further details about ATLAAS are discussed in section 2.4.3.

*1.10.2.3 Texture Analysis in Oesophageal Cancer*

Relatively few studies have investigated PET or CT texture analysis in OC. These have been small, retrospective studies investigating associations with stage, treatment response and survival. (van Rossum et al. 2016b)

PET texture analysis evaluates metabolic heterogeneity of the primary tumour. Hatt et al found that entropy, homogeneity, dissimilarity and zone percentage were preferred variables for predicting response in 50 patients with OC. These variables were robust for each segmentation method used. (Hatt et al. 2013)

Dong et al investigated the association of $SUV_{max}$ with histological grade, T-staging and 2 texture features (entropy and energy) in 40 pre-operative patients with SCC. (Dong et al. 2013) T-stage correlated significantly with entropy ($r=0.693$, $p<0.001$) and energy ($r= -0.469$, $p=0.002$). An entropy of >4.699 distinguished stage IIb from more advanced tumours ($p<0.001$).

Tixier et al investigated baseline FDG-PET texture to predict tumour response in 41 patients treated with dCRT. (Tixier et al. 2011) Thirty-eight different texture metrics were calculated. Patients were grouped into non-responders (NR), partial responders (PR) and complete responders (CR), defined by the RECIST criteria. (Eisenhauer et al. 2009) $SUV_{peak}$ ($p=0.045$), Entropy ($p<0.001$) and Intensity Variability ($p<0.001$) were significantly associated with treatment response. The RECIST criteria is considered a poor tool to assess treatment response in OC and limits these findings.

Yip et al investigated the association of CT texture analysis, OS and tumour response in patients treated with dCRT. (Yip et al. 2014) Histogram statistics were calculated for fine, medium and coarse textures using a software package called TexRad (TexRad Ltd, Cambridge, England). Post-treatment medium Entropy ($p<0.001$), coarse Entropy ($p<0.001$) and medium Uniformity ($p<0.001$) were associated with OS.

In another study from this group, histogram statistics were calculated from pre- and post-NACT CT examinations. Primary tumours became more homogenous following NACT, demonstrating a significant decrease in Entropy ($p<0.001$) and an increase in Uniformity ($p<0.001$). Histogram Standard Deviation showed a statistically significant association with pathological tumour response ($p=0.009$) and Histogram Skewness was associated with OS ($p<0.001$). (Yip et al. 2015) Further details of these metrics are found in section 2.4.5.

## 1.11 Prognostic research

The Oxford Medical Dictionary defines prognosis as “an assessment of the future course and outcome of a patient’s disease, based on knowledge of the course of the disease in other patients, together with the general health, age, and sex of the patient”. (Oxford University Press 2010) Its use is not confined to patients that are acutely unwell or suffering from cancer, and can be used to predict future events in healthy individuals. (Moons et al. 2009b)

### 1.11.1 Background

Prognostic research is different to aetiological research. The aim is to identify predictors, not to discover causality. An example is the ability of tumour markers to predict death. (Moons et al. 2009b) Markers of prediction are different to markers of prognosis. Whereas a marker of prediction can be used identify a select group of patients in which a treatment is most likely to be successful, a prognostic marker informs the likely course of that individual. This difference is also translated when comparing predictive and prognostic models.

Prognostic modelling can inform the probability of an individual’s outcome, can influence public health policy and assess the role of diagnostic tests. The Framingham cardiovascular risk score is widely used in primary care to inform prophylactic anti-hypertensive and cholesterol reducing medication use. (Kannel et al. 1976) Other commonly used prognostic models are the Nottingham prognostic index

in breast cancer and the Manchester triage system in emergency departments. (Haybittle et al. 1982; Mackway-Jones 1997) Prognostic models can be used to help design RCTs (Hernandez et al. 2004) and to compare performance across departments or hospitals in order to improve standards. (Jarman et al. 1999)

In cancer research, prognostic models can be developed to predict patient outcome and assess the additional benefit of new biomarkers. (Mallett et al. 2010b; Moons et al. 2012) More recently, studies investigating different combinations of imaging, histopathological and genomic data have attempted to develop clinical prognostic and prediction models. (Hemingway et al. 2013; Aerts et al. 2014)

A trend towards personalised or 'precision' medicine has utilised prognostic and predictive models, attempting to tailor treatment to individual patients. Predictive models could identify groups of patients that may significantly benefit from a specific treatment, or have no response at all. (Steyerberg et al. 2013) Examples of this include Tamoxifen therapy in women with HER-2 positive breast cancer and Imatinib use in AML patients with BCR-ABL mutations. (Hingorani et al. 2013)

Despite the abundance of prognostic research in the literature, its methodology has not been fully established and should be improved. (Riley et al. 2013) The importance of developing and validating prognostic models has been emphasised in several articles published by eminent statisticians. (Altman 2001; Altman et al. 2009; Moons et al. 2009a; Moons et al. 2009b; Royston et al. 2009; Mallett et al. 2010a; Moons et al. 2012; Hemingway et al. 2013; Hingorani et al. 2013; Riley et al. 2013;

Steyerberg et al. 2013) Validating a prognostic model is important and can be performed either in the same institution (internal) or in a different centre (external). Further details regarding model validation are discussed in section 2.6.2.1. Due to deficiencies in prognostic research methodology, it is likely that perceived important factors with unestablished significance continue to be investigated, populating the literature with potentially false-positive studies. (Altman 2001; Simon 2001; Ioannidis 2005)

Prognostic research should use evidence-based, standardised statistical methodology and researchers should collaborate internationally using electronic health records, standardised nomenclature and prospectively planned data collection tools, such as the REMARK criteria. (McShane et al. 2005) This minimises heterogeneity between studies allowing comprehensive and informative meta-analyses to be performed. (Hemingway et al. 2013) Providing the model is tested in developmental, validation and impact studies, it can be used by clinicians to guide treatment decisions. (Moons et al. 2009b; Steyerberg 2009)

## 1.12 Rationale, Aims and Hypotheses of Thesis

Accurate staging of OC is vital. However, each staging investigation has limitations which can affect its diagnostic accuracy. Using quantitative imaging techniques, it is possible to obtain additional prognostic variables from radiological staging investigations of OC.

The primary working hypothesis of this thesis was that additional variables could be quantified from radiological investigations that improve prediction of patient outcome compared to current staging methods.

**The primary aim of this thesis was therefore to assess the additional value of novel prognostic variables over and above the current staging system.**

These additional variables could improve patient risk-stratification and may influence treatment decisions, thereby improving OS in those undergoing radical treatment. Potentially unnecessary and likely ineffective therapy could be avoided in patients that may traditionally have been treated radically. Alternatively, these patients could be identified as potential participants in trials of novel chemotherapeutic agents.

The specific aims in this thesis were:

- **Chapter 3**. LNMs are a major prognostic indicator in OC. Therefore, is it vital that N-staging is accurate to inform treatment decisions and avoid over- or under-treatment. However, direct imaging of lymph nodes is challenging and largely evaluated by anatomical criteria. The concept of micro-metastases is now established. Therefore, the hypothesis of this chapter was that N-staging is suboptimal and that the incidence of LNMs is higher than expected.

  **The aim of this chapter was to evaluate the accuracy of OC N-staging and provide radiological-pathological correlation of LNMs.**

- **Chapter 4**. Results of a published study from our centre, which investigated the role of EUS in patients staged N0 on PET/CT, was internally validated. (Foley et al. 2014b) Validation is an important step in prognostic model development, however, this is often not performed. The hypothesis was that the model demonstrates continued benefit of EUS N-stage in patients staged N0 on PET/CT.

  **The primary aim was to internally validate the prognostic model. The secondary aim was to evaluate the prognostic significance of pathological LNMs.**

- **Chapter 5**. CRM involvement is now widely regarded as an important prognostic factor in patients undergoing surgical resection. As described in sections 1.7.4.3, OS significantly reduces following an R1 resection. It is hypothesised that larger, more FDG-avid tumours (defined by MTL, MTW and $SUV_{max}$) may have higher rates of CRM involvement. Better prediction of pathological CRM involvement would assist clinical decision-making and surgical planning, and may improve recurrence rates and OS.

  **The primary aim of the study was to investigate the additional value of PET-defined tumour variables for predicting pathological CRM involvement compared with CT and EUS.**

- **Chapter 6**. Treatment decision-making and planning are guided by radiological measurement of LoD in patients with OC. Discrepancies between PET and EUS LoD could have significant impact on deciding which patients are suitable for treatment, and which treatments are feasible.

  **The primary aim of the study was to investigate differences in PET and EUS LoD. The secondary aim was to investigate their prognostic significance.**

- **Chapter 7**. This study evaluated the additional value of PET texture analysis over the current OC staging system, by developing and internally validating a prognostic model, using appropriate statistical methodology. The model was tested against current staging methods. The main hypothesis was that texture analysis enhances staging by providing additional prognostic value compared to current methods.

  **The primary aim was to develop and internally validate a prognostic model incorporating PET texture analysis, which demonstrates additional prognostic performance compared to current staging methods.**

Chapter 8 provides a general discussion of the work undertaken in this thesis including results obtained and their implications for clinical practice, the limitations and strengths of the work, and suggestions for future research.

# Chapter 2. Materials & Methods

This chapter discusses common materials and methods used in this thesis. Specific patient selection criteria and methods unique to each results chapter are listed and discussed separately.

## 2.1 Clinical Database

A clinical database containing patient demographics, radiological, surgical, histopathological and outcome data was collated for this thesis. The clinical database was maintained in a Statistical Package for the Social Sciences (SPSS) file (IBM, Chicago, USA) and was manually populated with data obtained from the Cancer Network Information System Cymru (CaNISC) database (Velindre NHS Trust, Wales).

### 2.1.1 Patients

Patients diagnosed with potentially curable OC from September 2010 onwards were identified at the Regional South-East Wales Upper GI cancer MDT and added to the database. Patients were referred from four local health boards; Aneurin Bevan, Cwm Taf, Hywel Dda and Cardiff and Vale University Health Boards, serving a combined population of 1.4 million. All relevant staging investigations were discussed at the MDT, including endoscopic, radiological and pathological findings.

### 2.1.2 Clinical Management

A management plan was agreed by consensus following discussion at the MDT. Treatment was selected based on radiological staging, patient choice, and relevant co-morbidities, following algorithms used by the Regional Upper GI cancer network. (Crosby et al. 2004; Stephens et al. 2006; Gwynne et al. 2011) The agreed management plan was recorded in the CaNISC database. The management of OC is discussed in section 1.6.

All surgery was performed by specialist upper GI surgeons in a centralised tertiary referral unit. Trans-hiatal surgery was selected for patients with tumours of the distal oesophagus, in whom it was considered that a thoracotomy may carry an unacceptable risk of respiratory complications due to poor performance status. In general, fit patients with tumours pre-operatively staged as T3/T4a, N0/N1 were pre-operatively treated with NACT or NACRT. Less fit patients, or those with T1/2 N0 disease, had surgery alone. Patients deemed unsuitable for surgery due to co-morbidity and/or performance status, extensive loco-regional disease, or personal choice received dCRT.

Patients that received NACT had two cycles of 80mg/$m^2$ cisplatin and 1000 mg/$m^2$ of 5-fluorouracil (5-FU) for 4 days. A minority received four cycles of epirubicin (50 mg/$m^2$), cisplatin (60 mg/$m^2$) and 5-FU (200 mg/$m^2$). Patients treated with NACRT typically receive 2 cycles of oxaliplatin (130 mg/$m^2$) and capecitabine (625mg/$m^2$) as

induction NACT, followed by 45 Gy of radiotherapy administered in 25 fractions over a 5-week period with concurrent chemotherapy.

Radiotherapy planning was performed following direct comparison of all available imaging modalities and after consideration of the maximum LoD recorded. Occasionally, non-deformable fusion of the PET and planning CT was performed, provided the diagnostic PET/CT had been acquired in the radiotherapy planning position to allow accurate fusion.

### 2.1.3 Outcome Data

Outcome data were also recorded in the clinical database. This included dates of diagnosis and death, from which survival from diagnosis (recorded in months) is calculated. Dates of death were searched for using the CaNISC database, which is populated with data from the Office for National Statistics. Survival data used in this thesis was updated in July 2016, at the time of statistical analysis.

TRG was recorded in the clinical database in patients undergoing surgery following neo-adjuvant therapy. A TRG (Mandard et al. 1994) was classified by the reporting Consultant Pathologist. The Mandard Score is described in section 1.7.4.5.

Patients were followed-up every 3 months for the first year, then every 6 months thereafter. All patients were followed up for 5 years, or until death. No patients were lost to follow-up, unless they re-located outside of area.

### 2.1.4 Institutional Review Board and Ethics Approval

Institutional scientific review and ethical approval were granted for all work included in this thesis. Institutional approval for chapters 3-6 is included in Appendix B. Institutional and ethical approval for chapter 7 is included in Appendices C & D.

## 2.2 Radiological Data

Radiological staging was performed according to UICC TNM 7$^{th}$ edition. (Sobin et al. 2009) Reports of staging investigations are available on PACS (IMPAX 6.4.0.6010, AGFA Healthcare, Belgium, 2015) and CaNISC database.

### 2.2.1 Computed Tomography

All patients with potentially curable OC had a CT examination. CT in Cardiff and Vale University Health Board (CAVUHB) was performed with a GE HD 750 Discovery 64-slice scanner (GE Healthcare, Buckinghamshire, UK). CT images were acquired with helical acquisition, collimation of 40 mm, pitch 0.984:1 and tube rotation speed of 0.4 s. Tube output was 120 kVp with smart mA dose modulation between 60-600 mA. Slice thickness was 0.625 mm with 3 mm reconstructions using soft tissue and lung algorithms. Approximately 500 ml of water was given orally. Between 100-150 mls of Niopam 300 was given intravenously. CT examinations from external institutions were all reviewed at the Regional Upper GI cancer MDT and deemed an acceptable standard. Lymph nodes were considered involved on CT if located in the expected

distribution of disease, were round with loss of fatty hilum, demonstrated altered density or enhancement and the short axis diameter measurement was 10 mm or greater. (Botet et al. 1991)

### 2.2.2 Endoscopic Ultrasound

During the thesis period, EUS in South Wales was performed by 4 experienced endosonographers. Data from all 4 endosonographers have been used in this thesis to allow generalisability of results, except in Chapter 6, where only EUS measurements recorded by Dr Ashley Roberts were used. Data from a single operator was used in this chapter to eliminate inter-observer variability in measurements. The accuracy of Dr Roberts' EUS staging has previously been published. (Weaver et al. 2004)

An initial endoscopic examination was performed using a 9 mm diameter Olympus P-10 gastroscope (Key Med, Southend, UK) to assess the degree of oesophageal luminal stenosis. Patients with an estimated oesophageal luminal diameter less than 15 mm underwent examination using the smaller-diameter MH-908 oesophagoprobe (Key Med). Oesophageal dilation (Savary-Gilliard, Cook Medical, Bloomington, USA) was performed before endosonography for patients with oesophageal lumens < 9 mm. The type of echoendoscope used was at the discretion of the endoscopist. No significant difference in accuracy existed between the 2 echoendoscopes. (Twine et al. 2009a) The primary oesophageal tumour was assessed, together with an evaluation of the para-oesophageal anatomical structures. The criteria for malignant

lymphadenopathy specifies a hypo-echoic pattern, a spherical contour, the presence of a distinct border, and a short axis diameter of 6 mm or more. (Grimm et al. 1993; Weaver et al. 2004; Gleeson et al. 2009)

Patients with tumours too stenotic to be crossed at EUS were unable to be fully staged therefore final pre-treatment radiological stage relied on a combination of findings from PET/CT and CT investigations.

### 2.2.3 Positron Emission Tomography/Computed Tomography

All patients treated with radical intent underwent PET/CT. The Positron Emission Tomography Imaging Centre (PETIC) in Cardiff University opened in August 2010. To ensure that no patient was missed, a list of patients that underwent PET/CT was generated from PETIC, and cross-checked against the clinical database. This patient list was password-protected, temporarily stored on a Cardiff University computer and deleted following cross-checking. All patients in chapters 3, 5, 6 & 7 had PET/CT in PETIC, using the same scanner and protocol. In chapter 4, 47 patients had PET/CT at the Cobalt Imaging Centre, Cheltenham prior to the opening of PETIC. These patients were included in the original cohort, in which the prognostic model was developed.

Patients were fasted for at least 6 hours prior to tracer administration. Serum glucose levels were routinely checked and confirmed to be less than 7.0 mmol/L prior to imaging. Patients received an activity of 4 MBq of $^{18}$F-FDG/kg of body weight.

Activity uptake time was 90 minutes. $^{18}$F-FDG PET/CT imaging was performed with a GE 690 PET/CT scanner (GE Healthcare, Buckinghamshire, UK). PET images were acquired at 3 minutes per field of view. The length of the axial field of view was 15.7 cm. Images were reconstructed with the ordered subset expectation maximisation algorithm, with 24 subsets and 2 iterations. Matrix size was 256 x 256 pixels, using the VUE Point™ time of flight algorithm. CT images were acquired in a helical acquisition with a pitch of 0.98 and a tube rotation speed of 0.5 s. Tube output was 120 kVp with output modulation between 20 and 200 mA. Matrix size for the CT acquisition was 512 x 512 pixels with a 50 cm field of view. No oral or intravenous contrast was administered.

All PET/CT examinations were reported by experienced consultant radiologists with an interest in nuclear medicine, who have reported at least 600 examinations each in order to become certified. In general, PET/CT T-stage was assigned TX given the poor spatial resolution of PET. Nodes were classed as involved if identified on the CT component and showed FDG-uptake appreciably higher than background values. No specific threshold of FDG uptake was used to define positivity.

## 2.3 PET-Defined Tumour Variables

PET-defined tumour variables were measured and recorded using consistent methodology on a GE advantage windows 4.5 reporting workstation (GE Healthcare, Buckinghamshire, UK). (Fig. 2.3.1) MTL was measured using maximum intensity projection (MIP) images, which were rotated to visualise the greatest length. MTL was recorded in centimetres (cm). LoD was measured using the same method, also recorded in cm. Metabolic tumour width (MTW) was recorded by measuring the greatest perceived width of the tumour, perpendicular to the MTL, also recorded in cm. $SUV_{max}$ is obtained by placing a ROI over the primary tumour. The value is automatically returned by the software. This is a common method used in clinical reporting and has therefore been employed here. Identical settings were used to ensure consistent results; the field of view (FOV) was 88.1 cm and the SUV of the MIP display was maintained at 12 g/ml for each case.

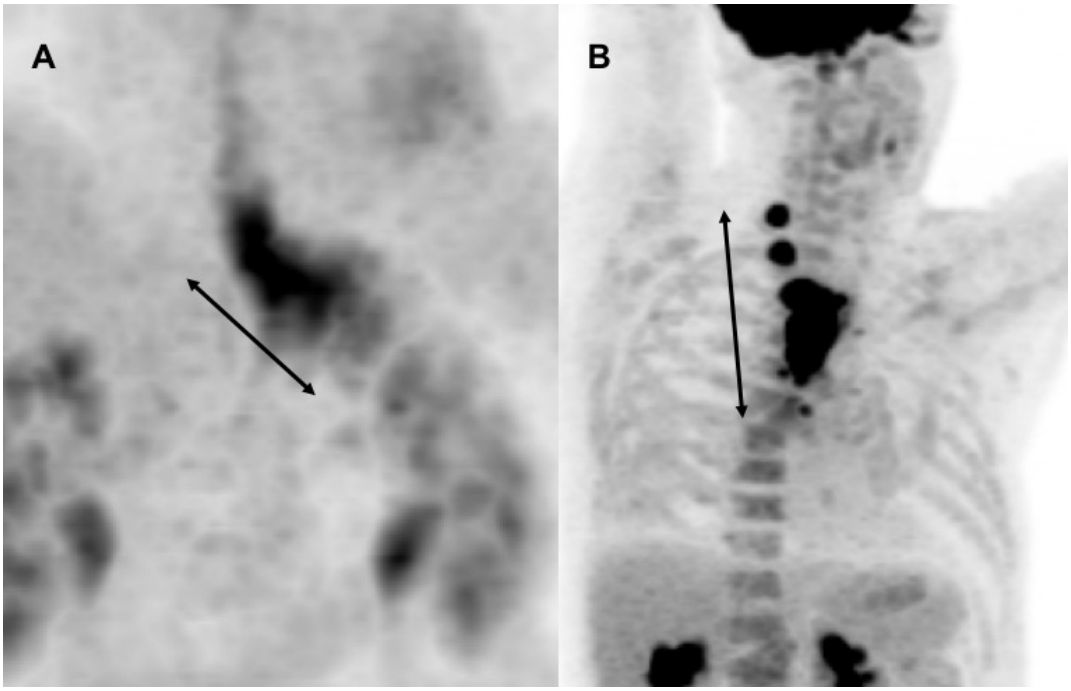



Figure 2.3.1. PET maximum intensity projection (MIP) images indicating measurements of MTL (A) and LoD (B). These selected images were obtained using a GE advantage windows 4.5 reporting workstation (GE Healthcare, Buckinghamshire, UK) and orientated to adequately represent the maximum dimension of MTL and LoD. Image A also demonstrates an 'avidity gradient' that can be seen in GOJ tumours involving the proximal stomach. FDG-avid LNMs seen in image B are included in the measurement of LoD.

## 2.4 Texture Analysis

Materials and methods of texture analysis used in this thesis are described below. More specific information regarding study design is included in Chapter 7.

### 2.4.1 PET-STAT Graphical User Interface

The PET-STAT graphical user interface (GUI) was developed at Cardiff University, between scientists based at PETIC, Velindre Cancer Centre and Cardiff School of Engineering. The development is part of a large series of investigations, funded by Cancer Research Wales (Grant No. 7061 and 2476). It was written and implemented in the MatLab-based (The Mathworks, Natick, USA) open-source software Computational Environment for Radiotherapy Research (CERR). (Deasy et al. 2003) PET-STAT allows DICOM files to be uploaded and converted into CERR files. PET-STAT also allows manual and automatic segmentation of CT and PET images. Several variables quantifying intensity values and texture features can be calculated from the segmented tumours and are listed in Section 2.4.5.

### 2.4.2 Data Preparation

Data preparation was performed by KF. PET/CT examinations were anonymised at source. Data were exported in DICOM format and stored in the restricted access, password-protected PETIC server. Both PET and CT series were imported into PET-STAT and converted in CERR files, then saved in MatLab file format.

Tumours must first be segmented to facilitate texture analysis. This requires creation of a bounding box using the PET-STAT GUI, indicating the region of the image to undergo segmentation. The bounding box does not need to follow a perfect contour of the tumour, as some background uptake should be included to facilitate segmentation (section 2.4.3). No pre-defined method of bounding box creation was used, and was performed on an individual case basis. Adjustment of window level and colour of displayed PET images was performed at the discretion of the user. No pre-defined window levels were used since these have no influence on the segmentation.

After a bounding box is created, a contour can be generated and is saved in the CERR planC file as an RTSTRUCT. An RTSTRUCT represents the contour of the tumour and is a binary mask of the image, representing the boundary of voxels to be included in analysis. A MatLab batch analysis process is implemented in PET-STAT which allows automatic analysis of multiple RTSTRUCTs in succession.

### 2.4.3 ATLAAS

ATLAAS is an automated Decision Tree-Based Learning Algorithm for Advanced Segmentation and is implemented in PET-STAT. (Berthon et al. 2016) All tumours have been segmented with ATLAAS in this thesis. (Fig. 2.4.1)

ATLAAS was designed to select the most accurate PET automatic segmentation (PET-AS) method for optimal segmentation in individual cases. ATLAAS accurately selected the best PET-AS method in 77% of phantom cases and 89% of simulated

data which led to a high average accuracy of 0.83 using the Dice Similarity Coefficient (DSC). (Dice 1945)

ATLAAS selects the best segmentation method from a list of PET-AS methods incorporated in its system. Several automatic segmentation methods are implemented and compared in ATLAAS, including AT, RG, K-means (KM) with 2, 3 & 4 clusters, fuzzy C-means (FCM), Gaussian C-means (GCM) with 3 & 4 clusters and WT. These PET-AS methods have been described in detail in previous publications. (Berthon et al. 2013; Berthon et al. 2014)

Three parameters are identified from the target object and evaluated by ATLAAS in order to select the best PET-AS method. The MTV of the target object is calculated and measured in millilitres (mL). Peak target background ratio ($TBR_{peak}$) is the ratio of target object $SUV_{peak}$ and the background SUV. $SUV_{peak}$ is defined as the mean uptake value in a 1 $cm^3$ sphere centred on the maximum SUV of the target object. The background SUV is the mean intensity in a 0.5 cm thick extension of the object contour and is the reason some background uptake should be included in the bounding box. The number of intensity levels (NI) in the target object is calculated by grey-level run length, which is obtained from methods described by Loh et al (Loh et al. 1988) and represents the number of intensity values following image re-sampling.

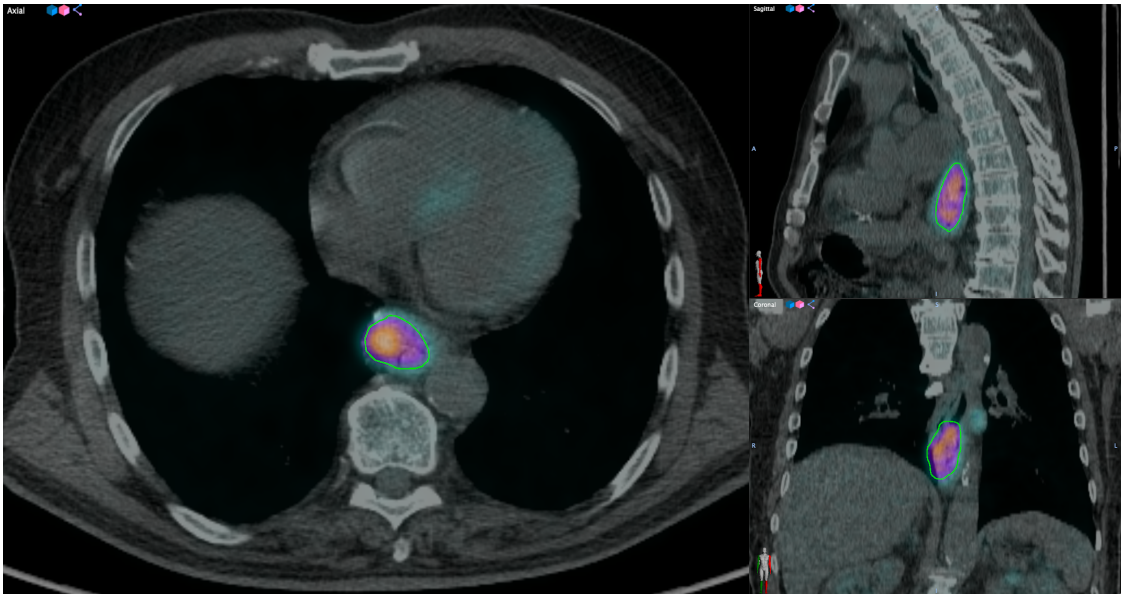



Figure 2.4.1. An example of axial, sagittal and coronal fused PET/CT images showing a FDG-avid mid-oesophageal tumour segmented with ATLAAS (green contour). PET images are co-registered with the corresponding CT images and the ATLAAS contour of the FDG-avid primary tumour is displayed.

### 2.4.4 Discretisation Methods

Discretisation of PET images, or image re-sampling, allows quicker calculation of the texture metrics over a variety of tumour intensity ranges. Different methods have been published including discrete numbers of intensity values (e.g. 64 or 128) and fixed bin-widths (e.g. 0.5 SUV). (Leijenaar et al. 2015) Image re-sampling with 0.5 SUV bin

widths has been used in this thesis, as this creates equal-sized bin widths and provides standardised methodology.

### 2.4.5 PET-STAT Texture Variables

Hundreds of texture features are described in the literature. (Lambin et al. 2012) Sixteen PET variables and texture metrics were selected for texture analysis in this thesis following demonstration of their prognostic and predictive significance in research studies published prior to 2015, when PET-STAT was being developed. (Table 2.4.1)

Following analysis, texture data were automatically compiled in a Microsoft Excel spreadsheet (Microsoft Corporation, Washington, USA). These data were then manually imported to the clinical database for statistical analysis.

Table 2.4.1. Image Metrics Implemented in PET-STAT

| PET-STAT metric | Equation/Comments |
|---|---|
| *First-order metrics* | |
| SUVmax | Maximum intensity value |
| SUVmean | Mean intensity value |
| Metabolic Tumour Volume (MTV) | Metabolic Tumour Volume (mL) |
| Total Lesion Glycolysis (TLG) | MTV * SUVmean |
| *Histogram metrics* | |
| Standard Deviation | $\sqrt{\frac{1}{N} * \sum_i (I(i) - \mu)^2}$ |
| Skewness | $\frac{1}{N} \frac{\sum_i (I(i) - \mu)^3}{\left(\sqrt{\frac{1}{N} * \sum_i (I(i) - \mu)^2}\right)^3}$ |
| Kurtosis | $\frac{\frac{1}{N}\sum_i (I(i) - \mu)^4}{\left(\frac{1}{N}\sum_i (I(i) - \mu)^2\right)^2}$ |
| Histogram Energy | $\sum_i \left(P(i)\right)^2$ |
| Histogram Entropy | $\sum_i (-P(i) * \ln(P(i)))$ |

| | |
|---|---|
| *Neighbourhood grey-tone difference matrix* | |
| Coarseness | $\left[\epsilon + \sum_{i=0}^{G_h} p_i s(i)\right]^{-1}$ |
| *Grey level co-occurrence matrix* | |
| Homogeneity | $\sum_{r,c} \frac{GLCM(r,c)}{1+\lvert r-c\rvert}$ |
| Entropy | $\sum_{k} (-GLCM(k) * \ln(GLCM(k)))$ |
| Dissimilarity | $\sum_{r,c} (GLCM(r,c) * \lvert r-c\rvert)$ |
| *Grey level size-zone matrix* | |
| Intensity Variability | $\sum_{r}\sum_{c} \left(\mathrm{GL}SZM(r,c)\right)^2$ |
| Large area emphasis | $\sum_{r}\sum_{c} \left(c^2 * \mathrm{GL}SZM(r,c)\right)$ |
| Zone Percentage | $\sum_{r}\sum_{c} \left(c * \mathrm{GL}SZM(r,c)\right)$ |

$\mu$ = mean; I = 3D image matrix; N = number of voxels in image; P(i) = Probability of intensity i in I where $P(i) = \frac{N_i}{N}$, with $N_i$ the number of voxels of intensity I; $G_h$ = the highest grey-tone value present in the image; ϵ = a small number used to avoid zero-division; s(i) = neighbourhood grey-tone difference matrix (NGTDM) value of intensity *i* calculated as Σ *i-Ai*; GLCM = Grey level co-occurrence matrix; k = number of iterations; GLCM(r,c) = element of GLCM matrix in row r (intensity value) and column c (intensity value); GLSZM = grey-level size-zone matrix; GLSZM(r,c) = element of GLSZM matrix in row r (intensity value) and column c (zone size).

## 2.5 Pathological Data

The RCPath standardised reporting template is discussed in section 1.7.4. OC specimens were reported by a Consultant Histopathologist with an interest in Upper GI cancer. Fewer, but more specialist pathologists now report the cancer specimens within the centralised service. Methods of pathological reporting were consistent throughout the thesis.

Thin cross-sections of tumour were routinely prepared for examination and the maximum depth of invasion evaluated. All lymph nodes removed from the specimen were identified and assessed for metastases. pT- and pN-stage were assigned accordingly. Reports were available on the Cardiff and Vale UHB Clinical Portal information system (Clinical Portal, CAVUHB, NHS Wales).

## 2.6 Statistical Methodology

All statistical analyses in this thesis were performed using SPSS v23 (IBM, Chicago, USA). A p-value of <0.05 was considered statistically significant throughout. Specific statistical methods will be described in each results chapter, as they were selected to test each hypothesis depending on the nature and distribution of the data. A more general discussion regarding prognostic model methodology is included below.

### 2.6.1 Prognostic models

The following section highlights important methodological considerations regarding the design of prognostic models.

#### *2.6.1.1 Objectives and outcomes*

A prognostic model must have a clearly defined clinical objective and the primary outcome must be identified at the outset of study design. The outcome should be well defined to ensure consistency and reproducibility. (Mallett et al. 2010a) OS was used as the primary outcome in this thesis, defined as the number of months survived from the date of diagnosis.

#### *2.6.1.2 Study design*

A prospective cohort study is the gold standard for prognostic modelling, but can be difficult to achieve in short periods of time. Pre-specified patient inclusion criteria, data collection methods and statistical analysis should be defined prior to data collection. Data analysis should only be performed once sufficient measurable events have occurred. (Moons et al. 2009b)

A combination of retrospective and prospectively collected data was used in this thesis to provide sufficient events to power each study. However, this design can introduce bias during data collection and relies on completeness of pre-recorded, historic data. A common weakness of retrospective studies is missing data. (Simon and Altman 1994)

Most published prognostic models in cancer are retrospective in design. A review of 47 published prognostic cancer models in 2005 revealed that 68% of studies were either constructed from a database, or had used data collected retrospectively. (Mallett et al. 2010a)

#### *2.6.1.3 Study participants*

The number of patients included in a prognostic model should be reported. (Moons et al. 2012) Patient inclusion criteria are listed in each results chapter and the numbers

of patients that met the criteria are provided. Selection bias is possible in retrospective studies. (Mallett et al. 2010a)

#### *2.6.1.4 Number of Events in model*

The prognostic models developed in this thesis were powered by the event-per-variable (EPV) ratio. An event is defined when the outcome of interest occurs. There should be at least 10 EPV, in order to start confidently estimating the coefficients of included variables. (Peduzzi et al. 1996) The number of events occurring within a model should be reported, as this allows interpretation of the model's strength. Around 30% of studies do not report the number of events and an inadequate number of EPV is reported in 30% of models, with insufficient information to calculate EPV in 40% of studies. (Mallett et al. 2010a)

#### *2.6.1.5 Selecting candidate predictors*

Royston et al describe two main approaches to selecting predictors; the full model approach and the elimination approach. (Royston et al. 2009) There is no agreed consensus on the best method. Firstly, a predictor or variable should be clearly defined thereby allowing standardisation and reproducibility of its measurement. (Simon and Altman 1994) Previously reported prognostic indicators should be considered for inclusion in the model. However, selection of reported predictors may introduce publication bias into the model. Kyzas et al discovered that 98.5% of cancer studies reported significant results or highlighted non-significant trends. (Kyzas

et al. 2007) Ultimately, those that have established clinical credibility and are well established should be used. (Mallett et al. 2010a) This approach was adopted in Chapter 7.

In the full model approach, all measured variables are included in the model. Although this may reduce selection bias and over-fitting, the researchers must have prior knowledge about potentially important predictors and initially choose variables to build the model. This may be difficult to perform, as models can be hard to define and often suffer from the inclusion of too many variables. (Harrell 2001; Moons et al. 2009b; Steyerberg 2009)

Backward elimination is preferable to the forward elimination approach. Forward elimination determines the best candidate predictors initially, and then builds up the model accordingly. (Mantel 1970) Backward elimination was used in Chapter 7 however forward selection was used in Chapter 4 in order to compare with the original study. (Foley et al. 2014b)

Approximately 49% of studies select predictors for multi-variate analysis based on their significance in uni-variate analysis. (Harrell 2001; Mallett et al. 2010a) However, this method can produce over-fitted models, which is more common in small size samples because the introduction of error is not adequately controlled for. Similarly, potentially important but non-significant variables may be rejected from inclusion in multivariate analysis by other weakly predictive variables. (Sun et al. 1996) These

less important, weakly significant predictors are then included in the final model, thus over-estimating their effect size. (Royston et al. 2009)

#### *2.6.1.6 Data handling decisions*

Variables were kept continuous as this ensures predictive information is retained. (Royston et al. 2006) Likewise, ordered categorical data were only grouped for specific reasons, such as combining small numbers of patients with similar stages of disease. (Royston et al. 2009) Continuous variables were assessed for a normal distribution and not assumed to have a linear association with the outcome being studied. Transformation of non-normal variables into a logarithmic scale was performed following visual assessment of their histogram distribution. (Bland and Altman 1996)

#### *2.6.1.7 Developing final model*

Development and internal validation of a prognostic model was performed in Chapter 7. Hazard ratios and parameter estimates represent the risk of an event in relation to the outcome. An individual's risk of outcome occurrence can be quantified using a model equation to calculate a prognostic score. The products of parameter estimates and variable included in the model are summated to produce an individual score. (Moons et al. 2012)

### 2.6.2 Prognostic Model Validation

Validation of prognostic models is essential for translation into a clinical setting. This process tests the model's application in different groups of patients, hospitals or countries. An inability to translate to another group of patients may be due to deficiencies in the initial design, statistical analysis errors or inadvertent exclusion of an important variable. (Altman et al. 2009) Model validation is not commonly performed; Mallett et al found that only 34% of the models studied had been validated, 11% externally. (Mallett et al. 2010b) This may explain why few of the prognostic models developed are routinely used in clinical practice. (Reilly and Evans 2006)

#### *2.6.2.1 Model Validation*

Prognostic (or predictive) models can be internally or externally validated. (Altman et al. 2009) Internal validation involves splitting a dataset into a development cohort and a validation cohort, the latter being tested on the former, and was performed in Chapters 4 & 7. Internal validation can give optimistic results if the two datasets are similar. (Mallett et al. 2010b). The same patient pathways and diagnostic work-up are performed, which can introduce the same errors and biases. (Altman et al. 2009)

The gold standard method is external validation, but this can take substantial time to co-ordinate and perform. External validation examines the model's generalisation by testing data from another centre. The dataset collected must match the original

model, but can be collected retrospectively, thereby allowing incorporation of crucial follow-up data to expedite the model's validation. (Altman et al. 2009) Following validation, the model should be tested in a clinical setting after being used to influence clinical decision-making, potentially affecting patient outcomes. (Reilly and Evans 2006)

#### *2.6.2.2 Measures of Model Performance*

Measures of model performance based on log-likelihood function such as Akaike Information Criterion (AIC) or Bayes Information Criterion (BIC) are preferred. (Whittle et al. 2017) AIC is calculated by 2k - 2*log(L), where k is the number of parameters and L is the likelihood of the model. (Akaike 1974) Lower values indicate a better fit and estimated model quality reduces if the number of included variables increase. (Cook 2007) AIC was used to estimate and compare the performance of the prognostic models developed in Chapter 7.

Techniques for describing the accuracy of a model include calibration and discrimination. Calibration is a measure of the agreement between the predicted probabilities and the actual observed risk. Calibration can be measured using the Hosmer-Lemesbow statistic. (Hosmer and Lemesbow 1980) Discrimination is defined as the ability of a model to predict which patients experience an event of interest compared to those without an event. (Moons et al. 2012) Kaplan-Meier analysis (Kaplan and Meier 1958) and log-rank tests are common methods to evaluate discrimination. (Mallett et al. 2010a) These methods do not provide the size of the

difference between groups, a similar problem to p-values. (Mallett et al. 2010b) Another common measure of model performance is the c-statistic, also known as the area under a receiver operator characteristic (ROC) curve. (Hanley and McNeil 1982) Limitations in use of the c-statistic have been highlighted. Inclusion of novel variables in prognostic or prediction models can lead to more accurate risk classification but with little change in the c-statistic. (Cook 2007)

# Chapter 3. Accuracy of Contemporary Oesophageal Cancer Lymph Node Staging with Radiological-Pathological Correlation

## 3.1 Introduction

Accurate staging optimises management plans and provides the best chance of survival for patients with potentially curable disease. As outlined in section 1.8, radiological staging is complex and time-consuming.

Management decisions are influenced by results of lymph node assessment on radiological staging investigations. If the MDT decide upon surgical management and radiological staging is ≥T3 and/or ≥N1, 2 cycles of NACT are given prior to resection. Differentiation of node-negative (N0) from node-positive (N+) disease is important because this ensures that patients avoid unnecessary chemotherapy if over-staged, and are not denied potentially beneficial NACT if under-staged.

However, the existence of small LNMs which cannot be directly visualised on any imaging modality, may cause inaccurate staging and subsequently progress, with a detrimental effect on patient outcome. (Koenig et al. 2009)

Therefore, the aim was to define the accuracy of CT, EUS and PET/CT N-stage in the modern era of radiological OC staging. The secondary aim was to investigate the prevalence of micro-metastases and size of LNMs in patients radiologically staged N0 but pathologically node-positive (pN+), using radiological-pathological correlation.

## 3.2 Materials and Methods

### 3.2.1 Patient Cohort

This retrospective cohort study included consecutive patients who underwent surgical resection of an oesophageal or GOJ tumour, over a 5-year period (November 2010 – December 2015) within a centralised service.

Inclusion criteria were a previously untreated, biopsy-proven oesophageal or GOJ tumour in patients who underwent surgery alone, or had a poor response (TRG 4) or no response (TRG 5) following either NACT or NACRT. (Mandard et al. 1994) All patients had completed CT, EUS and PET/CT staging investigations and pN assigned.

A total of 190 patients were considered for inclusion in the study. Patients with tumours that showed complete pathological response (pCR, TRG 1) or tumours with some response (TRG 2 & 3) following NACT or NACRT were excluded (n=16, n=13 and n=13, respectively), because the final pathology is not likely to be representative of pre-operative status. Incomplete radiological staging investigations, such as EUS examinations in which the operator was unable to traverse a stenotic tumour and fully classify N-stage, were excluded (n=14). Patients that underwent an 'open-and-close' procedure due to irresectable disease at the time of operation were also excluded (n=22). Following exclusions, 112 patients were included in the study.

### 3.2.2 Histopathological Methods

Histopathological reporting of OC specimens was performed according to the minimum requirements defined by the RCPath. (Mapstone 2007) All lymph nodes identified in the resection specimen were prepared in 3 mm slices for pathological evaluation. N-stage was then assigned depending on the number of LNMs identified. TRG of the primary tumour was assigned. (Mandard et al. 1994) All available resection specimens in discordant cases radiologically staged N0 but pathologically N+ were further evaluated. The maximum size of both involved lymph nodes and metastases within those lymph nodes, were retrospectively recorded. Maximum size was defined as the largest dimension on the glass slide, measured by a Consultant Pathologist. A micro-metastasis was defined as tumour deposit measuring ≤2 mm. (Weaver 2010) Furthermore, a metastasis to lymph node size ratio was calculated.

### 3.2.3 Statistical Analysis

Descriptive statistics were used to describe categorical and continuous variables. In this study, N-stage was separated into negative (N0) and N+ (N1, N2 or N3) groups. Accuracy was defined as the number of cases with correct N-stage divided by total number of investigations. Sensitivity and specificity of N+ disease were calculated for each modality. A Chi-square test assessed significant differences in under- or over-staging for each modality. Significant differences in under-staging between modalities were assessed with McNemar's test.

## 3.3 Results

### 3.3.1 Patient Cohort

The median age was 65 years (range 24-78) and the male: female ratio was 92 (82.1%): 20 (17.9%). Table 3.3.1 describes the characteristics of the patient cohort.

For CT, 75 patients (67.0%) were staged N0 and 37 (33.0%) were N+. For EUS, 72 patients (64.3%) were staged N0 and 40 (35.7%) were N+. For PET/CT, 84 (75.1%) were staged N0 and 28 (24.9%) were staged N+. Table 3.3.2 compares the frequency of radiological and pathological N-stages for CT, EUS and PET/CT.

Table 3.3.1. Patient Characteristics

| Characteristic | Frequency (%) |
|---|---|
| Tumour Location | |
| Oesophagus | 59 (52.7) |
| Mid | 10 (16.9) |
| Distal | 49 (83.1) |
| Gastro-oesophageal junction (GOJ) | 53 (47.3) |
| Siewert Type I | 19 (35.8) |
| Siewert Type II | 15 (28.4) |
| Siewert Type III | 19 (35.8) |
| Histology | |
| Adenocarcinoma | 100 (89.3) |
| SCC | 11 (9.8) |
| Neuro-endocrine | 1 (0.9) |
| Treatment | |
| NACT | 67 (59.8) |
| Surgery alone | 41 (36.6) |
| NACRT | 4 (3.6) |
| Tumour Regression Grade | n=71 |
| TRG 4 | 42 (59.2) |
| TRG 5 | 29 (40.8) |

Table 3.3.2. Comparison of N-stage Frequency Classified by CT, EUS, PET/CT and Pathology.

| CT N-stage | | | | | |
|---|---|---|---|---|---|
| Frequency (%) | N0 | N1 | N2 | N3 | Total |
| pN0 | 34 (30.4) | 8 (7.1) | 2 (1.7) | 0 (0.0) | 44 (39.3) |
| pN1 | 21 (18.8) | 4 (3.6) | 2 (1.7) | 0 (0.0) | 27 (24.1) |
| pN2 | 16 (14.3) | 10 (8.9) | 1 (0.9) | 0 (0.0) | 27 (24.1) |
| pN3 | 4 (3.6) | 7 (6.3) | 3 (2.7) | 0 (0.0) | 14 (12.5) |
| Total | 75 (67.0) | 29 (25.9) | 8 (7.1) | 0 (0.0) | 112 (100.0) |
| EUS N-Stage | | | | | |
| Frequency (%) | N0 | N1 | N2 | N3 | Total |
| pN0 | 33 (29.5) | 9 (8.0) | 1 (0.9) | 1 (0.9) | 44 (39.3) |
| pN1 | 20 (17.9) | 7 (6.3) | 0 (0.0) | 0 (0.0) | 27 (24.1) |
| pN2 | 13 (11.6) | 10 (8.9) | 4 (3.6) | 0 (0.0) | 27 (24.1) |
| pN3 | 6 (5.4) | 6 (5.4) | 1 (0.9) | 1 (0.9) | 14 (12.5) |
| Total | 72 (64.3) | 32 (28.6) | 6 (5.4) | 2 (1.7) | 112 (100.0) |

| | PET/CT N-stage | | | | |
|---|---|---|---|---|---|
| Frequency (%) | N0 | N1 | N2 | N3 | Total |
| pN0 | 40 (35.8) | 4 (3.6) | 0 (0.0) | 0 (0.0) | 44 (39.3) |
| pN1 | 23 (20.5) | 4 (3.6) | 0 (0.0) | 0 (0.0) | 27 (24.1) |
| pN2 | 15 (13.4) | 10 (8.9) | 2 (1.7) | 0 (0.0) | 27 (24.1) |
| pN3 | 6 (5.4) | 6 (5.4) | 2 (1.7) | 0 (0.0) | 14 (12.5) |
| Total | 84 (75.1) | 24 (21.5) | 4 (3.4) | 0 (0.0) | 112 (100.0) |

Overall, median time between radiological staging and surgery was 3 months (range 1-9 months), 1 month (range 0-3 months) in patients undergoing surgery alone and 4 months (range 3-4 months) in patients receiving NACT.

### 3.3.2 Accuracy, Sensitivity and Specificity

N0 vs N+ disease was correctly differentiated with CT, EUS and PET/CT in 61 cases (54.5%), 62 (55.4%) and 64 (57.1%), respectively. There was no significant difference between CT, EUS and PET/CT for detecting N+ disease ($X^2$ 0.169, df 2, p=0.919). The sensitivity and specificity for identifying N0 vs N+ disease with CT,

EUS and PET/CT was 39.7% and 77.3%, 42.6% and 75.0%, and 35.3% and 90.9%, respectively.

### 3.3.3 Under-staging vs Over-staging

All modalities were significantly more likely to under-stage nodal disease; CT ($X^2$ 32.890, df 1, p<0.001), EUS ($X^2$ 28.471, df 1, p<0.001) and PET/CT ($X^2$ 50.790, df 1, p<0.001). Comparing modalities, there was a borderline significant difference in under-staging between CT and EUS (p=0.063) but no difference between CT and PET/CT (p=1.000). However, there was a statistically significant difference between EUS and PET/CT (p=0.031), suggesting PET/CT may further under-stage nodal disease.

### 3.3.4 Pathological Lymph Node Measurement

Fifteen archived resection specimens in patients pre-operatively staged N0 were available for retrospective measurement of lymph nodes that contained metastases. In total, 50 involved lymph nodes were assessed. (Table 3.3.3) The median size of involved lymph nodes was 6 mm (range 2-15 mm) and the median metastasis size was 3 mm (0.5-13.5 mm). Twenty-two LNMs (44%) measured ≤2 mm, which are defined as micro-metastases. (Fig. 3.3.1) Forty-one LNMs (82%) were ≤6 mm and 46 LNMs (92%) were ≤10 mm. A metastasis to lymph node size ratio was calculated. Thirty-one (62%) of the lymph nodes examined were replaced with ≥50% metastatic

deposit, 19 (38%) were replaced with <50% metastatic deposit, with 12 (24%) replaced with <25% metastatic deposit, using maximum size criteria.

Table 3.3.3. Frequency and Distribution of Lymph Node and Metastasis Size when Separated into 2 mm Groups for Descriptive Purposes.

| Frequency (%) | Maximum Size (mm) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 0-2 | 2.1-4 | 4.1-6 | 6.1-8 | 8.1-10 | 10.1-12 | 12.1-14 | 14.1-16 |
| Lymph Node | 3 (6.0) | 11 (22.0) | 13 (26.0) | 12 (24.0) | 4 (8.0) | 3 (6.0) | 3 (6.0) | 1 (2.0) |
| Metastasis | 22 (44.0) | 9 (18.0) | 10 (20.0) | 3 (6.0) | 2 (4.0) | 2 (4.0) | 2 (4.0) | 0 (0.0) |

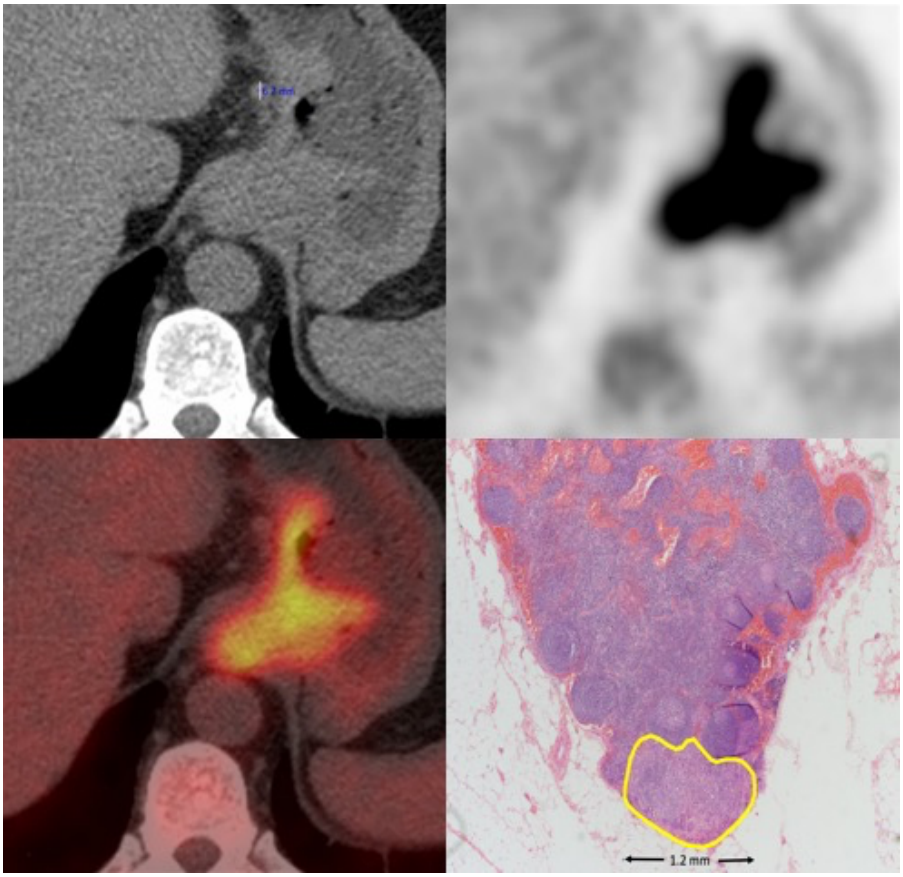



Figure 3.3.1. An example of a false-negative LNM from selected CT, MIP and fused PET/CT images. The lymph node measured 6.2 mm on CT (top left) but was non-avid compared to the GOJ tumour. A low-power magnification of the lymph node (bottom right) shows the presence of a micro-metastasis (highlighted with yellow outline). For reference, the lymph node measures 5 mm in maximum size and the micro-metastasis measures 1.2 mm. The image was acquired at a magnification of 25x. The slide was prepared following lymph node dissection. The lymph node was placed into a cassette block, then sections were cut with a thickness of 4 μm and stained with Haematoxylin & Eosin.

## 3.4 Discussion

This study found poor N-stage accuracy with CT, EUS and PET/CT. In general, all modalities were more likely to under-stage nodal disease, with PET/CT more likely to under-stage than EUS. Another important finding, is the prevalence of small LNMs (<6 mm) in the resection specimens of patients radiologically staged N0. Micro-metastases have been found in lymph nodes of early oesophageal tumours (Grotenhuis et al. 2010) but little has been published with radiological correlation. Studies investigating lung cancer have detected micro-metastases in patients radiologically staged N0 (Martin et al. 2016), although evidence in OC is lacking.

The majority of LNMs (82%) were <6 mm, which makes direct visualisation extremely challenging on current medical imaging techniques. This is likely to be a major reason for discrepancy between radiological and pathological staging. In addition, traditional radiological measurement of lymph nodes is taken in the short-axis, which further reduces the likelihood that LNMs are diagnosed. (Eisenhauer et al. 2009) Even with the improved contrast resolution of EUS compared to cross-sectional imaging, it is unlikely that a lymph node of this size would confidently be classified as metastatic. (Weaver et al. 2004) Similarly, there was a relatively high prevalence of micro-metastases (44%).

These results have significant implications for treatment decision-making processes and demonstrate that contemporary radiology techniques are inadequate for N-staging. Numerous studies have demonstrated the importance of LNMs. (Kayani et

al. 2011) Better evidence is required to understand the prognostic significance of micro-metastases, but they are generally felt to confer a worse prognosis. (Izbicki et al. 1997; Glickman et al. 1999)

There is evidence that a significant proportion of surgical patients have systemic micro-metastases at the time of resection. A study investigating resected ribs from these patients detected micro-metastases in up to 78%, although the incidence is dependent on the histological technique used. (Ryan et al. 2004) This is a higher detection rate than our study, but the results are comparable despite different techniques and tissues used between studies. The high rate of micro-metastases may be a reason that our results show significant under-staging of nodal disease, and perhaps clinicians could consider lowering the threshold for treating patients with systemic neo-adjuvant therapy.

Previous research from our institution has shown N-stage, lymph node metastasis count and volume of nodal disease to have prognostic significance in patients with OC. (Chan et al. 2013a; Foley et al. 2014a) Nodal disease in these studies may continue to be an important prognostic indicator, but the radiological staging is likely to have under-estimated the total nodal disease burden in those patient cohorts. Results of staging performance have also been published from our institution. These studies compared CT and EUS N-stage with pN. Blackshaw et al investigated the accuracy of N-staging in GOJ tumours and found significant differences in agreement, sensitivity and specificity between Siewert type II and type III tumours. (Blackshaw et al. 2008) Weaver et al found agreement (calculated using weighted kappa values),

sensitivity and specificity of N-staging was 0.603, 79% and 84% for CT and 0.610, 91% and 68% for EUS. (Weaver et al. 2004) The results of the current study show poorer agreement and sensitivity. There are a number of possible reasons for these findings, including disease evolution, greater inter-observer variability between reporters, and fewer, but more specialised upper GI cancer pathologists reporting the resection specimens, with increased rates of LNM detection. (Weaver 2010)

Accuracy of diagnosing N+ disease with CT, EUS and PET/CT was 54.5%, 55.4% and 57.1%, respectively. In a clinical context, these results are unsatisfactory given that the presence of LNMs is such a major prognostic indicator. (Kayani et al. 2011) Specificity results are comparable with past meta-analyses but sensitivity results are reduced for all modalities. Previously published literature states sensitivity for N-staging of CT, EUS and PET/CT is 50%, 80% and 57%, and specificity is 83%, 70% and 85%, respectively. (van Vliet et al. 2008) The rigour of the pathological evaluation is unknown for this meta-analysis, but micro-metastases may not have been routinely evaluated. This reason could explain the reduced sensitivity in this current study compared to this meta-analysis.

*Limitations*

Patients with an ‘open-and-close’ procedure were excluded because of the absence of final pathological stage. Ideally, these patients should be included on an intention-to-treat basis. An ‘open-and-close’ procedure can occur following radiological under-staging of disease. There are also limitations of pathological lymph node

examination. Approximately 3 mm sections are taken through lymph nodes once they are mounted in a cassette, but this may be performed with varying skill and consistency. Micro-metastases may be missed if not bisected during preparation, and this suggests that the true incidence of micro-metastases in this cohort of patients may be even greater. Although the RCPath define the minimum requirements for pathological reporting, there is currently no standardised method for lymph node preparation and assessment in OC. In addition, this is a single-centre study and results may not be representative of the national accuracy rate.

## 3.5 Conclusion

In conclusion, this evaluation of contemporary staging performance over a 5-year period in a centralised upper GI cancer service has shown poor N-staging accuracy for CT, EUS and PET/CT. Radiological-pathological correlation in patients staged N0 has shown many small LNMs (<6 mm) that are extremely challenging to diagnose directly from medical imaging. The findings of this study have significant implications for patient care, because radiological staging results largely influence treatment decisions made by the MDT.

# Chapter 4. Internal Validation of N-staging of Oesophageal and Junctional Carcinoma: is There Still a Role for EUS in Patients Staged N0 at PET/CT?

## 4.1 Introduction

This chapter aimed to validate the results of a study that investigated the role of EUS in patients staged N0 on PET/CT. (Foley et al. 2014b) The original study developed 2 models and demonstrated that both N-stage and N0 vs N+ were significantly and independently associated with OS.

Although results from Chapter 3 showed that staging investigations have poor accuracy for identifying LNMs, EUS was marginally more accurate. This is primarily due to the improved contrast resolution of EUS compared to PET and CT. Other factors for improved accuracy include the relatively poor spatial resolution of PET imaging, which does not allow differentiation of LNMs from the primary tumour, or detection of small LNMs. (Kapoor et al. 2004)

The primary aim of this study was to internally validate the results of the 2 models in a new, independent cohort of patients, known as the validation cohort. The methodology used in the published study was repeated to allow comparison and is discussed below. The secondary aim was to assess the prognostic significance of LNMs in the validation cohort of patients.

## 4.2 Materials & Methods

### 4.2.1 Patient Cohorts

Patients staged N0 with PET/CT between 1st January 2013 and 31st June 2015, were considered for the validation cohort (n=166). Patients with FDG-avid lymph nodes or distant metastases (n=4), or incomplete EUS staging (n=23) were excluded. Following exclusions, 139 patients were included in the study.

All patients in the validation cohort followed the usual staging pathway and had PET/CT in Cardiff using the same scanner and protocol described in section 2.2.3. All patients had complete EUS staging. Details of the EUS procedure are found in section 2.2.2. As in the published study, 2 variables were recorded for each patient; EUS T-stage (T1-4a) and EUS N-stage (N0-3). A third variable was derived from the EUS N-stage; EUS N0 vs N+ (N1, N2 or N3).

One-hundred and seventeen patients were included in the original patient cohort. All patients were staged N0 at PET/CT between 1st December 2008 and 31st May 2012. PET/CT examinations were performed across 2 sites; 47 in the first centre (Cheltenham) and 70 in the second centre (Cardiff). At the first centre, PET/CT examination was performed using a Philips 16 section Gemini GXL dedicated PET/CT system (Philips Medical Systems, Cleveland, USA). The activity uptake time was 60 min. A standard administered activity of 350 MBq FDG was injected. Reconstructions

were performed using a 3D acquisition with non-TOF acquisition for 4 min per bed position.

### 4.2.2 Statistical Analysis

Categorical data are expressed as frequency (percent). Differences in patient characteristics between original and validation cohort was assessed with the Chi-square test. Uni-variate analysis (Kaplan and Meier 1958) was performed for EUS T-stage, EUS N-stage and EUS N0 vs N+, and differences between groups assessed using the log-rank test. Two Cox regression models (Cox 1972) were constructed to assess the independent prognostic value of variables; model 1 included EUS T-stage and EUS N-stage and model 2 included EUS T-stage and EUS N0 vs N+. In addition, a log-rank test was used to assess survival differences between pN0 and pN+ groups in the sub-group of patients who underwent surgical resection in the validation cohort.

## 4.3 Results

### 4.3.1 Patient Cohorts

Table 4.3.1 details the baseline characteristics of original and validation patient cohorts. The median age of the original and validation cohorts was 67.0 years (range 24.0-82.0) and 66.0 years (39-84), respectively. The median follow-up period was 25.0 months in the original cohort (95% CI 23.1-26.9) and 26.0 months (95% CI 22.7-29.3) in the validation cohort. Mean survival times are presented as a 50% mortality

rate was not reached in either cohort. Two-year OS in the original cohort was 39.3% and 49.0% in the validation cohort. The mean OS of the original cohort was 33.1 months (95% CI 30.1-36.1) and 29.8 months (95% CI 27.1-35.2) in the validation cohort.

Table 4.3.1. Baseline Characteristics of Patients in Original and Validation Cohorts

| Frequency (%) | Original Cohort (n=117) | Validation Cohort (n=139) | p-value |
|---|---|---|---|
| Male: Female | 88 (75.2): 29 (24.8) | 108 (77.7): 31 (22.3) | p=0.640 |
| Tumour Location | | | Oesophagus vs Junction p=0.212 |
| Oesophagus | 73 (62.4) | 76 (54.7) | |
| Upper | 0 (0.0) | 2 (2.7) | |
| Mid | 20 (27.4) | 22 (28.9) | |
| Distal | 53 (72.6) | 52 (68.4) | |
| Junction | 44 (37.6) | 63 (45.3) | |
| Siewert type I | 5 (11.3) | 25 (39.7) | |
| Siewert type II | 12 (27.3) | 18 (28.6) | |
| Siewert type III | 27 (61.4) | 20 (31.7) | |
| Histology | | | p=0.228 |
| Adenocarcinoma | 98 (83.8) | 107 (77.0) | |
| SCC | 19 (16.2) | 26 (18.7) | |
| HGD | 0 (0.0) | 3 (2.2) | |
| Neuro-endocrine | 0 (0.0) | 2 (1.4) | |
| Undifferentiated | 0 (0.0) | 1 (0.7) | |

| | | | |
|---|---|---|---|
| EUS T-stage | | | p=0.682 |
| T1 | 18 (15.4) | 20 (14.4) | |
| T2 | 16 (13.7) | 18 (12.9) | |
| T3 | 75 (64.1) | 86 (61.9) | |
| T4a | 8 (6.8) | 13 (9.4) | |
| T4b | 0 (0.0) | 2 (1.4) | |
| EUS N-stage | | | p=0.027 |
| N0 | 78 (66.7) | 89 (64.0) | |
| N1 | 23 (19.7) | 42 (30.2) | |
| N2 | 9 (7.6) | 7 (5.1) | |
| N3 | 7 (6.0) | 1 (0.7) | |
| Treatment | | | Curative vs Palliative p=0.144 |
| Curative | 105 (89.7) | 116 (83.5) | |
| NACT | 40 (38.1) | 44 (37.9) | |
| dCRT | 29 (27.6) | 39 (33.6) | |
| Surgery alone | 32 (30.5) | 19 (16.4) | |
| NACRT | 1 (0.9) | 11 (9.5) | |
| EMR | 3 (2.9) | 3 (2.6) | |
| Palliative | 12 (10.3) | 23 (16.5) | |

### 4.3.2 Summary of Results from Original Study

Univariate analysis showed that EUS T-stage ($X^2$ 8.321, df 3, p=0.040), N-stage ($X^2$ 14.879, df 3, p=0.002) and N0 vs N+ ($X^2$ 11.325, df 1, p=0.001) were significantly associated with OS. When EUS T-stage and N-stage were entered in the first Cox regression model, only EUS N-stage was significantly and independently associated with duration of survival [hazard ratio (HR) 1.616-4.707, df 3, p=0.005]. When EUS T-stage and EUS N0 vs N+ were entered in a second Cox regression model, EUS N0 vs N+ was significantly and independently associated with OS (HR 3.105, 95% CI 1.543 – 6.247, p=0.001).

### 4.3.3 Univariate Analysis in Validation Cohort

EUS T-stage ($X^2$ 21.031, df 4, p<0.001) (Fig. 4.3.1) and EUS N0 vs N+ ($X^2$ 4.300, df 1, p=0.038) were significantly associated with OS. EUS N-stage did not show a statistically significant association with OS ($X^2$ 5.699, df 3, p=0.127).

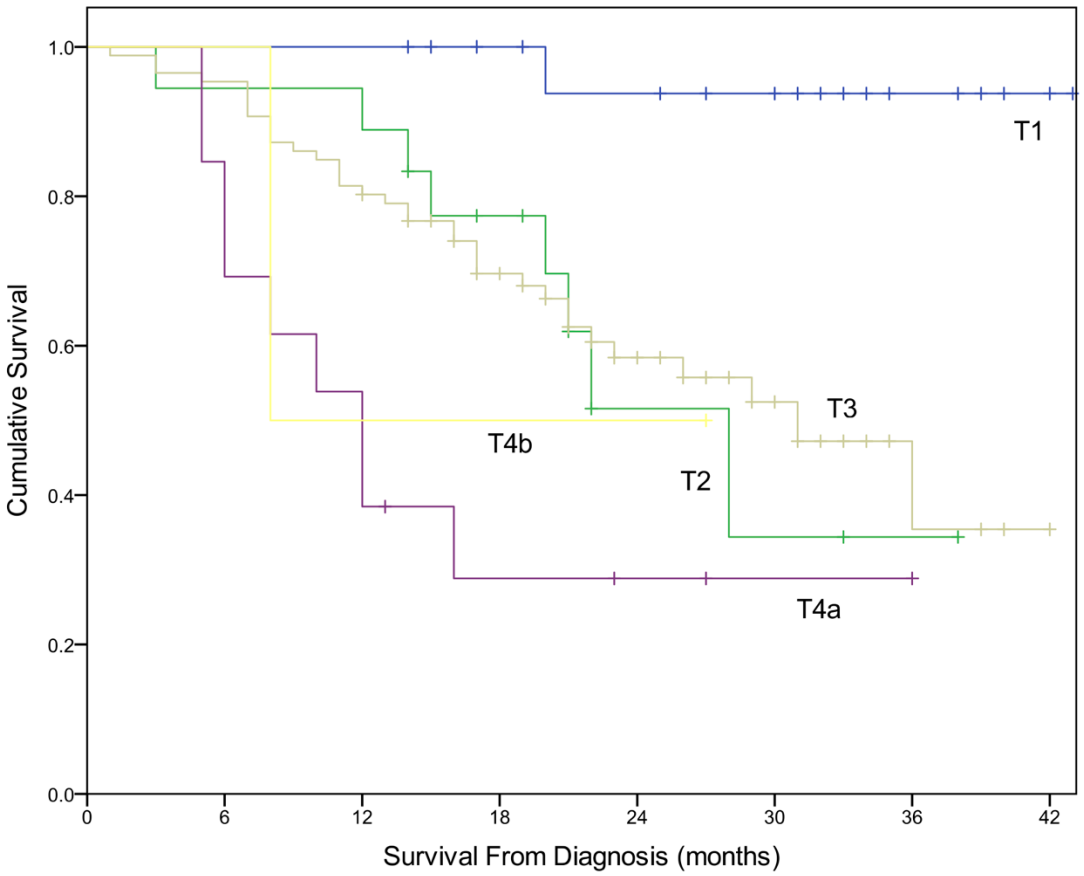



Figure 4.3.1. Significant difference in cumulative survival by EUS T-stage ($X^2$ 21.031, df 4, p<0.001). Patients with more advanced EUS T-stage have worse OS. EUS T1 (blue line), T2 (green line), T3 (gold line), T4a (purple line) and T4b (yellow line).

Table 4.3.2 provides mean survival data for patients classified by EUS T-stage, N-stage and EUS N0 vs N+.

Table 4.3.2. Survival Data of Patients in Validation Cohort Derived from Uni-variate Analysis.

| | | 95% Confidence Interval | |
|---|---|---|---|
| EUS Variable | Mean OS (months) | Lower | Upper |
| T-stage | | | |
| T1 | 41.563 | 38.834 | 44.291 |
| T2 | 25.830 | 20.062 | 31.598 |
| T3 | 27.908 | 24.484 | 31.332 |
| T4a | 16.846 | 9.842 | 23.851 |
| T4b | 17.500 | 4.334 | 30.666 |
| N-stage | | | |
| N0 | 31.853 | 28.735 | 34.971 |
| N1 | 25.625 | 21.246 | 30.004 |
| N2 | 16.857 | 11.873 | 21.841 |
| N3 | 17.000 | 17.000 | 17.000 |
| N+ | 24.924 | 20.966 | 28.882 |

### 4.3.4 Multivariate Analysis

Again, 2 alternative Cox Regression models were produced in the validation cohort.

#### *4.3.4.1 Model 1*

EUS T-stage and EUS N-stage were entered in model 1. EUS T-stage was significantly and independently associated with OS (HR 11.656-30.114, 95% CI 0.994-243.079, p=0.011). (Table 4.3.3)

Table 4.3.3. Results of Model 1 Multi-variate Analysis including EUS T-stage and EUS N-stage.

| | | | | 95% Confidence Interval | |
|---|---|---|---|---|---|
| Variable | p-value | Hazard Ratio | df | Lower | Upper |
| EUS T-Stage | 0.011 | | 4 | | |
| T2 | 0.018 | 12.482 | 1 | 1.528 | 101.957 |
| T3 | 0.016 | 11.656 | 1 | 1.570 | 86.548 |
| T4a | 0.001 | 30.114 | 1 | 3.731 | 243.079 |
| T4b | 0.050 | 16.270 | 1 | 0.994 | 266.273 |
| EUS N-Stage | 0.553 | | 3 | | |
| N1 | 0.560 | 1.192 | 1 | 0.660 | 2.154 |
| N2 | 0.353 | 1.653 | 1 | 0.572 | 4.772 |
| N3 | 0.260 | 3.176 | 1 | 0.425 | 23.716 |

*4.3.4.2 Model 2*

EUS T-stage and EUS N0 vs N+ were entered in model 2. Again, EUS T-stage was significantly and independently associated with OS (HR 11.714-29.631, 95% CI 1.067-238.959, p=0.012). (Table 4.3.4)

Table 4.3.4. Results of Model 2 Multi-variate Analysis including EUS T-stage and EUS N0 vs N+.

| | | | | 95% Confidence Interval | |
|---|---|---|---|---|---|
| Variable | p-value | Hazard Ratio | df | Lower | Upper |
| EUS T-Stage | 0.012 | | 4 | | |
| T2 | 0.020 | 11.977 | 1 | 1.469 | 97.620 |
| T3 | 0.016 | 11.714 | 1 | 1.579 | 86.902 |
| T4a | 0.001 | 29.631 | 1 | 3.674 | 238.959 |
| T4b | 0.045 | 17.243 | 1 | 1.067 | 278.714 |
| EUS N0 vs N+ | 0.359 | 1.292 | 1 | 0.747 | 2.235 |

### 4.3.5 Effect of Including Cheltenham Patients on Original Cohort Models

A post-hoc analysis was performed to determine the effect of including patients scanned in Cheltenham on the original cohort models. To perform this analysis, Cheltenham patients were excluded in attempt to control the comparison of the original Cardiff cohort with the validation cohort. Seventy patients were originally scanned at Cardiff. Of these, 53 patients (75.7%) were staged EUS N0, 11 (15.7%) were EUS N1, 3 (4.3%) were N2 and 3 (4.3%) were N3. Both EUS N-stage (HR 2.365-32.757, 95% CI 0.476-223.922, p=0.005) and EUS N0 vs N+ (HR 3.783, 1.141-12.539, p=0.03) were again independent predictors of OS, in keeping with findings from the original study. Therefore, inclusion of Cheltenham patients had little effect on the prognostic models. Confidence intervals are wide, likely due to the small numbers in N2 and N3 groups.

An assessment of differences between patient cohorts that may explain the lack of validation was performed. A comparison of the proportion of patients who were staged N0 and considered for inclusion during both study periods was made. Post-hoc review of the Upper GI database revealed that 117 of 207 patients (56.5%) were staged N0 on PET/CT in the study period of the original cohort and 139 of 317 (43.8%) were staged N0 in the study period of the validation cohort. This difference was statistically significant ($X^2$ 8.049, df 1, p=0.005). A significantly higher proportion of Cardiff patients were staged N0 in the original cohort, suggesting different proportions of patients were eligible for inclusion, which could affect the results of the prognostic models.

### 4.3.6 Prognostic Significance of Pathological Lymph Nodes

In total, 74 patients from the validation cohort underwent surgical resection. Thirty-nine patients (52.7%) were classified pN0 and 35 patients (47.3%) were classified pN+. There was a significant difference in OS between pN0 and pN+ groups ($X^2$ 13.315, df 1, p<0.001). (Fig. 4.3.2) Mean OS for the pN0 group was 40.091 months (95 % CI 36.931-43.251) and 26.538 (22.123-30.953) for the pN+ group.

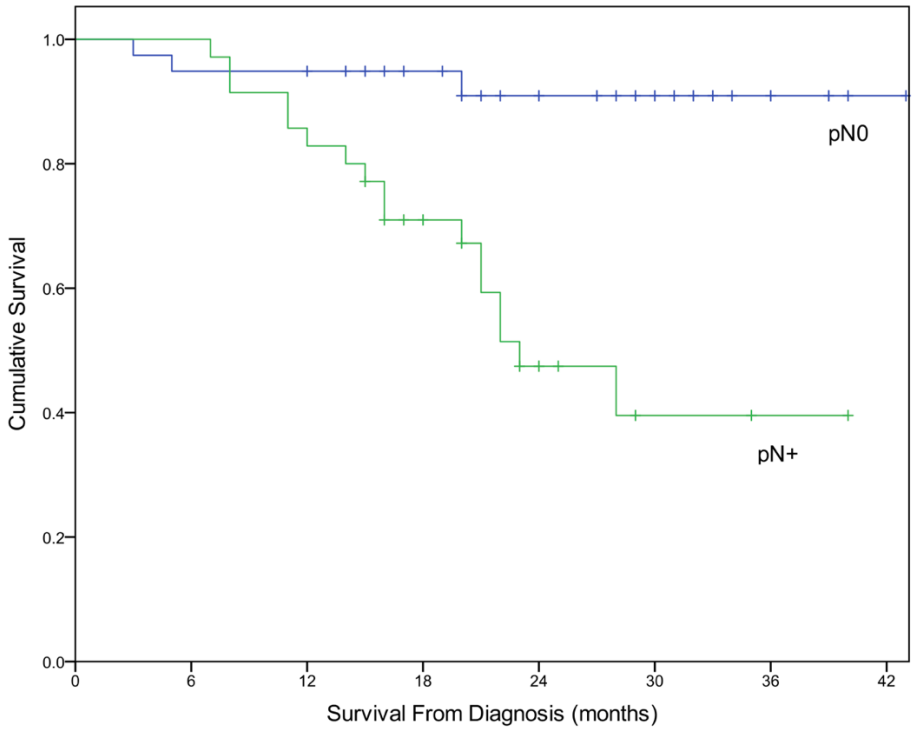



Figure 4.3.2. Significant difference in cumulative survival between pN0 and pN+ groups (X2 13.315, df 1, p<0.001). Patients with positive pathological lymph nodes have worse OS. Negative pathological lymph nodes (pN0, blue line) and positive pathological lymph nodes (pN+, green line).

## 4.4 Discussion

This study could not replicate the results of the previous models and therefore does not validate the published work. In the validation cohort, EUS N-stage and EUS N0 vs N+ did not have prognostic significance in multi-variate analysis, although EUS N0 vs N+ was statistically significantly in univariate analysis (p=0.038). EUS N-stage was not associated with OS in the validation cohort. However, this study shows that EUS T-stage is significantly and independently associated with OS, which supports relatively few data from other studies. (Paraf et al. 1995; Hiele et al. 1997; Reid et al. 2012)

When patients scanned in Cheltenham were removed in post-hoc analysis, EUS N-stage and EUS N0 vs N+ remained independent predictors of OS in the original Cardiff cohort. This suggests that the Cheltenham PET/CT scanner had little effect on the original study results. One reason for the inability to validate the original model could be a statistically significant difference in proportions of patients staged N0 during both study periods.

An important issue in validation studies is the extent to which the cohorts differ. This can result in validation failure unless appropriately adjusted for. (Debray et al. 2015) A higher proportion of patients were staged as EUS N0 in the original Cardiff cohort and relatively more staged N+ in the validation cohort. Reporting trends over time have not been assessed in this thesis, but context bias has been shown to influence radiologists image interpretation. (Egglin and Feinstein 1996) EUS operators were

not blinded to the results of the PET/CT, which may influence the interpretation of N-stage.

Prior to the opening of the PETIC in 2010, patients were scanned at the Cobalt Imaging Centre, Cheltenham using a Philips 16 section Gemini GXL dedicated PET/CT system (Philips Medical Systems, Cleveland, USA). The 2 sites used different scanners and protocols, and patients were scanned at different activity uptake times. Patients were scanned at 60 minutes of uptake time in Cheltenham and after 90 minutes in Cardiff. Longer uptake times lead to higher tumour to background avidity and can therefore increase the conspicuity of LNMs. Secondly, the Cardiff scanner had a TOF algorithm but the Cheltenham scanner did not. TOF reconstructions improve signal-to-noise ratio, detection and anatomical localisation of LNMs by allowing more precise measurement of the time difference between detections. (Surti 2015) Finally, images were acquired for 4 minutes per bed position in Cheltenham, whereas the acquisition was 3 minutes per bed position in Cardiff. Some improvement in image quality may be expected in Cheltenham with longer acquisition times, provided the patient remained still. However, the results of this validation study assume that longer acquisition did not affect the models.

The additional sub-group analysis conducted in this study confirms the presence of LNMs as a major prognostic indicator. (Kayani et al. 2011) There was a highly significant difference in OS between pN0 and pN+ groups. This finding highlights the importance of accurate pre-treatment lymph node staging in OC. However, as

demonstrated in Chapter 3, the challenge of inaccurate radiological staging with CT, EUS and PET remains.

*Limitations*

This validation study replicated methods used in the original study and as such, reproduced its limitations. This remains a relatively heterogeneous cohort of patients. Dissemination of LNMs is dependent on T-stage and to a lesser extent, histological cell type of the primary tumour. (Rice et al. 1998; Siewert et al. 2001) Stratification of T-stage or histological cell type was not performed in the original study but is important because the incidence of LNMs increase in SCC and with advanced T-stage. The EUS operators may have been influenced by results of the PET/CT or CT reports, resulting in inadvertent changes in lymph node assessment over time. In addition, the author has become more aware of appropriate prognostic research methods since the original publication. (Moons et al. 2009b) The statistical methodology used in the original study could have been improved.

## 4.5 Conclusion

Validation studies are important in prognostic research. (Altman and Royston 2000) This validation study did not replicate the results of the original work. However, this validation study has shown the continued benefit of EUS T-stage in patients staged N0 on PET/CT. EUS remains a valuable component of a multi-modality approach to OC staging. The presence of pathological LNMs have a significant detrimental effect on OS.

# Chapter 5. Radiological Prediction of Positive Circumferential Resection Margin in Oesophageal Cancer

## 5.1 Introduction

The impact of CRM involvement on patient outcome in OC has been widely reported. (Sagar et al. 1993; Chan et al. 2013b; Salih et al. 2013) Although studies have failed to demonstrate the prognostic significance of pathological CRM involvement (Pultrum et al. 2010; Harvin et al. 2012), it is now widely accepted that a positive resection margin is important. (Okada et al. 2016)

Survival statistics from the USA Intergroup 113 RCT were presented in section 1.6.2.2. The trial also investigated the effect of CRM status on survival. Thirty-two percent of patients with a R0 resection were alive and disease-free at 5 years, compared to only 5% survival in those with a R1 resection.

Prediction of pathological CRM involvement could influence treatment selection, potentially improving OS and recurrence rates. Clinicians may have a lower threshold for offering neo-adjuvant therapy to patients at risk. MRI accurately identifies a threatened CRM in rectal cancer (Brown et al. 2003), however early MRI studies in OC encountered initial difficulties because the examination is technically challenging. (Riddell et al. 2006) Alternative methods are required to improve CRM prediction in OC.

PET/CT is predominately used to exclude distant metastases not demonstrated on CT. PET/CT is also used for treatment planning and PET-defined tumour variables including MTL, MTW and $SUV_{max}$ are prognostic indicators of survival and treatment response. (Roedl et al. 2009; Hatt et al. 2011b) There is currently limited evidence investigating the association between PET-defined tumour variables and CRM involvement. PET-defined tumour variables may provide additional predictive value when assessing the CRM.

Therefore, this study investigated the additional value of MTL, MTW and $SUV_{max}$ to predict pathological CRM involvement, compared with EUS and CT T-stage. The prognostic significance of a positive CRM was also assessed.

## 5.2 Materials and Methods

### 5.2.1 Patient Cohort

A retrospective cohort study was conducted in consecutive patients with biopsy-proven OC treated between 1st March 2012 and 31st July 2015. Patients were identified for inclusion at the Regional Upper GI Cancer MDT and deemed to have potentially curable disease. All patients underwent PET/CT examination in PETIC and had surgical resection (with or without neo-adjuvant therapy) in the centralised service. Patients were excluded from the study if the patient had incomplete staging (n=11), salvage oesophagectomy after dCRT (n=3) or an 'open-and-close' procedure

(n=12). Following exclusions, 117 patients were included in the study. Details of the clinical management are found in section 2.1.2.

### 5.2.2 Radiological Staging

Radiological staging was classified by the TNM 7$^{th}$ edition. Protocols for CT, EUS and PET/CT examinations can be found in section 2.2.

### 5.2.3 PET-Defined Tumour Variables

Methods for quantifying PET-defined tumour variables are described in section 2.3. Non-avid tumours were recorded with a value of 0.

### 5.2.4 Histopathological Analysis

CRM status was assessed using the RCPath definition in section 1.7.4.3. (Mapstone 2007) (Fig. 5.2.1) Only the RCPath definition is used in the UK, and a comparison with the American CAP definition is not performed in this study. TRG was assigned according to the Mandard Classification in patients who received neo-adjuvant therapy. (Mandard et al. 1994)

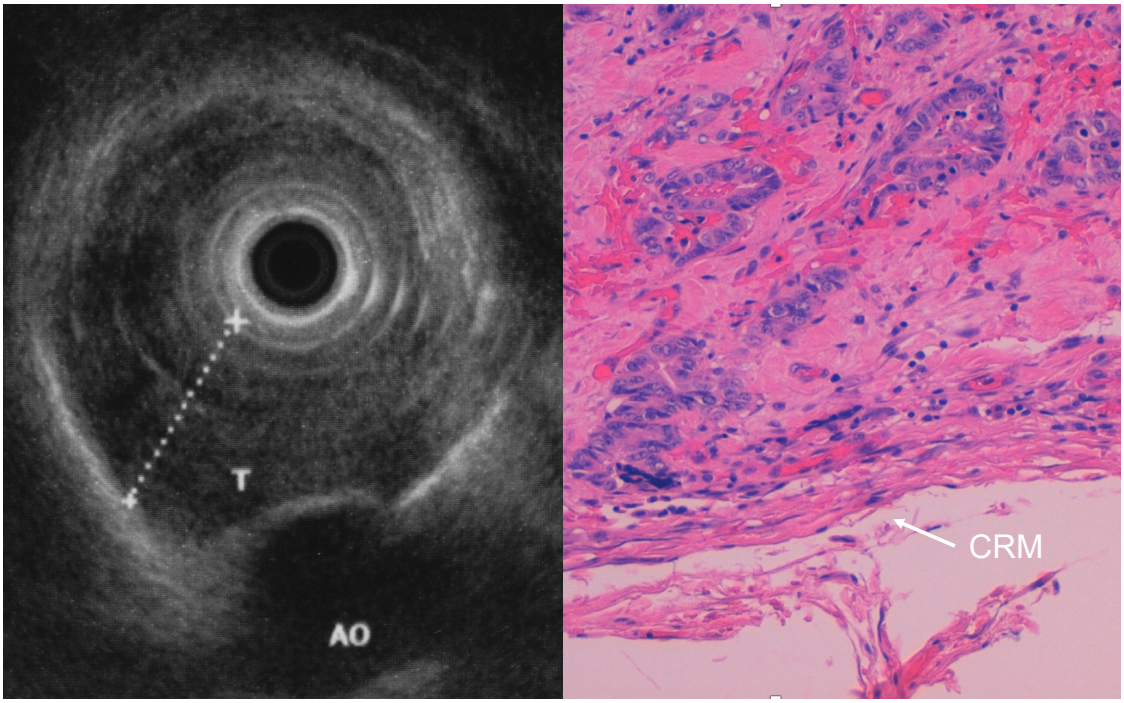



Figure 5.2.1. A selected radial EUS image (left) showing a mid-oesophageal tumour (T) in close proximity to the anterior surface of the descending thoracic aorta (AO). The EUS image was acquired using an Olympus MH-908 blind oesophagoprobe (Key Med, Southend, UK). The tumour thickness is marked with calipers. A corresponding histopathological slide of resected tumour (right) shows malignant cells at the CRM, stained with Haematoxylin & Eosin, indicating an R1 resection. The image was acquired at medium-power magnification (50x).

### 5.2.5 Statistical analysis

Categorical variables were summarised as frequency (percentage) and continuous variables as median (range). Chi-square tests assessed differences between EUS T-stage, CT T-stage, TRG and treatment type with CRM status. EUS and CT T-stage was separated into ≤T2 vs ≥T3 prior to analysis given that relatively few patients present with T1 and T2 tumours. Mann-Whitney U tests assessed differences between MTL, MTW and $SUV_{max}$ with CRM status. Multi-variate analysis was performed by entering the 5 variables into a binary logistic regression model. The model was powered by EPV ratio, with an event defined as a positive CRM. (Peduzzi et al. 1996) A log-rank test assessed differences in OS between CRM status.

## 5.3 Results

Patient characteristics are detailed in Table 5.3.1. The median age of the cohort was 64.0 years (range 24-78). Median survival was 36.0 months (95% CI 24.1-47.9) and 2-year OS was 55.4%. The radiological and pathological TN classification of disease is detailed in Table 5.3.2. No patients in the cohort were classified as having M1 disease following radiological staging investigations.

Table 5.3.1. Characteristics of Patient Cohort

| Demographic | Frequency (%) |
|---|---|
| Gender | |
| Male | 102 (87.2) |
| Female | 15 (12.8) |
| Histology | |
| Adenocarcinoma | 110 (94.0) |
| Squamous Cell Carcinoma | 6 (5.1) |
| Neuro-endocrine | 1 (0.9) |
| Degree of Differentiation | |
| Well | 13 (11.1) |
| Moderate | 40 (34.2) |
| Poor | 57 (48.7) |
| GX | 7 (6.0) |
| Radiological T-stage | |
| T1 | 12 (10.3) |
| T2 | 13 (11.1) |
| T3 | 83 (70.9) |
| T4a | 9 (7.7) |
| Radiological N-stage | |
| N0 | 58 (49.6) |
| N1 | 41 (35.0) |
| N2 | 14 (12.0) |
| N3 | 4 (3.4) |

| | |
|---|---|
| Treatment Type | |
| NACT | 67 (57.3) |
| Surgery Alone | 31 (26.5) |
| NACRT | 19 (16.2) |
| Tumour Regression Grade | |
| TRG 1 | 9 (10.5) |
| TRG 2 | 10 (11.6) |
| TRG 3 | 11 (12.8) |
| TRG 4 | 32 (37.2) |
| TRG 5 | 24 (27.9) |
| Operation Type | |
| Trans-hiatal oesophagectomy | 56 (47.9) |
| Ivor-Lewis oesophagectomy | 35 (29.9) |
| Total gastrectomy | 22 (18.8) |
| 3-stage oesophagectomy | 3 (2.5) |
| Oesophago-gastrectomy | 1 (0.9) |
| Circumferential Resection Margin | |
| Negative | 66 (56.4) |
| Positive | 51 (43.6) |

Table 5.3.2. Radiological and Pathological TN Staging Classification

| Frequency (%) | CT | EUS | PET/CT | Pathology |
|---|---|---|---|---|
| T0 | 0 (0.0) | 0 (0.0) | 0 (0.0) | 9 (7.7) |
| T1 | 12 (10.3) | 12 (10.3) | 0 (0.0) | 18 (15.4) |
| T2 | 22 (18.8) | 13 (11.1) | 0 (0.0) | 9 (7.7) |
| T3 | 73 (62.4) | 83 (70.9) | 0 (0.0) | 71 (60.7) |
| T4a | 10 (8.5) | 9 (7.7) | 0 (0.0) | 10 (8.5) |
| TX | 0 (0.0) | 0 (0.0) | 117 (100.0) | 0 (0.0) |
| Total | 117 (100.0) | 117 (100.0) | 117 (100.0) | 117 (100.0) |
| N0 | 70 (59.8) | 67 (57.2) | 81 (69.2) | 48 (41.0) |
| N1 | 31 (26.5) | 36 (30.8) | 27 (23.1) | 26 (22.2) |
| N2 | 12 (10.3) | 12 (10.3) | 8 (6.8) | 30 (25.6) |
| N3 | 2 (1.7) | 2 (1.7) | 1 (0.9) | 13 (11.1) |
| NX | 2 (1.7) | 0 (0.0) | 0 (0.0) | 0 (0.0) |
| Total | 117 (100.0) | 117 (100.0) | 117 (100.0) | 117 (100.0) |

The positive CRM rate in patients treated with NACT, surgery alone and NACRT was 50.7%, 38.7% and 26.3%, respectively. There was no significant difference in positive CRM rates between these treatments ($X^2$ 4.001, df 2, p=0.135).

Most tumours were staged ≥T3 by EUS and CT (78.6% and 70.9%, respectively) with relatively few early cancers (T1 & T2). (Table 5.3.2) The median MTL was 4.8 cm (range 0.0-8.8), the median MTW was 2.4 cm (0.0-4.8) and the median $SUV_{max}$ was 11.1 (0.0-70.9).

A chi-square test demonstrated EUS ≥T3 tumours were more likely to have a positive CRM than EUS ≤T2 tumours ($X^2$ 4.962, df 1, p=0.026). (Table 5.3.3) CT ≤T2 vs ≥T3, MTL, MTW and $SUV_{max}$ were not significantly associated with CRM status (p=0.161, 0.852, 0.605 and 0.413, respectively). In addition, the TRG was significantly associated with CRM status ($X^2$ 14.042, df 4, p=0.007).

Table 5.3.3. Association of EUS T-stage Groups and CRM involvement

| Frequency (%) | CRM negative | CRM positive | Total |
|---|---|---|---|
| EUS ≤T2 | 19 (16.2) | 6 (5.1) | 25 (21.3) |
| EUS ≥T3 | 47 (40.3) | 45 (38.4) | 92 (78.7) |
| Total | 66 (56.5) | 51 (43.5) | 117 (100.0) |

EUS ≤T2 vs ≥T3, CT ≤T2 vs ≥T3, MTL and MTW and $SUV_{max}$ were entered in a binary logistic regression model. (Table 5.3.4) The EPV ratio was 10.2. EUS ≤T2 vs ≥T3 was significantly and independently associated with CRM involvement (HR 5.188, 95% CI 1.265-21.273, p=0.022).

Table 5.3.4. Results of Multi-Variate Binary Logistic Regression Model

| Variable | Hazard Ratio (95% CI) | p-value |
|---|---|---|
| EUS ≤T2 vs ≥T3 | 5.188 (1.265-21.273) | 0.022 |
| CT ≤T2 vs ≥T3 | 1.163 (0.398-3.397) | 0.783 |
| PET MTL | 0.633 (0.684-1.273) | 0.652 |
| PET MTW | 0.836 (0.445-1.570) | 0.578 |
| $SUV_{max}$ | 0.989 (0.940-1.040) | 0.655 |

There was a significant difference in OS for CRM status ($X^2$ 4.920, df 1, p=0.027). (Fig. 5.3.1) The mean OS for patients with negative CRM resections was 39.6 months (95% CI 34.5-44.7) compared to 30.9 months (25.6-36.2) for those with a positive CRM.

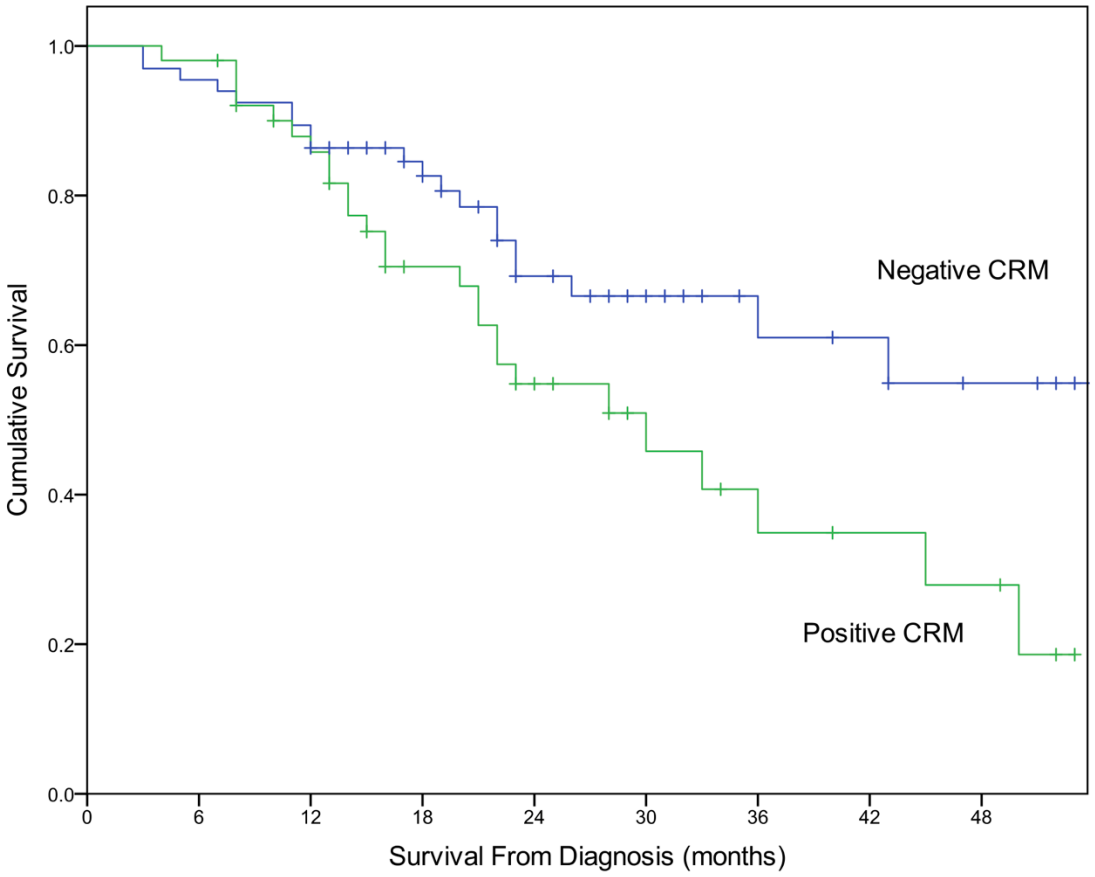



Figure 5.3.1. Significant difference in cumulative survival according to CRM status ($X^2$ 4.920, df 1, p=0.027). Patients with a positive CRM have worse OS. The blue line represents patients with a negative CRM and the green line patients with a positive CRM.

## 5.4 Discussion

EUS ≥T3 is a significant, independent predictor of CRM involvement. These results highlight the continued benefit of EUS in the OC staging pathway and support previously published results from our centre, which demonstrated that EUS ≥T3 has an increased risk of CRM involvement compared to tumours ≤T2. (Reid et al. 2012) The recruitment period for Reid et al ended in February 2012, therefore these results provide some internal validation in a new, independent cohort of patients. Studies investigating the association of radiological staging investigations and CRM are limited in frequency. This study has shown that PET-defined tumour variables may not have any additional value for predicting pathological CRM involvement.

EUS provides the most accurate T-stage assessment. (Tangoku et al. 2012) Pooled sensitivities of 82-92% and an accuracy of 83% are described. (van Vliet et al. 2007; Puli et al. 2008) EUS benefits from superior contrast resolution compared to PET and CT. PET is unlikely to provide sufficient detail to predict CRM involvement, due to its inherently limited spatial resolution. (Kinahan and Fletcher 2010) Similarly, CT is poor at differentiating individual layers of the oesophageal wall. CT T-stage was not significantly associated with CRM status in this study.

Identification of a threatened CRM from radiological staging investigations is likely to benefit patient outcome. In the UK, PET/CT and EUS are generally only performed prior to treatment initiation and are not repeated post-neo-adjuvant therapy as in other countries such as the USA. Two strategies for individualising treatment exist; the first

is the decision to use neo-adjuvant therapy if the CRM is threatened and the second is the decision to operate post-neo-adjuvant therapy. One study did not show any benefit when surgical resection was performed following a complete response on PET/CT. (Monjazeb et al. 2010) The former strategy of predicting CRM involvement prior to treatment initiation is more suited to the UK staging pathway, given that PET/CT and EUS are often not repeated. Furthermore, results of this study have shown that patients with a good response are significantly less likely to have a positive CRM following resection.

Treatment selection can influence the positive CRM rate. NACRT may increase the number of R0 resections compared to NACT, with R0 rates of 87.5% and 92.0% described in the literature. (Reid et al. 2012; van Hagen et al. 2012) This current study did not demonstrate a significantly different positive CRM rate between treatments, but there were small numbers of patients in the NACRT group.

The positive CRM rate in this patient cohort is high but comparable to those quoted in the National Oesophago-Gastric Cancer Audit. (National Oesophago-Gastric Cancer Audit 2016) A trans-hiatal approach was employed in 47.9% of patients in this cohort. This is significantly higher than the national rate of 4%, which could in turn explain the relatively high positive CRM rate. The surgeons in Cardiff employ this technique because the population of patients on which they operate tend to have significant co-morbidities attributable to the effects of poor lifestyle and deprivation. (Blake et al. 2017) As described in section 1.6.2.1, morbidity rates are reduced following trans-hiatal resection compared to a trans-thoracic approach.

*Limitations*

A confounding factor in this analysis is the technique of the surgeon and the approach employed. The surgical approach used was considered the best for the patient and most likely to result in a positive outcome. In line with national data, the two most common types of oesophagectomy were used; trans-thoracic (Ivor-Lewis) and trans-hiatal. The surgeons work together within a centralised system so their techniques are likely to be similar. Individual surgical data has not been included in this thesis.

There is significant heterogeneity in this patient cohort. Patients with differing stage of disease, treatment type, histology and response to treatment were included. In general, OC patients are a heterogeneous cohort and introduction of some sample heterogeneity into research studies can be unavoidable. This reflects the intention-to-treat basis of clinical research, but can introduce potential bias into the results. Three treatment types are included in this study, which have differing effects on CRM status, as evidenced by the CROSS trial. (van Hagen et al. 2012) However, only 19 patients had a complete (TRG 1, n=9) or excellent (TRG 2, n=10) response to neo-adjuvant therapy, which suggests that the disease did not change significantly in the majority of cases. EUS examinations were performed by different endosonographers, which may cause variability in T-stage accuracy, but again adds weight to the generalisability of the results.

## 5.5 Conclusion

A positive CRM is an important prognostic indicator of survival. Prior knowledge of a threatened CRM would assist clinicians with management decisions. This study has shown that EUS ≥T3 is a significant independent predictor of a positive CRM and validates previous research from our institution, but PET-defined tumour variables are unlikely to add additional predictive value regarding CRM status.

# Chapter 6. Comparison of PET and EUS Length of Disease and Potential Impact of Difference on Treatment Planning in Patients with Oesophageal Cancer

## 6.1 Introduction

Primary tumour length (TL) is commonly reported following upper GI endoscopy, CT, EUS and PET/CT staging investigations. (Allum et al. 2011) Of more critical importance is the estimated LoD, defined as the cranio-caudal length of primary tumour plus any LNMs. Assessment of treatment options, including suitability for dCRT, relies on assessment of LoD at staging. A discrepancy in LoD between imaging modalities could affect clinical decision-making and subsequent treatment planning. Inappropriate radical treatment may be initiated in unsuitable patients, or potentially beneficial therapy could be withheld from those that may respond.

There is now significant interest in the use of PET imaging to assist radiotherapy planning, particularly in OC. (Ward et al. 2016) Localisation of the gross tumour volume (GTV) in radiotherapy planning relies on accurate localisation of the TL and LoD. Moreover, there has been a decline in EUS use nationally, making delineation of the GTV more reliant on PET and CT. (National Oesophago-Gastric Cancer Audit 2016)

Therefore, this study tested the hypothesis that significant differences exist between PET and EUS LoD. These differences could impact on clinical decision-making and

treatment planning, especially in cases where EUS is not performed. The primary aim of this study was to investigate differences in PET and EUS LoD in patients with OC. The secondary aim was to assess their prognostic significance.

## 6.2 Materials and Methods

### 6.2.1 Patient Cohort

Consecutive patients staged between January 1st 2011 and December 31st 2014 with biopsy proven oesophageal or GOJ tumours were considered for this study. All EUS examinations were performed by the same operator. In total, 222 patients were considered for inclusion. Exclusion criteria were a non FDG-avid primary tumour (n=30), a tumour too stenotic to be passed with the endoscope (n=13), LoD not recorded in the EUS report (n=18) and patients lost to follow-up (n=1). Following exclusions, 160 patients were included in the study.

### 6.2.2 PET and EUS LoD

Protocols for PET/CT and EUS can be found in section 2.2. The method of PET LoD measurement is found in section 2.3. All EUS examinations were performed by the same highly experienced operator (AR) with a published track-record, to ensure consistency in LoD measurement. EUS LoD was recorded by AR during the procedure, and documented in the final report. EUS LoD was calculated as the

length of endoscope insertion relative to the incisors between proximal and distal tumour and/or LNMs, recorded in cm. (Rice and Roberts 2003)

### 6.2.3 Outcome Data

The secondary outcome of the study is OS. Methods of survival data collection are described in section 2.1.3.

### 6.2.4 Statistical Analysis

Continuous data are expressed as median (range) and categorical data as frequency (percent). A Bland-Altman analysis was used to assess the level of agreement between PET and EUS LoD. (Bland and Altman 1986) The mean difference (PET minus EUS) and 95% limits of agreement (LA) were calculated. A difference of more than 2 cm between PET and EUS LoD is considered clinically significant for radiotherapy planning, therefore the proportion of cases with a clinically significant difference was also calculated. (Crosby et al. 2013) A non-parametric Wilcoxon signed rank test was used to assess differences between PET and EUS LoD. Univariate survival analysis was performed with the log-rank test according to the life-table method of Kaplan-Meier. (Kaplan and Meier 1958) Multi-variate analysis was performed by entering age (years), stage group (I, II, III or IV), treatment (curative vs palliative), PET LoD (cm) and EUS LoD (cm) into a Cox Regression model. (Cox 1972) Model power was based on the EPV ratio.

## 6.3 Results

Patient Characteristics are detailed in Table 6.3.1. The median age of the cohort was 66.0 years (range 24-83). The median OS of the cohort was 20.0 months (95% CI 16.2-23.8) and median follow-up was 40.0 months (35.1-44.9).

Table 6.3.1. Baseline Characteristics of Patient Cohort

| Patient Characteristic | Frequency (%) |
|---|---|
| Gender | |
| Male | 124 (77.5) |
| Female | 36 (22.5) |
| Histology | |
| Adenocarcinoma | 115 (71.9) |
| Squamous Cell Carcinoma | 41 (25.6) |
| High-grade Dysplasia | 2 (1.3) |
| Neuro-endocrine | 1 (0.6) |
| Undifferentiated | 1 (0.6) |
| Tumour Location | |
| Oesophagus | 96 (60.0) |
| GOJ | 64 (40.0) |

| | |
|---|---|
| EUS T-stage | |
| T1 | 5 (3.1) |
| T2 | 14 (8.8) |
| T3 | 97 (60.6) |
| T4a | 33 (20.6) |
| T4b | 11 (6.9) |
| EUS N-stage | |
| N0 | 54 (33.8) |
| N1 | 49 (30.6) |
| N2 | 35 (21.8) |
| N3 | 22 (13.8) |
| PET/CT N-stage | |
| N0 | 81 (50.6) |
| N1 | 51 (31.8) |
| N2 | 22 (13.8) |
| N3 | 6 (3.8) |
| PET/CT M-stage | |
| M0 | 144 (90.0) |
| M1 | 14 (8.8) |
| MX | 2 (1.2) |
| Treatment Type | |
| NACT | 37 (23.1) |
| dCRT | 35 (21.9) |
| Surgery alone | 17 (10.6) |
| NACRT | 14 (8.8) |
| EMR | 1 (0.6) |
| Palliative | 56 (35.0) |

Boxplot representation of the measurements showed that PET tended to yield smaller LoD measurements compared to EUS. (Fig 6.3.1)

Additionally, a Wilcoxon signed rank test demonstrated a significant difference between PET and EUS LoD (Z= -7.021, p<0.001). EUS LoD was more than 2 cm longer than PET LoD in 61 cases (38.1%). In 8 cases (5.0%), PET LoD was more than 2 cm longer than EUS LoD.

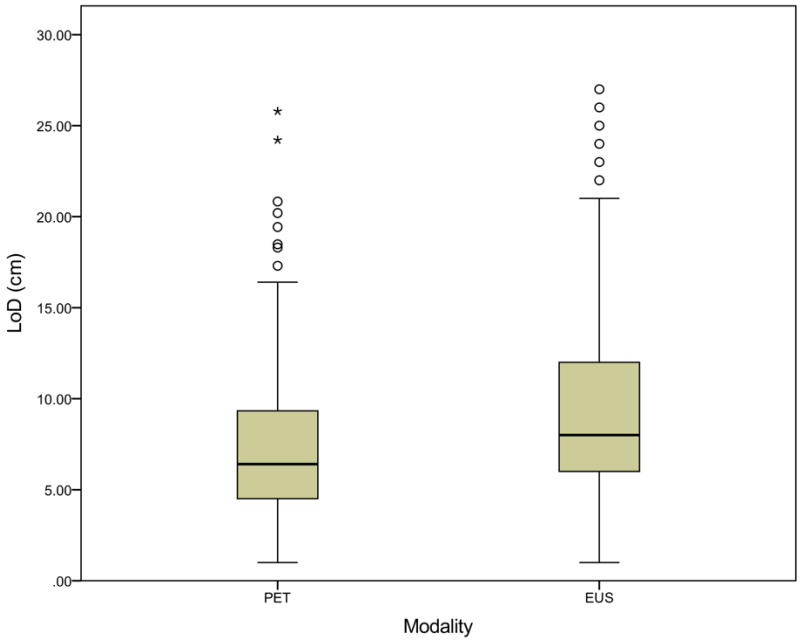



Figure 6.3.1. Boxplot representation of PET and EUS LoD measurements. The median PET LoD was 6.4 cm (standard deviation (SD) 4.5, interquartile range (IQR) 4.5-9.4, range 1.0-25.8), respectively. The median EUS LoD was 8.0 cm (SD 5.7, IQR 6.0-12.0, range 1.0-27.0), respectively.

Bland-Altman analysis demonstrated substantial variation in PET and EUS LoD measurements. (Fig. 6.3.2) The Bland Altman analysis indicates that the 95% LA between PET and EUS represented a level of disagreement that is potentially clinically significant, suggesting that PET and EUS measurements should not be used inter-changeably. (Fig. 6.3.3)

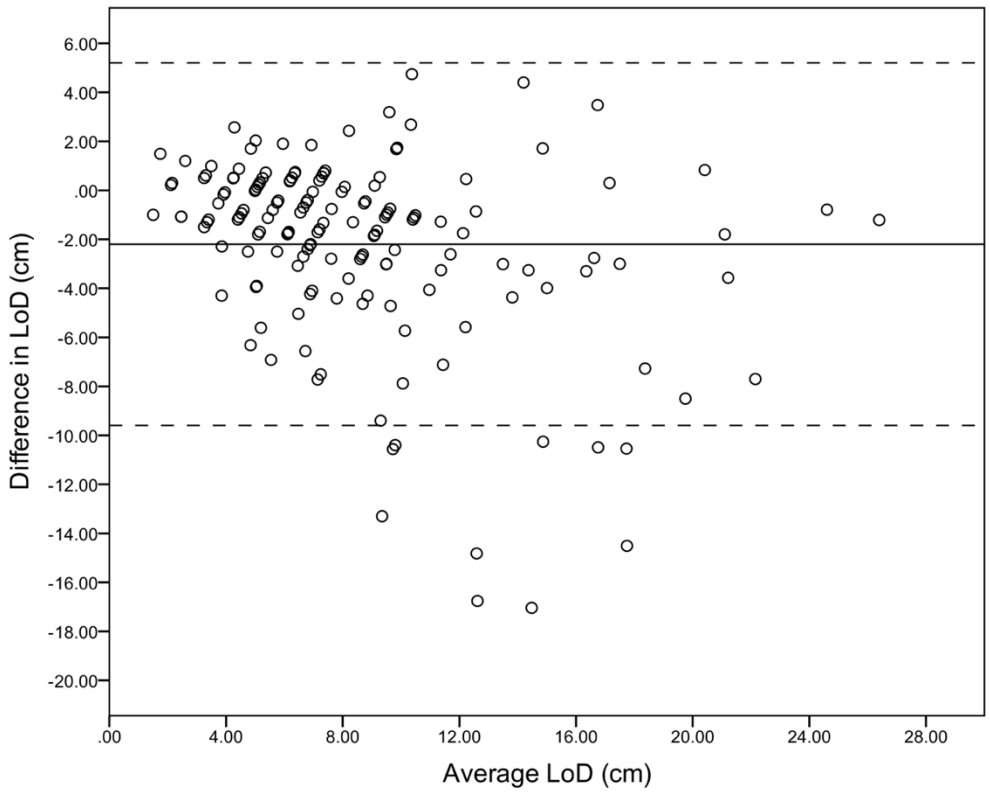



Figure 6.3.2. Bland-Altman plots demonstrating limited agreement in PET and EUS measurements of LoD. Mean difference (solid line) and 95% LA (dashed lines) are displayed. The mean difference between PET and EUS was -2.2 cm (SD 3.8, 95% LA -9.6 to 5.2), suggesting PET yielded smaller LoD measurements than EUS.

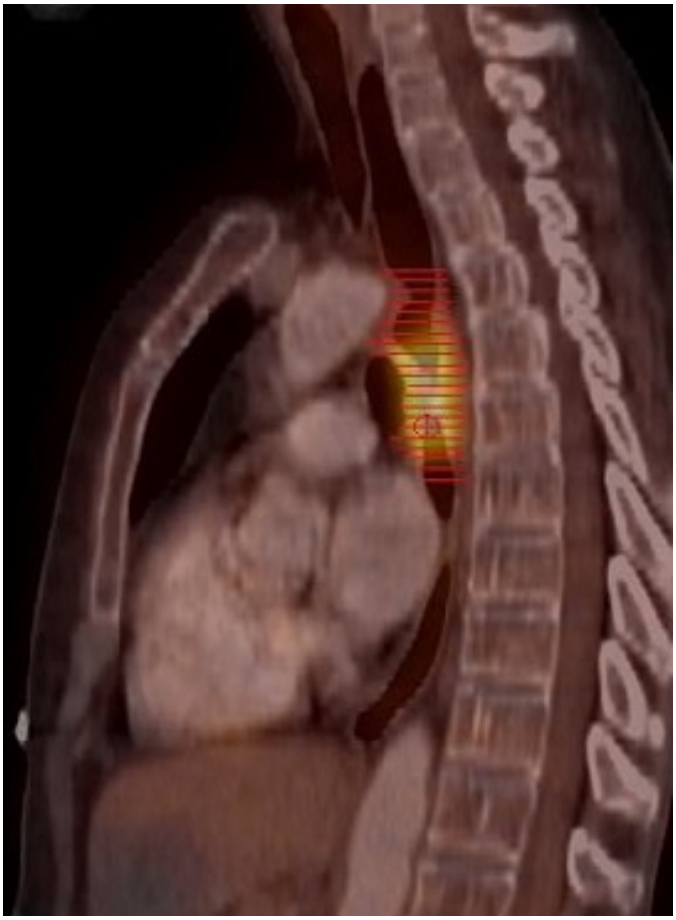


Figure 6.3.3. Selected sagittal fused PET/CT radiotherapy planning image demonstrating a FDG-avid mid-oesophageal SCC staged T3 N1 with EUS. A peri-tumoural LNM did not increase the total LoD measurement. The horizontal red lines delineate the EUS LoD measurement, with each line representing a 5 mm interval. The PET LoD measured 6.5 cm, whereas the EUS LoD was recorded as 10 cm, indicating non-FDG avid tumour at proximal and distal margins.

In univariate analysis, PET LoD (HR 1.076, 95% CI 1.037-1.115, p<0.001) and EUS LoD (HR 1.059, 95% CI 1.028-1.091, p<0.001) were significantly associated with OS. There were significant differences in OS between upper and lower quartiles of PET LoD (13.0 months if >9.4 cm and 29.0 months if <4.5 cm, p<0.001) and EUS LoD (13.0 months if >12.0 cm and 29.0 months if <6.0 cm, p=0.002).

However, in multivariate analysis, PET and EUS LoD were not independently associated with OS. (Table 6.3.2) The EPV ratio was 22.2.

Table 6.3.2. Results of the Multivariate Cox Regression Model

| | | | 95% Confidence Interval | |
|---|---|---|---|---|
| Variable | p-value | Hazard Ratio | Lower | Upper |
| Age | 0.026 | 1.024 | 1.003 | 1.045 |
| Stage Group | 0.002 | 1.728 | 1.227 | 2.433 |
| Treatment | <0.001 | 0.414 | 0.265 | 0.648 |
| PET LoD | 0.787 | 0.992 | 0.933 | 1.054 |
| EUS LoD | 0.996 | 1.000 | 0.950 | 1.053 |

## 6.4 Discussion

This study has demonstrated significant differences between PET and EUS LoD in patients with OC. These results are important for treatment option assessment, which can be complex in OC. Both PET and EUS LoD were significantly associated with OS on univariate analysis, but were not independent predictors, findings which concur with the few published data available. (Twine et al. 2010; Hatt et al. 2011b; Davies et al. 2012; Foley et al. 2014a).

Selection of patients for surgical management, neo-adjuvant treatments or dCRT partly relies on accurate assessment of disease extent, often gained from PET/CT and EUS. The LoD is an important measurement that can influence these decisions. These results suggest that PET tends to under-measure LoD compared to EUS.

An accepted maximum LoD for consideration of radiotherapy is 10 cm as described in the SCOPE trial series protocols. (Mukherjee et al. 2015) There is often more concern about length of irradiated volume in the neoadjuvant setting however, leading to a more conservative approach in this scenario. Inaccuracies in LoD estimation could affect patient selection for NACRT. (Hurt et al. 2011)

In terms of radiotherapy planning, a difference of more than 2 cm between PET and EUS LoD is considered clinically significant. (Crosby et al. 2013) Most modern oesophageal radiotherapy planning protocols allow a margin of 2 cm from GTV to clinical target volume (CTV) to allow for microscopic spread along the oesophagus.

Differences in LoD of more than 2 cm could lead to a significant risk of a geographical miss if the PET measurement alone had been used. In this study, up to 38.1% of cases were at risk of a geographical miss.

Delineation of target volumes for radiotherapy planning in OC is increasingly guided by metabolic activity of the primary tumour and regional nodes on PET/CT. (Mukherjee et al. 2015) In addition to clinical information, PET images are most commonly viewed alongside the planning CT. The oesophageal GTV can be difficult to define on CT alone because of submucosal spread, the propensity for skip lesions and poor differentiation of tumour from normal oesophagus. Accurate definition of GTV has become even more important given the growing trend for reduced margins combined with increased conformity of treatment volumes and use of advanced techniques such as Volumetric Modulated Arc Therapy (VMAT). Some centres use fusion techniques, but inaccuracies can be introduced if patient positioning differs between diagnostic and planning examinations.

Centres that utilise EUS for radiotherapy planning have reported satisfactory recurrence rates with few edge-of-field relapses. (Button et al. 2009) However, EUS is occasionally unavailable at the time of radiotherapy planning, often due to non-traversable tumour, patient choice or increasing service pressures. Limited information can still be acquired from a non-traversable tumour, such as the proximal extent of tumour and assessment of visible lymph nodes, but the maximum LoD may not be fully appreciated in these cases, thus increasing the risk of edge-of-field recurrence.

If PET alone is relied upon to guide delineation of GTV, all available diagnostic information, including the upper GI endoscopy report, diagnostic CT and PET/CT images, should be used together to plan radiotherapy. The temptation to outline FDG-avid regions of disease alone should be resisted because it is vital to include disease identified on all available imaging modalities. Usually, the most recent imaging is the radiotherapy planning CT and areas of adjacent, non-avid oesophageal wall thickening should be included in the GTV. This approach is also recommended in the recent SCOPE2 trial radiotherapy planning protocol. (SCOPE2 2016)

EUS assesses local disease extent more accurately than PET due to its superior contrast and spatial resolution and is preceded by video endoscopy, which provides the opportunity to visualise subtle areas of proximal or distal extension of disease that would not normally be detected on CT or PET. Submucosal infiltration is also better assessed with EUS. (Thosani et al. 2012) Physiological FDG-uptake in the oesophagus or stomach is often located adjacent to the tumour, creating an 'avidity gradient' which can cause error in measurement. (Fig. 2.3.1) Another limitation of PET is the suboptimal differentiation of adjacent peri-tumoural lymph node metastases from the primary tumour. (Konski et al. 2005) However, PET/CT can add useful information in patients with non-traversable tumours, or in cases where there is involvement of the GOJ. Identification of nodal disease distant to the primary tumour can also be assessed. Overall, these results support the combined use of PET and EUS in radiation treatment planning of OC.

It has been suggested that EUS use should be more focused in OC. EUS is an invasive procedure with risk of serious complications and is operator dependent. In many centres access to EUS is limited, which can impact on patient pathways and time to treatment. This is supported by evidence that EUS use is declining. According to the National Oesophago-Gastric Cancer Audit (NOGCA) data, 47.5% of patients with OC had a staging EUS completed in 2016, compared to 62% reported in 2013. (National Oesophago-Gastric Cancer Audit 2016) A large single-centre study showed minimal benefit of EUS versus the potential risk of complications in the majority of patients staged T2-T4a on CT. (Findlay et al. 2015) The authors suggest that EUS use should be limited to early stage OC and the assessment of resectability in more advanced cases. The additional utility of EUS for accurate radiotherapy planning was not discussed in this paper and should be an additional consideration given the increasing use of NACRT in recent years.

Interestingly, LoD is also used to define patient eligibility criteria into clinical radiotherapy trials. Examples of these criteria include the NEOSCOPE (Mukherjee et al. 2015) and SCOPE1 (Hurt et al. 2011) trials, both of which stipulated a maximum LoD (8 cm and 10 cm, respectively). Currently, the Neo-AEGIS trial (Keegan et al. 2014) differs by stipulating a maximum TL of 8 cm on any imaging modality.

*Limitations*

As for other studies investigating tumour length measurements on imaging, the true pathological length is unknown, making accurate comparison of imaging modalities

difficult. Cancer resections specimens can shrink up to 50% in size which is an important consideration when comparing radiological and pathological measurements. (Siu et al. 1986) Only measurements from single observers for both PET and EUS were analysed in this study, which maintains consistent methodology, but does not allow assessment of inter-observer variability. Future research should focus on the impact of inter-observer variability on treatment decision-making in patients with OC. Identical settings were used when measuring LoD on the PET MIP images. Some tumours with high intensity variation may not have displayed optimally, which potentially introduced error in measurement. However, this methodology was adopted to ensure consistency between patients. In addition, the patient population was relatively heterogeneous, which reflects the observational nature of the study. As a result, the patients included in this study received different treatments. Treatment was included in the multi-variate analysis as curative and palliative groups only. Curative therapies were combined as the numbers in some treatment groups were relatively small.

## 6.5 Conclusion

This retrospective study has demonstrated significant differences in PET and EUS LoD measurements recorded from OC staging investigations. These measurements showed prognostic significance on univariate analysis but were not independent predictors of survival. Differences in these measurements could potentially impact on clinical-decision making and radiotherapy treatment planning. These results highlight the continued benefit of EUS in the OC staging and treatment pathway, particularly adding information in patients requiring radiotherapy.

# Chapter 7. Development and Validation of a Prognostic Model Incorporating Texture Analysis Derived from Standardised Segmentation of PET in Patients with Oesophageal Cancer

## 7.1 Introduction

Multiple sub-clonal populations of cells are known to co-exist within tumours. (Gerlinger et al. 2012) As described in section 1.10.1, texture analysis data could act as a surrogate markers of underlying tumour heterogeneity. This, in combination with traditional staging methods, may improve clinical decision tools and optimise treatment pathways. (Aerts et al. 2014)

Retrospective studies have investigated the ability of PET texture analysis to predict treatment response and survival in different solid cancers including lung, oesophageal, cervical and head & neck. (Orlhac et al. 2014; Hatt et al. 2015; van Rossum et al. 2016a) A large multi-centre study including 1,019 patients with lung and head & neck cancer conducted retrospective radiomic analysis on external datasets and demonstrated the additional benefit of CT texture analysis in the staging pathway. Radiomic data were combined with genomic data to produce a prognostic signature resulting in improved prognostic performance compared to traditional TNM staging alone. (Aerts et al. 2014)

This study aimed to investigate the additional prognostic value of PET texture analysis compared with the current staging methods, by developing a prognostic model in

patients with OC. This study also aimed to calculate a prognostic score which can stratify patients accordingly and perform internal validation of the model in an independent cohort of patients.

## 7.2 **Materials and Methods**

### *7.2.1* Patient Cohort

This is a retrospective cohort study of consecutive patients with biopsy-proven OC, including GOJ tumours, radiologically staged between 16th September 2010 and 31st July 2015.

Overall, 550 patients were considered for inclusion. Exclusion criteria were non- or poorly FDG-avid tumours [$SUV_{max}$ < 3 (n=60)], an MTV < 5mL (n=52), histology other than adenocarcinoma or SCC (n=21), a synchronous primary malignancy (n=7) or an oesophageal stent in situ (n=7).

Following exclusions, 403 patients were included and chronologically separated into 2 independent cohorts. The first (development) cohort included 302 patients radiologically staged between 16th September 2010 and 15th September 2014. The second (validation) cohort included 101 patients radiologically staged between 16th September 2014 and 31st July 2015.

### *7.2.2* Treatment Protocols

Patients either had surgery alone, NACT or NACRT prior to surgery, dCRT or palliative therapy. Further details are found in section 2.1.2.

### 7.2.3 Data Preparation and PET Segmentation

Data preparation and PET segmentation with ATLAAS was performed as detailed in sections 2.4.2 and 2.4.3.

### *7.2.4* Prognostic Variables

Nineteen variables were included in the Cox regression model. Age (number of years) and stage group (I A or B = 1, II A or B = 2, III A, B or C = 3, IV = 4), were included. Treatment was divided into curative (=1) and palliative (=2) groups prior to data analysis. PET-STAT metrics were analysed as described in section 2.4.5.

### *7.2.5* Transformation of Variables

Visual inspection of continuous variable histograms was performed before model development to assess for normal distribution and skewness. Specific normality tests were not used but logarithmic transformation of variables with significant long-tails was performed prior to analysis to reduce the leverage created from outlying data. Four texture variables were transformed; TLG [log(TLG)] Histogram Energy

[log(Energy)], Coarseness [log(Coarseness)] and Homogeneity [log(Homogeneity)]. Repeat inspection of the transformed histograms revealed the 4 variables had a more normalised distribution. (Fig. 7.2.1)

### *7.2.6* Metabolic Tumour Volume and Texture Metrics

An important consideration in texture analysis is the range of tumour volumes that are assessed. Tumours with small volumes may provide redundant texture information due to highly correlated variables. (Wu et al. 2016) Some authors have suggested excluding tumours with MTV less than 5 mL. (Orlhac et al. 2014) Therefore, patients with MTVs < 5 mL were excluded from analysis.

### *7.2.7* Outcome Data

The primary outcome of the study is OS. Details of outcome data are found in section 2.1.3.

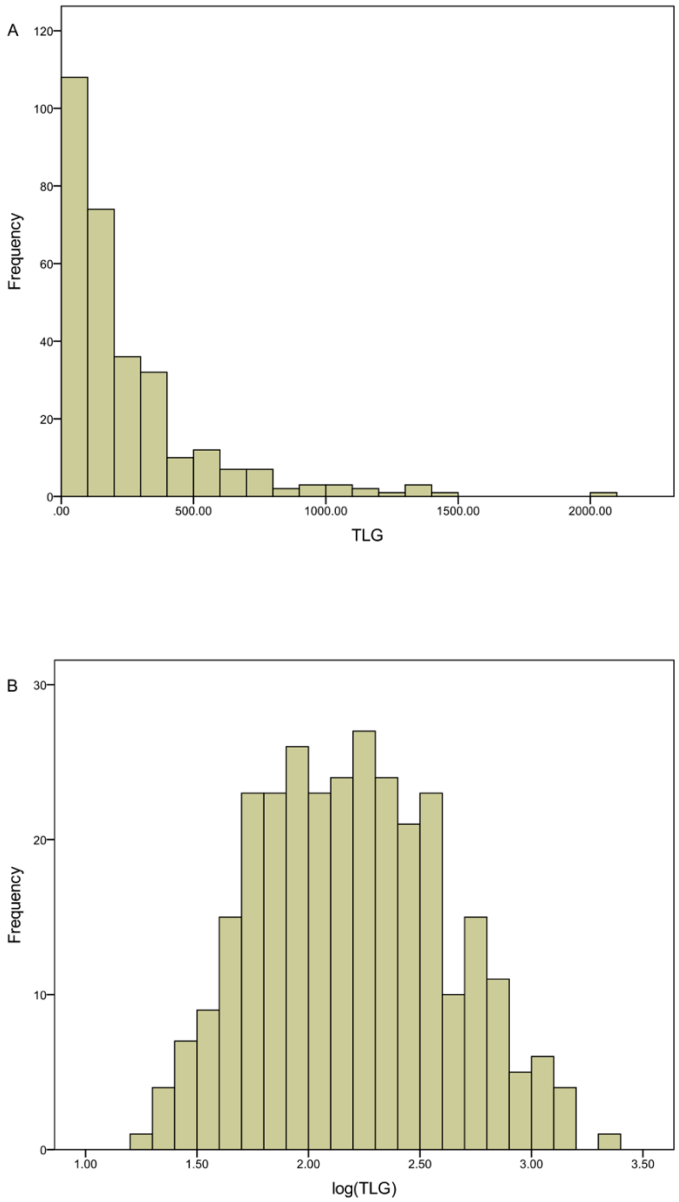



Figure 7.2.1. An example of logarithmic transformation of a variable. In this case, the histogram of TLG (A) in the development cohort showed a non-normal distribution with positive skew. Logarithmic transformation of TLG into log(TLG) was performed (B). log(TLG) showed a more normal distribution so was included in the Cox Regression model.

### *7.2.8* Statistical Analysis

Categorical variables are described as frequency (percent) and continuous variables as median (range) and differences assessed with appropriate non-parametric tests. Differences between cohorts were assessed using either a Chi-square or Mann-Whitney U test. Cumulative survival was calculated by the Kaplan-Meier life-table method. A Cox Regression Model with backward conditional method was constructed by an experienced medical statistician. Model power was based on an event-to-variable ratio (EPV), recommended to be a minimum level of 10. (Peduzzi et al. 1996) EPV is defined as the ratio of number of patient deaths compared to number of variables in the model. The prognostic score was calculated by summation of the products of variables and their corresponding parameter estimate. Using this, patients were separated into quartiles and a log-rank test evaluated significant differences in OS. The effect of curative or palliative treatment on the performance of the prognostic score was assessed with a test of interaction. Furthermore, the Akaike information criterion (AIC) statistic evaluated the estimated quality of 3 incremental models; 1) a model including age, radiological stage group and treatment; 2) a model including these variables plus newer prognostic indicators $SUV_{max}$, $SUV_{mean}$ and MTV; and 3) a model including the additional texture metrics. (Akaike 1974) The model with a lowest AIC value is considered the better model. Internal validation of the prognostic model was performed retrospectively in a separate cohort of patients.

## 7.3 Results

### 7.3.1 Patient Cohorts

Baseline characteristics of patients included in the development and validation cohorts are detailed in Table 7.3.1. The median OS of the development and validation cohorts was 16.0 months (95% CI 13.8-18.2) and 14.0 months (95% CI 10.4-17.6), respectively. Median follow-up was 43.0 months (95% CI 35.3-50.7) in the development cohort and 17.0 months (95% CI 15.7-18.3) in the validation cohort. Overall 1- and 2-year survival in the development cohort was 66.9% and 33.3%, respectively and 1-year OS in the validation cohort was 57.4%. Classification of radiological EUS and PET/CT TNM stage are detailed in Table 7.3.2.

All EUS examinations were performed in 3 centres by 4 experienced endosonographers. Patients with tumours too stenotic to be crossed at EUS were unable to be fully staged, therefore final pre-treatment radiological stage relied on a combination of findings from PET/CT and CT investigations. EUS was not attempted if the decision to treat the patient palliatively was made after PET/CT. In the development cohort, EUS T-stage was assigned in 227 patients (75.2%). EUS N-stage was assigned in 221 patients (73.2%). EUS was not attempted due to M1 disease on PET/CT, or was incomplete due to a non-traversable tumour in 81 cases. In the validation cohort, EUS staging was completed in 78 patients (77.2%). EUS was not attempted due to M1 disease on PET/CT, or was incomplete due to a non-traversable tumour in 23 cases.

Table 7.3.1. Baseline Characteristics of Patients in Development and Validation Cohorts

| Frequency (%) | Development Cohort (n=302) | Validation Cohort (n=101) | p-value |
|---|---|---|---|
| Median Age | 67.0 years (Range 39-83) | 69.0 years (Range 39-84) | 0.179 |
| Gender (M: F) | 227 (75.2): 75 (24.8) | 78 (77.2): 23 (22.8) | 0.676 |
| Histology | | | 0.956 |
| Adenocarcinoma | 237 (78.5) | 79 (78.2) | |
| SCC | 65 (21.5) | 22 (21.8) | |
| Tumour Location | | | 0.003 |
| Oesophagus | 192 (63.6) | 47 (46.5) | |
| Upper Third | 6 (3.1) | 3 (6.4) | |
| Middle Third | 53 (27.6) | 10 (21.3) | |
| Lower Third | 133 (69.3) | 34 (72.3) | |
| GOJ | 110 (36.4) | 54 (53.5) | |
| Siewert I | 41 (37.3) | 24 (44.5) | |
| Siewert II | 30 (27.3) | 18 (33.3) | |
| Siewert III | 39 (35.4) | 12 (22.2) | |
| Stage Groups | | | 0.238 |
| Stage I | 17 (5.6) | 2 (2.0) | |
| Stage II | 56 (18.5) | 24 (23.8) | |
| Stage III | 160 (53.1) | 57 (56.4) | |
| Stage IV | 69 (22.8) | 18 (17.8) | |

| | | | |
|---|---|---|---|
| Treatment | | | 0.624 |
| Curative | 158 (52.3) | 50 (49.5) | |
| Surgery Alone | 24 (15.2) | 4 (8.0) | |
| NACT | 67 (42.4) | 23 (46.0) | |
| NACRT | 13 (8.2) | 7 (14.0) | |
| dCRT | 54 (34.2) | 16 (32.0) | |
| Palliative | 144 (47.7) | 51 (50.5) | |
| | | | |
| Overall Survival | | | <0.001 |
| Alive | 70 (23.2) | 43 (42.6) | |
| Dead | 232 (76.8) | 58 (57.4) | |

Table 7.3.2. TNM Classification of PET/CT and EUS Staging Investigations in Development and Validation Cohorts

| Frequency (%) | Development Cohort (n=302) | Validation Cohort (n=101) | p-value |
|---|---|---|---|
| EUS T-stage | | | 0.656 |
| T1 | 3 (1.3) | 1 (1.3) | |
| T2 | 15 (6.6) | 4 (5.1) | |
| T3 | 161 (71.0) | 57 (73.1) | |
| T4a | 35 (15.4) | 15 (19.2) | |
| T4b | 12 (5.3) | 1 (1.3) | |
| TX | 1 (0.4) | 0 (0.0) | |
| Total | 227 (100.0) | 78 (100.0) | |
| EUS N-stage | | | 0.003 |
| N0 | 87 (39.4) | 28 (35.9) | |
| N1 | 59 (26.7) | 36 (46.2) | |
| N2 | 47 (21.3) | 12 (15.4) | |
| N3 | 28 (12.7) | 2 (2.5) | |
| Total | 221 (100.0) | 78 (100.0) | |
| PET/CT N-stage | | | 0.519 |
| N0 | 126 (41.7) | 46 (45.5) | |
| N1 | 96 (31.8) | 25 (24.8) | |
| N2 | 62 (20.5) | 25 (24.8) | |
| N3 | 18 (6.0) | 5 (4.9) | |
| Total | 302 (100.0) | 101 (100.0) | |
| PET M-stage | | | 0.556 |
| M0 | 228 (75.5) | 81 (80.2) | |
| M1 | 72 (23.8) | 19 (18.8) | |
| MX | 2 (0.7) | 1 (1.0) | |
| Total | 302 (100.0) | 101 (100.0) | |

### *7.3.2* Prognostic Model Development

The final step of the prognostic model is presented in Table 7.3.3. Descriptive statistics for all calculated PET metrics in the development cohort are detailed in Table 7.3.4. There were 232 events and 19 variables in the model, providing 12.2 EPV. In addition to known important prognostic factors in OC (age, radiological stage and treatment), the model identified 3 texture metrics that were independently and significantly associated with survival. The significant variables were log(TLG), log(Histogram Energy) and Histogram Kurtosis. Their inclusion in the model illustrates their additional prognostic value compared with current prognostic factors. TLG is calculated as the product of $SUV_{mean}$ and MTV. (Wahl et al. 2009) Histogram Energy (Orlhac et al. 2014) was calculated using Equation 1:

$$Histogram\ Energy = \sum_{i} \big(P(i)\big)^2$$

Eq.1

where $P(i) = \frac{N_i}{N}$, with $N_i$ the number of voxels of intensity I, and N, the total number of voxels.

Histogram Kurtosis (Orlhac et al. 2014) was calculated using Equation 2:

$$Histogram\ Kurtosis = \frac{\frac{1}{N}\sum_i (I(i) - \mu)^4}{\left(\frac{1}{N}\sum_i (I(i) - \mu)^2\right)^2}$$

Eq.2

where N is the number of voxels in the image, I(i) is the positive intensity value in the 3D matrix and μ is the mean intensity value.

Table 7.3.3. Results of the Cox Regression Model

| Prognostic Variable | p-value | Parameter Estimate | Hazard Ratio | 95 % Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower | Upper |
| TNM Stage | <0.001 | 0.397 | 1.49 | 1.20 | 1.84 |
| Treatment | <0.001 | -1.094 | 0.34 | 0.24 | 0.47 |
| Age | 0.001 | 0.024 | 1.02 | 1.01 | 1.04 |
| log(Histogram Energy) | 0.011 | -1.320 | 0.27 | 0.10 | 0.74 |
| log(TLG) | 0.013 | 1.748 | 5.74 | 1.44 | 22.83 |
| Histogram Kurtosis | 0.017 | 0.198 | 1.22 | 1.04 | 1.44 |

Table 7.3.4. Results of PET Variables and Texture Metrics in Development Cohort

| Metric | Mean | 95% Confidence Interval | | Minimum Value | Maximum Value |
|---|---|---|---|---|---|
| | | Lower | Upper | | |
| $SUV_{max}$ | 16.55 | 15.57 | 17.54 | 3.56 | 59.97 |
| $SUV_{mean}$ | 9.14 | 8.58 | 9.71 | 2.07 | 35.06 |
| MTV | 25.80 | 23.33 | 28.27 | 5.04 | 132.47 |
| log(TLG) | 2.20 | 2.15 | 2.24 | 1.30 | 3.30 |
| Standard Deviation | 2.51 | 2.35 | 2.66 | 0.46 | 9.73 |
| Histogram Entropy | 3.82 | 3.81 | 3.84 | 2.94 | 4.09 |
| log(Histogram Energy) | 4.74 | 4.68 | 4.81 | 3.52 | 6.37 |
| Histogram Skewness | 0.58 | 0.53 | 0.62 | -0.31 | 2.82 |
| Histogram Kurtosis | 2.81 | 2.70 | 2.92 | 1.72 | 12.69 |
| log(Coarseness) | -1.96 | -1.99 | -1.93 | -2.78 | -1.29 |
| log(Homogeneity) | -0.53 | -0.55 | -0.52 | -1.03 | -0.21 |
| Entropy | 5.34 | 5.25 | 5.43 | 3.10 | 7.25 |
| Dissimilarity | 5.05 | 4.73 | 5.38 | 0.97 | 23.18 |
| Intensity Variability | 18.04 | 16.59 | 19.48 | 2.76 | 90.81 |
| Large Area Emphasis | 258.77 | 78.03 | 439.51 | 1.46 | 20512.36 |
| Zone Percentage | 42.46 | 40.43 | 44.50 | 1.73 | 87.86 |

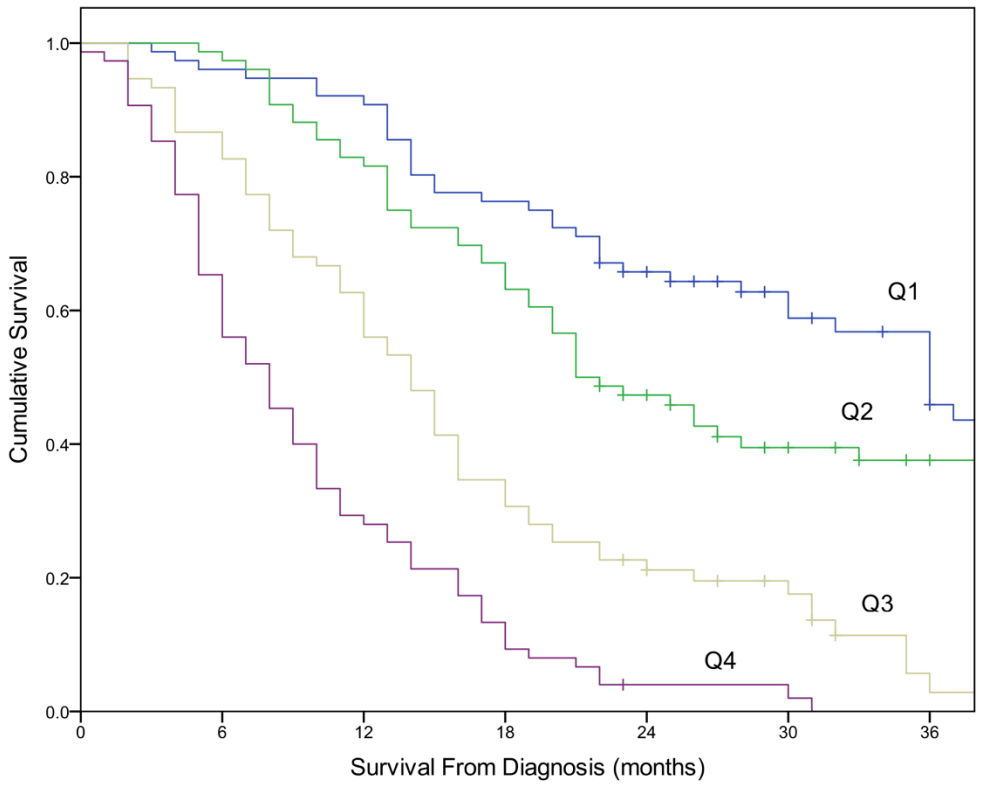
### *7.3.3* Prognostic Score Calculation

Equation 3 was used to calculate the prognostic score in the development cohort.

(Stage Group*0.397) - (Treatment*1.094) + (Age*0.024) - (log(Histogram Energy)*1.320) + (log(TLG)*1.748) + (Histogram Kurtosis*0.198)

Eq.3

The median score of quartile 1 was -0.73 (n=76, range -1.66 to -0.45), quartile 2 was -0.14 (n=76, -0.45 to 0.29), quartile 3 was 0.76 (n=75, 0.31 to 1.06) and quartile 4 was 1.38 (n=75, 1.08 to 2.15). There was a significant difference in OS between quartiles ($X^2$ 143.14, df 3, $p<0.001$). (Fig 7.3.1) Median OS of quartiles 1 to 4 was 36.0 months (95% CI 31.1-40.9), 21.0 months (16.1-25.9), 14.0 months (11.7-16.3) and 8.0 months (5.9-10.1), respectively. The interaction test revealed no statistical difference in performance of the prognostic score between curative and palliative treatments ($X^2$ 1.344, df 1, $p=0.246$).




Patients at Risk, n

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Total | 302 | 250 | 194 | 135 | 92 | 61 | 38 |
| Quartile 1 | 76 | 72 | 69 | 57 | 46 | 30 | 20 |
| Quartile 2 | 76 | 74 | 62 | 48 | 32 | 22 | 17 |
| Quartile 3 | 75 | 62 | 42 | 23 | 13 | 9 | 1 |
| Quartile 4 | 75 | 42 | 21 | 7 | 1 | 0 | 0 |

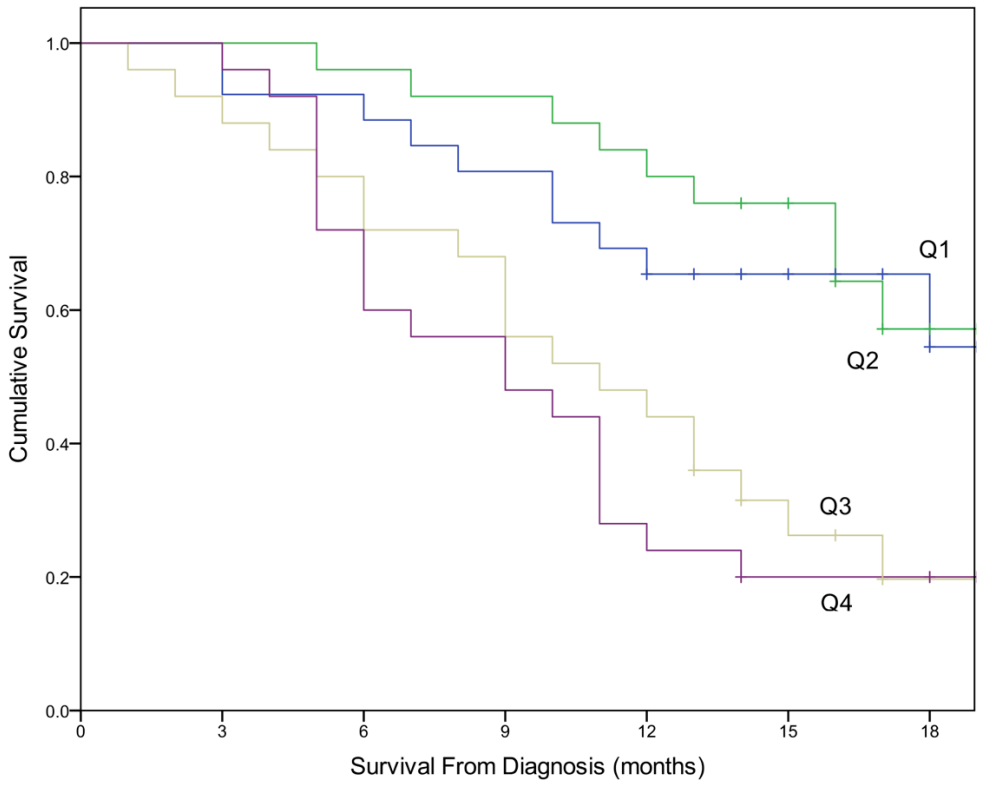
Figure 7.3.1. Kaplan-Meier plot demonstrating cumulative survival curves of prognostic score quartiles in the development cohort ($X^2$ 143.142, df 3, p<0.001). Patients in Q1 had better OS and patients in Q4 had the worst OS. Q1 quartile; Q2 quartile 2; Q3 quartile 3; Q4 quartile 4.

### *7.3.4* Comparison of Estimated Model Performance

The AIC of the traditional model including radiological stage group, treatment and age was 2247.693. The AIC of the model which also included $SUV_{max}$, $SUV_{mean}$ and MTV was also 2247.693. The AIC of the development prognostic model including additional texture metrics is 2238.007, which is the lowest value. This suggests that incorporation of PET variables and texture metrics improves current prognostic models in OC.

### *7.3.5* Internal Validation of Prognostic Model

The prognostic model was applied to the validation cohort. Again, there was a significant difference in OS between patient quartiles ($X^2$ 20.621, df 3, $p<0.001$). (Fig. 7.3.2) Results of PET metrics obtained from the validation cohort are detailed in Table 7.3.5. Mean OS of patients in quartiles 1 and 2 was 16.6 months (95% CI 13.9-19.3) and 17.4 months (95% CI 15.4-19.4), respectively. Patients in quartile 1 had lower mean OS than those in quartile 2, but the difference between quartiles was not significant ($X^2$ 0.219, df 1, $p=0.640$). The median OS for quartile 3 and 4 was 11.0 months (6.1-15.9) and 9.0 months (4.1-13.9). Three of 26 (11.5%) patients were treated with palliative intent in quartile 2, and 2 of 25 (8.0%) patients were treated with curative intent in quartile 3. The AIC of the validation model including PET variables and texture metrics was lower (464.671) than models including radiological stage group, treatment and age (470.420), and $SUV_{max}$, $SUV_{mean}$ and MTV (470.420), respectively.




Patients at Risk, n

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Total | 101 | 94 | 79 | 67 | 52 | 31 | 13 |
| Quartile 1 | 26 | 24 | 23 | 19 | 15 | 10 | 4 |
| Quartile 2 | 25 | 24 | 23 | 22 | 20 | 13 | 5 |
| Quartile 3 | 25 | 22 | 18 | 14 | 11 | 5 | 2 |
| Quartile 4 | 25 | 24 | 15 | 12 | 6 | 3 | 2 |

Figure 7.3.2. Kaplan-Meier plot demonstrating cumulative survival curves of prognostic score quartiles in the validation cohort ($X^2$ 20.621, df 3, $p<0.001$). Q1 quartile; Q2 quartile 2; Q3 quartile 3; Q4 quartile 4.

Table 7.3.5. Results of PET Variables and Texture Metrics in Validation Cohort

| Metric | Mean | 95% Confidence Interval | | Minimum Value | Maximum Value |
|---|---|---|---|---|---|
| | | Lower | Upper | | |
| $SUV_{max}$ | 17.13 | 15.02 | 19.23 | 4.57 | 70.97 |
| $SUV_{mean}$ | 9.69 | 8.45 | 10.93 | 2.97 | 39.59 |
| MTV | 25.35 | 21.20 | 29.50 | 5.13 | 100.27 |
| log(TLG) | 2.20 | 2.12 | 2.29 | 1.31 | 3.44 |
| Standard Deviation | 2.54 | 2.21 | 2.87 | 0.27 | 11.11 |
| Histogram Entropy | 3.83 | 3.80 | 3.85 | 3.27 | 4.05 |
| log(Histogram Energy) | 4.76 | 4.64 | 4.88 | 3.45 | 6.49 |
| Histogram Skewness | 0.63 | 0.56 | 0.70 | -0.03 | 1.90 |
| Histogram Kurtosis | 2.79 | 2.65 | 2.93 | 1.85 | 6.14 |
| log(Coarseness) | -1.98 | -2.04 | -1.93 | -2.66 | -1.52 |
| log(Homogeneity) | -0.54 | -0.57 | -0.52 | -1.04 | -0.27 |
| Entropy | 5.37 | 5.21 | 5.52 | 3.45 | 7.37 |
| Dissimilarity | 5.44 | 4.69 | 6.18 | 1.62 | 26.47 |
| Intensity Variability | 18.00 | 15.40 | 20.61 | 4.43 | 75.37 |
| Large Area Emphasis | 107.65 | 38.78 | 176.51 | 1.47 | 2467.89 |
| Zone Percentage | 43.21 | 39.51 | 46.91 | 6.55 | 88.13 |

## 7.4 Discussion

This study has developed a prognostic model which provides new and important results for OC staging. Internal validation of the model demonstrated a continued difference in OS ($p<0.001$) between quartiles, in an independent cohort of patients. The results of this study show that PET texture analysis may enhance the prognostic TNM staging model in OC.

The prognostic model has identified 3 PET metrics; log(TLG), log(Histogram Energy) and Histogram Kurtosis, that are significantly and independently associated with OS. These metrics have added value over and above currently known prognostic factors; age, radiological stage and treatment. These findings indicate the additional value of novel texture analysis methods in modern staging pathways, which was confirmed with the AIC statistic. Improved risk-stratification could identify sub-groups of patients in which a certain treatment may improve OS (Moons et al. 2009b), or where a therapeutic intervention may be ineffective or harmful. (Blazeby et al. 2000)

According to the model, patients with increased log(TLG) and Histogram Kurtosis, and reduced log(Histogram Energy), have an increased likelihood of mortality. Raised TLG represents larger, more FDG-avid tumours. The correlation of Histogram Kurtosis and log(Histogram Energy) suggest that tumours with less intensity variation have a worse prognosis. This is an unexpected finding, since it is thought that tumours with more intensity variation result in poorer outcome. Further studies

correlating texture features with underlying tumour biology are required to fully understand the interpretation of these metrics. (Orlhac et al. 2016)

The AIC was identical for traditional TNM and models including SUV and MTV in both the development and validation cohorts. This suggests that SUV and MTV have no additional prognostic value over current staging methods. However, this study has not been designed to test this hypothesis and cannot draw this conclusion.

Our findings concur with other studies in which texture metrics derived from histograms demonstrated significant associations with OS, stage of disease and likelihood of treatment response in OC. (Tixier et al. 2011; Ganeshan et al. 2012; Yip et al. 2015) However, such studies included smaller sample sizes and used different texture analysis software packages.

In this study, the texture metrics were derived using the ATLAAS algorithm and a standardised workflow was implemented to ensure reproducible and consistent methods. The benefit of ATLAAS is that the best fitting PET-AS method is selected in each individual case from a range of segmentation methods that are built into the ATLAAS algorithm. Commonly used PET-AS methods built into ATLAAS algorithm Adaptive Thresholding, Fuzzy C-means (FCM) and Region-Growing (RG) methods. (Berthon et al. 2016)

*Limitations*

As this study is retrospective, treatment was included in the model and simplified into 2 groups; curative and palliative. However, the test for interaction showed that the prognostic score could be used in both curative and palliative cohorts with no significant difference in performance. This prognostic model excludes patients with a MTV of less than 5 mL because the quality of the additional data obtained from these models is uncertain. (Wu et al. 2016) This criterion excludes 11.6% of potential patients from this study. Another prognostic model including small tumour volumes should be developed for these patients but this model is applicable to many patients with FDG-avid oesophageal tumours.

ATLAAS was originally designed and tested on patients with FDG-avid head & neck tumours. It is also applicable to other FDG-avid tumour sites and validation studies are on-going at our institution. Although a new version of ATLAAS had not specifically been designed for this prognostic OC model, visual inspection of the segmented tumour was performed in each case to ensure an appropriate contour had been produced.

Texture metrics are dependent on several parameters. (Galavis et al. 2010) Standardisation of texture analysis is essential for multi-centre comparison studies and development of externally validated prognostic models, but have not been established yet. (Lambin et al. 2015; Leijenaar et al. 2015) The technical implementation of each metric, segmentation method used, scan acquisition, image

smoothing, influence of quantisation and reconstruction parameters all influence texture analysis results. (Doumou et al. 2015; Leijenaar et al. 2015; Gillies et al. 2016) There are also limitations specific to PET images, given the relatively large voxel volume and presence of noise artefact. (Cook et al. 2014)

Limitations of the PET-STAT software may exist. PET-STAT was developed with CERR and MatLab based functions, which are validated and commonly used image-processing tools. Comparison with other open-source and commercially available texture analysis software packages has not been performed in this thesis. Different software packages calculate a variety of image features. In addition, the nomenclature and implementation of each metric can vary.

## 7.5 Conclusion

This large study has developed and validated a prognostic model that demonstrates the additional value of PET texture analysis in OC staging. Three PET metrics; log(TLG), log(Histogram Energy) and Histogram Kurtosis were identified as potentially important variables. These metrics were derived using ATLAAS, a novel machine-learning method designed to optimise and standardise image segmentation. This prognostic model requires further internal and external validation but may be used as a 'bench-mark' for further studies investigating the value of PET texture analysis in OC. This study highlights the additional benefit of quantitative imaging techniques in cancer staging, which have the potential to improve patient risk stratification.

# Chapter 8. General Discussion

As highlighted in section 1.3, the prognosis of OC is poor. There has been a limited increase in survival rates in recent decades. Significant improvements in staging and treatment are required to reduce the burden of this disease in the population. This cancer, identified by Cancer Research UK as one of 4 cancers of substantial unmet need (Cancer Research UK 2017), requires vital research to be conducted in an attempt to improve survival rates.

Radiological staging investigations are extremely important in OC, largely influencing clinical decision-making, patient selection and treatment planning. The focus of this thesis was to investigate these staging investigations, assessing the additional value of novel prognostic variables over and above the current staging system. This research is important because improved assessment of disease status at diagnosis will enhance treatment decisions, and ultimately improve OS.

The rationales, aims and hypotheses of this thesis were presented in section 1.12 and are re-visited in this chapter. To highlight the significance and original contributions of this thesis, a summary of each results chapter is provided below. The importance of the research, implications to clinical practice, the limitations and strengths of the research and suggestions for future work are discussed.

## 8.1 Significance of Results

Chapter 3

This chapter aimed to evaluate the accuracy of N-staging and provide radiological-pathological correlation of LNMs. This research is necessary because treatment is largely affected by the diagnosis of LNMs and the sensitivity and specificity of staging investigations are known to be suboptimal. (van Vliet et al. 2008) The important finding in this chapter was that the accuracy of CT, EUS and PET/CT N-staging was poor. Current staging investigations are unreliable for differentiating N0 from N+ disease. Further analysis in patients with discordant N-staging (patients staged N0 radiologically but N+ pathologically) demonstrated that the poor accuracy was most likely attributable to a significant number of small LNMs in this group, which cannot be detected by conventional imaging. Eighty-two percent of LNMs measured < 6 mm and 44% were < 2 mm (defined as micro-metastases). These findings have substantial clinical implications given the apparent under-staging of disease and highlight a major requirement for improved LNM prediction. Future research should focus on new methods of predicting the likelihood of LNMs. Diagnostic decisions could be optimised by incorporating imaging biomarkers, such as texture analysis of the primary tumour, into LNM prediction models. MRI may provide an alternative N-staging modality. Research studies have demonstrated variable diagnostic ability, with sensitivity, specificity and accuracy ranging between 38-62%, 68-85% and 64-77%, respectively. These results are comparable to CT, EUS and PET/CT but continued developments in functional MRI scanner technology may yield further

diagnostic improvements. (Wu et al. 2003; Nishimura et al. 2006) Given the significance of the chapter results, clinicians may be more inclined to offer neo-adjuvant therapy prior to surgery to treat potentially undetected nodal disease.

Chapter 4

The results of a published prognostic model in OC patients staged N0 by PET/CT underwent internal validation in Chapter 4. (Foley et al. 2014b) Validation of prognostic models is important but not commonly performed (section 2.6.2). The model was originally developed because EUS use in the UK is declining (section 6.4) with utilisation and reliance on PET/CT increasing. However, the limited spatial resolution of PET provides difficulty differentiating peri-tumoural nodes from the primary tumour. Therefore, the role of EUS in patients staged N0 on PET/CT was assessed. In addition, LNMs are known to be a major prognostic indicator in the general OC population but less evidence exists in this N0 sub-group. Therefore, the prognostic significance of pLNMs was also investigated.

The findings of the original study could not be validated because EUS N-stage and EUS N0 vs N+ were no longer significantly associated with OS on multi-variate analysis. However, EUS T-stage was independently and significantly associated with OS. This finding is important because the data demonstrates that EUS use in the OC staging pathway should continue. The benefit of EUS in the OC staging pathway is further supported by evidence from chapters 5 and 6, which are discussed below.

There was a significant difference in OS between pN0 and pN+ groups in patients staged N0 on PET/CT, which confirms the importance of accurate N-staging.

Chapter 5

CRM involvement is regarded as an important prognostic factor in patients undergoing surgical resection. Better prediction of CRM involvement would greatly assist oncologists with treatment decisions, and surgeons with resection planning. The prediction of CRM involvement using MRI in rectal cancer staging has been widely adopted by the international community. (Brown and Daniels 2005) Staging MRI has been investigated in OC but is technically challenging and results to date have not matched the performance of MRI in rectal cancer. Despite this, prediction of a threatened CRM would greatly benefit clinicians, as the CROSS trial showed that NACRT significantly reduces the R1 resection rate. (van Hagen et al. 2012) Given the results of the CROSS trial, patients with a threatened CRM are more likely to be offered NACRT in the future.

Therefore, chapter 5 investigated the additional value of PET-defined tumour variables to predict pathological CRM involvement, compared with EUS and CT. As PET/CT is already part of the staging pathway, simple metrics measured from the examination that predict CRM involvement would benefit patients at minimal extra cost. There was no additional predictive value of PET-defined tumour variables, but the study confirmed that EUS ≥T3 was an independent predictor of CRM involvement. These findings validate results of a previously published study from our institution

(Reid et al. 2012) and add further evidence that EUS should continue to be utilised as part of the multi-modality staging pathway.

Chapter 6

As described in section 6.4, EUS use in the UK is declining. As a result, treatment planning (surgery and radiotherapy) is becoming more reliant on CT and PET/CT. Given the poor N-stage accuracy and low sensitivity of staging investigations (particularly PET/CT) demonstrated in chapter 3, the risk of missing undetected LNMs during treatment planning is significant. An important consideration during radiotherapy planning is the LoD. Traditionally, a 2 cm expansion of GTV is performed during radiotherapy planning to include microscopic spread of disease along the oesophagus. (Crosby et al. 2013) The risk of edge of field relapse increases if a geographical miss exists. Therefore, the difference between PET and EUS LoD measurements was assessed, which is an important consideration if PET/CT is used for radiotherapy planning alone.

A significant difference was found between PET and EUS LoD measurements. In addition, there was substantial variation between measurements. Both variables had prognostic significance on univariate analysis, but were not independent predictors of OS. Inter-observer variability of LoD measurements was not investigated in this thesis, but could potentially affect treatment planning. Future work should evaluate the influence and impact of inter-observer variability on clinical decision-making and treatment planning. A multi-modality approach to staging provides complementary,

but occasionally conflicting data, which can introduce uncertainty for clinicians. Automatically acquired measurements may reduce inter-observer variability, but require accurate tumour and disease delineation.

The differences between PET and EUS LoD had potentially significant implications for decisions regarding patient selection and treatment planning, particularly in patients considered for radiotherapy. These data again support the utilisation of EUS in the staging pathway and treatment planning.

Chapter 7

The additional value of imaging biomarkers has been widely investigated in cancer research. (O'Connor et al. 2017) Validation of imaging biomarkers must be performed using structured methodology to implement their use in clinical practice. As mentioned in section 1.10.2, texture analysis has been applied in several primary tumour sites. The hypothesis is that additional data can be extracted from the tumour and used to improve patient risk-stratification, assisting patients and clinicians with decision-making. The great advantage of texture analysis is that the tumour is imaged in 3D, rather than relying on a small volume of biopsy tissue to characterise a potentially heterogeneous tumour. This technique could have significant clinical benefit.

A prognostic model incorporating PET texture analysis was developed in Chapter 7. The model was then internally validated. The aim of the study was to compare the

additional prognostic performance of a model incorporating texture analysis against current staging methods. The developed model demonstrated 3 image features; log(TLG), log(Histogram Energy) and Histogram Kurtosis, that were independently associated with OS. Internal validation in a new, independent cohort of patients again demonstrated additional prognostic value over and above current staging methods. Statistical methods appropriate for prognostic research were utilised following review of the literature. Sufficient EPV was ensured to reduce the risk of over-fitted models, providing greater confidence in the results of the prognostic models. Results of uni-variate analyses did not affect variable selection for multi-variate regression models. (Moons et al. 2009b) The prognostic model, developed and validated in 403 patients, is currently one of the largest studies of its kind in the literature. These novel results highlight the additional value of advanced quantitative imaging techniques in cancer staging and could have a significant impact on future OC staging techniques.

The 3 image features above require reliability and reproducibility testing. (Yip and Aerts 2016) This is true of all texture analysis studies, a subject which is relatively new and has generated significant interest, prompting an exponential rise in the number of publications and software platforms for generating quantitative data. Multi-centre, international collaborations such as the Image Biomarker Standardisation Initiative (Zwanenburg et al. 2016) are required to standardise texture metrics before incorporating the features into routine clinical practice.

More research is required to understand the underlying biological explanation of imaging phenotype. As discussed throughout this thesis, direct comparison of

imaging with pathology is challenging. It is possible that some imaging features, including histogram metrics, have no biological equivalent. (O'Connor et al. 2017) As molecular characterisation of OC improves, well-designed studies correlating imaging features with underlying tumour biology are required. (Cancer Genome Atlas Research Network 2017)

The limitations of biological validation could be circumvented by associating image features with clinical outcome. (O'Connor et al. 2017) Provided these surrogate imaging biomarkers correlate precisely with outcome, models with clinically important endpoints such as overall survival could be developed using large patient datasets assembled by international multi-centre collaborations. The developed prognostic model will not immediately change clinical practice since further validation is required, but the added value of quantitative imaging variables that enhance current staging methods has been demonstrated. Prediction models incorporating imaging biomarkers may arguably be more clinically useful than prognostic models, as they can inform and guide diagnostic or treatment decisions, rather than simply confirming that prognosis is poor. Platforms to extract imaging biomarkers from staging investigations should be integrated into PACS reporting workstations to facilitate a change in clinical practice. These platforms need to be standardised to ensure reproducible metrics are generated. Clinicians are busy and under pressure from increasing clinical demands, so the technology must be simple and quick to use if hoped to be adopted widely.

## 8.2 General Limitations

Specific limitations relevant to each results chapter are described above, but there are general limitations of the work included in this thesis.

OC patients are a relatively heterogeneous cohort, presenting with different stages of disease, different histological cell types, and who receive varying treatment protocols which are occasionally tailored to individual patient requirements. This heterogeneity affects selection criteria for clinical studies and has influenced inclusion criteria in each results chapter in this thesis. Some sample cohort heterogeneity is often unavoidable when testing hypotheses related to the general population of OC patients. In addition, the retrospective nature of the work can introduce selection and observational bias into the data.

In ideal settings, the patient cohorts used in this thesis would have been more tightly controlled. The most controlled sources of data are RCTs, although trial inclusion criteria occasionally do not reflect 'real-life' scenarios. One method of controlling the heterogeneity of patient cohorts would be to conduct research in the setting of a RCT. For instance, imaging may not be the primary subject of the trial, but a sub-study within an RCT could investigate the predictive performance of texture analysis.

Multiple reporters classified the TNM stage of disease in this thesis, which may have affected the accuracy of staging results. Diagnostic accuracy is difficult to ascertain in OC because surgery is only performed in a minority of patients. Therefore, many

radiological staging results used in this thesis lack a gold-standard comparator. There are several difficulties associated with radiological-pathological comparison studies in OC, including the time-period between pre-treatment imaging and final pathology. Tumour progression and LNM development can occur in patients with no response to neo-adjuvant therapy. Conversely, tumours that have a good or complete response to neo-adjuvant therapy can morphologically regress from the pre-treatment imaging. Given these reasons, radiological-pathological correlation is challenging.

### 8.3 Strengths of Thesis

Despite its limitations, this thesis has many strengths. The staging pathway has not altered during the thesis period which ensures consistent radiological staging techniques. All patients were discussed at the Regional MDT and the management plan was decided by consensus. The Regional MDT covers a large population of over 1.4 million people and is highly experienced in the management of OC. (Karran et al. 2014) With the exception of 47 patients scanned in Cheltenham, all PET/CT examinations were performed in Cardiff using the same scanner and protocol. Histopathological examination was performed by consultant GI pathologists according to guidelines defined by the RCPath. (Mapstone 2007) Resections were all performed as part of a centralised Upper GI cancer service, comprising a group of surgeons that worked together throughout the thesis period.

Tumour segmentation was performed using ATLAAS, a novel machine learning tool that provides accurate standardised PET segmentation and eliminates inter-observer

variability. (Berthon et al. 2016) ATLAAS ensures that the tumour is outlined optimally in each case. A limitation of many PET-AS methods is that they do not work effectively each time, producing contours that are not representative of the primary tumour.

## 8.4 TNM 8th Edition

A new consideration in OC is the introduction of the TNM 8th edition, which was published in 2017 and generally took effect in clinical practice on 1st January 2018. (Rice et al. 2017) The introduction of the 8th edition was not incorporated into this thesis but its influence on staging and patient outcome requires evaluation. Table 8.4.1 details differences between the 7th and 8th editions. (Rice et al. 2017) Clinical TNM (cTNM) (Table 8.4.2) and pathological TNM (pTNM) stage groups have been separated to reflect the fact the pre-treatment staging is largely performed without pathological data and that survival differences exist between the two. As described throughout this thesis, there are limitations when directly comparing cTNM and pTNM, given that a minority of patients undergo surgery and most of these receive some form of neo-adjuvant therapy. (National Oesophago-Gastric Cancer Audit 2016) This may improve accuracy of cTNM but direct comparison studies will remain challenging.

Table 8.4.1. Summary of Differences Between TNM 7th and 8th Editions.

| Stage Categories | Changes from 7th edition |
|---|---|
| pTNM | |
| T | T1 subcategorised as T1a and T1b producing stage subgroups IA and IC for adenocarcinoma and IA and IB for SCC |
| | T2 SCC. Location removed as staging category |
| | T4a includes direct invasion of peritoneum |
| G | G4 “undifferentiated” category eliminated. |
| Location | Siewert type III tumours re-classified as gastric cancer |
| Stage Groups | |
| III | Subgroup IIIC removed |
| IV | Sub-grouped as IVA and IVB |
| ypTNM | Introduced in 8th edition. Not shared with pTNM. Identical groupings for adenocarcinoma and SCC |
| cTNM | Introduced in 8th edition. Not shared with pTNM. Separate groupings for adenocarcinoma and SCC |

Table 8.4.2. Clinical TNM 8th Edition Stage Groups

| Adenocarcinoma | | | | SCC | | | |
|---|---|---|---|---|---|---|---|
| Stage Group | T | N | M | Stage Group | T | N | M |
| 0 | Tis | N0 | M0 | 0 | Tis | N0 | M0 |
| I | T1 | N0 | M0 | I | T1 | N0, N1 | M0 |
| IIA | T1 | N1 | M0 | II | T2 | N0, N1 | M0 |
| IIB | T2 | N0 | M0 | | T3 | N0 | M0 |
| III | T1 | N2 | M0 | III | T1, T2 | N2 | M0 |
| | T2 | N1, N2 | M0 | | T3 | N1, N2 | M0 |
| | T3, T4a | N0, N1 | M0 | | | | |
| IVA | T4b | N0, N1 | M0 | IVA | T4a, T4b | Any N | M0 |
| | Any T | N2, N3 | M0 | | Any T | N3 | M0 |
| IVB | Any T | Any N | M1 | IVB | Any T | Any N | M1 |

One important change from the 7th edition is the re-classification of Siewert type III tumours into gastric cancers. Type III tumours are defined as cancers more than 5 cm distal to the GOJ, even if the GOJ is involved. Type I and II GOJ tumours will continue to be staged using the oesophageal classification. Siewert type III tumours are re-classified as gastric cancers until comprehensive genetic analysis can identify the cells of origin, rather than relying on tumour location. (Hayakawa et al. 2016)

A new regional lymph node map has been produced. (Rice et al. 2017) It was felt the previous lymph node station classification map was problematic because it included lung lymph node stations, some of which were not regional oesophageal lymph nodes (stations 5 (aorto-pulmonary window), 6 (anterior mediastinum) and 10 (hilar)). This re-classification may assist clinical radiologists with N-staging. A higher proportion of patients with lymph nodes in the mediastinum may be up-staged to M1 disease. This could have consequences for clinical management, with fewer patients being offered radical therapy. The clinical effect of the new lymph node map will need to be evaluated.

Patients with N3 disease have been re-classified. These patients arguably have as poor an outcome as those with M1 disease due to the substantial lymph node burden. This revision may also anticipate changes to the new lymph node map. Re-classification of N3 disease from stage IIIC in the 7$^{th}$ edition to IVA in the 8$^{th}$ edition may serve to artificially increase survival statistics in both stage groups. This is coined the 'Will Rogers Phenomenon', after an American comedian described the movement of people from Oklahoma to California, resulting in an increase in average intelligence in both states. (Feinstein et al. 1985)

Another important change in the 8$^{th}$ edition is the introduction of post-neoadjuvant therapy pathologic stage groups (ypTNM). Survival varies depending on the response of the primary tumour to the neo-adjuvant therapy. (Mandard et al. 1994) Assessment of treatment response is the focus of increasing numbers of research studies and clinical trials, with adaptive treatment protocols being introduced. For

example, PET/CT is being used to quantify early chemotherapy response in attempt to guide further treatment in patients with advanced Hodgkin's lymphoma. (Johnson et al. 2016)

## 8.5 Summary

This thesis investigated OC staging and the additional value of novel imaging variables compared to current staging methods. Accuracy of lymph node staging is poor and must be improved. Incorporation of quantitative PET image features added prognostic value during staging. This thesis demonstrates significant radiological prognostic variables that add value in OC management. Future work must focus on improving radiological OC staging techniques. The work in this thesis has contributed towards this aim.

## Publications Resulting from this Work

Foley, K. G. et al. 2017. Impact of Positron Emission Tomography and Endoscopic Ultrasound Length of Disease Difference on Treatment Planning in Patients with Oesophageal Cancer. *Clin Oncol (R Coll Radiol)* 29(11), pp. 760-766.

Foley, K. G. et al. 2018. Development and validation of a prognostic model incorporating texture analysis derived from standardised segmentation of PET in patients with oesophageal cancer. *Eur Radiol* 28(1), pp. 428-436.

* Foley, K. G. et al. 2017. Accuracy of contemporary oesophageal cancer lymph node staging with radiological-pathological correlation. *Clin Radiol* 72(8), pp. e691-e697.

* Foley, K. G. et al. 2016. The 100 most cited articles investigating the radiological staging of oesophageal and junctional cancer: a bibliometric analysis. *Insights Imaging* 7(4), pp. 619-628.

* Foley, K. G. et al. 2014. N-staging of oesophageal and junctional carcinoma: is there still a role for EUS in patients staged N0 at PET/CT? *Clinical Radiology* 69(9), pp. 959-964.

** Included at back of thesis*

## Prizes

* Royal College of Radiologists Ellis-Barnett Prize in 2014 for:

Foley, K. G. et al. 2014. N-staging of oesophageal and junctional carcinoma: is there still a role for EUS in patients staged N0 at PET/CT? *Clinical Radiology* 69(9), pp. 959-964.

** Included at back of thesis*

## Conference Proceedings and Poster Presentations

Foley, K. G. et al. Development of a Prognostic Model Incorporating Texture Analysis Derived from Standardised Segmentation of PET in Patients with Oesophageal Cancer: PV-0323. ESTRO 2017.

Foley, K. G. et al. Prognostic significance of PET texture analysis in oesophageal cancer staging. ESGAR 2016.

Foley, K. G. et al. Positron emission tomography-defined tumour variables to predict lymph node metastases in patients with oesophageal cancer. ESGAR 2016.

Foley, K. G. et al. Positron emission tomography-defined tumour variables to predict pathological T-stage in patients with oesophageal cancer. ESGAR 2016.

Foley, K. G. et al. The use of pet texture analysis to predict lymph node metastases in patients with oesophageal cancer: PO-0702. ESTRO 2016.

## Bibliography

Abraham, S. C. et al. 2007. Duplication of the muscularis mucosae in Barrett esophagus: an underrecognized feature and its implication for staging of adenocarcinoma. *Am J Surg Pathol* 31(11), pp. 1719-1725.

Adams, H. L. and Jaunoo, S. S. 2014. Clinical significance of incidental findings on staging positron emission tomography for oesophagogastric malignancies. *Ann R Coll Surg Engl* 96(3), pp. 207-210.

Aerts, H. J. et al. 2014. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 5, p. 4006.

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Trans Automat Contr* 19(6), pp. 716-723.

Alderson, D. et al. 2015. Neoadjuvant chemotherapy for resectable oesophageal and junctional adenocarcinoma: results from the UK Medical Research Council randomised OEO5 trial (ISRCTN 01852072). *J Clin Oncol* 33(suppl; abstr 4002).

Allisy-Roberts, P. and Williams, J. 2008. Farr's Physics for Medical Imaging. 2nd ed. London: Saunders

Allum, W. H. et al. 2011. Guidelines for the management of oesophageal and gastric cancer. *Gut* 60(11), pp. 1449-1472.

Allum, W. H. et al. 2009. Long-term results of a randomized trial of surgery with or without preoperative chemotherapy in esophageal cancer. *J Clin Oncol* 27(30), pp. 5062-5067.

Altman, D. G. 2001. Systematic reviews of evaluations of prognostic variables. *BMJ* 323(7306), pp. 224-228.

Altman, D. G. and Royston, P. 2000. What do we mean by validating a prognostic model? *Stat Med* 19(4), pp. 453-473.

Altman, D. G. et al. 2009. Prognosis and prognostic research: validating a prognostic model. *BMJ* 338, p. b605.

Arnold, M. et al. 2015. Global incidence of oesophageal cancer by histological subtype in 2012. *Gut* 64(3), pp. 381-387.

Atri, M. et al. 2015. Utility of preoperative ferumoxtran-10 MRI to evaluate retroperitoneal lymph node metastasis in advanced cervical cancer: Results of ACRIN 6671/GOG 0233. *Eur J Radiol Open* 2, pp. 11-18.

Bahmanyar, S. and Ye, W. 2006. Dietary patterns and risk of squamous-cell carcinoma and adenocarcinoma of the esophagus and adenocarcinoma of the gastric cardia: a population-based case-control study in Sweden. *Nutr Cancer* 54(2), pp. 171-178.

Basu, S. et al. 2011. Evolving role of molecular imaging with PET in detecting and characterizing heterogeneity of cancer tissue at the primary and metastatic sites, a plausible explanation for failed attempts to cure malignant disorders. *Eur J Nucl Med Mol Imaging* 38(6), pp. 987-991.

Bedenne, L. et al. 2007. Chemoradiation followed by surgery compared with chemoradiation alone in squamous cancer of the esophagus: FFCD 9102. *J Clin Oncol* 25(10), pp. 1160-1168.

Belhassen, S. and Zaidi, H. 2010. A novel fuzzy C-means algorithm for unsupervised heterogeneous tumor quantification in PET. *Med Phys* 37(3), pp. 1309-1324.

Berthon, B. et al. 2013. Influence of cold walls on PET image quantification and volume segmentation: a phantom study. *Med Phys* 40(8), p. 082505.

Berthon, B. et al. 2014. Evaluation of advanced automatic PET segmentation methods using nonspherical thin-wall inserts. *Med Phys* 41(2), p. 022502.

Berthon, B. et al. 2016. ATLAAS: an automatic decision tree-based learning algorithm for advanced image segmentation in positron emission tomography. *Phys Med Biol* 61(13), pp. 4855-4869.

Biehl, K. J. et al. 2006. 18F-FDG PET definition of gross tumor volume for radiotherapy of non-small cell lung cancer: is a single standardized uptake value threshold approach appropriate? *J Nucl Med* 47(11), pp. 1808-1812.

Blackshaw, G. et al. 2008. Prospective comparison of endosonography, computed tomography, and histopathological stage of junctional oesophagogastric cancer. *Clin Radiol* 63(10), pp. 1092-1098.

Blake, P. A. et al. 2017. Prognostic Significance of Deprivation in Upper Gastrointestinal Cancer. *Gastrointestinal Cancer: Research & Therapy* 2(2), p. 1018.

Bland, J. M. and Altman, D. G. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1(8476), pp. 307-310.

Bland, J. M. and Altman, D. G. 1996. Transforming data. *BMJ* 312(7033), p. 770.

Blazeby, J. M. et al. 2000. A prospective longitudinal study examining the quality of life of patients with esophageal carcinoma. *Cancer* 88(8), pp. 1781-1787.

Blencowe, N. S. et al. 2013. Evaluating the role of fluorodeoxyglucose positron emission tomography-computed tomography in multi-disciplinary team recommendations for oesophago-gastric cancer. *Br J Cancer* 109(6), pp. 1445-1450.

Blom, R. L. et al. 2012. External ultrasonography of the neck does not add diagnostic value to integrated positron emission tomography-computed tomography (PET-CT) scanning in the diagnosis of cervical lymph node metastases in patients with esophageal carcinoma. *Dis Esophagus* 25(6), pp. 555-559.

Bogoevski, D. et al. 2008. Is it time for a new TNM classification in esophageal carcinoma? *Ann Surg* 247(4), pp. 633-641.

Botet, J. F. et al. 1991. Preoperative staging of esophageal cancer: comparison of endoscopic US and dynamic CT. *Radiology* 181(2), pp. 419-425.

Brown, G. and Daniels, I. R. 2005. Preoperative staging of rectal cancer: the MERCURY research project. *Recent Results Cancer Res* 165, pp. 58-74.

Brown, G. et al. 2003. Preoperative assessment of prognostic factors in rectal cancer using high-resolution magnetic resonance imaging. *Br J Surg* 90(3), pp. 355-364.

Button, M. R. et al. 2009. Study to determine adequate margins in radiotherapy planning for esophageal carcinoma by detailing patterns of recurrence after definitive chemoradiotherapy. *Int J Radiat Oncol Biol Phys* 73(3), pp. 818-823.

Cancer Genome Atlas Research Network. 2017. Integrated genomic characterization of oesophageal carcinoma. *Nature* 541(7636), pp. 169-175.

Cancer Research UK. 2016a. Oesophageal Cancer Incidence Statistics [Online]. Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/oesophageal-cancer/incidence [Accessed: December 20th 2016].

Cancer Research UK. 2016b. Oesophageal Cancer Statistics [Online]. Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/oesophageal-cancer [Accessed: November 22nd 2016].

Cancer Research UK. 2017. Tackle cancers with substantial unmet need: our research strategy [Online]. Available at: http://www.cancerresearchuk.org/funding-for-researchers/our-research-strategy/tackle-cancers-with-substantial-unmet-need [Accessed: January 21st 2018].

Carstens, H. et al. 2007. A randomised trial of chemoradiotherapy versus surgery alone in patients with resectable esophageal cancer. *J Clin Oncol* 25(suppl; abstr 4530).

Chan, D. S. et al. 2013a. Prognostic significance of 18-FDG PET/CT and EUS-defined tumour characteristics in patients with oesophageal cancer. *Clin Radiol* 68(4), pp. 352-357.

Chan, D. S. et al. 2013b. Systematic review and meta-analysis of the influence of circumferential resection margin involvement on survival in patients with operable oesophageal cancer. *Br J Surg* 100(4), pp. 456-464.

Cheng, K. K. et al. 1992. Pickled vegetables in the aetiology of oesophageal cancer in Hong Kong Chinese. *Lancet* 339(8805), pp. 1314-1318.

Choi, J. et al. 2010. Comparison of endoscopic ultrasonography (EUS), positron emission tomography (PET), and computed tomography (CT) in the preoperative locoregional staging of resectable esophageal cancer. *Surg Endosc* 24(6), pp. 1380-1386.

Choi, J. Y. et al. 2000. Improved detection of individual nodal involvement in squamous cell carcinoma of the esophagus by FDG PET. *J Nucl Med* 41(5), pp. 808-815.

Chong, Y. et al. 2014. Quantitative CT variables enabling response prediction in neoadjuvant therapy with EGFR-TKIs: are they different from those in neoadjuvant concurrent chemoradiotherapy? *PLoS One* 9(2), p. e88598.

College of American Pathologists. 2005. Surgical Pathology Cancer Case Summary: Esophagus. Northfield: College of American Pathologists.

Cook, G. J. R. et al. 2014. Radiomics in PET: principles and applications. *Clin Transl Imaging* 2, pp. 269-276.

Cook, N. R. 2007. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation* 115(7), pp. 928-935.

Cooper, S. C. et al. 2009. The influence of deprivation and ethnicity on the incidence of esophageal cancer in England. *Cancer Causes Control* 20(8), pp. 1459-1467.

Cox, D. R. 1972. Regression Models and Life-Tables. *Journal of the Royal Statistical Society* 34(2), pp. 187-220.

Crawford, J. M. 2003. Principles of Anatomy. In: Rustgi, A.K. and Crawford, J.M. eds. *Gastrointestinal Cancers: a Companion to Sleisenger and Fordtran's Gastrointestinal and Liver Disease.* Philedelphia: Saunders, p. 121.

Crosby, T. et al. 2013. Chemoradiotherapy with or without cetuximab in patients with oesophageal cancer (SCOPE1): a multicentre, phase 2/3 randomised trial. *Lancet Oncol* 14(7), pp. 627-637.

Crosby, T. et al. 2017. Long-term results and recurrence patterns from SCOPE-1: a phase II/III randomised trial of definitive chemoradiotherapy +/- cetuximab in oesophageal cancer. *Br J Cancer* 116(6), pp. 709-716.

Crosby, T. D. et al. 2004. Definitive chemoradiation in patients with inoperable oesophageal carcinoma. *Br J Cancer* 90(1), pp. 70-75.

Crosby, T. D. and Evans, M. 2009. The management of locally advanced carcinoma of the oesophagus - SCOPE for improvement? *Clin Oncol (R Coll Radiol)* 21(4), pp. 281-284.

Cunningham, D. et al. 2006. Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. *N Engl J Med* 355(1), pp. 11-20.

Cunningham, D. et al. 2008. Capecitabine and oxaliplatin for advanced esophagogastric cancer. *N Engl J Med* 358(1), pp. 36-46.

Dave, U. R. et al. 2004. Esophageal cancer staging with endoscopic MR imaging: pilot study. *Radiology* 230(1), pp. 281-286.

Davies, A. R. et al. 2014a. Factors associated with early recurrence and death after esophagectomy for cancer. *J Surg Oncol* 109(5), pp. 459-464.

Davies, A. R. et al. 2014b. Surgical resection strategy and the influence of radicality on outcomes in oesophageal cancer. *Br J Surg* 101(5), pp. 511-517.

Davies, L. et al. 2012. Prognostic significance of total disease length in esophageal cancer. *Surg Endosc* 26(10), pp. 2810-2816.

Day, E. et al. 2009. A region growing method for tumor volume segmentation on PET images for rectal and anal cancer patients. *Med Phys* 36(10), pp. 4349-4358.

de Geus-Oei, L. F. et al. 2007. Biological correlates of FDG uptake in non-small cell lung cancer. *Lung Cancer* 55(1), pp. 79-87.

De Stefani, E. et al. 1993. Black (air-cured) and blond (flue-cured) tobacco and cancer risk. III: Oesophageal cancer. *Eur J Cancer* 29A(5), pp. 763-766.

Deasy, J. O. et al. 2003. CERR: A Computational Environment for Radiotherapy Research. *Med Phys* 30(5), pp. 979-985.

Debray, T. P. et al. 2015. A new framework to enhance the interpretation of external validation studies of clinical prediction models. *J Clin Epidemiol* 68(3), pp. 279-289.

DeNardi, F. G. and Riddell, R. H. 1991. The normal esophagus. *Am J Surg Pathol* 15(3), pp. 296-309.

Dice, L. 1945. Measures of the Amount of Ecologic Association Between Species. *Ecology* 26, pp. 297-302.

Dong, X. et al. 2013. Three-dimensional positron emission tomography image texture analysis of esophageal squamous cell carcinoma: relationship between tumor 18F-fluorodeoxyglucose uptake heterogeneity, maximum standardized uptake value, and tumor stage. *Nucl Med Commun* 34(1), pp. 40-46.

Doumou, G. et al. 2015. The precision of textural analysis in (18)F-FDG-PET scans of oesophageal cancer. *Eur Radiol* 25(9), pp. 2805-2812.

Drever, L. et al. 2007. Iterative threshold segmentation for PET target volume delineation. *Med Phys* 34(4), pp. 1253-1265.

Edgren, G. et al. 2013. A global assessment of the oesophageal adenocarcinoma epidemic. *Gut* 62(10), pp. 1406-1414.

Egger, J. et al. 2013. GBM volumetry using the 3D Slicer medical image computing platform. *Sci Rep* 3, p. 1364.

Egglin, T. K. and Feinstein, A. R. 1996. Context bias. A problem in diagnostic radiology. *JAMA* 276(21), pp. 1752-1755.

Eisenhauer, E. A. et al. 2009. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 45(2), pp. 228-247.

Eloubeidi, M. A. et al. 2002. Prognostic factors for the survival of patients with esophageal carcinoma in the U.S.: the importance of tumor length and lymph node status. *Cancer* 95(7), pp. 1434-1443.

Feinstein, A. R. et al. 1985. The Will Rogers phenomenon. Stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. *N Engl J Med* 312(25), pp. 1604-1608.

Ferlay, J. et al. 2015. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 136(5), pp. E359-386.

Findlay, J. M. et al. 2015. Pragmatic staging of oesophageal cancer using decision theory involving selective endoscopic ultrasonography, PET and laparoscopy. *Br J Surg* 102(12), pp. 1488-1499.

Fiorica, F. et al. 2004. Preoperative chemoradiotherapy for oesophageal cancer: a systematic review and meta-analysis. *Gut* 53(7), pp. 925-930.

Fitzgerald, R. C. et al. 2014. British Society of Gastroenterology guidelines on the diagnosis and management of Barrett's oesophagus. *Gut* 63(1), pp. 7-42.

Foley, K. G. et al. 2014a. Prognostic significance of novel 18F-FDG PET/CT defined tumour variables in patients with oesophageal cancer. *Eur J Radiol* 83(7), pp. 1069-1073.

Foley, K. G. et al. 2014b. N-staging of oesophageal and junctional carcinoma: is there still a role for EUS in patients staged N0 at PET/CT? *Clin Radiol* 69(9), pp. 959-964.

Foley, K. G. et al. 2016. The 100 most cited articles investigating the radiological staging of oesophageal and junctional cancer: a bibliometric analysis. *Insights Imaging* 7(4), pp. 619-628.

Ford, E. C. et al. 2006. Tumor delineation using PET in head and neck cancers: threshold contouring and lesion volumes. *Med Phys* 33(11), pp. 4280-4288.

Friis, S. et al. 2015. Low-Dose Aspirin or Nonsteroidal Anti-inflammatory Drug Use and Colorectal Cancer Risk: A Population-Based, Case-Control Study. *Ann Intern Med* 163(5), pp. 347-355.

Galavis, P. E. et al. 2010. Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. *Acta Oncol* 49(7), pp. 1012-1016.

Gammon, M. D. et al. 1997. Tobacco, alcohol, and socioeconomic status and adenocarcinomas of the esophagus and gastric cardia. *J Natl Cancer Inst* 89(17), pp. 1277-1284.

Ganeshan, B. et al. 2012. Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: preliminary evidence of an association with tumour metabolism, stage, and survival. *Clin Radiol* 67(2), pp. 157-164.

Geets, X. et al. 2007. A gradient-based method for segmenting FDG-PET images: methodology and validation. *Eur J Nucl Med Mol Imaging* 34(9), pp. 1427-1438.

Gerlinger, M. et al. 2012. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 366(10), pp. 883-892.

Gillies, R. J. et al. 2016. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* 278(2), pp. 563-577.

Gillies, R. S. et al. 2011. Additional benefit of (1)(8)F-fluorodeoxyglucose integrated positron emission tomography/computed tomography in the staging of oesophageal cancer. *Eur Radiol* 21(2), pp. 274-280.

Gleeson, F. C. et al. 2009. Prospective assessment of EUS criteria for lymphadenopathy associated with rectal cancer. *Gastrointest Endosc* 69(4), pp. 896-903.

Glickman, J. N. et al. 1999. The prognostic significance of lymph node micrometastasis in patients with esophageal carcinoma. *Cancer* 85(4), pp. 769-778.

Goh, V. et al. 2011. Assessment of response to tyrosine kinase inhibitors in metastatic renal cell cancer: CT texture as a predictive biomarker. *Radiology* 261(1), pp. 165-171.

Goldblum, J. R. and Lee, R. G. 2004. Esophagus. In: Mills, S.E. et al. eds. *Sternberg's Diagnostic Surgical Pathology*. Vol. 2. Philedelphia: Lippincott Williams & Wilkins, pp. 1399-1433.

Griffin, S. M. et al. 2011. Lymph node metastasis in early esophageal adenocarcinoma. *Ann Surg* 254(5), pp. 731-736; discussion 736-737.

Griffith, J. F. et al. 2000. Neck ultrasound in staging squamous oesophageal carcinoma - a high yield technique. *Clin Radiol* 55(9), pp. 696-701.

Grimm, H. et al. 1993. Endosonography for preoperative locoregional staging of esophageal and gastric cancer. *Endoscopy* 25(3), pp. 224-230.

Grotenhuis, B. A. et al. 2010. Lymphatic micrometastases in patients with early esophageal adenocarcinoma. *J Surg Oncol* 102(7), pp. 863-867.

Gwynne, S. et al. 2011. Definitive chemoradiation for oesophageal cancer--a standard of care in patients with non-metastatic oesophageal cancer. *Clin Oncol (R Coll Radiol)* 23(3), pp. 182-188.

Hanley, J. A. and McNeil, B. J. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143(1), pp. 29-36.

Haralick, R. M. et al. 1973. Textural Features For Image Classification. *IEEE Trans Syst Man Cybern* 3, pp. 610-621.

Harrell, F. E., Jr. 2001. Regression Modelling Strategies with Applications to Linear Models, Logistic Regression, and Survival Analysis. New York: Springer.

Harvin, J. A. et al. 2012. Neoadjuvant chemoradiotherapy followed by surgery for esophageal adenocarcinoma: significance of microscopically positive circumferential radial margins. *J Thorac Cardiovasc Surg* 143(2), pp. 412-420.

Hatt, M. et al. 2011a. PET functional volume delineation: a robustness and repeatability study. *Eur J Nucl Med Mol Imaging* 38(4), pp. 663-672.

Hatt, M. et al. 2017. Classification and evaluation strategies of auto-segmentation approaches for PET: Report of AAPM task group No. 211. *Med Phys* 44(6), pp. e1-e42.

Hatt, M. et al. 2015. 18F-FDG PET Uptake Characterization Through Texture Analysis: Investigating the Complementary Nature of Heterogeneity and Functional Tumor Volume in a Multi-Cancer Site Patient Cohort. *J Nucl Med* 56(1), pp. 38-44.

Hatt, M. et al. 2013. Robustness of intratumour 18F-FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma. *Eur J Nucl Med Mol Imaging* 40(11), pp. 1662-1671.

Hatt, M. et al. 2011b. Prognostic value of 18F-FDG PET image-based parameters in oesophageal cancer and impact of tumour delineation methodology. *Eur J Nucl Med Mol Imaging* 38(7), pp. 1191-1202.

Hayakawa, Y. et al. 2016. Oesophageal adenocarcinoma and gastric cancer: should we mind the gap? *Nat Rev Cancer* 16(5), pp. 305-318.

Haybittle, J. L. et al. 1982. A prognostic index in primary breast cancer. *Br J Cancer* 45(3), pp. 361-366.

Hemingway, H. et al. 2013. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. *BMJ* 346, p. e5595.

Henriksson, E. et al. 2007. 2-Deoxy-2-[18F] fluoro-D-glucose uptake and correlation to intratumoral heterogeneity. *Anticancer Res* 27(4B), pp. 2155-2159.

Hernandez, A. V. et al. 2004. Covariate adjustment in randomized controlled trials with dichotomous outcomes increases statistical power and reduces sample size requirements. *J Clin Epidemiol* 57(5), pp. 454-460.

Hiele, M. et al. 1997. Relation between endoscopic ultrasound findings and outcome of patients with tumors of the esophagus or esophagogastric junction. *Gastrointest Endosc* 45(5), pp. 381-386.

Higgins, K. A. et al. 2012. Analysis of pretreatment FDG-PET SUV parameters in head-and-neck cancer: tumor SUVmean has superior prognostic value. *Int J Radiat Oncol Biol Phys* 82(2), pp. 548-553.

Hingorani, A. D. et al. 2013. Prognosis research strategy (PROGRESS) 4: stratified medicine research. *BMJ* 346, p. e5793.

Holscher, A. H. et al. 1995. Prognosis of early esophageal cancer. Comparison between adeno- and squamous cell carcinoma. *Cancer* 76(2), pp. 178-186.

Hosmer, D. W. and Lemesbow, S. 1980. Goodness of fit tests for the multiple logistic regression model. *Communications in Statistics - Theory and Methods* 9(10), pp. 1043-1069.

Howard, J. M. and Johnston, C. 2013. Patterns of lymphatic drainage and lymph node involvement in esophageal cancer. *Abdom Imaging* 38(2), pp. 233-243.

Hu, Y. et al. 2010. How does the number of resected lymph nodes influence TNM staging and prognosis for esophageal carcinoma? *Ann Surg Oncol* 17(3), pp. 784-790.

Huang, B. et al. 2012. Nasopharyngeal carcinoma: investigation of intratumoral heterogeneity with FDG PET/CT. *AJR Am J Roentgenol* 199(1), pp. 169-174.

Hulscher, J. B. et al. 2002. Extended transthoracic resection compared with limited transhiatal resection for adenocarcinoma of the esophagus. *N Engl J Med* 347(21), pp. 1662-1669.

Humbert, O. et al. 2016. 18F-FDG PET-Derived Tumor Blood Flow Changes After 1 Cycle of Neoadjuvant Chemotherapy Predicts Outcome in Triple-Negative Breast Cancer. *J Nucl Med* 57(11), pp. 1707-1712.

Hurt, C. N. et al. 2011. SCOPE1: a randomised phase II/III multicentre clinical trial of definitive chemoradiation, with or without cetuximab, in carcinoma of the oesophagus. *BMC Cancer* 11, p. 466.

Hvid-Jensen, F. et al. 2011. Incidence of adenocarcinoma among patients with Barrett's esophagus. *N Engl J Med* 365(15), pp. 1375-1383.

I, H. S. et al. 2012. Predictive value of metabolic tumor volume measured by 18F-FDG PET for regional lymph node status in patients with esophageal cancer. *Clin Nucl Med* 37(5), pp. 442-446.

Ide, H. et al. 1994. Esophageal squamous cell carcinoma: pathology and prognosis. *World J Surg* 18(3), pp. 321-330.

Inui, K. et al. 1995. Endoscopic MRI: preliminary results of a new technique for visualization and staging of gastrointestinal tumors. *Endoscopy* 27(7), pp. 480-485.

Ioannidis, J. P. 2005. Why most published research findings are false. *PLoS Med* 2(8), p. e124.

Islami, F. et al. 2009. Tea drinking habits and oesophageal cancer in a high risk area in northern Iran: population based case-control study. *BMJ* 338, p. b929.

Izbicki, J. R. et al. 1997. Prognostic value of immunohistochemically identifiable tumor cells in lymph nodes of patients with completely resected esophageal cancer. *N Engl J Med* 337(17), pp. 1188-1194.

Jarman, B. et al. 1999. Explaining differences in English hospital death rates using routinely collected data. *BMJ* 318(7197), pp. 1515-1520.

Javeri, H. et al. 2009. Influence of the baseline 18F-fluoro-2-deoxy-D-glucose positron emission tomography results on survival and pathologic response in patients with gastroesophageal cancer undergoing chemoradiation. *Cancer* 115(3), pp. 624-630.

Jiao, X. et al. 2003. Immunohistochemistry analysis of micrometastasis in pretreatment lymph nodes from patients with esophageal cancer. *Ann Thorac Surg* 76(4), pp. 996-999.

Johnson, P. et al. 2016. Adapted Treatment Guided by Interim PET-CT Scan in Advanced Hodgkin's Lymphoma. *N Engl J Med* 374(25), pp. 2419-2429.

Julesz, B. 1975. Experiments in the visual perception of texture. *Sci Am* 232(4), pp. 34-43.

Kannel, W. B. et al. 1976. A general cardiovascular risk profile: the Framingham Study. *Am J Cardiol* 38(1), pp. 46-51.

Kantor, E. D. et al. 2012. Use of statin medications and risk of esophageal adenocarcinoma in persons with Barrett's esophagus. *Cancer Epidemiol Biomarkers Prev* 21(3), pp. 456-461.

Kaplan, E. and Meier, P. 1958. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 53, pp. 457-481.

Kapoor, V. et al. 2004. An introduction to PET-CT imaging. *Radiographics* 24(2), pp. 523-543.

Karran, A. et al. 2014. Propensity score analysis of oesophageal cancer treatment with surgery or definitive chemoradiotherapy. *Br J Surg* 101(5), pp. 502-510.

Kato, H. et al. 2008. The additional value of integrated PET/CT over PET in initial lymph node staging of esophageal cancer. *Oncol Rep* 20(4), pp. 857-862.

Kayani, B. et al. 2011. Lymph node metastases and prognosis in oesophageal carcinoma-a systematic review. *Eur J Surg Oncol* 37(9), pp. 747-753.

Keegan, N. et al. 2014. ICORG 10-14: Neo-AEGIS: A randomized clinical trial of neoadjuvant and adjuvant chemotherapy (modified MAGIC regimen) versus neoadjuvant chemoradiation (CROSS protocol) in adenocarcinoma of the esophagus and esophagogastric junction. *J Clin Oncol* 32(suppl; abstr 4145).

Kelsen, D. P. et al. 1998. Chemotherapy followed by surgery compared with surgery alone for localized esophageal cancer. *N Engl J Med* 339(27), pp. 1979-1984.

Kelsen, D. P. et al. 2007. Long-term results of RTOG trial 8911 (USA Intergroup 113): a random assignment trial comparison of chemotherapy followed by surgery compared with surgery alone for esophageal cancer. *J Clin Oncol* 25(24), pp. 3719-3725.

Khan, O. A. et al. 2003. Prognostic significance of circumferential resection margin involvement following oesophagectomy for cancer. *Br J Cancer* 88(10), pp. 1549-1552.

Kidd, E. A. and Grigsby, P. W. 2008. Intratumoral metabolic heterogeneity of cervical cancer. *Clin Cancer Res* 14(16), pp. 5236-5241.

Kinahan, P. E. and Fletcher, J. W. 2010. Positron emission tomography-computed tomography standardized uptake values in clinical practice and assessing response to therapy. *Semin Ultrasound CT MR* 31(6), pp. 496-505.

Koenig, A. M. et al. 2009. Strong impact of micrometastatic tumor cell load in patients with esophageal carcinoma. *Ann Surg Oncol* 16(2), pp. 454-462.

Konski, A. et al. 2005. The integration of 18-fluoro-deoxy-glucose positron emission tomography and endoscopic ultrasound in the treatment-planning process for esophageal carcinoma. *Int J Radiat Oncol Biol Phys* 61(4), pp. 1123-1128.

Kulling, D. et al. 1998. Local staging of esophageal cancer using endoscopic magnetic resonance imaging: prospective comparison with endoscopic ultrasound. *Endoscopy* 30(9), pp. 745-749.

Kyzas, P. A. et al. 2007. Almost all articles on cancer prognostic markers report statistically significant results. *Eur J Cancer* 43(17), pp. 2559-2579.

Lagarde, S. M. et al. 2006. Extracapsular lymph node involvement in node-positive patients with adenocarcinoma of the distal esophagus or gastroesophageal junction. *Am J Surg Pathol* 30(2), pp. 171-176.

Lagergren, J. et al. 1999. Association between body mass and adenocarcinoma of the esophagus and gastric cardia. *Ann Intern Med* 130(11), pp. 883-890.

Lambin, P. et al. 2012. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 48(4), pp. 441-446.

Lambin, P. et al. 2015. Modern clinical research: How rapid learning health care and cohort multiple randomised clinical trials complement traditional evidence based medicine. *Acta Oncol* 54(9), pp. 1289-1300.

Lee, C. H. et al. 2007. Carcinogenetic impact of alcohol intake on squamous cell carcinoma risk of the oesophagus in relation to tobacco smoking. *Eur J Cancer* 43(7), pp. 1188-1199.

Leijenaar, R. T. et al. 2015. The effect of SUV discretization in quantitative FDG-PET Radiomics: the need for standardized methodology in tumor texture analysis. *Sci Rep* 5, p. 11075.

Levin, C. S. 2012. Promising new photon detection concepts for high-resolution clinical and preclinical PET. *J Nucl Med* 53(2), pp. 167-170.

Li, H. et al. 2013. Computed tomography scan as a tool to predict tumor T category in resectable esophageal squamous cell carcinoma. *Ann Thorac Surg* 95(5), pp. 1749-1755.

Li, Y. M. et al. 2014. Pre-treatment metabolic tumor volume and total lesion glycolysis are useful prognostic factors for esophageal squamous cell cancer patients. *Asian Pac J Cancer Prev* 15(3), pp. 1369-1373.

Lieberman, M. D. et al. 1995. Carcinoma of the esophagus. Prognostic significance of histologic type. *J Thorac Cardiovasc Surg* 109(1), pp. 130-138; discussion 139.

Liu, Y. P. et al. 2010. Prognostic value of lymph node metastases and lymph node ratio in esophageal squamous cell carcinoma. *Eur J Surg Oncol* 36(2), pp. 155-159.

Loh, H. H. et al. 1988. The analysis of natural textures using run length features. *IEEE Trans Ind Electron* 35(2), pp. 323-328.

Luketich, J. D. et al. 1998. Detection of micrometastases in histologically negative lymph nodes in esophageal cancer. *Ann Thorac Surg* 66(5), pp. 1715-1718.

Mackway-Jones, K. 1997. Emergency Triage. 1st ed. London: BMJ Publishing Group.

Mallett, S. et al. 2010a. Reporting methods in studies developing prognostic models in cancer: a review. *BMC Med* 8, p. 20.

Mallett, S. et al. 2010b. Reporting performance of prognostic models in cancer: a review. *BMC Med* 8, p. 21.

Mandard, A. M. et al. 1994. Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma. Clinicopathologic correlations. *Cancer* 73(11), pp. 2680-2686.

Mantel, N. 1970. Why Stepdown Procedures in Variable Selection? *Technometrics* 12, pp. 621-625.

Mapstone, N. 1998. Minimum Dataset for Oesophageal Carcinoma Histopathology Reports. 1st ed. London: Royal College of Pathologists.

Mapstone, N. 2007. Dataset for the Histopathological Reporting of Oesophageal Carcinoma. 2nd ed. London: Royal College of Pathologists.

Mariette, C. et al. 2003. Extent of oesophageal resection for adenocarcinoma of the oesophagogastric junction. *Eur J Surg Oncol* 29(7), pp. 588-593.

Mariette, C. et al. 2008. The number of metastatic lymph nodes and the ratio between metastatic and examined lymph nodes are independent prognostic factors in esophageal cancer regardless of neoadjuvant chemoradiation or lymphadenectomy extent. *Ann Surg* 247(2), pp. 365-371.

Martin, L. W. et al. 2016. Detection of Occult Micrometastases in Patients With Clinical Stage I Non-Small-Cell Lung Cancer: A Prospective Analysis of Mature Results of CALGB 9761 (Alliance). *J Clin Oncol* 34(13), pp. 1484-1491.

McShane, L. M. et al. 2005. REporting recommendations for tumour MARKer prognostic studies (REMARK). *Br J Cancer* 93(4), pp. 387-391.

Medical Research Council Oesophageal Cancer Working Group. 2002. Surgical resection with or without preoperative chemotherapy in oesophageal cancer: a randomised controlled trial. *Lancet* 359(9319), pp. 1727-1733.

Minsky, B. D. et al. 2002. INT 0123 (Radiation Therapy Oncology Group 94-05) phase III trial of combined-modality therapy for esophageal cancer: high-dose versus standard-dose radiation therapy. *J Clin Oncol* 20(5), pp. 1167-1174.

Monjazeb, A. M. et al. 2010. Outcomes of patients with esophageal cancer staged with 18F-fluorodeoxyglucose positron emission tomography (FDG-PET): can postchemoradiotherapy FDG-PET predict the utility of resection? *J Clin Oncol* 28(31), pp. 4714-4721.

Moon, S. H. et al. 2014. Prediction of occult lymph node metastasis by metabolic parameters in patients with clinically N0 esophageal squamous cell carcinoma. *J Nucl Med* 55(5), pp. 743-748.

Moons, K. G. et al. 2009a. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ* 338, p. b606.

Moons, K. G. et al. 2012. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart* 98(9), pp. 683-690.

Moons, K. G. et al. 2009b. Prognosis and prognostic research: what, why, and how? *BMJ* 338, p. b375.

Moore, K. L. and Agur, A. M. R. 2002. Essential Clinical Anatomy. 2nd ed. Philedelphia: Lippincott Williams & Wilkins.

Morgan, M. A. et al. 2008. Prognostic significance of failure to cross esophageal tumors by endoluminal ultrasound. *Dis Esophagus* 21(6), pp. 508-513.

Mukherjee, S. et al. 2015. NEOSCOPE: a randomised Phase II study of induction chemotherapy followed by either oxaliplatin/capecitabine or paclitaxel/carboplatin based chemoradiation as pre-operative regimen for resectable oesophageal adenocarcinoma. *BMC Cancer* 15, p. 48.

National Cancer Intelligence Network. 2008. Cancer Incidence and Mortality by Cancer Network, UK, 2005. London: NCIN.

National Institute for Health and Clinical Excellence. 2015. Referral Guidelines for Suspected Cancer [Online]. Available at: http://www.nice.org.uk/guidance/NG12/chapter/1-recommendations#upper-gastrointestinal-tract-cancers [Accessed: August 4th 2015].

National Oesophago-Gastric Cancer Audit. 2016. An Audit of the Care Received by People with Oesophago-Gastric Cancer in England and Wales Annual Report. [Online]. Available from: http://content.digital.nhs.uk/og. [Accessed: 13th October 2016].

Ng, F. et al. 2013. Assessment of tumor heterogeneity by CT texture analysis: can the largest cross-sectional area be used as an alternative to whole tumor analysis? *Eur J Radiol* 82(2), pp. 342-348.

Nishimura, H. et al. 2006. Preoperative esophageal cancer staging: magnetic resonance imaging of lymph node with ferumoxtran-10, an ultrasmall superparamagnetic iron oxide. *J Am Coll Surg* 202(4), pp. 604-611.

O'Connor, J. P. et al. 2017. Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol* 14(3), pp. 169-186.

O'Farrell, N. J. et al. 2013. Appraisal of staging endoscopic ultrasonography in a modern high-volume esophageal program. *World J Surg* 37(7), pp. 1666-1672.

O'Riordan, J. M. et al. 2007. Impact of solitary involved lymph node on outcome in localized cancer of the esophagus and esophagogastric junction. *J Gastrointest Surg* 11(4), pp. 493-499.

Okada, M. et al. 2009. Integrated FDG-PET/CT compared with intravenous contrast-enhanced CT for evaluation of metastatic regional lymph nodes in patients with resectable early stage esophageal cancer. *Ann Nucl Med* 23(1), pp. 73-80.

Okada, N. et al. 2016. The prognostic significance of the positive circumferential resection margin in pathologic T3 squamous cell carcinoma of the esophagus with or without neoadjuvant chemotherapy. *Surgery* 159(2), pp. 441-450.

Omloo, J. M. et al. 2011. FDG-PET parameters as prognostic factor in esophageal cancer patients: a review. *Ann Surg Oncol* 18(12), pp. 3338-3352.

Omloo, J. M. et al. 2009. Additional value of external ultrasonography of the neck after CT and PET scanning in the preoperative assessment of patients with esophageal cancer. *Dig Surg* 26(1), pp. 43-49.

Orlhac, F. et al. 2014. Tumor texture analysis in 18F-FDG PET: relationships between texture parameters, histogram indices, standardized uptake values, metabolic volumes, and total lesion glycolysis. *J Nucl Med* 55(3), pp. 414-422.

Orlhac, F. et al. 2016. Multiscale Texture Analysis: From 18F-FDG PET Images to Histologic Images. *J Nucl Med* 57(11), pp. 1823-1828.

Oxford University Press. 2010. Concise Medical Dictionary. 8th ed. Oxford: Oxford University Press.

Pan, L. et al. 2009. Prognostic significance of SUV on PET/CT in patients with esophageal cancer: a systematic review and meta-analysis. *Eur J Gastroenterol Hepatol* 21(9), pp. 1008-1015.

Paraf, F. et al. 1995. Surgical pathology of adenocarcinoma arising in Barrett's esophagus. Analysis of 67 cases. *Am J Surg Pathol* 19(2), pp. 183-191.

Peduzzi, P. et al. 1996. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 49(12), pp. 1373-1379.

Prieto, E. et al. 2012. Twelve automated thresholding methods for segmentation of PET images: a phantom study. *Phys Med Biol* 57(12), pp. 3963-3980.

Puli, S. R. et al. 2008. Staging accuracy of esophageal cancer by endoscopic ultrasound: a meta-analysis and systematic review. *World J Gastroenterol* 14(10), pp. 1479-1490.

Pultrum, B. B. et al. 2010. A critical appraisal of circumferential resection margins in esophageal carcinoma. *Ann Surg Oncol* 17(3), pp. 812-820.

Rajendran, J. G. et al. 2006. Tumor hypoxia imaging with [F-18] fluoromisonidazole positron emission tomography in head and neck cancer. *Clin Cancer Res* 12(18), pp. 5435-5441.

Reid, T. D. et al. 2012. Prognostic significance of circumferential resection margin involvement following oesophagectomy for cancer and the predictive role of endoluminal ultrasonography. *Br J Cancer* 107(12), pp. 1925-1931.

Reilly, B. M. and Evans, A. T. 2006. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. *Ann Intern Med* 144(3), pp. 201-209.

Rice, P. F. and Roberts, S. A. 2003. Discrepancy of echo-endoscopic measurements, which could lead to inaccurate recording of levels during staging of esophageal cancer. *Endoscopy* 35(4), p. 376.

Rice, T. W. 2010. Staging of esophageal cancer: TNM and beyond. *Esophagus* 7, pp. 189-195.

Rice, T. W. et al. 2017. Cancer of the Esophagus and Esophagogastric Junction: An Eighth Edition Staging Primer. *J Thorac Oncol* 12(1), pp. 36-42.

Rice, T. W. et al. 1998. Esophageal carcinoma: depth of tumor invasion is predictive of regional lymph node status. *Ann Thorac Surg* 65(3), pp. 787-792.

Richards, D. G. et al. 2000. Endoscopic ultrasound in the staging of tumours of the oesophagus and gastro-oesophageal junction. *Ann R Coll Surg Engl* 82(5), pp. 311-317.

Riddell, A. M. et al. 2007. The appearances of oesophageal carcinoma demonstrated on high-resolution, T2-weighted MRI, with histopathological correlation. *Eur Radiol* 17(2), pp. 391-399.

Riddell, A. M. et al. 2006. The development and optimization of high spatial resolution MRI for imaging the oesophagus using an external surface coil. *Br J Radiol* 79(947), pp. 873-879.

Riley, R. D. et al. 2013. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. *PLoS Med* 10(2), p. e1001380.

Riquet, M. et al. 1993. Lymphatic drainage of the esophagus in the adult. *Surg Radiol Anat* 15(3), pp. 209-211.

Rizk, N. et al. 2006. The prognostic importance of the number of involved lymph nodes in esophageal cancer: implications for revisions of the American Joint Committee on Cancer staging system. *J Thorac Cardiovasc Surg* 132(6), pp. 1374-1381.

Roberts, S. A. and Kay, C. 2014. Oesophagus and stomach cancers. In: Nicholson, T. ed. *Recommendations for cross-sectional imaging in cancer management*. London: Royal College of Radiologists.

Robey-Cafferty, S. S. et al. 1991. Prognostic factors in esophageal squamous carcinoma. A study of histologic features, blood group expression, and DNA ploidy. *Am J Clin Pathol* 95(6), pp. 844-849.

Roedl, J. B. et al. 2009. Metabolic tumor width parameters as determined on PET/CT predict disease-free survival and treatment response in squamous cell carcinoma of the esophagus. *Mol Imaging Biol* 11(1), pp. 54-60.

Royston, P. et al. 2006. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med* 25(1), pp. 127-141.

Royston, P. et al. 2009. Prognosis and prognostic research: Developing a prognostic model. *BMJ* 338, p. b604.

Ryan, P. et al. 2004. Prevalence of bone marrow micrometastases in esophagogastric cancer patients with and without neoadjuvant chemoradiotherapy. *J Surg Res* 117(1), pp. 121-126.

Sagar, P. M. et al. 1993. Significance of circumferential resection margin involvement after oesophagectomy for cancer. *Br J Surg* 80(11), pp. 1386-1388.

Salih, T. et al. 2013. Prognostic significance of cancer within 1 mm of the circumferential resection margin in oesophageal cancer patients following neo-adjuvant chemotherapy. *Eur J Cardiothorac Surg* 43(3), pp. 562-567.

Schmidt, T. et al. 2015. Value of functional imaging by PET in esophageal cancer. *J Natl Compr Canc Netw* 13(2), pp. 239-247.

SCOPE2. 2016. SCOPE2 - A Randomised Phase II/III Trial to Study Radiotherapy Dose Escalation in Patients With Oesophageal Cancer Treated With Definitive Chemo-radiation With an Embedded Phase II Trial for Patients With a Poor Early

Response Using Positron Emission Tomography (PET). Identifier: NCT02741856. [Online]. Available from: https://clinicaltrials.gov/ct2/show/NCT02741856?term=SCOPE+2&rank=1.

Shapiro, J. et al. 2015. Neoadjuvant chemoradiotherapy plus surgery versus surgery alone for oesophageal or junctional cancer (CROSS): long-term results of a randomised controlled trial. *Lancet Oncol* 16(9), pp. 1090-1098.

Shi, W. et al. 2013. Meta-analysis of 18FDG PET-CT for nodal staging in patients with esophageal cancer. *Surgical Oncology* 22(2), pp. 112-116.

Siewert, J. et al. 2000. Adenocarcinoma of the esophagogastric junction: results of surgical therapy based on anatomical/topographic classification in 1,002 consecutive patients. *Ann Surg* 232(3), pp. 353-361.

Siewert, J. R. et al. 2001. Histologic tumor type is an independent prognostic parameter in esophageal cancer: lessons from more than 1,000 consecutive resections at a single center in the Western world. *Ann Surg* 234(3), pp. 360-367; discussion 368-369.

Simon, R. 2001. Evaluating Prognostic Factor Studies. In: Gospodarowicz, M.K. ed. *Prognostic Factors in Cancer*. New York: Wiley, pp. 49-56.

Simon, R. and Altman, D. G. 1994. Statistical aspects of prognostic factor studies in oncology. *Br J Cancer* 69(6), pp. 979-985.

Singh, S. et al. 2014. Acid-suppressive medications and risk of oesophageal adenocarcinoma in patients with Barrett's oesophagus: a systematic review and meta-analysis. *Gut* 63(8), pp. 1229-1237.

Siu, K. F. et al. 1986. Shrinkage of the esophagus after resection for carcinoma. *Ann Surg* 203(2), pp. 173-176.

Sobin, L. H. et al. 2009. UICC TNM Classification of Malignant Tumours. 7th ed. New York: Wiley.

Sobin, L. H. and Wittekind, C. 1997. UICC TNM Classification of Malignant Tumours. 5th ed. New York: Wiley.

Sobin, L. H. and Wittekind, C. 2002. UICC TNM Classification of Malignant Tumours. 6th ed. New York: Wiley.

Stahl, M. et al. 2005. Chemoradiation with and without surgery in patients with locally advanced squamous cell carcinoma of the esophagus. *J Clin Oncol* 23(10), pp. 2310-2317.

Stephens, M. R. et al. 2006. Multidisciplinary team management is associated with improved outcomes after surgery for esophageal cancer. *Dis Esophagus* 19(3), pp. 164-171.

Stevens, A. and Lowe, J. S. 1997. Human Histology. 2nd ed. Philadelphia: Mosby.

Steyerberg, E. W. 2009. Clinical Prediction Models - A Practical Approach to Development, Validation, and Updating. New York: Springer.

Steyerberg, E. W. et al. 2013. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med* 10(2), p. e1001381.

Sun, G. W. et al. 1996. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *J Clin Epidemiol* 49(8), pp. 907-916.

Surti, S. 2015. Update on time-of-flight PET imaging. *J Nucl Med* 56(1), pp. 98-105.

Surveillance Epidemiology and End Results Program (SEER). 2006-2012. Cancer Stat Facts: Esophageal Cancer [Online]. Available at: https://seer.cancer.gov/statfacts/html/esoph.html [Accessed: March 8th 2017].

Takashima, S. et al. 1991. Carcinoma of the esophagus: CT vs MR imaging in determining resectability. *AJR Am J Roentgenol* 156(2), pp. 297-302.

Takeshita, K. et al. 1997. Endoscopic treatment of early oesophageal or gastric cancer. *Gut* 40(1), pp. 123-127.

Tanabe, T. et al. 2003. Immunohistochemically detected micrometastasis in lymph nodes from superficial esophageal squamous cell carcinoma. *J Surg Oncol* 82(3), pp. 153-159.

Tangoku, A. et al. 2012. The new era of staging as a key for an appropriate treatment for esophageal cancer. *Ann Thorac Cardiovasc Surg* 18(3), pp. 190-199.

Thosani, N. et al. 2012. Diagnostic accuracy of EUS in differentiating mucosal versus submucosal invasion of superficial esophageal cancers: a systematic review and meta-analysis. *Gastrointest Endosc* 75(2), pp. 242-253.

Tixier, F. et al. 2011. Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med* 52(3), pp. 369-378.

Twine, C. P. et al. 2009a. Prospective comparison of optic versus blind endoscopic ultrasound in staging esophageal cancer. *Surg Endosc* 23(12), pp. 2778-2784.

Twine, C. P. et al. 2009b. The assessment of prognosis of surgically resected oesophageal cancer is dependent on the number of lymph nodes examined pathologically. *Histopathology* 55(1), pp. 46-52.

Twine, C. P. et al. 2010. Prognostic significance of endoluminal ultrasound-defined disease length and tumor volume (EDTV) for patients with the diagnosis of esophageal cancer. *Surg Endosc* 24(4), pp. 870-878.

Urschel, J. D. and Vasan, H. 2003. A meta-analysis of randomized controlled trials that compared neoadjuvant chemoradiation and surgery to surgery alone for resectable esophageal cancer. *Am J Surg* 185(6), pp. 538-543.

van Hagen, P. et al. 2012. Preoperative chemoradiotherapy for esophageal or junctional cancer. *N Engl J Med* 366(22), pp. 2074-2084.

van Rossum, P. S. et al. 2016a. The Incremental Value of Subjective and Quantitative Assessment of 18F-FDG PET for the Prediction of Pathologic Complete Response to Preoperative Chemoradiotherapy in Esophageal Cancer. *J Nucl Med* 57(5), pp. 691-700.

van Rossum, P. S. et al. 2013. Imaging strategies in the management of oesophageal cancer: what's the role of MRI? *Eur Radiol* 23(7), pp. 1753-1765.

van Rossum, P. S. et al. 2016b. The emerging field of radiomics in esophageal cancer: current evidence and future potential. *Transl Cancer Res* 5(4), pp. 410-423.

van Vliet, E. P. et al. 2007. Publication bias does not play a role in the reporting of the results of endoscopic ultrasound staging of upper gastrointestinal cancers. *Endoscopy* 39(4), pp. 325-332.

van Vliet, E. P. et al. 2006. Staging of esophageal carcinoma in a low-volume EUS center compared with reported results from high-volume centers. *Gastrointest Endosc* 63(7), pp. 938-947.

van Vliet, E. P. et al. 2008. Staging investigations for oesophageal cancer: a meta-analysis. *Br J Cancer* 98(3), pp. 547-557.

Vogt, K. et al. 2006. Preoperative chemotherapy for resectable thoracic esophageal cancer. *Cochrane Database of Systematic Reviews* [Online] (3). Available from: http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD001556.pub2/abstract.

Wahl, R. L. et al. 2009. From RECIST to PERCIST: Evolving Considerations for PET Response Criteria in Solid Tumours. *J Nucl Med* 50(Suppl 1), pp. 1-50.

Ward, G. et al. 2016. Superiority of Deformable Image Co-registration in the Integration of Diagnostic Positron Emission Tomography-Computed Tomography to the Radiotherapy Treatment Planning Pathway for Oesophageal Carcinoma. *Clin Oncol (R Coll Radiol)* 28(10), pp. 655-662.

Watanabe, H. et al. 1990. Histological Typing of Oesophageal and Gastric Tumours. 2nd ed. Berlin: Springer-Verlag.

Wayman, J. et al. 2002. The pattern of recurrence of adenocarcinoma of the oesophago-gastric junction. *Br J Cancer* 86(8), pp. 1223-1229.

Weaver, D. L. 2010. Pathology evaluation of sentinel lymph nodes in breast cancer: protocol recommendations and rationale. *Mod Pathol* 23 Suppl 2, pp. S26-32.

Weaver, S. R. et al. 2004. Comparison of special interest computed tomography, endosonography and histopathological stage of oesophageal cancer. *Clin Radiol* 59(6), pp. 499-504.

Whittle, R. et al. 2017. Prognosis research ideally should measure time-varying predictors at their intended moment of use. *Diagnostic and Prognostic Research* 1(1).

Wilson, M. et al. 2008. Prognostic significance of lymph node metastases and ratio in esophageal cancer. *J Surg Res* 146(1), pp. 11-15.

Win, T. et al. 2013. Tumor heterogeneity and permeability as measured on the CT component of PET/CT predict survival in patients with non-small cell lung cancer. *Clin Cancer Res* 19(13), pp. 3591-3599.

Wu, L. F. et al. 2003. Preoperative TN staging of esophageal cancer: comparison of miniprobe ultrasonography, spiral CT and MRI. *World J Gastroenterol* 9(2), pp. 219-224.

Wu, W. et al. 2016. Exploratory Study to Identify Radiomics Classifiers for Lung Cancer Histology. *Front Oncol* 6, p. 71.

Yamada, I. et al. 2014. Esophageal carcinoma: ex vivo evaluation with diffusion-tensor MR imaging and tractography at 7 T. *Radiology* 272(1), pp. 164-173.

Yamada, I. et al. 2001. Superficial esophageal carcinoma: an in vitro study of high-resolution MR imaging at 1.5T. *J Magn Reson Imaging* 13(2), pp. 225-231.

Yamada, I. et al. 1997. Staging of esophageal carcinoma in vitro with 4.7-T MR imaging. *Radiology* 204(2), pp. 521-526.

Yip, C. et al. 2015. Assessment of changes in tumor heterogeneity following neoadjuvant chemotherapy in primary esophageal cancer. *Dis Esophagus* 28(2), pp. 172-179.

Yip, C. et al. 2014. Primary esophageal cancer: heterogeneity as potential prognostic biomarker in patients treated with definitive chemotherapy and radiation therapy. *Radiology* 270(1), pp. 141-148.

Yip, S. S. and Aerts, H. J. 2016. Applications and limitations of radiomics. *Phys Med Biol* 61(13), pp. R150-r166.

Yokoyama, A. and Omori, T. 2003. Genetic polymorphisms of alcohol and aldehyde dehydrogenases and risk for esophageal and head and neck cancers. *Jpn J Clin Oncol* 33(3), pp. 111-121.

Yuan, S. et al. 2006. Additional value of PET/CT over PET in assessment of locoregional lymph nodes in thoracic esophageal squamous cell cancer. *J Nucl Med* 47(8), pp. 1255-1259.

Zaidi, H. and El Naqa, I. 2010. PET-guided delineation of radiation therapy treatment volumes: a survey of image segmentation techniques. *Eur J Nucl Med Mol Imaging* 37(11), pp. 2165-2187.

Zhang, S. et al. 2014. Cyclooxygenase inhibitors use is associated with reduced risk of esophageal adenocarcinoma in patients with Barrett's esophagus: a meta-analysis. *Br J Cancer* 110(9), pp. 2378-2388.

Zhu, W. Q. et al. 2012. Oesophageal squamous cell carcinoma: relationship between fluorine-18 fludeoxyglucose positron emission tomography CT maximum standardised uptake value, metabolic tumour volume, and tumour, node and metastasis classification. *Br J Radiol* 85(1016), pp. e383-387.

zum Buschenfelde, C. M. et al. 2011. (18)F-FDG PET-guided salvage neoadjuvant radiochemotherapy of adenocarcinoma of the esophagogastric junction: the MUNICON II trial. *J Nucl Med* 52(8), pp. 1189-1196.

Zwanenburg, A. et al. 2016. Image biomarker standardisation initiative - feature definitions [Online]. Available at: https://arxiv.org/abs/1612.07003v3 [Accessed: March 20th 2017].

## Appendix A

### NATIONAL DATASET FOR
### OESOPHAGEAL CARCINOMA HISTOPATHOLOGY REPORTS

Surname ..... Forenames ..... Date of birth .....

Hospital ..... Hospital no ..... NHS no .....

Date of receipt ..... Date of reporting ..... Report no .....

Pathologist ..... Surgeon ..... Sex .....

**Shaded data items = 'non core' data**

**GROSS DESCRIPTION**

Maximum length of specimen: ..... mm
Length of oesophagus: ..... mm
Length of stomach:..... mm
Length of tumour..... mm
Width of tumour: ..... mm

Tumour edge to nearest distal margin: ..... mm
Tumour edge to nearest proximal margin: ..... mm
Type of tumour ☐ Polypoid ☐ Other ☐ Pinned ☐ Not pinned
Siewert tumour type (cardiac cancers only) ☐ 1 ☐ 2

**HISTOLOGY**

**Type of tumour**
☐ Squamous ☐ Adenocarcinoma
☐ Other (specify) .....
Differentiation by worst area:
☐ Well ☐ Moderately ☐ Poorly differentiated

**Depth of invasion**
☐ Tis high-grade dysplasia
☐ T1 invasion of lamina propria/submucosa
☐ T2 invasion of muscularis propria
☐ T3 invasion beyond muscularis propria
☐ T4 invasion of adjacent structures
☐ Yes ☐ No – serosal involvement:

**Proximal margin**
☐ Normal ☐ Dysplasia ☐ Carcinoma ☐ Barrett's

**Distal margin**
☐ Normal ☐ Dysplasia ☐ Carcinoma

**Circumferential margin**
Involvement (<1 mm): ☐ Yes ☐ No ☐ N/A
(If no: distance of carcinoma to nearest circumferential margin ..... mm)

**Other features**
Vascular invasion ☐ Yes ☐ No
Barrett's metaplasia adjacent to tumour ☐ Yes ☐ No

**Lymph nodes**
Number examined..... Number positive.....
(N0 if no nodes positive, otherwise N1)

**Distant metastases**
Coeliac axis node positive ☐ Yes ☐ No
(M1a if lower thoracic carcinoma, otherwise M1b)
Cervical node positive ☐ Yes ☐ No
(M1a if upper thoracic carcinoma, otherwise M1b)
Other distant metastasis (M1b) ☐ Yes ☐ No

**COMMENTS**

**PATHOLOGICAL STAGING**

Complete resection ☐ Yes(R0) ☐ No(R1 or R2) (y) pT..... pN..... pM..... TNM 5th edition

(y) pT..... pN.....(i +/-) pM..... TNM 6th edition

# Appendix B

GIG CYMRU NHS WALES
Bwrdd Iechyd Prifysgol Caerdydd a'r Fro
Cardiff and Vale University Health Board

**Ysbyty Athrofaol Cymru**
**University Hospital of Wales**

Heath Park,
Cardiff, CF14 4XW
Phone 029 2074 7747
Fax 029 2074 3838
Minicom 029 2074 3632

Parc Y Mynydd Bychan,
Caerdydd, CF14 4XW
Ffôn 029 2074 7747
Ffacs 029 20743838
Minicom 029 2074 3632

Tel: 029 20746986
Fax: 029 20745311
CAV_Research.Development@wales.nhs.uk

From: Professor C Fegan
Acting R&D Director
R&D Office, 2nd Floor TB2
University Hospital of Wales
Cardiff
CF14 4XW

02 December 2013

Dr Ashley Roberts
Radiology Department
University Hospital of Wales
Heath Park
Cardiff
CF14 4XW

Dear Dr Roberts

**Cardiff and Vale UHB Ref and Study Title : 13//DMD5769 : Prognostic Significance of Radiological Parameters in Patients with Oesophageal and Oesophago-Gastric Cancer**

**IRAS Project ID:** 141324

The above project was forwarded to Cardiff and Vale University Health Board R&D Office by the NISCHR Permissions Coordinating Unit. A Governance Review has now been completed on the project.

Documents approved for use in this study are:

| Document | Version | Date |
|---|---|---|
| NHS RD Form | 3.5 | Rec'd 07/11/13 |
| SSI Form | 3.5 | Rec'd 07/11/13 |
| Protocol | 1.0 | 16/08/13 |

I am pleased to inform you that your study has been classed as pathway-to-portfolio and that the UHB has no objection to your proposal.

Please accept this letter as confirmation of sponsorship by Cardiff and Vale University Local Health Board under the Research Governance Framework for Health and Social Care, and permission for the project to begin within this UHB.

## Appendix C

GIG CYMRU NHS WALES
Bwrdd Iechyd Prifysgol Caerdydd a'r Fro
Cardiff and Vale University Health Board

Ysbyty Athrofaol Cymru
University Hospital of Wales

Heath Park,
Cardiff, CF14 4XW
Phone 029 2074 7747
Fax 029 2074 3838
Minicom 029 2074 3632

Parc Y Mynydd Bychan,
Caerdydd, CF14 4XW
Ffôn 029 2074 7747
Ffacs 029 20743838
Minicom 029 2074 3632

Tel: 029 20746986
Fax: 029 20745311
CAV_Research.Development@wales.nhs.uk

From: Professor C Fegan
R&D Director
R&D Office, 2nd Floor TB2
University Hospital of Wales
Cardiff
CF14 4XW

19 November 2014

Dr Ashley Roberts
Consultant Radiologist
Radiology Department
University Hospital of Wales
Heath Park
Cardiff

Dear Dr Roberts

**Cardiff and Vale UHB Ref and Study Title : 14/DMD/5948 : The Relationship Between Tumour Markers And Radiological Staging Investigations In Patients With Oesophageal Cancer**

**IRAS Project ID:** 158679

The above project was forwarded to Cardiff and Vale University Health Board R&D Office by the NISCHR Permissions Coordinating Unit. A Governance Review has now been completed on the project.

Documents approved for use in this study are:

| Document | Version | Date |
|---|---|---|
| Protocol | 1 | 05/06/2014 |
| R&D Form | 3.5 | Received 14/08/14 |
| SSI Form | 3.5 | Received 14/08/14 |
| Protocol | 1 | 06/06/2014 |

I am pleased to inform you that your study has been classed as pathway-to-portfolio and that the UHB has no objection to your proposal.

# Appendix D

Part of the research infrastructure for Wales funded by the National Institute for Social Care and Health Research, Welsh Government.
Yn rhan o seilwaith ymchwil Cymru a ariannir gan y Sefydliad Cenedlaethol ar gyfer Ymchwil Gofal Cymdeithasol ac Iechyd, Llywodraeth Cymru

**Wales Research Ethics Committee 1**
**Sixth Floor, Churchill House**
**17 Churchill Way**
**Cardiff CF10 2TW**

**Telephone: 029 2037 6823**

**E-mail: jagit.sidhu@wales.nhs.uk**

**Website: www.hra.nhs.uk**

07 November 2014

Dr Ashley Roberts
Consultant Radiologist
Cardiff and Vale University Health Board
Department of Radiology
University Hospital of Wales, Heath Park
Cardiff
CF14 4XW

Dear Dr Roberts

| | |
|---|---|
| **Study title:** | **The Relationship Between Tumour Markers and Radiological Staging Investigations in Patients with Oesophageal and Gastro-oesophageal Junction Cancer** |
| **REC reference:** | **14/WA/1208** |
| **IRAS project ID:** | **158679** |

The Research Ethics Committee reviewed the above application at their meeting held on 06 November 2014. The Committee were grateful to yourself, Dr Foley and Dr Christian for attending to discuss the application.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details, unless you expressly withhold permission to do so. Publication will be no earlier than three months from the date of this favourable opinion letter. Should you wish to provide a substitute contact point, require further information, or wish to withhold permission to publish, please contact the REC Manager Mrs Jagjit Sidhu, jagit.sidhu@wales.nhs.uk.

**DECISION**

**The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.**

**Conditions of the favourable opinion**

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Clinical Radiology 72 (2017) 693.e1–693.e7

Contents lists available at ScienceDirect

Clinical Radiology

journal homepage: www.clinicalradiologyonline.net

# Accuracy of contemporary oesophageal cancer lymph node staging with radiological-pathological correlation

K.G. Foley [a,*], A. Christian [b], P. Fielding [c], W.G. Lewis [d], S.A. Roberts [e]

[a] Division of Cancer & Genetics, Cardiff University, Cardiff, UK
[b] Department of Pathology, University Hospital of Wales, Cardiff, UK
[c] Wales Research & Diagnostic PET Imaging Centre, Cardiff, UK
[d] Department of Upper GI Surgery, University Hospital of Wales, Cardiff, UK
[e] Department of Clinical Radiology, University Hospital of Wales, Cardiff, UK

ARTICLE INFORMATION

*Article history:*
Received 31 October 2016
Received in revised form
10 February 2017
Accepted 27 February 2017

AIM: To evaluate the accuracy of contemporary N-staging and provide radiological–pathological correlation in patients with lymph node metastases (LNMs) that were radiologically staged N0.

MATERIALS AND METHODS: One hundred and twelve patients were included who underwent surgery alone ($n$=41) or neoadjuvant therapy ($n$=71) between October 2010 and December 2015. Contrast-enhanced computed tomography (CECT), endoscopic ultrasound (EUS), and combined positron-emission tomography (PET) and CT N-stage were compared to pathological N-stage [node-negative (N0) versus node-positive (N+) groups]. Fifty LNMs from 15 patients preoperatively staged as N0 were measured and the maximum size recorded.

RESULTS: Accuracy, sensitivity, and specificity of N0 versus N+ disease with CECT, EUS, and PET/CT was 54.5%, 39.7% and 77.3%, 55.4%, 42.6% and 75%, and 57.1% 35.3%, and 90.9%, respectively. All techniques were more likely to under-stage nodal disease; CECT ($X^2$ 32.890, df=1, $p$<0.001), EUS ($X^2$ 28.471, df=1, $p$<0.001), and PET/CT ($X^2$ 50.790, df=1, $p$<0.001). PET/CT was more likely to under-stage nodal disease than EUS ($p$=0.031). Median LNM size was 3 mm, with 41 (82%) of LNMs measuring <6 mm and 22 (44%) classified as micro-metastases (≤2 mm).

CONCLUSION: This study has demonstrated poor N-staging accuracy in the modern era of radiological staging. Eighty-two percent of LNMs measured <6 mm, making direct identification extremely challenging on medical imaging. Future research should focus on investigating and developing alternative surrogate markers to predict the likelihood of LNMs.



## Introduction

Contemporary radiological staging of oesophageal cancer (OC) involves a multi-technique approach. In the UK, patients have initial contrast-enhanced (CE) computed

* Guarantor and correspondent: K. G. Foley, Division of Cancer & Genetics, School of Medicine, Cardiff University, Heath Park, Cardiff, Wales, CF14 4XN, UK. Tel.: +44 2920 747747; fax: +44 2920743029.
*E-mail address:* foleykg@cardiff.ac.uk (K.G. Foley).

http://dx.doi.org/10.1016/j.crad.2017.02.022

tomography (CT) of the thorax and abdomen following histological confirmation to assess the potential resectability of the tumour, or any distant metastatic disease, which may preclude radical therapy.

If the patient is deemed suitable for radical treatment, either in the form of definitive chemoradiotherapy (dCRT) or surgery (+/− neoadjuvant therapy), positron-emission tomography (PET) combined with CT and endoscopic ultrasound (EUS) are performed for a more detailed assessment of disease stage.[1] PET/CT has greater sensitivity for distant metastatic disease than CECT,[2] whereas EUS is regarded as the reference standard investigation for defining T- and N-stage, whilst also assisting surgical and radiotherapy planning.[3]

This staging process is complex and time-consuming, but necessary, because each technique has limitations for lymph node staging. CECT provides anatomical information only, relies on size criteria, and involves radiation. PET/CT also involves radiation but provides additional functional metabolic data and improves the positive predictive value (PPV) of lymph node metastases (LNMs).[4] The differentiation of peri-tumoural LNMs from adjacent avid tumour can be challenging on PET images.[5] This may increase false-negative rates therefore under-staging the extent of nodal disease. EUS has better sensitivity compared to CECT and PET/CT due to its superior contrast resolution.

The prognosis of OC is poor, with 5-year survival approximately 13%.[6] Many patients present with advanced disease and the incidence is increasing.[7] The presence of LNMs is a major prognostic indicator, therefore it is vital to stage nodal disease accurately.[8] Accurate staging optimises management plans and provides the best chance of survival for patients with potentially curable disease. If the multidisciplinary team (MDT) decide upon surgical management and radiological staging is ≥T3 or ≥N1, two cycles of neoadjuvant chemotherapy (NACT) are given prior to resection. This is currently considered best practice in the UK, because overall survival was shown to improve compared to surgery alone.[9]

Management decisions are influenced by the results of lymph node assessment based on findings of radiological staging investigations. Differentiation of node-negative (N0) from node-positive (N+) disease is important, because this should ensure that patients avoid unnecessary chemotherapy if over-staged, and are not denied potentially beneficial NACT if under-staged; however, the existence of small LNMs (<6 mm), which cannot be directly visualised on any imaging technique, are likely to cause inaccurate staging and have a subsequent detrimental effect on patient outcome.[10]

Therefore, this study was undertaken to review the accuracy of CECT, EUS, and PET/CT N-staging in the modern era of radiological OC staging. In addition, the prevalence of micro-metastases and size of LNMs was investigated in patients staged N0 on imaging, but node positive (pN+) at histopathology, by providing radiological–pathological correlation.

## Materials and methods

This retrospective cohort study includes consecutive patients who underwent surgical resection of an oesophageal or gastro-oesophageal (GOJ) tumour, over a 5-year period (November 2010 to December 2015) within a centralised service. Radiological and pathological staging data were obtained from the Cancer Network Information System Cymru database (CaNISC) following regional upper gastrointestinal (GI) cancer MDT discussion. Institutional review board (IRB) approval was granted (ref 14/WA/1208). The requirement for informed consent was waived.

Inclusion criteria were a previously untreated, biopsy-proven oesophageal or GOJ tumour in patients who underwent surgery alone, or had a poor Mandard tumour regression grade (TRG 4) or no response (TRG 5) following either NACT or neoadjuvant chemoradiotherapy (NACRT).[11] All patients had completed CECT, EUS, and PET/CT staging investigations and were classified according to the International Union Against Cancer (UICC) Tumour Node Metastasis (TNM), 7$^{th}$ edition.[12] All patients also had a full pathological N-stage (pN), also defined by the TNM 7$^{th}$ edition.

Patients with tumours that showed complete pathological response (pCR, TRG 1) or tumours with some response (TRG 2 and 3) following NACT or NACRT were excluded because the final pathology is not likely to be representative of preoperative status. Incomplete radiological staging investigations in particular, EUS examinations, in which the operator was unable to traverse a stenotic tumour in order to fully classify N-stage, were excluded. Patients that underwent an “open-and-close” procedure due to irresectable disease at surgery, were also excluded.

### *CECT acquisition protocol*

CECT was performed either in the host institution of the centralised service (University Hospital of Wales) or in local referring hospitals prior to surgery, according to Royal College of Radiologists guidelines.[1] All CECT examinations were reviewed at the regional upper GI MDT, and deemed to be of a satisfactory technical standard. The technique used at the host institution was as follows: GE HD 750 Discovery 64-section CT system (GE Healthcare, Pollards Wood, Buckinghamshire, UK); helical acquisition with collimation of 40 mm, pitch 0.984:1, and tube rotation speed of 0.4 seconds; tube output of 120kVp with smart current dose modulation between 60–600 mA; section thickness of 0.625mm; up to 500 ml water orally and 100–150 ml iopamidol (300 mg iodine/ml; Niopam 300, Bracco, High Wycombe, UK) intravenously with bolus tracking. Lymph nodes were considered involved on CECT if the short axis measurement was ≥1 cm, located in the expected distribution of disease, round with loss of fatty hilum, and demonstrated altered density or enhancement.

### *EUS protocol*

All EUS examinations were performed in three centres by four endosonographers. At the host institution, an initial

endoscopic examination was performed using a 9 mm diameter Olympus Paediatric gastroscope (Olympus, Southend, UK) to assess the degree of oesophageal luminal stenosis. Patients with an estimated oesophageal luminal diameter <15 mm underwent examination using the smaller-diameter MH-908 oesophagoprobe, and where there was no luminal stenosis, the standard UM-2000 echoendoscope was used (Olympus, Southend, UK). The type of echoendoscope used was at the discretion of the endoscopist. The primary oesophageal tumour was assessed, together with an evaluation of peri-oesophageal and peri-gastric structures as described previously.[13] The criteria for malignant lymphadenopathy specified a hypoechoic pattern, spherical contour, distinct border, and short axis diameter of ≥6 mm.

### *PET/CT acquisition protocol*

Patients were fasted for at least 6 hours prior to tracer administration. Serum glucose levels were routinely checked and confirmed to be <7 mmol/l prior to proceeding with imaging. Patients received a dose of 4 MBq of 2-[$^{18}$F]-fluoro-2-deoxy-D-glucose ($^{18}$F-FDG) per kilogram of body weight. Uptake time was 90 minutes, which the standard used. $^{18}$F-FDG PET/CT imaging was performed using a GE 690 PET/CT system (GE Healthcare). CT images were acquired in a helical acquisition with a pitch of 0.98 and a tube rotation speed of 0.5 seconds. Tube output was 120 kVp with output modulation between 20 and 200 mA. Matrix size for the CT acquisition was 512×512 pixels with a 50 cm field of view. No oral or intravenous contrast medium was administered. PET images were acquired at 3 minutes per field of view. The length of the axial field of view was 15.7 cm. Images were reconstructed with the ordered subset expectation maximisation algorithm, with 24 subsets and two iterations. Matrix size was 256×256 pixels, using the VUE Point time of flight algorithm. Nodes were classed as involved on PET/CT if identified on the CT component and showed FDG uptake appreciably higher than background values. No specific standardised uptake value was used for the inclusion of regional nodes. Lymph nodes considered physiological or related to an alternative aetiology were excluded from the N-stage.

### *Histopathological methods*

Histopathological reporting of OC specimens was performed according to the minimum requirements defined by the Royal College of Pathologists (RCPath).[14] All lymph nodes identified in the resection specimen were prepared in 3 mm sections for histopathological evaluation. N-stage was then assigned depending on the number of LNMs identified. TRG of the primary tumour was assigned according to the degree of fibrosis compared to residual tumour cells.[11] In discordant cases, all available resection specimens that were radiologically staged N0 but pathologically N+ were further evaluated. All available specimens were retrieved and reviewed from the archive. Due to the retrospective nature of analysis, some of the older cases were archived off-site, and were unavailable at the time of evaluation. The maximum size (long axis) of both involved lymph nodes and metastases within those lymph nodes, were retrospectively recorded. Maximum size was defined as the largest dimension on the glass slide measured by a consultant pathologist. A micrometastasis was defined as a tumour deposit measuring ≤2 mm.[15] Furthermore, a metastasis: lymph node ratio was calculated.

### *Statistical analysis*

Descriptive statistics were used to describe categorical and continuous variables. In this study, N-stage was separated into negative (N0) and N+ (N1, N2 or N3) groups. Accuracy was defined as the number of correct investigations divided by the total number of investigations. Sensitivity and specificity of N+ disease were calculated for each technique. A chi-square test assessed significant differences in under- or over-staging for each technique. Significant differences in under-staging between techniques were assessed with McNemar's test. A $p$-value <0.05 was considered statistically significant. Statistical analysis was performed with SPSS v23 (IBM, Chicago, IL, USA).

## Results

A total of 190 patients were considered for inclusion in the study. Seventy-eight patients (41.1%) were excluded from the study; 22 were "open-and-close" procedures, 16 were TRG 1, 13 were TRG 2, 13 were TRG 3 following neoadjuvant treatment, and 14 had incomplete EUS staging.

Following exclusions, 112 patients were included in the study. The median age was 65 years (range 24–78 years) and the male: female ratio was 92 (82.1%): 20 (17.9%). Fifty-nine tumours (52.7%) were located in the oesophagus; 10 in the mid-oesophagus, and 49 in the distal oesophagus. Fifty-three tumours (47.3%) were located at the GOJ; 19 Siewert (Sw) type I, 15 Sw type II and 19 Sw type III.

One hundred tumours (89.3%) were adenocarcinoma, with 11 squamous cell carcinoma (SCC) (9.8%) and one neuroendocrine (0.9%). Forty-one patients (36.6%) were treated with surgery alone, 67 (59.8%) treated with NACT, and four (3.6%) treated with NACRT. Of the 71 treated with neoadjuvant therapy, 42 were TRG 4 and 29 were TRG 5.

For CECT, 75 patients (67%) were staged N0 and 37 (33%) were N+. For EUS, 72 patients (64.3%) were staged N0 and 40 (35.7%) were N+. For PET/CT, 84 (75.1%) were staged N0 and 28 (24.9%) were staged N+. Table 1 compares the frequency of radiological and pathological N-stages for CECT, EUS, and PET/CT.

Overall, the median time between radiological staging and surgery was 3 months (range 1–9 months), 1 month (range 0–3 months) in patients undergoing surgery alone and 4 months (range 3–4 months) in patients receiving NACT.

**Table 1**
Comparison of N-stage frequency classified by contrast-enhanced computed tomography (CECT), endoscopic ultrasound (EUS), combined positron-emission tomography and computed tomography (PET/CT), and pathology.

| CECT N-stage | | | | | |
|---|---|---|---|---|---|
| Frequency (%) | N0 | N1 | N2 | N3 | Total |
| pN0 | 34 (30.4) | 8 (7.1) | 2 (1.7) | 0 (0.0) | 44 (39.3) |
| pN1 | 21 (18.8) | 4 (3.6) | 2 (1.7) | 0 (0.0) | 27 (24.1) |
| pN2 | 16 (14.3) | 10 (8.9) | 1 (0.9) | 0 (0.0) | 27 (24.1) |
| pN3 | 4 (3.6) | 7 (6.3) | 3 (2.7) | 0 (0.0) | 14 (12.5) |
| Total | 75 (67.0) | 29 (25.9) | 8 (7.1) | 0 (0.0) | 112 (100.0) |
| EUS N-Stage | | | | | |
| Frequency (%) | N0 | N1 | N2 | N3 | Total |
| pN0 | 33 (29.5) | 9 (8.0) | 1 (0.9) | 1 (0.9) | 44 (39.3) |
| pN1 | 20 (17.9) | 7 (6.3) | 0 (0.0) | 0 (0.0) | 27 (24.1) |
| pN2 | 13 (11.6) | 10 (8.9) | 4 (3.6) | 0 (0.0) | 27 (24.1) |
| pN3 | 6 (5.4) | 6 (5.4) | 1 (0.9) | 1 (0.9) | 14 (12.5) |
| Total | 72 (64.3) | 32 (28.6) | 6 (5.4) | 2 (1.7) | 112 (100.0) |
| PET/CT N-stage | | | | | |
| Frequency (%) | N0 | N1 | N2 | N3 | Total |
| pN0 | 40 (35.8) | 4 (3.6) | 0 (0.0) | 0 (0.0) | 44 (39.4) |
| pN1 | 23 (20.5) | 4 (3.6) | 0 (0.0) | 0 (0.0) | 27 (24.1) |
| pN2 | 15 (13.4) | 10 (8.9) | 2 (1.7) | 0 (0.0) | 27 (24.1) |
| pN3 | 6 (5.4) | 6 (5.4) | 2 (1.7) | 0 (0.0) | 14 (12.5) |
| Total | 84 (75.1) | 24 (21.4) | 4 (3.6) | 0 (0.0) | 112 (100.0) |

### *Accuracy, sensitivity, and specificity of CECT, EUS, and PET/CT N-stage*

N0 versus N+ disease was correctly identified with CECT, EUS, and PET/CT in 61 (54.5%), 62 (55.4%), and 64 (57.1%) cases, respectively. There was no significant difference between CECT, EUS, and PET/CT for detecting N+ disease ($X^2$ 0.169, df=2, $p$=0.919). The sensitivity and specificity for identifying N0 versus N+ disease with CECT, EUS, and PET/CT was 39.7% and 77.3%, 42.6% and 75%, and 35.3% and 90.9%, respectively.

### *Under-staging versus over-staging*

All techniques were significantly more likely to under-stage nodal disease; CECT ($X^2$ 32.890, df=1, $p$<0.001), EUS ($X^2$ 28.471, df=1, $p$<0.001), and PET/CT ($X^2$ 50.790, df=1, $p$<0.001). Comparing technique, there was a borderline significant difference in under-staging between CECT and EUS ($p$=0.063), but no difference between CECT and PET/CT ($p$=1.000); however, there was a statistically significant difference between EUS with PET/CT ($p$=0.031), suggesting PET/CT may further under-stage nodal disease.

### *Pathological lymph node measurement*

Fifteen archived resection specimens in patients staged N0 preoperatively were available for retrospective measurement of the lymph nodes and their respective metastases. In total, 50 involved lymph nodes were assessed. (Table 2) The median size of involved lymph nodes was 6 mm (range 2–15 mm) and the median metastasis size was 3 mm (0.5–13.5 mm). Twenty-two (44%) LNMs measured ≤2 mm, which are defined as micro-metastases (Fig 1). Forty-one (82%) LNMs were ≤6 mm and 46 (92%) LNMs were ≤10 mm. A metastasis: lymph node size ratio was calculated. Thirty-one (62%) of the lymph nodes examined were replaced with ≥50% metastatic deposit, 19 (38%) were replaced with <50% metastatic deposit, with 12 (24%) replaced with <25% metastatic deposit, using maximum size criteria.

## Discussion

This study has found poor N-stage accuracy with CECT, EUS, and PET/CT. In general, all modalities were more likely to under-stage nodal disease, with PET/CT more likely to under-stage than EUS. Another important finding, is the prevalence of small LNMs (<6 mm) in the resection specimens of patients radiologically staged N0. Micro-metastases have been found in lymph nodes of early oesophageal tumours,[16] but little has been published with radiological correlation. Studies investigating lung cancer have detected micro-metastases in patients radiologically staged N0,[17] although evidence in OC is lacking.

The majority of LNMs (82%) were <6 mm, which makes direct visualisation extremely challenging on current medical imaging techniques and is likely to be the main reason for discrepancy between radiological and pathological staging. In addition, traditional radiological measurement of lymph nodes is taken in the short-axis,[18] which further reduces the likelihood that LNMs are diagnosed. Even with the improved contrast resolution of EUS compared to cross-sectional imaging, it is unlikely that a lymph node of this size would confidently be classified as involved.[13] Similarly, there was a relatively high prevalence of micro-metastases (44%).

These results have significant implications for treatment decision-making processes and demonstrate that contemporary radiology techniques are inadequate for N-staging. Numerous studies have demonstrated the importance of LNMs, which have a significant effect on overall survival.[8] Better evidence is required to understand the prognostic

**Table 2**
Frequency of and distribution of lymph node and metastasis size when separated in groups of 2 mm for descriptive purposes.

| Frequency (%) | Maximum size (mm) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 0–2 | 2.1–4 | 4.1–6 | 6.1–8 | 8.1–10 | 10.1–12 | 12.1–14 | 14.1–16 |
| Lymph node | 3 (6) | 11 (22) | 13 (26) | 12 (24) | 4 (8) | 3 (6) | 3 (6) | 1 (2) |
| Metastasis | 22 (44) | 9 (18) | 10 (20) | 3 (6) | 2 (4) | 2 (4) | 2 (4) | 0 (0) |

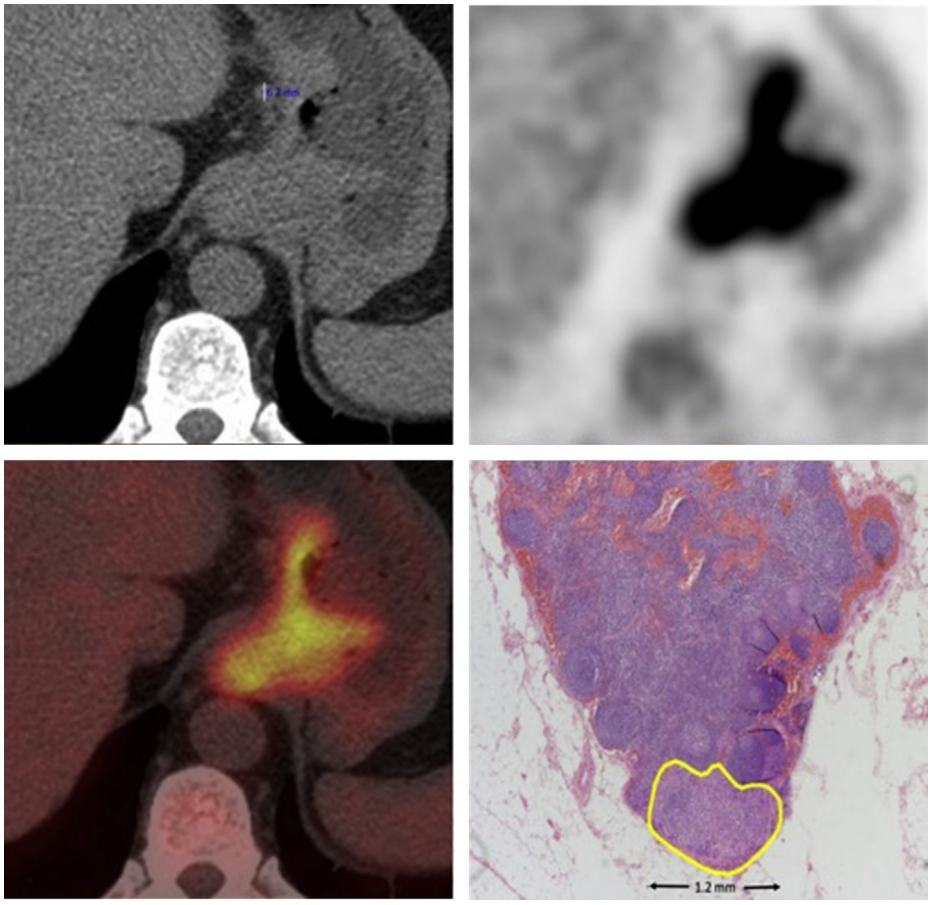



**Figure 1** CT (with callipers), PET, and fused PET/CT images of a "false-negative" left gastric lymph node in a patient with junctional adenocarcinoma. A low-power magnification of the lymph node shows a micro-metastasis. For reference, the lymph node measured 5 mm in maximum size and the micro-metastasis (highlighted with yellow outline) measured 1.2 mm.

significance of micro-metastases, but they are generally felt to confer a worse prognosis.[19,20]

There is evidence that a significant proportion of surgical patients have systemic micro-metastases at the time of resection. In one study, micro-metastases were detected in the resected rib in up to 78% of cases, and was dependent on the histological technique used.[21] This is a higher detection rate than the current study, but the results are comparable due to different techniques and tissues used between the studies. The high rate of micro-metastases may be a reason that the present results show significant under-staging of nodal disease, and perhaps clinicians could consider lowering the threshold for treating patients with systemic neoadjuvant therapy.

Previously published research from our centre has shown N-stage, LNM count, and volume of nodal disease to have prognostic significance in patients with OC.[22,23] Nodal disease in these studies probably continues to be an important prognostic indicator, but the radiological staging is likely to have under-estimated the total nodal disease burden in those patient cohorts. Results of staging performance have also been published from our centre. These studies compared CECT and EUS with pN-staging. Blackshaw *et al.*[24] focused on the accuracy of N-staging in GOJ tumours and found significant differences in agreement, sensitivity, and specificity between Sw type II and type III tumours. Weaver *et al.*[13] found agreement, sensitivity, and specificity of N-staging was 0.603, 79%, and 84% for CECT and 0.610, 91%, and 68% for EUS. The results of the current study show poorer agreement and sensitivity. There are a number of reasons for these findings, including disease evolution, greater interobserver variability between reporters, and fewer, but more specialised upper GI cancer pathologists reporting the resection specimens, with possibly higher rates of LNM detection.[15] The accuracy of diagnosing N+ disease with CECT, EUS, and PET/CT was 54.5%, 55.4%, and 57.1%, respectively. In a clinical context, these results are unsatisfactory given that the presence of LNMs is such a major prognostic indicator.[8] The sensitivity and specificity for identifying N0 versus N+ disease with CECT, EUS, and PET/CT was 39.7% and 77.3%, 42.6% and 75%, and 35.3% and 90.9%. Specificity results are comparable with past meta-analyses, but sensitivity results are lower for all techniques. Previously published literature states sensitivity for N-staging of CECT, EUS and PET/CT is 50%, 80% and 57%, and specificity is 83%, 70% and 85%, respectively.[2] However, this meta-analysis was conducted prior to this centralisation of many upper GI cancer services. The

reduced sensitivity of staging investigations is supported by the current results, which demonstrate that under-staging is more common for all techniques.

As current investigations are unreliable for differentiating N0 from N+ disease, future research should focus on investigating and developing new methods of predicting the likelihood of lymph node involvement. Surrogate markers of LNMs, such as texture analysis of the primary tumour and other non-invasive quantitative imaging techniques, may allow better risk stratification of patients, provide more powerful prognostic data, and further inform optimum treatment decisions.[25,26] Magnetic resonance imaging (MRI) may provide an alternative staging technique. Research studies have demonstrated variable diagnostic ability, with the sensitivity, specificity, and accuracy ranging between 38–62%, 68–85% and 64–77%, respectively. These current results are comparable to CT, EUS, and PET/CT, but continuing improvements in functional MRI technology may yield further developments.[27,28]

### *Strengths of study*

This study provides radiological–pathological correlation in a group of OC patients with discordant nodal staging. Radiological–pathological correlation is essential for understanding the limitations of staging techniques and identifies areas requiring further research. All patients were discussed at the regional MDT and the management plan for each individual was decided upon in consensus. The regional MDT covers a large population of over 1.4 million people and is highly experienced in the management of OC. Histopathological examination was performed by consultant GI pathologists according the guidelines defined by the RCPath.[14] Strict criteria were implemented to control the selection of patients for the current study, which compares imaging findings to reference standard pathological staging. The majority of patients received neoadjuvant therapy, which can alter the stage of disease between pre-treatment imaging and surgical resection. To control for this, only patients with Mandard TRG 4 or 5 were included, which should allow a more direct comparison with the final pathological resection specimen. The majority of patients tend to have a TRG 4 or 5 response.[29]

### *Limitations*

As a result of neoadjuvant therapy, there is a time lag between radiological staging and surgical resection, which could allow for tumour progression and LNM development; however, the median time period in the present study was 3 months. In addition, patients with an “open-and-close” procedure were excluded, which further demonstrates radiological disease under-staging. There are also known limitations of pathological lymph node examination. Approximately 3 mm sections are taken through lymph nodes once they are mounted in a cassette, but this may be performed with varying skill and consistency. Micro-metastases may be missed if not bisected during preparation, and this suggests that the true incidence of micro-metastases in this cohort of patients may be even greater. Although the RCPath define the minimum requirements for pathological reporting, there is no recommended, standardised method for lymph node preparation and assessment in OC, at present. The centralised upper GI cancer service receives patients referred from several local NHS trusts. As a result, multiple readers from different hospitals report the staging CECT examinations. During this period, four endosonographers performed the EUS examinations in three different hospitals. All PET/CT examinations were performed using the same system and protocol and were reported by four different consultant radiologists; however, all staging was performed according to the TNM 7$^{th}$ edition.

In conclusion, this evaluation of contemporary staging performance over a 5-year period in a centralised upper GI cancer service has shown poor N-staging accuracy for CECT, EUS, and PET/CT. Radiological–pathological correlation in patients staged N0 has shown a large number of small LNMs (<6 mm) that are extremely challenging to diagnose directly from medical imaging. The findings of the current study have significant implications for patient care, because radiological staging results influence treatment decisions made by the MDT. Future research should focus on prediction of the likelihood of lymph node involvement as current lymph node imaging is inadequate.

## Acknowledgements

K.F. received funding from Tenovus Cancer Care, UK (Grant reference TIG2016/04). We also acknowledge the members of the South East Wales Upper GI cancer MDT.

## References

1. Roberts SA, Kay C. Oesophagus and stomach cancers. In: Nicholson T, editor. *Recommendations for cross-sectional imaging in cancer management*. 2nd edn. London: Royal College of Radiologists; 2014.
2. van Vliet EP, Heijenbrok-Kal MH, Hunink MG, *et al*. Staging investigations for oesophageal cancer: a meta-analysis. *Br J Cancer* 2008;**98**(3):547–57.
3. Tangoku A, Yamamoto Y, Furukita Y, *et al*. The new era of staging as a key for an appropriate treatment for esophageal cancer. *Ann Thorac Cardiovasc Surg* 2012;**18**(3):190–9.
4. Okada M, Murakami T, Kumano S, *et al*. Integrated FDG-PET/CT compared with intravenous contrast-enhanced CT for evaluation of metastatic regional lymph nodes in patients with resectable early stage esophageal cancer. *Ann Nucl Med* 2009;**23**(1):73–80.
5. Foley KG, Lewis WG, Fielding P, *et al*. N-staging of oesophageal and junctional carcinoma: is there still a role for EUS in patients staged N0 at PET/CT? *Clin Radiol* 2014;**69**(9):959–64.
6. Cancer Research UK. *Oesophageal Cancer Statistics*. 2016. Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/oesophageal-cancer. Accessed 22 November 2016.
7. Cancer Research UK. *Oesophageal Cancer Incidence Statistics*. 2016. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/oesophageal-cancer/incidence. Accessed 20 December 2016.
8. Kayani B, Zacharakis E, Ahmed K, *et al*. Lymph node metastases and prognosis in oesophageal carcinoma—a systematic review. *Eur J Surg Oncol* 2011;**37**(9):747–53.

9. Medical Research Council Oesophageal Cancer Working Group. Surgical resection with or without preoperative chemotherapy in oesophageal cancer: a randomised controlled trial. *Lancet* 2002;**359**(9319):1727–33.
10. Koenig AM, Prenzel KL, Bogoevski D, *et al*. Strong impact of micrometastatic tumor cell load in patients with esophageal carcinoma. *Ann Surg Oncol* 2009;**16**(2):454–62.
11. Mandard AM, Dalibard F, Mandard JC, *et al*. Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma. Clinicopathologic correlations. *Cancer* 1994;**73**(11):2680–6.
12. Sobin L, Gospodarowicz M, Wittekind C. *UICC TNM Classification of Malignant Tumours*. 7th edn. New York: Wiley; 2009.
13. Weaver SR, Blackshaw GRJC, Lewis WG, *et al*. Comparison of special interest computed tomography, endosonography and histopathological stage of oesophageal cancer. *Clin Radiol* 2004;**59**(6):499–504.
14. Mapstone N. *Dataset for the Histopathological Reporting of Oesophageal Carcinoma*. 2nd edn. London: Royal College of Pathologists; 2007.
15. Weaver DL. Pathology evaluation of sentinel lymph nodes in breast cancer: protocol recommendations and rationale. *Mod Pathol* 2010;**23**(Suppl. 2):S26–32.
16. Grotenhuis BA, van Heijl M, Wijnhoven BP, *et al*. Lymphatic micrometastases in patients with early esophageal adenocarcinoma. *J Surg Oncol* 2010;**102**(7):863–7.
17. Martin LW, D'Cunha J, Wang X, *et al*. Detection of occult micrometastases in patients with clinical stage I non-small-cell lung cancer: a prospective analysis of mature results of CALGB 9761 (Alliance). *J Clin Oncol* 2016;**34**(13):1484–91.
18. Eisenhauer EA, Therasse P, Bogaerts J, *et al*. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009;**45**(2):228–47.
19. Glickman JN, Torres C, Wang HH, *et al*. The prognostic significance of lymph node micrometastasis in patients with esophageal carcinoma. *Cancer* 1999;**85**(4):769–78.
20. Izbicki JR, Hosch SB, Pichlmeier U, *et al*. Prognostic value of immunohistochemically identifiable tumor cells in lymph nodes of patients with completely resected esophageal cancer. *N Engl J Med* 1997;**337**(17):1188–94.
21. Ryan P, McCarthy S, Kelly J, *et al*. Prevalence of bone marrow micrometastases in esophagogastric cancer patients with and without neoadjuvant chemoradiotherapy. *J Surg Res* 2004;**117**(1):121–6.
22. Chan DS, Fielding P, Roberts SA, *et al*. Prognostic significance of 18-FDG PET/CT and EUS-defined tumour characteristics in patients with oesophageal cancer. *Clin Radiol* 2013;**68**(4):352–7.
23. Foley KG, Fielding P, Lewis WG, *et al*. Prognostic significance of novel $^{18}$F-FDG PET/CT defined tumour variables in patients with oesophageal cancer. *Eur J Radiol* 2014;**83**(7):1069–73.
24. Blackshaw G, Lewis WG, Hopper AN, *et al*. Prospective comparison of endosonography, computed tomography, and histopathological stage of junctional oesophagogastric cancer. *Clin Radiol* 2008;**63**(10):1092–8.
25. Ganeshan B, Skogen K, Pressney I, *et al*. Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: preliminary evidence of an association with tumour metabolism, stage, and survival. *Clin Radiol* 2012;**67**(2):157–64.
26. Dong X, Xing L, Wu P, *et al*. Three-dimensional positron emission tomography image texture analysis of esophageal squamous cell carcinoma: relationship between tumor $^{18}$F-fluorodeoxyglucose uptake heterogeneity, maximum standardized uptake value, and tumor stage. *Nucl Med Commun* 2013;**34**(1):40–6.
27. Wu LF, Wang BZ, Feng JL, *et al*. Preoperative TN staging of esophageal cancer: comparison of miniprobe ultrasonography, spiral CT and MRI. *World J Gastroenterol* 2003;**9**(2):219–24.
28. Nishimura H, Tanigawa N, Hiramatsu M, *et al*. Preoperative esophageal cancer staging: magnetic resonance imaging of lymph node with ferumoxtran-10, an ultrasmall superparamagnetic iron oxide. *J Am Coll Surg* 2006;**202**(4):604–11.
29. Noble F, Nolan L, Bateman AC, *et al*. Refining pathological evaluation of neoadjuvant therapy for adenocarcinoma of the esophagus. *World J Gastroenterol* 2013;**19**(48):9282–93.

Clinical Radiology 69 (2014) 959–964

Contents lists available at ScienceDirect

Clinical Radiology

journal homepage: www.clinicalradiologyonline.net

# N-staging of oesophageal and junctional carcinoma: Is there still a role for EUS in patients staged N0 at PET/CT?

K.G. Foley [a,*], W.G. Lewis [b], P. Fielding [c], A. Karran [b], D. Chan [b], P. Blake [b], S.A. Roberts [a]

[a] Department of Radiology, University Hospital of Wales, Cardiff, UK
[b] Department of Surgery, University Hospital of Wales, Cardiff, UK
[c] Department of Wales Research & Diagnostic Positron Emission Tomography Imaging Centre (PETIC), University Hospital of Wales, Cardiff, UK

ARTICLE INFORMATION

*Article history:*
Received 23 December 2013
Received in revised form
16 April 2014
Accepted 25 April 2014

AIM: To assess whether separate endoscopic ultrasound (EUS) lymph node (N)-staging is still of prognostic value in those staged node negative (N0) at combined positron-emission tomography/computed tomography (PET/CT) in patients with oesophageal cancer (OC).

MATERIALS AND METHODS: One hundred and seventeen consecutive patients [median age 67 years; 88 male; 98 cases of adenocarcinoma, 19 cases of squamous cell carcinoma (SCC)] staged as N0 at PET/CT underwent EUS to record tumour (T)- and N-stage. The patients were subsequently separated into two groups: EUS N0 ($n = 78$) and EUS N+ ($n = 39$). Survival analysis using Kaplan–Meier and Cox's proportional hazard methods was performed. Primary outcome was overall survival from diagnosis.

RESULTS: EUS N-stage and EUS N0 versus EUS N+ ($p = 0.005$ and $p = 0.001$, respectively) were found to be significantly and independently associated with survival in two models of multivariate analysis, in patients staged N0 at PET/CT. EUS T-stage was significantly associated with survival on univariate analysis.

CONCLUSION: EUS N-staging still has prognostic value in patients staged N0 at PET/CT. There is a significant difference in survival between EUS N0 and positive nodal EUS status in those staged N0 at PET/CT, suggesting PET/CT is unreliable for local staging. PET/CT and EUS continue to have complimentary roles in OC staging.



## Introduction

Accurate assessment of lymph node status is vital in patients with oesophageal cancer (OC) as prognosis remains poor, with overall 5-year survival rates approximately 13%.[1] More than 8000 patients each year are diagnosed with OC, accounting for 3% of all cancers in the UK.[2]

Currently, routine staging of OC is performed with a combination of computed tomography (CT), endoscopic ultrasonography (EUS) and 2-[$^{18}$F]-fluoro-2-deoxy-D-glucose (FDG) positron-emission tomography (PET) combined with computed tomography (PET/CT). Treatment strategies are chosen for OC patients after consideration of

* Guarantor and correspondent: K.G. Foley, Department of Radiology, University Hospital of Wales, Heath Park Campus, Heath, Cardiff CF14 4XW, UK. Tel.: +44 2920 747747; fax: +44 2920 743029.
*E-mail address:* kfoley@doctors.org.uk (K.G. Foley).

0009-9260/$ – see front matter 
http://dx.doi.org/10.1016/j.crad.2014.04.023

the radiological staging examinations and thorough discussion with the patient.

Patients with potentially curative disease can be managed in different ways, and this decision is significantly influenced by the lymph node (N)-stage of disease. These management choices include surgery alone, neoadjuvant chemotherapy prior to surgery, or definitive chemoradiotherapy (dCRT); however, the evidence regarding which modality provides the greatest survival benefit is still lacking. It is largely agreed that there are two main options: dCRT or neo-adjuvant chemotherapy followed by surgery.[3]

Studies including the RTOG 85-01 trial have demonstrated a greater survival benefit with dCRT compared to definitive chemotherapy (dCT), although the incidence of toxicity was greater.[4,5] The Medical Research Council Oesophageal Cancer Working Group published data from a randomized, controlled trial showing that preoperative chemotherapy improves survival without adverse effects in patients with resectable OC.[6]

Other studies have been unable to demonstrate the benefit of surgical resection following neo-CRT.[7,8] Likewise, no significant difference in survival between patients receiving dCRT or surgery alone has been shown in studies, such as the CURE trial.[9–11]

Until a multicentre, randomized, controlled trial demonstrates the optimum treatment strategy, patients will continue to receive neoadjuvant chemotherapy prior to surgery in disease staged $\geq$T3 or $\geq$N1. Therefore, it is important that the patient is staged accurately, and N0 is differentiated from N1 disease. This will ensure that patients avoid unnecessary chemotherapy with potential complications of toxicity if over-staged, or are withheld potentially beneficial neoadjuvant chemotherapy if under-staged.[6]

EUS has been shown to be the most accurate method of assessing nodal status. On meta-analysis, the sensitivity and specificity of detecting regional lymph node metastases is 80% and 70%, respectively,[12] and the accuracy of overall N-staging in 66%, compared with 68% at PET/CT.[13] With the increased utilization of PET/CT in the modern investigation of OC, PET/CT may be used for complete assessment of nodal and distant metastases, as comparable accuracy in assessing overall nodal status with EUS has been shown.

The aim of the present study was to investigate whether there is any added prognostic information obtained at EUS in those patients staged N0 at PET/CT.

## Materials and methods

Patients with OC and oesophago-gastric junction (OGJ) tumours referred to the Regional Upper GI cancer network staged N0 at PET/CT examination between December 2008 and May 2012, were studied. The PET/CT examinations were performed in two sites and were double reported by experienced consultant radiologists. A consensus was reached if there were discrepancies between reports. Staging of OC was performed according to TNM 7th edition.[14]

EUS was performed at two sites by three experienced endosonographers. Again, staging was performed according to the TNM 7th edition.[14] Patients were separated into two groups for analysis: EUS N0 or EUS N+, if more than one regional lymph node was involved.

PET/CT was performed in patients treated with potentially curative intent. Patients were included in the study if staged N0 at PET/CT, and EUS was subsequently performed. The primary outcome was measured as overall survival from diagnosis.

### *PET/CT acquisition protocol*

PET/CT examinations were performed at two centres. At the first centre, 51 patients underwent PET/CT examinations performed using a Philips 16 section Gemini GXL dedicated PET/CT system (Philips Medical Systems, Cleveland, OH, USA). The uptake time was 60 min. A standard dose of 350 MBq FDG was injected. Reconstructions were performed using a three-dimensional (3D) acquisition with non-time-of-flight acquisition for 4 min per bed position. At the second centre, 72 patients underwent FDG PET/CT examination using a GE 690 PET/CT machine (GE Healthcare, Pollards Wood, Buckinghamshire, UK). Serum glucose levels were routinely checked and confirmed to be $<$7 mmol/l prior to proceeding with imaging. Patients received a dose of 4 MBq FDG per kilogram of body weight. Uptake time was 90 min. CT images were acquired in a helical acquisition with a pitch of 0.98 and a tube rotation speed of 0.5 s. Tube output was 120 kVp with output modulation between 20 and 200 mA. The matrix size for the CT acquisition was $512 \times 512$ pixels with a 50 cm field of view. PET images were acquired at 3 min per field of view. The length of the axial field of view was 15.7 cm. Images were reconstructed with the ordered subset expectation maximization algorithm, with 24 subsets and two iterations. The matrix size was $256 \times 256$ pixels, using the VUE Point™ time-of-flight algorithm. At both centres, patients were starved for 6 h prior to tracer administration and no oral or intravenous contrast medium was administered. In all cases, regional lymph nodes were assigned as “positive” if they showed discernable increased tracer uptake compared to the background. No particular threshold of FDG uptake was used to define positivity on the PET/CT examinations. For inclusion in the study, all patients were reported as negative for nodal and distant metastatic disease at PET/CT. No further PET/CT variables were used in the subsequent survival analysis.

### *Details of EUS*

An initial endoscopic examination was performed using a 9 mm diameter Olympus P-10 gastroscope (Olympus Medical, Southend, UK) to assess the degree of oesophageal luminal stenosis. Patients with an estimated oesophageal luminal diameter $<$15 mm underwent EUS using the smaller-diameter MH-908 oesophagoprobe (Olympus Medical). Oesophageal dilation (Savary-Gilliard, Cook Medical, Bloomington, IN, USA) was performed before endosonography for patients with oesophageal lumens

<9 mm. If the luminal diameter was >15 mm, the standard radial echoendoscope was used (UM-20, Olympus Medical). The primary oesophageal tumour was assessed, together with an evaluation of the para-oesophageal anatomical structures as described previously.[15] The criteria for malignant lymphadenopathy specified a hypoechoic pattern, a spherical contour, the presence of a distinct border, and a short axis diameter of 6 mm or more. For each patient, two variables were recorded; EUS T-stage (T1–4) and EUS N-stage (N0–3) according to TNM 7th edition.[14] A third variable was derived from the EUS N-stage, N0 versus N+, a variable describing positive nodal status.

### *Treatment*

An appropriate management plan was selected based on radiological staging, patient choice, and relevant co-morbidity according to algorithms used by the Regional Upper GI cancer network.[16–18]

### *Follow-up and survival*

Patients were followed-up every 3 months for the first year, then every 6 months thereafter. No patients were lost to follow-up and death certification was obtained from the Office for National Statistics.

### *Ethical approval*

Scientific review by the Research Review Board was performed, and institutional research and development approval was obtained. The Review Board confirmed that formal ethical approval was not required for this study.

### *Statistical methods*

Grouped data were expressed as median (range) and statistical analysis for non-parametric data was used. Cumulative survival was calculated according to the life-table method of Kaplan and Meier,[19] and differences between groups were analysed with the log-rank test. Cox's proportional hazards model was used to assess the prognostic value of individual categorical variables.[20] Two models of data analysis were performed. The first model included EUS T-stage and N-stage; the second model included EUS T-stage and the N-stage was simplified to N0 versus N+, to assess the influence of positive EUS nodal status in those staged N0 at PET/CT. Statistical tests were two-sided and the level of significance taken as $p < 0.05$. Data analysis was performed using SPSS version 18 (SPSS, Chicago, IL, USA).

**Table 1**
Details of patient demographics, histology, and tumour location in this population.

| | |
|---|---|
| Sex; male:female (%) | 88:29 (75.2:24.8) |
| Median age, years (range) | 67 (24–82) |
| Histology (%) | |
| Adenocarcinoma | 98 (83.8) |
| Squamous cell carcinoma | 19 (16.2) |
| Tumour location (%) | |
| Middle third oesophagus | 20 (17.1) |
| Lower third oesophagus | 53 (45.3) |
| Oesophago-gastric junction | 44 (37.6) |
| Siewert type I | 5 (4.3) |
| Siewert type II | 12 (10.3) |
| Siewert type III | 27 (23.1) |

**Table 2**
Results of first multivariate analysis [including endoscopic ultrasound (EUS) T-stage and N-stage].

| EUS staging | *p*-Value | Hazard ratio | Degrees of freedom | 95% Confidence intervals |
|---|---|---|---|---|
| N-stage | 0.005 | | 3 | |
| N1 | 0.005 | 3.055 | 1 | 1.392–6.707 |
| N2 | 0.002 | 4.707 | 1 | 1.778–12.459 |
| N3 | 0.529 | 1.616 | 1 | 0.363–7.190 |
| T-stage | 0.194 | | 3 | |

## Results

One hundred and seventeen patients staged N0 at PET/CT and examined with EUS between December 2008 and May 2012 were studied. PET/CT examinations were performed in two sites, 47 in the first site and 70 in the second. Table 1 details the demographics of the patient population.

The EUS T-stage of patients was T1 ($n = 18$, 15.4%), T2 ($n = 16$, 13.7%), T3 ($n = 75$, 64.1%), and T4a ($n = 8$, 6.8%). The N-stage of patients at EUS was N0 ($n = 78$, 66.7%), N1 ($n = 23$, 19.7%), N2 ($n = 9$, 7.7%), and N3 ($n = 7$, 6%). Patients were then separated into two groups to derive the third variable N0 versus N+; EUS N0 ($n = 78$, 66.7%) and EUS N+ ($n = 39$, 33.3%). All patients with metastases at PET/CT were excluded.

One hundred and five patients were treated with curative intent, with 12 receiving palliation, as they were not fit for curative treatment following discussion by the multidisciplinary team (MDT). Seventy-three underwent surgical resection with 40 receiving neoadjuvant chemotherapy and one neoadjuvant chemoradiotherapy. Twenty-nine patients

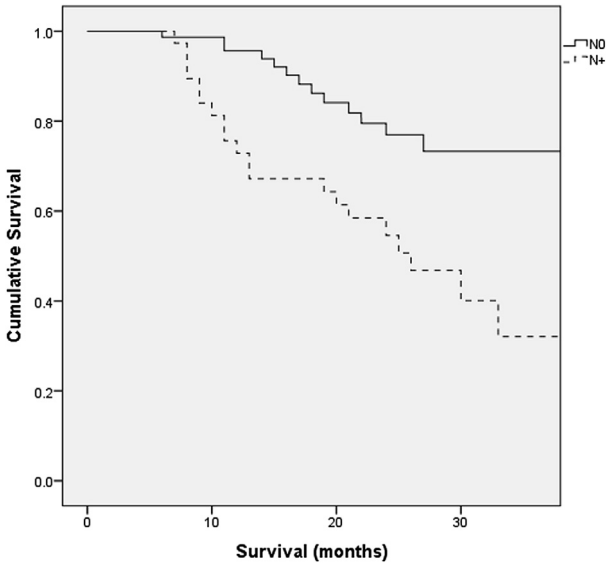



**Figure 1** Cumulative survival related to N0 versus N+.

had definitive chemoradiotherapy (dCRT) and three patients were treated with endoscopic mural resection (EMR).

At the time of analysis, 84 patients were alive (71.8%) and 33 had died (28.2%). Overall 1 year survival was 74.4% (87/117) with 39.3% (46/117) 2 year survival. Ninety-nine patients were followed up for at least 1 year (84.6%), and 72 patients (61.5%) for at least 2 years, or until death.

### *Univariate analysis*

Univariate analysis of EUS T-stage, N-stage, and N0 versus N+ was performed using log-rank analysis of the Kaplan–Meier method.[19] All variables were significantly associated with overall survival from diagnosis; EUS T-stage (T1–4) ($X^2$ 8.321, df 3, $p = 0.040$), EUS N-stage (N0–3) ($X^2$ 14.879, df 3, $p = 0.002$), and EUS N0 versus N+ ($X^2$ 11.325, df 1, $p = 0.001$).

### *Multivariate analysis*

A multivariate analysis of the factors significant on univariate analysis were entered into two alternative Cox's proportional hazards model.[20]

*Model 1*

When EUS T-stage and N-stage were entered into Cox's proportional hazards model, only EUS N-stage was significantly and independently associated with duration of survival (Table 2).

*Model 2*

When EUS T-stage and N0 versus N+ stage were entered into Cox's proportional hazard model, N0 versus N+ was significantly and independently associated with duration of survival (HR 3.105, 95% CI: 1.543–6.247, $p = 0.001$; Fig 1).

## Discussion

The present study has confirmed the importance, and shown the continued benefit, of fully assessing nodal status with EUS when staging patients with OC. EUS N-stage is an independent predictor of survival ($p = 0.005$), as shown in previous studies.[21] Separation of patients into EUS N0 and EUS N+ groups, in patients staged N0 at PET/CT, was also significantly and independently associated with survival ($p = 0.001$), and highlights the prognostic value of EUS nodal assessment in those with negative nodal status at PET/CT.

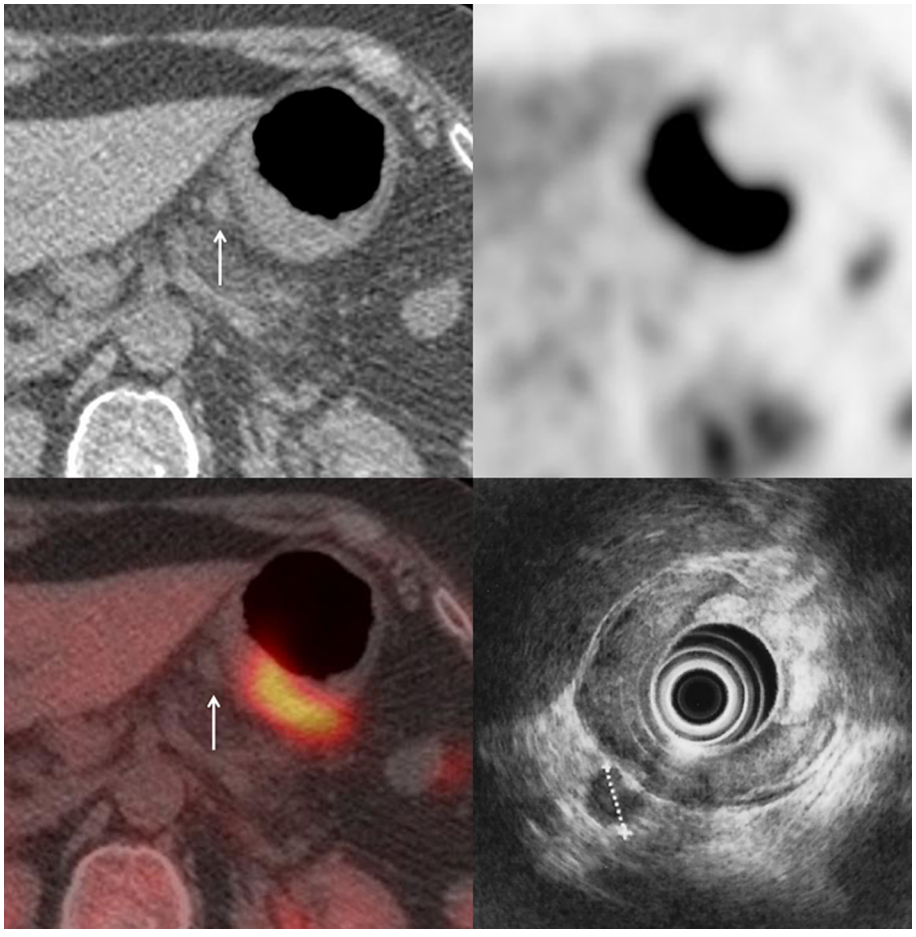


**Figure 2** Axial CT, PET, integrated PET/CT, and EUS images in a patient with a gastro-oesophageal junction tumour staged N0 at PET/CT and N3 at EUS. This example shows a non-FDG-avid regional lymph node metastasis (arrow). Following resection, the histopathological N-stage was pN3.

This adds to further evidence supporting N-staging with EUS using TNM 7th edition.[14,21] Staging of involved regional lymph nodes with EUS has been shown to have good agreement when compared to histopathological stage.[22] The examination is operator dependent, but reliable when performed by experienced endosonographers.

PET/CT has previously been shown to have good diagnostic accuracy in the detection of involved lymph node metastases. Okada et al.[23] found the sensitivity, specificity, accuracy, positive, and negative predictive values to be 60%, 99.5%, 94.8%, 93.8%, and 94.8% respectively, in 18 patients with a total of 210 lymph nodes excised following radical oesophagectomy.[23] Similarly, a meta-analysis demonstrated the sensitivity and specificity of regional lymph node metastases on PET/CT to be 57% and 85%, respectively; however, the study concluded that EUS is more sensitive for local staging.[12] The results of the present study concur with these studies, as they have demonstrated that PET/CT is unreliable in excluding regional lymph node metastases, with a significant difference in survival between N0 and N+ in the present cohort.

The significant difference between PET/CT and EUS is likely to be explained by the inability of PET/CT to differentiate peri-tumoural nodes from the primary tumour. The poor spatial resolution of PET/CT, approximately 4 mm, remains a limiting factor in accurate assessment of local lymph nodes (Fig 2).

Accurate assessment of nodal status in patients with OC is vital to inform optimum treatment strategies. The differentiation of N0 from N+ has significant implications for the patient and will guide the MDT when considering the appropriateness of neoadjuvant therapy. This study has confirmed the need for concurrent N-staging with EUS, as positive nodal status is a significant prognostic indicator of survival.

The first model shows EUS N-stage to be independently associated with survival for EUS N1 and N2, therefore demonstrating that PET/CT is unreliable in detection of regional lymph node metastases and that those with positive nodal status on EUS have a poorer outcome. The absence of statistical significance of the N3 group in this model is likely to be a reflection of the small numbers. Whereas the second model demonstrated N0 versus N+ disease as an independent predictor of survival, this result is biased; the groups N0 versus N+ are derived from the EUS N-stage and this exaggerates the statistical significance by grouping N1, N2, and N3 together. This results in an apparent poorer prognosis. However, in essence, both models support the hypothesis.

### *Limitations of method*

The present study has potential limitations. The PET/CT examinations were performed across two sites using different scanners, protocols, and uptake times. Comparison between examinations from different sites may need to be analysed with caution, as the above factors may affect image quality and diagnostic accuracy.

The accuracy of EUS examination may also be a limitation. All examinations were performed by experienced clinicians and data regarding their accuracy has been published showing good agreement with histopathological status.[24] The experience of the endosonographers may be a limitation of the study, as the findings may not be generalizable to other institutions. Advances in pathological analysis in recent years have meant a greater detection of micrometastases in lymph nodes that may appear morphologically normal and therefore, will affect agreement.[25] This implies the true prevalence of N+ disease in this cohort may be higher than the imaging suggests.

As this is a retrospective study of routine data obtained as part of the OC pathway, the relationship between these novel parameters and overall survival are described following treatment influenced by management decisions. This may not necessarily reflect the natural history of the disease, but is unavoidable in this context.

### *Strengths of study*

Despite its limitations, the present study has several strengths. All patient data were prospectively maintained on a dedicated database, collected from a well-defined geographical area with an established Upper GI cancer network working in a dedicated MDT. This represents a large patient cohort of more than 100 patients. The survival data are especially robust as no patients were lost to follow-up and causes and exact dates of death were obtained from death certificates provided by the Office for National Statistics.

In conclusion, EUS and PET/CT remain complementary in staging OC and should continue to be used in a multi-technique approach to the staging of OC. EUS N-stage and positive nodal status are significantly and independently associated with survival on multivariate analysis in patients staged N0 on PET/CT. The present study has shown that T- and N-staging using EUS remains an important prognostic indicator, informs clinicians, and influences management decisions.

## Acknowledgements

The authors acknowledge the contributions of the South East Wales Upper Gastrointestinal Cancer MDT, and colleagues from the Cheltenham PET/CT Centre.

## References

1. Cancer Research UK. Statistics and outlook for oesophageal cancer. Published June 17th, 2013. Available at: http://www.cancerresearchuk.org/cancer-help/type/oesophageal-cancer/treatment/statistics-and-outlook-for-oesophageal-cancer. [accessed 29.10.13].
2. Cancer Research UK. Oesophageal cancer incidence statistics. Published December 8th, 2011, Updated June 18th, 2013. Available at: http://www.cancerresearchuk.org/cancer-info/cancerstats/types/oesophagus/incidence/. [accessed 15.04.14].
3. Crosby TD, Evans M. The management of locally advanced carcinoma of the oesophagus — SCOPE for improvement? *Clin Oncol* 2009;**21**:281–4.
4. Cooper JS, Guo MD, Herskovic A, et al. Chemoradiotherapy of locally advanced esophageal cancer: long-term follow-up of a prospective

randomized trial (RTOG 85-01). Radiation Therapy Oncology Group. *JAMA* 1999;**281**:1623–7.
5. Coia LR, Minsky BD, Berkey BA, et al. Outcome of patients receiving radiation for cancer of the esophagus: results of the 1992–1994 Patterns of Care Study. *J Clin Oncol* 2000;**18**:455–62.
6. Medical Research Council Oesophageal Cancer Working Group. Surgical resection with or without preoperative chemotherapy in oesophageal cancer: a randomised controlled trial. *Lancet* 2002;**359**:1727–33.
7. Bedenne L, Michel P, Bouche O, et al. Chemoradiation followed by surgery compared with chemoradiation alone in squamous cancer of the esophagus: FFCD 9102. *J Clin Oncol* 2007;**25**:1160–8.
8. Stahl M, Walz MK, Stuschke M, et al. Phase III comparison of preoperative chemotherapy compared with chemoradiotherapy in patients with locally advanced adenocarcinoma of the esophagogastric junction. *J Clin Oncol* 2009;**27**:851–6.
9. Carstens H, Albertsson M, Friesland S, et al. A randomised trial of chemoradiotherapy versus surgery alone in patients with resectable esophageal cancer. *J Clin Oncol* 2007;**25**:A4530.
10. Chan A, Wong A. Is combined chemotherapy and radiation therapy equally effective as surgical resection in localized esophageal carcinoma? *Int J Radiat Oncol Biol Phys* 1999;**45**:265–70.
11. Chiu PW, Chan AC, Leung SF, et al. Multicenter prospective randomized trial comparing standard esophagectomy with chemoradiotherapy for treatment of squamous esophageal cancer: early results from the Chinese University Research Group for Esophageal Cancer (CURE). *J Gastrointest Surg* 2005;**9**:794–802.
12. van Vliet EP, Heijenbrok-Kal MH, Hunink MG, et al. Staging investigations for oesophageal cancer: a meta-analysis. *Br J Cancer* 2008;**98**:547–57.
13. Choi J, Kim SG, Kim JS, et al. Comparison of endoscopic ultrasonography (EUS), positron emission tomography (PET), and computed tomography (CT) in the preoperative locoregional staging of resectable esophageal cancer. *Surg Endosc* 2010;**24**:1380–6.
14. Sobin LH, Gospodarowicz MK, Wittekind CH. *UICC TNM classification of malignant tumours*. 7th ed. New York: Wiley; 2009.
15. Bowrey DJ, Clark GW, Roberts SA, et al. Endosonographic staging of 100 consecutive patients with esophageal carcinoma: introduction of the 8-mm esophagoprobe. *Dis Esophagus* 1999;**12**:258–63.
16. Crosby TD, Brewster AE, Borley A, et al. Definitive chemoradiation in patients with inoperable oesophageal carcinoma. *Br J Cancer* 2004;**90**:70–5.
17. Stephens MR, Lewis WG, Brewster AE, et al. Multidisciplinary team management is associated with improved outcomes after surgery for esophageal cancer. *Dis Esophagus* 2006;**19**:164–71.
18. Gwynne S, Hurt C, Evans M, et al. Definitive chemoradiation for oesophageal cancer—a standard of care in patients with non-metastatic oesophageal cancer. *Clin Oncol* 2011;**23**:182–8.
19. Kaplan E, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958;**53**:457–81.
20. Cox D. Regression models and life-tables. *J Roy Stat Soc* 1972;**34**:187–220.
21. Twine CP, Roberts SA, Lewis WG, et al. Prognostic significance of endoluminal ultrasound-defined disease length and tumor volume (EDTV) for patients with the diagnosis of esophageal cancer. *Surg Endosc* 2010;**24**:870–8.
22. Blackshaw G, Lewis WG, Hopper AN, et al. Prospective comparison of endosonography, computed tomography, and histopathological stage of junctional oesophagogastric cancer. *Clin Radiol* 2008;**63**:1092–8.
23. Okada M, Murakami T, Kumano S, et al. Integrated FDG-PET/CT compared with intravenous contrast-enhanced CT for evaluation of metastatic regional lymph nodes in patients with resectable early stage esophageal cancer. *Ann Nucl Med* 2009;**23**:73–80.
24. Weaver SR, Blackshaw GR, Lewis WG, et al. Comparison of special interest computed tomography, endosonography and histopathological stage of oesophageal cancer. *Clin Radiol* 2004;**59**:499–504.
25. Jiao X, Eslami A, Ioffe O, et al. Immunohistochemistry analysis of micrometastasis in pretreatment lymph nodes from patients with esophageal cancer. *Ann Thorac Surg* 2003;**76**:996–9.

Insights Imaging
DOI 10.1007/s13244-016-0505-6

ORIGINAL ARTICLE

# The 100 most cited articles investigating the radiological staging of oesophageal and junctional cancer: a bibliometric analysis

**Kieran G. Foley**[1] · **Arfon Powell**[2] · **Wyn G. Lewis**[2] · **Stuart Ashley Roberts**[3]

Received: 23 March 2016 / Revised: 18 May 2016 / Accepted: 24 May 2016


**Abstract**
*Objectives* Accurate staging of oesophageal cancer (OC) is vital. Bibliometric analysis highlights key topics and publications that have shaped understanding of a subject. The 100 most cited articles investigating radiological staging of OC are identified.
*Methods* The Thomas Reuters Web of Science database with search terms including "CT, PET, EUS, oesophageal and gastro-oesophageal junction cancer" was used to identify all English language, full-script articles. The 100 most cited articles were further analysed by topic, journal, author, year and institution.
*Results* A total of 5,500 eligible papers were returned. The most cited paper was Flamen et al. (n = 306), investigating the utility of positron emission tomography (PET) for the staging of patients with potentially operable OC. The most common research topic was accuracy of staging investigations (n = 63). The article with the highest citation rate (38.00), defined as the number of citations divided by the number of complete years published, was Tixier et al. investigating PET texture analysis to predict treatment response to neo-adjuvant chemo-radiotherapy, cited 114 times since publication in 2011.
*Conclusion* This bibliometric analysis has identified key publications regarded as important in radiological OC staging. Articles with the highest citation rates all investigated PET imaging, suggesting this modality could be the focus of future research.
*Main Messages*
- *This study identifies key articles that investigate radiological staging of oesophageal cancer.*
- *The most common topic was accuracy of staging investigations.*
- *The article with the highest citation rate investigated the use of texture analysis in PET images.*

**Keywords** Bibliometric analysis · Oesophageal cancer · Gastro-oesophageal junction cancer · TNM staging · Citation

## Introduction

Bibliometric analysis assesses the number of times that an article is cited in the literature, and in which particular journal. A citation is received when an article references another peer-reviewed publication. An article that is felt to have greater importance and higher impact by the scientific community is more likely to be cited and therefore may be more influential on current healthcare practice. Articles and journals can be ranked based on the number of citations they receive. Bibliometric analysis also reveals topics of current interest, identifies potential novel techniques and shows historical developments in that subject [1]. Medical researchers have used bibliometric analysis to identify the most influential papers in their clinical specialties, including orthopaedic surgery [2] and oncology [3].

Worldwide, the prognosis of oesophageal cancer, including gastro-oesophageal junction cancer (OC), is poor. Overall 5-year survival in the UK is approximately 13 % [4]. As a part of the diagnostic pathway, patients undergo a variety of staging investigations to assess the extent of disease. Radiological

✉ Kieran G. Foley
foleykg@cardiff.ac.uk

[1] Division of Cancer & Genetics, Cardiff University, Cardiff, UK
[2] Department of Surgery, University Hospital of Wales, Cardiff, UK
[3] Department of Radiology, University Hospital of Wales, Cardiff, UK

Published online: 08 June 2016

staging is performed to further inform management decisions by the multi-disciplinary team (MDT). Accurate radiological staging is vital to ensure the most appropriate treatment is selected. Currently, OC is staged according to the International Union Against Cancer (UICC) Tumour Node Metastasis (TNM) 7th edition [5].

In the UK, patients with OC are initially staged with CT of the thorax and abdomen to exclude irresectable disease or distant metastases. If the patient is deemed suitable for radical treatment, either in the form of definitive chemo-radiotherapy (dCRT) or surgery (± neo-adjuvant therapy), positron emission tomography combined with computed tomography (PET/CT) and endoscopic ultrasound (EUS) are performed for a more detailed assessment of disease stage [6].

This bibliometric analysis of OC staging investigations aims to identify key research that has influenced staging methods, the institutions leading this research, studies that may change staging methods in the future and imaging modalities being focused upon.

## Materials and methods

The Thomas Reuters Web of Science citation indexing database was used to perform the search. The following search terms were used in order to capture the variety of imaging modalities and the different nomenclature of tumours: (oesophag* AND (neoplas* OR cancer* OR carcin* OR tumo* OR malig*)) OR (esophag* AND (neoplas* OR cancer* OR carcin* OR tumo* OR malig*)) OR (gastro-oesophageal junction AND (neoplas* OR cancer* OR carcin* OR tumo* OR malig*)) OR (oesophago-gastric junction AND (neoplas* OR cancer* OR carcin* OR tumo* OR malig*) AND (computed tomography OR CT OR CAT) OR (positron emission tomography OR PET OR F18 OR FDG OR fluorodeoxyglucose) OR (endoscopic ultrasonography OR endoscopic ultrasound OR endosonographic OR EUS) OR (magnetic resonance imaging OR MRI or diffusion weight* OR DWI) AND (stag* OR TNM OR lymph node OR metasta*). The search was performed on 18 September 2015.

All databases within the Thomas Reuters Web of Science were searched. The results were filtered to include only full script articles written in the English language, throughout all available years. The results were sorted by number of citations, with the article with most citations analysed first. The method was developed by Paladugu et al. [7].

The title and abstract of the returned articles were manually assessed to ensure that their relevance and content were in keeping with this field. The inclusion criterion was that the article investigated the use of a single or combination of radiological modalities in patients with OC. This criterion was pre-specified and defined prior to data collection.

Articles were excluded from the list if the content was not relevant to radiological OC staging. The 100 most cited articles were identified and further analysed.

The articles were further evaluated for the publishing journal, names of the first and senior author, the institution and department to which the first author was affiliated, the country of origin, year of publication, the radiological investigation(s) being studied, the topic of the article and the number of citations according to Web of Science. Rank within the top 100 articles was also recorded.

Articles have the opportunity to accrue more citations if they have been published for longer. To adjust for this, a citation rate was calculated, defined as the number of citations divided by the number of complete years published. A list of the ten articles with the highest citation rates is provided.

In addition, the individual and 5-year impact factors in 2014 were recorded for the publishing journal. The overall median 2014 and 5-year impact factor for all journals was calculated.

## Results

The Web of Science search returned 5,500 full articles, written in English language. The 100 most cited articles are listed in Table 1 [8–107].

The article with the highest number of citations (n = 306) was Flamen et al. [8], entitled 'Utility of positron emission tomography for the staging of patients with potentially operable oesophageal carcinoma'. The article ranked 100 in the list was Wu et al. [107], entitled 'Preoperative TN staging of oesophageal cancer: comparison of miniprobe ultrasound, spiral CT and MRI', with 46 citations.

The oldest article was published in 1979 by Daffner et al. [71] 'CT of the Oesophagus. 2. Carcinoma'. Tixier et al. [47] published the most recent paper in the list in 2011, entitled 'Intratumor Heterogeneity Characterized by Textural Features on Baseline F-18-FDG PET Images Predicts Response to Concomitant Radiochemotherapy in Esophageal Cancer', which has been cited 114 times.

The journal with the highest number of published articles was Gastrointestinal Endoscopy (Table 2). Fourteen articles were published with a total of 1675 citations [11, 16, 34, 38, 48, 59, 62, 63, 66, 76, 80, 82, 83, 85]. The 2014 impact factor of the Gastrointestinal Endoscopy was 5.369, with 5-year impact factor 5.225. The journal with the highest impact factor was the Journal of Clinical Oncology (JCO), which had a total of 1,258 citations [8, 12, 15, 18, 23, 70]. Five of these six articles were investigating PET imaging. The 2014 impact factor of JCO was 18.428, with 5-year impact factor 16.996. Overall, the median 2014 impact factor of the journals was 5.238 and median 5-year impact factor was 5.225.

**Table 1** The 100 most cited articles in radiological staging of oesophageal and junctional cancer

| Rank | Number of Citations | First author |
|---|---|---|
| 1 | 306 | Flamen P [8] |
| 2 | 294 | Tio TL [9] |
| 3 | 290 | Kinkel K [10] |
| 4 | 273 | Catalano MF [11] |
| 5 | 271 | Wieder HA [12] |
| 6 | 261 | Botet JF [13] |
| 7 | 242 | Skinner DB [14] |
| 8 | 229 | Ott K [15] |
| 9 | 221 | Bhutani MS [16] |
| 10 | 201 | Flamen P [17] |
| 11 | 193 | Downey RJ [18] |
| 12 | 192 | Flanagan FL [19] |
| 13 | 189 | Picus D [20] |
| 14 | 184 | Rosch T [21] |
| 15 | 183 | Kelly S [22] |
| 16 | 168 | van Westreenen HL [23] |
| 17 | 165 | May A [24] |
| 18 | 164 | Lerut T [25] |
| 19 | 160 | Kato H [26] |
| 20 | 159 | Block MI [27] |
| 20 | 159 | Swisher SG [28] |
| 22 | 157 | Swisher SG [29] |
| 23 | 150 | Luketich JD [30] |
| 24 | 146 | Rice TW [31] |
| 24 | 146 | Vilgrain V [32] |
| 26 | 145 | van Vliet EPM [33] |
| 27 | 143 | Rosch T [34] |
| 28 | 137 | Ziegler K [35] |
| 29 | 136 | Kole AC [36] |
| 30 | 135 | Luketich JD [37] |
| 31 | 133 | Buskens CJ [38] |
| 32 | 132 | Watt I [39] |
| 33 | 130 | Moss AA [40] |
| 34 | 123 | Zuccaro G [41] |
| 35 | 122 | Grimm H [42] |
| 36 | 121 | Vazquez-Sequeiros E [43] |
| 37 | 116 | Yoon YC [44] |
| 38 | 115 | Tio TL [45] |
| 39 | 114 | Cerfolio RJ [46] |
| 39 | 114 | Tixier F [47] |
| 41 | 113 | Vazquez-Sequeiros E [48] |
| 42 | 112 | Dittler HJ [49] |
| 43 | 111 | Rasanen JV [50] |
| 44 | 108 | Quint LE [51] |
| 45 | 107 | Reed CE [52] |
| 46 | 105 | Choi JY [53] |
| 46 | 105 | Rankin SC [54] |
| 46 | 105 | Thompson WM [55] |
| 46 | 105 | Wallace MB [56] |
| 50 | 104 | Quint LE [57] |
| 51 | 103 | Kato H [58] |
| 51 | 103 | Larghi A [59] |
| 53 | 102 | Takashima S [60] |
| 54 | 100 | van Westreenen HL [61] |
| 55 | 98 | Eloubeidi MA [62] |
| 55 | 98 | Hasegawa N [63] |
| 57 | 96 | Levine EA [64] |
| 57 | 96 | Rice TW [65] |
| 59 | 92 | Isenberg G [66] |
| 59 | 92 | Leong T [67] |
| 59 | 92 | Puli SR [68] |
| 62 | 91 | Flamen P [69] |
| 62 | 91 | Lightdale CJ [70] |
| 64 | 89 | Daffner RH [71] |
| 64 | 89 | Kim K [72] |
| 64 | 89 | Meltzer CC [73] |
| 64 | 89 | Vrieze O [74] |
| 68 | 88 | Jones DR [75] |
| 69 | 86 | Chak A [76] |
| 70 | 84 | Hyun SH [77] |
| 70 | 84 | Quint LE [78] |
| 72 | 83 | Murata Y [79] |
| 73 | 82 | Hiele M [80] |
| 74 | 80 | Wakelin SJ [81] |
| 75 | 79 | Scotiniotis IA [82] |
| 76 | 78 | Fockens P [83] |
| 77 | 76 | Beseth BD [84] |
| 77 | 76 | Catalano MF [85] |
| 77 | 76 | Giovannini M [86] |
| 80 | 75 | Rice TW [87] |
| 81 | 74 | Heeren PAM [88] |
| 81 | 74 | Rizk N [89] |
| 83 | 72 | Kostakoglu L [90] |
| 83 | 72 | Moureau-Zabotto L [91] |
| 85 | 70 | Pech O [92] |
| 86 | 69 | Choi JY [93] |
| 86 | 69 | Lehr L [94] |
| 86 | 69 | Lightdale CJ [95] |
| 89 | 68 | Kobori D [96] |
| 89 | 68 | Luketich JD [97] |
| 91 | 67 | Lowe VJ [98] |
| 92 | 62 | Song SY [99] |
| 92 | 62 | Yuan S [100] |
| 94 | 61 | Bar-Shalom R [101] |
| 94 | 61 | McAteer D [102] |
| 94 | 61 | van Westreenen HL [103] |
| 97 | 60 | Eloubeidi MA [104] |

**Table 1** (continued)

| Rank | Number of Citations | First author |
|---|---|---|
| 97 | 60 | Konski A [105] |
| 99 | 59 | Meyers BF [106] |
| 100 | 46 | Wu LF [107] |

Twenty-nine of the 100 articles were published in a radiology-related journal, including nuclear medicine and radiation oncology journals. Thirty-five of the first authors were affiliated to a radiology, nuclear medicine or radiation oncology department, according to the Thomas Reuters Web of Science citation indexing database. Three radiology-related journals, with 5-year impact factor >5.00, published 16 articles in the top 100. These were Radiology (5-year impact factor 7.259; n = 6), Journal of Nuclear Medicine (5-year impact factor 6.280; n = 8) and European Journal of Nuclear Medicine and Molecular Imaging (5-year impact factor 5.090; n = 1).

Researchers from the USA published the greatest number of articles in the 100 most cited (n = 47) [11, 13, 14, 16, 18–20, 27–31, 37, 40, 41, 43, 46, 48, 51, 52, 55–57, 59, 62, 64–66, 68, 70, 71, 73, 75, 76, 78, 82, 84, 85, 87, 89, 90, 95, 97, 98, 104–106], jointly followed by Germany [12, 15, 21, 24, 34, 35, 42, 49, 92, 94] and The Netherlands [9, 23, 33, 36, 38, 45, 61, 83, 88, 103] (n = 10, each) (Table 3). The Technical University of Munich, Germany, was the institution with the joint highest number of publications in the 100 Most Cited (n = 6) and the highest number of citations (1,008) [12, 15, 21, 34, 49, 94]. The University Hospital Gasthuisberg, Leuven, Belgium, also had 6 published articles, with a total of 933 citations [8, 17, 25, 69, 74, 80]. The most cited article was from this institution [8] and written by Dr Patrick Flamen (first author) with Prof. Luc Mortelmans as senior author. Dr Flamen has 3 first author articles in the 100 most cited [8, 17, 69] and a total of 598 citations. Prof. Mortelmans has 4 articles published as senior author [8, 17, 25, 69] and a total of 762 citations.

**Table 3** Number of articles per country of origin in 100 most cited

| Country | Total number of articles |
|---|---|
| USA | 47 |
| Germany | 10 |
| The Netherlands | 10 |
| Belgium | 6 |
| Japan | 6 |
| South Korea | 6 |
| UK | 5 |
| France | 4 |
| China | 2 |
| Australia | 1 |
| Finland | 1 |
| Israel | 1 |
| Switzerland | 1 |

The most common researched topic was the accuracy of radiological staging investigations (n = 63) (Table 4). Several of the study themes overlapped but accuracy of staging was commonly compared between different modalities (n = 29). The investigation of lymph node metastases (n = 15) and radiological response to treatment (n = 14) were also commonly cited topics.

EUS was the most common modality investigated (n = 51), with PET/CT (n = 48) and CT (n = 46) following. The combination of CT, EUS and PET/CT was commonly investigated

**Table 2** Journals with ≥2 articles in 100 most cited

| Journal | Number of articles | 2014 Impact factor | 5-Year impact factor | Total number of citations |
|---|---|---|---|---|
| Gastrointestinal Endoscopy | 14 | 5.369 | 5.225 | 1,675 |
| Annals of Thoracic Surgery | 9 | 3.849 | 4.104 | 1,038 |
| Journal of Nuclear Medicine | 8 | 6.16 | 6.280 | 657 |
| Cancer | 7 | 5.238 | 5.517 | 830 |
| American Journal of Roentgenology | 6 | 2.731 | 3.302 | 764 |
| Journal of Clinical Oncology | 6 | 18.428 | 16.966 | 1,258 |
| Radiology | 6 | 6.867 | 7.259 | 1,044 |
| Endoscopy | 5 | 5.053 | 4.855 | 494 |
| Journal of Thoracic and Cardiovascular Surgery | 5 | 4.168 | 4.068 | 428 |
| Annals of Surgery | 3 | 8.327 | 8.844 | 502 |
| Gut | 3 | 14.66 | 12.553 | 485 |
| International Journal of Radiation Oncology Biology Physics | 3 | 4.258 | 4.359 | 194 |
| American Journal of Gastroenterology | 2 | 10.755 | 9.145 | 193 |
| Annals of Surgical Oncology | 2 | 3.93 | 4.532 | 195 |
| British Journal of Cancer | 2 | 4.836 | 5.305 | 281 |
| Gastroenterology | 2 | 16.716 | 13.811 | 415 |
| Radiotherapy and Oncology | 2 | 4.363 | 4.502 | 181 |
| World Journal of Gastroenterology | 2 | 2.369 | 2.671 | 138 |

together, which is the recommended staging pathway for potentially curable disease in the UK (n = 11) [8, 25, 28, 31, 50, 53, 88, 98, 105]. MRI (n = 5), bone scintigraphy (n = 2), PET alone (n = 1), EUS-FNA (n = 1), US (n = 1) and laparoscopic US (n = 1) were also cited.

The article with the highest citation rate (38.00) was Tixier et al. [47], published in 2011 and investigated texture analysis in OC. The ten articles with the highest citation rates were published between 2002 and 2011 and all involved investigation of PET images (Table 5). Four of the articles investigated treatment response [12, 15, 18, 47]. An international collaboration collecting data that informed the International Union Against Cancer (UICC) Tumour Node and Metastasis (TNM) 7th edition [31] had the second highest citation rate (36.50). Five of the ten articles with the highest citation rates were published in the Journal of Clinical Oncology, which had the highest impact factor (5-year 16.971).

## Discussion

OC is the eighth most common malignancy worldwide, resulting in around 400,000 deaths per annum [108]. This study demonstrates that accuracy of staging was the most frequently studied topic (n = 63) (Table 4). Accurate staging investigations are vital to inform appropriate treatment decisions, providing the best chance of survival for the patient whilst minimising harm from over- or under-treatment. The most cited article was Flamen et al. [8], which investigated the use of PET in potentially operable OC.

**Table 4** Most frequently cited topics of investigation (numbers do not add up to 100 as there are different combinations of topics in the articles)

| Topic | Number of articles |
|---|---|
| Accuracy of staging | 63 |
| Comparison of imaging modalities | 29 |
| Lymph node metastases | 15 |
| Treatment response | 14 |
| Review of staging | 9 |
| Imaging features of malignancy | 9 |
| Prognosis | 7 |
| Distant metastases | 5 |
| Treatment planning | 4 |
| Early cancer | 3 |
| Cost-effectiveness | 1 |
| Restaging | 1 |
| Staging recurrent cancer | 1 |
| Correlation with tumour markers | 1 |
| Synchronous tumours | 1 |
| Texture analysis | 1 |

The OC staging pathway is complex, utilising various modalities with different strengths and weaknesses. PET/CT is superior to CT for detection of distant metastases and influences the change of MDT management decisions in up to 38 % of patients [109], whereas EUS is superior to CT for T-staging [110]. Comparison of techniques allows a modality to be tested against the perceived "gold-standard" staging investigation. This may reflect the desire for a simplified staging pathway with fewer investigations or the desire to increase evidence and awareness of a particular modality, thus introducing potential publication bias.

Influential articles are more likely to be cited by the scientific and clinical community. These citations form the basis of a journal's impact factor. The impact factor quantifies the average number of citations per manuscript published within that journal during a specific time period. Therefore, journals with a higher impact factor are recognised as being of higher quality and more likely to contain influential articles.

Radiological OC staging appeals to specialist radiologists and other members of the upper gastro-intestinal (GI) cancer MDT, and its clinical impact is great. The overall median 2014 and 5-year impact factors were 5.238 and 5.225, respectively, demonstrating that this field of research, often producing novel results, in a specific cancer population is not likely to be published in high-impact journals. The Journal of Clinical Oncology (JCO) had the highest 5-year impact factor (16.971) of articles in the 100 most cited.

In total, only 29 of the 100 most cited articles were published in radiology-related journals. This could represent the desire to achieve publication in a high-impact journal. The majority of radiology-related journals have impact factors <5.00. Only 16 % of the top 100 articles were published in radiology-related journals with a 5-year impact factor >5.00 (Radiology, Journal of Nuclear Medicine and European Journal of Nuclear Medicine and Molecular Imaging). It may also reflect a lack of research conducted by radiologists, which is supported by evidence from a National Cancer Research Institute (NCRI) survey in 2012, which commented upon the lack of academic radiologists [111].

Many of the first authors (n = 65) are not affiliated to radiology departments, according to Thomas Reuters Web of Science citation indexing database. It is possible the authors work closely with a radiologist as part of the specialist Upper GI cancer MDT or have a clinical radiologist as a named co-author.

EUS was the most commonly investigated modality overall. This may be a reflection of the current reliance and importance of EUS for T and N staging, considered the current "gold standard" [110].

Despite EUS being the most frequently investigated modality, the ten articles with the highest citation rates all investigated functional PET imaging. The CT component of the PET/CT examinations provided attenuation correction for

**Table 5** Ten articles with the highest citation rates

| Rank | Year | Number of citations | Citation rate | First author | Senior author | Title | Journal |
|---|---|---|---|---|---|---|---|
| 1 | 2011 | 114 | 38.00 | Tixier F [47] | Visvikis D | Intratumour heterogeneity characterized by textural features on baseline 18 F-FDG PET images predicts response to concomitant radiochemotherapy in oesophageal cancer | Journal of Nuclear Medicine |
| 2 | 2010 | 146 | 36.50 | Rice TW [31] | Blackstone EH | Cancer of the Oesophagus and Esophagogastric Junction Data-Driven Staging for the Seventh Edition of the American Joint Committee on Cancer/International Union Against Cancer Cancer Staging Manual | Cancer |
| 3 | 2006 | 229 | 28.63 | Ott K [15] | Siewert JR | Metabolic imaging predicts response, survival, and recurrence in adenocarcinomas of the esophagogastric junction | Journal of Clinical Oncology |
| 4 | 2004 | 271 | 27.10 | Wieder HA [12] | Weber WA | Time course of tumour metabolic activity during chemoradiotherapy of oesophageal squamous cell carcinoma and response to treatment | Journal of Clinical Oncology |
| 5 | 2002 | 290 | 24.17 | Kinkel K [10] | Thoeni RF | Detection of hepatic metastases from cancers of the gastrointestinal tract by using noninvasive imaging methods (US, CT, MR imaging, PET): A meta-analysis | Radiology |
| 5 | 2008 | 145 | 24.17 | van Vliet EPM [33] | Siersema PD | Staging investigations for oesophageal cancer: a meta-analysis | British Journal of Cancer |
| 7 | 2000 | 306 | 21.86 | Flamen P [8] | Mortelmans L | Utility of positron emission tomography for the staging of patients with potentially operable oesophageal carcinoma | Journal of Clinical Oncology |
| 8 | 2010 | 84 | 21.00 | Hyun SH [77] | Kim BT | Prognostic value of metabolic tumour volume measured by 18 F-fluorodeoxyglucose positron emission tomography in patients with oesophageal carcinoma | Annals of Surgical Oncology |
| 9 | 2003 | 193 | 17.55 | Downey RJ [18] | Rusch V | Whole body (18)FDG-PET and the response of oesophageal cancer to induction therapy: Results of a prospective trial | Journal of Clinical Oncology |
| 10 | 2004 | 168 | 16.80 | van Westreenen HL [23] | Plukker JTM | Systematic review of the staging performance of 18 F-fluorodeoxyglucose positron emission tomography in oesophageal cancer | Journal of Clinical Oncology |

PET data. Many PET/CT topics of research are relatively novel and have been described in other types of cancer. One of these topics, texture analysis, is the subject of the article with the highest citation rate [47]. Novel subjects are less likely to be published in high-impact journals, but may well be considered influential and provide the catalyst for future research.

Four of the PET/CT articles with the highest citation rates [12, 15, 18, 47] investigated its use in assessing treatment response. There is significant interest in the capability of metabolic imaging to assess for early treatment response, but these techniques have not been standardised for use outside of clinical research studies [112]. PET/CT scanning is expensive, and costly research could potentially only produce marginal long-term benefits for patients. The paradox of healthcare is that innovation increases expense, rather than producing more cost-effective and efficient processes, as is the case in industry [113]. These articles however are likely to be highly influential in forthcoming years, as the use of PET/CT increases in cancer imaging.

This bibliometric analysis has a number of limitations. Citation rates can be misleading because of various biases, e.g. institutional, language or self-citation bias. Older articles tend to accrue more citations compared to newer research. We attempted to adjust for this by calculating a citation rate, which may provide information regarding the importance and potential influence that the research has. This in itself has limitations as the likelihood of citation rises with increasing numbers of published articles in peer-reviewed journals. Only articles written in English were included, which may have excluded some frequently cited research in other languages. Also, this study concentrated on radiological staging rather than other techniques such as endoscopy and laparoscopy.

The expanding volume of published literature has increased significantly over the past few decades. Between 1978 to 1985 and 1994 to 2001, the annual number of Medline articles increased by 46 %, particularly in the area of clinical research [114]. The annual rate of publication growth in PubMed Medline was 5.6 % between 1997 and 2006, equating to a “doubling time” of 13 years [115]. This may explain the higher citation rate of PET/CT compared to that of EUS, as PET/CT is a more recent innovation. Overall, there are now more articles published per annum compared to previous years. This would therefore potentially increase the citation rate as a matter of course, not necessarily reflecting higher importance.

As expected, the older articles only described CT and perhaps a review of the last 10 years of literature only may be more reflective of contemporary staging practice. Another limitation is that only the first and senior authors of the articles were included in the current analysis. It is possible that these authors contributed to other articles in this list, but would not have been counted during analysis. These authors may therefore be under-represented in terms of published article numbers and have had a greater influence on current OC staging.

Of the 29 articles comparing imaging techniques or modalities, 17 studies correlated imaging findings and histopathological diagnosis, widely regarded as the "gold standard". Limitations exist in radiological studies that compare new findings against a potentially inaccurate alternative imaging test. In this current study, articles that did not compare against pathological results included those investigating radiotherapy planning techniques and the diagnosis of distant metastases. In these studies, tissue was not necessarily sampled. There are several reasons why pathological confirmation is not possible. These include patients undergoing non-surgical management, which is true of the majority of cases of OC, and in situations where it would be unethical to obtain tissue purely for research purposes, such as in patients with unequivocal distant metastases.

There are further limitations of studies comparing imaging findings to histopathological specimens. Comparison of pre-treatment imaging characteristics in patients receiving neo-adjuvant therapy prior to surgery can be inaccurate, as the chemotherapy or radiotherapy may alter the morphology of the tumour. In this situation, a direct comparison is often not possible.

## Conclusion

This bibliometric analysis describes the 100 most cited articles in the field of radiological OC staging investigations. Common topics of investigation include the accuracy of staging, comparison of modalities, treatment response and assessment of lymph nodes for metastases. The majority of articles are published outside of radiology-related journals, which may reflect the desire for high-impact publications or perhaps a lack of radiologists conducting imaging research. This study provides an understanding of research that has influenced current OC staging and citation rates may suggest important topics for future research, particularly validation studies of innovative techniques in larger patient populations.

**Acknowledgments** KF would like to acknowledge the support of the Wales Clinical Academic Training (WCAT) fellowship programme. No grant funding was used for this study and there are no conflicts of interest.



## References

1. De Bellis N (2009) Bibliometrics and citation analysis: from the science citation index to cybermetrics. Scarecrow Press, Lanham, Maryland, USA
2. Kelly JC, Glynn RW, O'Briain DE, Felle P, McCabe JP (2010) The 100 classic papers of orthopaedic surgery: a bibliometric analysis. J Bone Joint Surg (Br) 92(10):1338–1343
3. Tas F (2014) An analysis of the most-cited research papers on oncology: which journals have they been published in? Tumour Biol 35(5):4645–4649
4. Cancer Research UK. Statistics and outlook for oesophageal cancer (2014) [updated December 8th 2014; accessed 11 June 2015] Available from: http://www.cancerresearchuk.org/cancer-help/type/oesophageal-cancer/treatment/statistics-and-outlook-for-oesophageal-cancer
5. Sobin L, Gospodarowicz M, Wittekind C (2009) UICC TNM classification of malignant tumours, 7th edn. Wiley, New York
6. Roberts SA, Kay C (2014) Oesophagus and stomach cancers. In: Nicholson T (ed) Recommendations for cross-sectional imaging in cancer management, 2nd edn. Royal College of Radiologists, London
7. Paladugu R, Schein M, Gardezi S, Wise L (2002) One hundred citation classics in general surgical journals. World J Surg 26(9):1099–1105
8. Flamen P, Lerut A, Van Cutsem E, De Wever W, Peeters M, Stroobants S et al (2000) Utility of positron emission tomography for the staging of patients with potentially operable esophageal carcinoma. J Clin Oncol 18(18):3202–3210
9. Tio TL, Cohen P, Coene PP, Udding J, Jager F, Tytgat GNJ (1989) Endosonography and computed-tomography of esophageal-carcinoma—preoperative classification compared to the new (1987) TNM system. Gastroenterology 96(6):1478–1486
10. Kinkel K, Lu Y, Both M, Warren RS, Thoeni RF (2002) Detection of hepatic metastases from cancers of the gastrointestinal tract by using noninvasive imaging methods (US, CT, MR imaging, PET): a meta-analysis. Radiology 224(3):748–756
11. Catalano MF, Sivak MV, Rice T, Gragg LA, Vandam J (1994) Endosonographic features predictive of lymph-node metastasis. Gastrointest Endosc 40(4):442–446
12. Wieder HA, Brucher B, Zimmermann F, Becker K, Lordick F, Beer A et al (2004) Time course of tumor metabolic activity during chemoradiotherapy of esophageal squamous cell carcinoma and response to treatment. J Clin Oncol 22(5):900–908
13. Botet JF, Lightdale CJ, Zauber AG, Gerdes H, Urmacher C, Brennan MF (1991) Preoperative staging of esophageal cancer—comparison of endoscopic US and dynamic CT. Radiology 181(2):419–425
14. Skinner DB, Little AG, Ferguson MK, Soriano A, Staszak VM (1986) Selection of operation for esophageal cancer based on staging. Ann Surg 204(4):391–401
15. Ott K, Weber WA, Lordick F, Becker K, Busch R, Herrmann K et al (2006) Metabolic imaging predicts response, survival, and recurrence in adenocarcinomas of the esophagogastric junction. J Clin Oncol 24(29):4692–4698
16. Bhutani MS, Hawes RH, Hoffman BJ (1997) A comparison of the accuracy of echo features during endoscopic ultrasound (EUS)

and EUS-guided fine-needle aspiration for diagnosis of malignant lymph node invasion. Gastrointest Endosc 45(6):474–479
17. Flamen P, Van Cutsem E, Lerut A, Cambier JP, Haustermans K, Bormans G et al (2002) Positron emission tomography for assessment of the response to induction radiochemotherapy in locally advanced oesophageal cancer. Ann Oncol 13(3):361–368
18. Downey RJ, Akhurst T, Ilson D, Ginsberg R, Bains MS, Gonen M et al (2003) Whole body (18)FDG-PET and the response of esophageal cancer to induction therapy: results of a prospective trial. J Clin Oncol 21(3):428–432
19. Flanagan FL, Dehdashti F, Siegel BA, Trask DD, Sundaresan SR, Patterson GA et al (1997) Staging of esophageal cancer with F-18-fluorodeoxyglucose positron emission tomography. AJR Am J Roentgenol 168(2):417–424
20. Picus D, Balfe DM, Koehler RE, Roper CL, Owen JW (1983) Computed-tomography in the staging of esophageal-carcinoma. Radiology 146(2):433–438
21. Rosch T (1995) Endosonographic staging of esophageal cancer: a review of literature results. Gastrointest Endosc Clin N Am 5(3): 537–547
22. Kelly S, Harris KM, Berry E, Hutton J, Roderick P, Cullingworth J et al (2001) A systematic review of the staging performance of endoscopic ultrasound in gastro-oesophageal carcinoma. Gut 49(4):534–539
23. van Westreenen HL, Westerterp M, Bossuyt PM, Pruim J, Sloof GW, van Lanschot JJ et al (2004) Systematic review of the staging performance of 18F-fluorodeoxyglucose positron emission tomography in esophageal cancer. J Clin Oncol 22(18):3805–3812
24. May A, Gunter E, Roth F, Gossner L, Stolte M, Vieth M et al (2004) Accuracy of staging in early oesophageal cancer using high resolution endoscopy and high resolution endosonography: a comparative, prospective, and blinded trial. Gut 53(5):634–640
25. Lerut T, Flamen P, Ectors N, Van Cutsem E, Peeters M, Hiele M et al (2000) Histopathologic validation of lymph node staging with FDG-PET scan in cancer of the esophagus and gastroesophageal junction—a prospective study based on primary surgery with extensive lymphadenectomy. Ann Surg 232(6):743–751
26. Kato H, Kuwano H, Nakajima M, Miyazaki T, Yoshikawa M, Ojima H et al (2002) Comparison between positron emission tomography and computed tomography in the use of the assessment of esophageal carcinoma. Cancer 94(4):921–928
27. Block MI, Patterson GA, Sundaresan RS, Bailey MS, Flanagan FL, Dehdashti F et al (1997) Improvement in staging of esophageal cancer with the addition of positron emission tomography. Ann Thorac Surg 64(3):770–776
28. Swisher SG, Maish M, Erasmus JJ, Correa AM, Ajani JA, Bresalier R et al (2004) Utility of PET, CT, and EUS to identify pathologic responders in esophageal cancer. Ann Thorac Surg 78(4):1152–1160
29. Swisher SG, Erasmus J, Maish M, Correa AM, Macapinlac H, Ajani JA et al (2004) 2-fluoro-2-deoxy-D-glucose positron emission tomography imaging is predictive of pathologic response and survival after preoperative chemoradiation in patients with esophageal carcinoma. Cancer 101(8):1776–1785
30. Luketich JD, Friedman DM, Weigel TL, Meehan MA, Keenan RJ, Townsend DW et al (1999) Evaluation of distant metastases in esophageal cancer: 100 consecutive positron emission tomography scans. Ann Thorac Surg 68(4):1133–1136
31. Rice TW, Rusch VW, Ishwaran H, Blackstone EH, Collaboration WEC (2010) Cancer of the esophagus and esophagogastric junction data-driven staging for the seventh edition of the American Joint Committee on Cancer/International Union Against Cancer Cancer Staging Manuals. Cancer 116(16):3763–3773
32. Vilgrain V, Mompoint D, Palazzo L, Menu Y, Gayet B, Ollier P et al (1990) Staging of esophageal-carcinoma—comparison of results with endoscopic sonography and CT. AJR Am J Roentgenol 155(2):277–281
33. van Vliet EP, Heijenbrok-Kal MH, Hunink MG, Kuipers EJ, Siersema PD (2008) Staging investigations for oesophageal cancer: a meta-analysis. Br J Cancer 98(3):547–557
34. Rosch T, Lorenz R, Zenker K, Vonwichert A, Dancygier H, Hofler H et al (1992) Local staging and assessment of resectability in carcinoma of the esophagus, stomach, and duodenum by endoscopic ultrasonography. Gastrointest Endosc 38(4):460–467
35. Ziegler K, Sanft C, Zeitz M, Friedrich M, Stein H, Haring R et al (1991) Evaluation of endosonography in TN staging of esophageal cancer. Gut 32(1):16–20
36. Kole AC, Plukker JT, Nieweg OE, Vaalburg W (1998) Positron emission tomography for staging of oesophageal and gastroesophageal malignancy. Br J Cancer 78(4):521–527
37. Luketich JD, Schauer PR, Meltzer CC, Landreneau RJ, Urso GK, Townsend DW et al (1997) Role of positron emission tomography in staging esophageal cancer. Ann Thorac Surg 64(3):765–769
38. Buskens CJ, Westerterp M, Lagarde SM, Bergman J, ten Kate FJW, van Lanschot JJB (2004) Prediction of appropriateness of local endoscopic treatment for high-grade dysplasia and early adenocarcinoma by EUS and histopathologic features. Gastrointest Endosc 60(5):703–710
39. Watt I, Stewart I, Anderson D, Bell G, Anderson JR (1989) Laparoscopy, ultrasound and computed-tomography in cancer of the esophagus and gastric cardia—a prospective comparison for detecting intra-abdominal metastases. Br J Surg 76(10):1036–1039
40. Moss AA, Schnyder P, Thoeni RF, Margulis AR (1981) Esophageal-carcinoma—pre-therapy staging by computed-tomography. AJR Am J Roentgenol 136(6):1051–1056
41. Zuccaro G, Rice TW, Goldblum J, Medendorp SV, Becker M, Pimentel R et al (1999) Endoscopic ultrasound cannot determine suitability for esophagectomy after aggressive chemoradiotherapy for esophageal cancer. Am J Gastroenterol 94(4):906–912
42. Grimm H, Binmoeller KF, Hamper K, Koch J, Hennebruns D, Soehendra N (1993) Endosonography for preoperative locoregional staging of esophageal and gastric-cancer. Endoscopy 25(3):224–230
43. Vazquez-Sequeiros E, Wiersema MJ, Clain JE, Norton ID, Levy MJ, Romero Y et al (2003) Impact of lymph node staging on therapy of esophageal carcinoma. Gastroenterology 125(6):1626–1635
44. Yoon YC, Lee KS, Shim YM, Kim BT, Kim K, Kim TS (2003) Metastasis to regional lymph nodes in patients with esophageal squamous cell carcinoma: CT versus FDG PET for presurgical detection—prospective study. Radiology 227(3):764–770
45. Tio TL, Tytgat GN (1984) Endoscopic ultrasonography in the assessment of intramural and transmural infiltration of tumors in the esophagus, stomach and papilla of vater and in the detection of extraesophageal lesions. Endoscopy 16(6):203–210
46. Cerfolio RJ, Bryant AS, Ohja B, Bartolucci AA, Eloubeidi MA (2005) The accuracy of endoscopic ultrasonography with fine-needle aspiration, integrated positron emission tomography with computed tomography, and computed tomography in restaging patients with esophageal cancer after neoadjuvant chemoradiotherapy. J Thorac Cardiovasc Surg 129(6):1232–1241
47. Tixier F, Le Rest CC, Hatt M, Albarghach N, Pradier O, Metges JP et al (2011) Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. J Nucl Med 52(3):369–378
48. Vazquez-Sequeiros E, Norton ID, Clain JE, Wang KK, Affi A, Allen M et al (2001) Impact of EUS-guided fine-needle aspiration on lymph node staging in patients with esophageal carcinoma. Gastrointest Endosc 53(7):751–757
49. Dittler HJ, Siewert JR (1993) Role of endoscopic ultrasonography in esophageal-carcinoma. Endoscopy 25(2):156–161
50. Rasanen JV, Sihvo EI, Knuuti MJ, Minn HR, Luostarinen ME, Laippala P et al (2003) Prospective analysis of accuracy of positron emission tomography, computed tomography, and endoscopic ultrasonography in staging of adenocarcinoma of the esophagus and the esophagogastric junction. Ann Surg Oncol 10(8):954–960

51. Quint LE, Hepburn LM, Francis IR, Whyte RI, Orringer MB (1995) Incidence and distribution of distant metastases from newly-diagnosed esophageal-carcinoma. Cancer 76(7):1120–1125
52. Reed CE, Mishra G, Sahai AV, Hoffman BJ, Hawes RH (1999) Esophageal cancer staging: improved accuracy by endoscopic ultrasound of celiac lymph nodes. Ann Thorac Surg 67(2):319–321
53. Choi JY, Lee KH, Shim YM, Lee KS, Kim JJ, Kim SE et al (2000) Improved detection of individual nodal involvement in squamous cell carcinoma of the esophagus by FDG PET. J Nucl Med 41(5): 808–815
54. Rankin SC, Taylor H, Cook GJ, Mason R (1998) Computed tomography and positron emission tomography in the pre-operative staging of oesophageal carcinoma. Clin Radiol 53(9):659–665
55. Thompson WM, Halvorsen RA, Foster WL, Williford ME, Postlethwait RW, Korobkin M (1983) Computed-tomography for staging esophageal and gastroesophageal cancer—reevaluation. AJR Am J Roentgenol 141(5):951–958
56. Wallace MB, Nietert PJ, Earle C, Krasna MJ, Hawes RH, Hoffman BJ et al (2002) An analysis of multiple staging management strategies for carcinoma of the esophagus: computed tomography, endoscopic ultrasound, positron emission tomography, and thoracoscopy/laparoscopy. Ann Thorac Surg 74(4):1026–1032
57. Quint LE, Glazer GM, Orringer MB, Gross BH (1985) Esophageal-carcinoma—CT findings. Radiology 155(1):171–175
58. Kato H, Miyazaki T, Nakajima M, Takita J, Kimura H, Faried A et al (2005) The incremental effect of positron emission tomography on diagnostic accuracy in the initial staging of esophageal carcinoma. Cancer 103(1):148–156
59. Larghi A, Lightdale CJ, Memeo L, Bhagat G, Okpara N, Rotterdam H (2005) EUS followed by EMR for staging of high-grade dysplasia and early cancer in Barrett's esophagus. Gastrointest Endosc 62(1):16–23
60. Takashima S, Takeuchi N, Shiozaki H, Kobayashi K, Morimoto S, Ikezoe J et al (1991) Carcinoma of the esophagus: CT vs MR imaging in determining resectability. AJR Am J Roentgenol 156(2):297–302
61. van Westreenen HL, Cobben DCP, Jager PL, van Dullemen HM, Wesseling J, Elsinga PH et al (2005) Comparison of F-18-FLT PET and F-18-FDG PET in esophageal cancer. J Nucl Med 46(3):400–404
62. Eloubeidi MA, Wallace MB, Reed CE, Hadzijahic N, Lewin DN, Van Velse A et al (2001) The utility of EUS and EUS-guided fine needle aspiration in detecting celiac lymph node metastasis in patients with esophageal cancer: a single-center experience. Gastrointest Endosc 54(6):714–719
63. Hasegawa N, Niwa Y, Arisawa T, Hase S, Goto H, Hayakawa T (1996) Preoperative staging of superficial esophageal carcinoma: comparison of an ultrasound probe and standard endoscopic ultrasonography. Gastrointest Endosc 44(4):388–393
64. Levine EA, Farmer MR, Clark P, Mishra G, Ho C, Geisinger KR et al (2006) Predictive value of 18-fluoro-deoxy-glucose-positron emission tomography (F-18-FDG-PET) in the identification of responders to chemoradiation therapy for the treatment of locally advanced esophageal cancer. Ann Surg 243(4):472–478
65. Rice TW, Boyce GA, Sivak MV (1991) Esophageal ultrasound and the preoperative staging of carcinoma of the esophagus. J Thorac Cardiovasc Surg 101(3):536–544
66. Isenberg G, Chak A, Canto MI, Levitan N, Clayman J, Pollack BJ et al (1998) Endoscopic ultrasound in restaging of esophageal cancer after neoadjuvant chemoradiation. Gastrointest Endosc 48(2):158–163
67. Leong T, Everitt C, Yuen K, Condron S, Hui A, Ngan SYK et al (2006) A prospective study to evaluate the impact of FDG-PET on CT-based radiotherapy treatment planning for oesophageal cancer. Radiother Oncol 78(3):254–261
68. Puli SR, Reddy JBK, Bechtold ML, Antillon D, Ibdah JA, Antillon MR (2008) Staging accuracy of esophageal cancer by endoscopic ultrasound: a meta-analysis and systematic review. World J Gastroenterol 14(10):1479–1490
69. Flamen P, Lerut A, van Custem E, Cambier JP, Maes A, De Wever W et al (2000) The utility of positron emission tomography for the diagnosis and staging of recurrent esophageal cancer. J Thorac Cardiovasc Surg 120(6):1085–1092
70. Lightdale CJ, Kulkarni KG (2005) Role of endoscopic ultrasonography in the staging and follow-up of esophageal cancer. J Clin Oncol 23(20):4483–4489
71. Daffner RH, Halber MD, Postlethwait RW, Korobkin M, Thompson WM (1979) CT of the esophagus. 2. Carcinoma. AJR Am J Roentgenol 133(6):1051–1055
72. Kim K, Park SJ, Kim BT, Lee KS, Shim YM (2001) Evaluation of lymph node metastases in squamous cell carcinoma of the esophagus with positron emission tomography. Ann Thorac Surg 71(1): 290–294
73. Meltzer CC, Luketich JD, Friedman D, Charron M, Strollo D, Meehan M et al (2000) Whole-body FDG positron emission tomographic imaging for staging esophageal cancer—comparison with computed tomography. Clin Nucl Med 25(11):882–887
74. Vrieze O, Haustermans K, De Wever W, Lerut T, Van Cutsem E, Ectors N et al (2004) Is there a role for FGD-PET in radiotherapy planning in esophageal carcinoma? Radiother Oncol 73(3): 269–275
75. Jones DR, Parker LA, Detterbeck FC, Egan TM (1999) Inadequacy of computed tomography in assessing patients with esophageal carcinoma after induction chemoradiotherapy. Cancer 85(5):1026–1032
76. Chak A, Canto M, Gerdes H, Lightdale CJ, Hawes RH, Wiersema MJ et al (1995) Prognosis of esophageal cancers preoperatively staged to be locally invasive (T4) by endoscopic ultrasound (EUS)—a multicenter retrospective cohort study. Gastrointest Endosc 42(6):501–506
77. Hyun SH, Choi JY, Shim YM, Kim K, Lee SJ, Cho YS et al (2010) Prognostic value of metabolic tumor volume measured by 18F-fluorodeoxyglucose positron emission tomography in patients with esophageal carcinoma. Ann Surg Oncol 17(1):115–122
78. Quint LE, Glazer GM, Orringer MB (1985) Esophageal imaging by MR and CT: study of normal anatomy and neoplasms. Radiology 156(3):727–731
79. Murata Y, Muroi M, Yoshida M, Ide H, Hanyu F (1987) Endoscopic ultrasonography in the diagnosis of esophageal carcinoma. Surg Endosc 1(1):11–16
80. Hiele M, DeLeyn P, Schurmans P, Lerut A, Huys S, Geboes K et al (1997) Relation between endoscopic ultrasound findings and outcome of patients with tumors of the esophagus or esophagogastric junction. Gastrointest Endosc 45(5):381–386
81. Wakelin SJ, Deans C, Crofts TJ, Allan PL, Plevris JN, Paterson-Brown S (2002) A comparison of computerised tomography, laparoscopic ultrasound and endoscopic ultrasound in the preoperative staging of oesophago-gastric carcinoma. Eur J Radiol 41(2):161–167
82. Scotiniotis IA, Kochman ML, Lewis JD, Furth EE, Rosato EF, Ginsberg GG (2001) Accuracy of EUS in the evaluation of Barrett's esophagus and high-grade dysplasia or intramucosal carcinoma. Gastrointest Endosc 54(6):689–696
83. Fockens P, Van den Brande JHM, van Dullemen HM, van Lanschot JJB, Tytgat GNJ (1996) Endosonographic T-staging of esophageal carcinoma: a learning curve. Gastrointest Endosc 44(1):58–62
84. Beseth BD, Bedford R, Isacoff WH, Holmes EC, Cameron RB (2000) Endoscopic ultrasound does not accurately assess pathologic stage of esophageal cancer after neoadjuvant chemoradiotherapy. Am Surg 66(9):827–831

85. Catalano MF, Alcocer E, Chak A, Nguyen CC, Raijman I, Geenen JE et al (1999) Evaluation of metastatic celiac axis lymph nodes in patients with esophageal carcinoma: accuracy of EUS. Gastrointest Endosc 50(3):352–356
86. Giovannini M, Monges G, Seitz JF, Moutardier V, Bernardini D, Thomas P et al (1999) Distant lymph node metastases in esophageal cancer: impact of endoscopic ultrasound-guided biopsy. Endoscopy 31(7):536–540
87. Rice TW (2000) Clinical staging of esophageal carcinoma. CT, EUS, and PET. Chest Surg Clin N Am 10(3):471–485
88. Heeren PAM, Jager PL, Bongaerts F, van Dullemen H, Sluiter W, Plukker JTM (2004) Detection of distant metastases in esophageal cancer with F-18-FDG PET. J Nucl Med 45(6):980–987
89. Rizk N, Downey RJ, Akhurst T, Gonen M, Bains MS, Larson S et al (2006) Preoperative 18[F]-fluorodeoxyglucose positron emission tomography standardized uptake values predict survival after esophageal adenocarcinoma resection. Ann Thorac Surg 81(3): 1076–1081
90. Kostakoglu L, Goldsmith SJ (2004) PET in the assessment of therapy response in patients with carcinoma of the head and neck and of the esophagus. J Nucl Med 45(1):56–68
91. Moureau-Zabotto L, Touboul E, Lerouge D, Deniaud-Alexandre E, Grahek D, Foulquier JN et al (2005) Impact of CT and F-18-deoxyglucose positron emission tomography image fusion for conformal radiotherapy in esophageal carcinoma. Int J Radiat Oncol Biol Phys 63(2):340–345
92. Pech O, May A, Guenter E, Gossner L, Ell C (2006) The impact of endoscopic ultrasound and computed tomography on the TNM staging of early cancer in Barrett's esophagus. Am J Gastroenterol 101(10):2223–2229
93. Choi JY, Jang HJ, Shim YM, Kim K, Lee KS, Lee KH et al (2004) 18F-FDG PET in patients with esophageal squamous cell carcinoma undergoing curative surgery: prognostic implications. J Nucl Med 45(11):1843–1850
94. Lehr L, Rupp N, Siewert JR (1988) Assessment of resectability of esophageal cancer by computed tomography and magnetic resonance imaging. Surgery 103(3):344–350
95. Lightdale CJ (1992) Endoscopic ultrasonography in the diagnosis, staging and follow-up of esophageal and gastric-cancer. Endoscopy 24:297–303
96. Kobori D, Kirihara Y, Kosaka N, Hara T (1999) Positron emission tomography of esophageal carcinoma using C-11-choline and F-18-fluorodeoxyglucose—a novel method of preoperative lymph node staging. Cancer 86(9):1638–1648
97. Luketich JD, Schauer P, Landreneau R, Nguyen N, Urso K, Ferson P et al (1997) Minimally invasive surgical staging is superior to endoscopic ultrasound in detecting lymph node metastases in esophageal cancer. J Thorac Cardiovasc Surg 114(5):817–821
98. Lowe VJ, Booya F, Fletcher JG, Nathan M, Jensen E, Mullan B et al (2005) Comparison of positron emission tomography, computed tomography, and endoscopic ultrasound in the initial staging of patients with esophageal cancer. Mol Imaging Biol 7(6):422–430
99. Song SY, Kim JH, Ryu JS, Lee GTH, Kim SB, Park SI et al (2005) FDG-PET in the prediction of pathologic response after neoadjuvant chemoradiotherapy in locally advanced, resectable esophageal cancer. Int J Radiat Oncol Biol Phys 63(4):1053–1059
100. Yuan S, Yu Y, Chao KS, Fu Z, Yin Y, Liu T et al (2006) Additional value of PET/CT over PET in assessment of locoregional lymph nodes in thoracic esophageal squamous cell cancer. J Nucl Med 47(8):1255–1259
101. Bar-Shalom R, Guralnik L, Tsalic M, Leiderman M, Frenkel A, Gaitini D et al (2005) The additional value of PET/CT over PET in FDG imaging of oesophageal cancer. Eur J Nucl Med Mol Imaging 32(8):918–924
102. McAteer D, Wallis F, Couper G, Norton M, Welch A, Bruce D et al (1999) Evaluation of F-18-FDG positron emission tomography in gastric and oesophageal carcinoma. Br J Radiol 72(858): 525–529
103. van Westreenen HL, Westerterp M, Jager PL, van Dullemen HM, Sloof GW, Comans EFI et al (2005) Synchronous primary neoplasms detected on F-18-FDG PET in staging of patients with esophageal cancer. J Nucl Med 46(8):1321–1325
104. Eloubeidi MA, Wallace MB, Hoffman BJ, Leveen MB, Van Velse A, Hawes RH et al (2001) Predictors of survival for esophageal cancer patients with and without celiac axis lymphadenopathy: impact of staging endosonography. Ann Thorac Surg 72(1):212–218
105. Konski A, Doss M, Milestone B, Haluszka O, Hanlon A, Freedman G et al (2005) The integration of 18-fluoro-deoxyglucose positron emission tomography and endoscopic ultrasound in the treatment-planning process for esophageal carcinoma. Int J Radiat Oncol Biol Phys 61(4):1123–1128
106. Meyers BF, Downey RJ, Decker PA, Keenan RJ, Siegel BA, Cerfolio RJ et al (2007) The utility of positron emission tomography in staging of potentially operable carcinoma of the thoracic esophagus: results of the American College of Surgeons Oncology Group Z0060 trial. J Thorac Cardiovasc Surg 133(3): 738–745
107. Wu LF, Wang BZ, Feng JL, Cheng WR, Liu GR, Xu XH et al (2003) Preoperative TN staging of esophageal cancer: comparison of miniprobe ultrasonography, spiral CT and MRI. World J Gastroenterol 9(2):219–224
108. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M et al (2015) Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 136(5):E359–386
109. Blencowe NS, Whistance RN, Strong S, Hotton EJ, Ganesh S, Roach H et al (2013) Evaluating the role of fluorodeoxyglucose positron emission tomography-computed tomography in multidisciplinary team recommendations for oesophago-gastric cancer. Br J Cancer 109(6):1445–1450
110. Tangoku A, Yamamoto Y, Furukita Y, Goto M, Morimoto M (2012) The new era of staging as a key for an appropriate treatment for esophageal cancer. Ann Thorac Cardiovasc Surg 18(3):190–199
111. National Cancer Research Institute (NCRI) (2012) Imaging research in the UK: An NCRI survery [accessed 16 Dec 2015]; Available from: http://www.ncri.org.uk/publication/imaging-research-in-the-uk-an-ncri-survey-2/
112. Schmidt T, Lordick F, Herrmann K, Ott K (2015) Value of functional imaging by PET in esophageal cancer. J Natl Compr Cancer Netw 13(2):239–247
113. Skinner JS (2013) The costly paradox of health-care technology. MIT technology review business report [accessed 16 Dec 2015] Available from: http://www.technologyreview.com/news/518876/the-costly-paradox-of-health-care-technology/
114. Druss BG (2005) Growth and decentralization of the medical literature: implications for evidence-based medicine. J Med Libr Assoc 93(4):499–501
115. Larsen PO, von Ins M (2010) The rate of growth in scientific publication and the decline in coverage provided by science citation index. Scientometrics 84:575–603

The Royal College of Radiologists

63 LINCOLN'S INN FIELDS, LONDON WC2A 3JW
T. +44(0)20 7405 1282 enquiries@rcr.ac.uk
www.rcr.ac.uk

Dr Kieran George Foley
31 Ilton Road
Penylan
Cardiff
CF23 5DU
UNITED KINGDOM

26 August 2015

Dear Dr Foley

**2014 ELLIS BARNETT PRIZE**

On behalf of the Editor of *Clinical Radiology*, I am very pleased to inform you that your paper *N-staging of oesophageal and junctional carcinoma: Is there still a role for EUS in patients staged N0 at PET/CT?* has been selected as the winner of the 2014 Ellis Barnett Prize.

The prize is awarded to the most outstanding paper with an ultrasound content published in *Clinical Radiology*. It is awarded to you as first author on behalf of you and your co-authors, and takes the form of a cheque for £300.00.

We will send the prize to you at the above address. If you would prefer to do so you can accept the prize in person from the President of the Royal College of Radiologists at a ceremony for the Admission of New Fellows. The next ceremony will be held in London on Friday 27 November 2015 at 3pm. If you would like to attend the ceremony, please let me know so that I may send you further details. Otherwise, please let me know that you would like us to send the award to you at the above address.

I look forward to hearing from you. Please do not hesitate to contact me if you have any queries, on ceremony@rcr.ac.uk

With many congratulations,

Yours sincerely

**Sarah Coulson**
Training Administrator