DOCTOR OF PHILOSOPHY

Economic evaluation of endoscopic ultrasound in gastro-oesophageal cancer staging: Exploring novel methodology in QALY estimations alongside clinical trials in the UK

Yeo, Seow Tien

Award date:
2024

Awarding institution:
Bangor University

Economic evaluation of endoscopic ultrasound in gastro-oesophageal cancer staging: Exploring novel methodology in QALY estimations alongside clinical trials in the UK

Seow Tien Yeo

Thesis submitted to Bangor University
in fulfilment of the requirements for the
Degree of Doctor of Philosophy (PhD) in Health Economics

Supervisors: Professor Rhiannon Tudor Edwards,
Dr Zoë Hoare & Dr Hasan Haboubi

Centre for Health Economics and Medicines Evaluation
School of Medical and Health Sciences
Bangor University

19th December 2023

Contents

List of Figures

List of Tables

Thesis Abstract

Background

The sensitivity of endoscopic ultrasound (EUS) in gastro-oesophageal cancer (GOC) staging has been studied; however, the cost-effectiveness of EUS staging has not been evaluated until the inception of the HTA-funded COGNATE trial (ISRCTN1444215). This thesis aimed to explore the cost-effectiveness of EUS in GOC staging alongside the COGNATE trial, the economic evidence and current practice of the utilisation of EUS in GOC staging in the UK, and the potential use of disease-specific measures in cost-utility analysis in clinical trials.

Method

From an NHS perspective, a cost-effectiveness analysis alongside the COGNATE trial (Chapter 2), using QALY as the measure of effect, was conducted to evaluate whether adding EUS to the usual staging strategy is cost-effective in staging patients with GOC.

A systematic review was undertaken retrospectively (Chapter 3) following the completion of the COGNATE trial to identify the economic evidence of EUS staging in patients with GOC.

An online survey of UK GOC healthcare professionals (Chapter 4) was conducted through Bristol Online Survey (BOS) platform to explore the utilisation of EUS in GOC staging and the current clinical practice in the UK.

Given that disease-specific measures are usually collected alongside EQ-5D, a generic preference-based health-related quality of life measure, in clinical trials and seeing that EQ-5D has been argued to be not always sensitive enough to pick up changes in individual's quality of life, the potential use of disease-specific measures in cost-utility analysis was explored (Chapter 5), aiming to further exploring the transferability/generalisability of the hybrid QALY technique first tried in the MORTISE trial (on which I was a Research Officer in Health Economics working with the Trial Statistician, Dr Daphne Russell) in other disease areas (e.g. cancer and ophthalmology) in cost-utility analysis. Data from two large clinical trials, each with a 12-month follow-up (the COGNATE trial and the NIHR-EME-funded CLARITY trial (ISRCTN32207582)), were used. Regression models between disease-specific measures and conventional QALYs (measured solely by EQ-5D) for both trials were established, and results were compared.

Results

In the COGNATE trial, on average, EUS was found to cost £2,860 less per patient (95% bootstrapped CI: –£7,987 to £2,192) (2008 price year) [2019 price year: £3,490 less per patient, 95% bootstrapped CI: –£9,746 to £2,675] and improved QALYs by 0.1969 years (equivalent to 72 days in perfect health). Combining these savings and benefits showed that the probability of EUS being cost-effective exceeds 95% at the NICE threshold range of £20,000 to £30,000 per QALY (Chapter 2).

The systematic review findings (Chapter 3), based on the six identified economic articles, suggested that use of EUS as a complementary staging technique to other staging techniques for GOCs appears to be cost saving and offers greater QALYs.

From the healthcare professionals survey (Chapter 4), results showed that the majority support the use of EUS in GOC staging (n=89; 90.8%), have experience in the field of EUS either by requesting, performing or both (n=81; 82.7%) with most of them felt that EUS is more useful for staging oesophageal (n=78; 96.3%) and gastro-oesophageal junction cancer (n=78; 96.3%) than gastric cancer (n=58; 71.6%). Interestingly, 'attend Upper GI cancer MDT meeting' and 'clinician's age' were found to be important factors ($p<0.05$) associated with referral for EUS. Attendance at MDT meetings is likely to increase EUS referral, and younger clinicians are less likely to refer GOC patients for EUS.

From the exploratory study of assessing the potential use of disease-specific measure in QALY estimations (hybrid QALY approach) for cost-utility analysis in clinical trials (Chapter 5), disease-specific measures collected in trials were found to be potentially useful for cost-utility analysis. In both trials, cost-utility analysis results showed not only more certainty around the estimates of incremental cost-effectiveness ratios when conventional QALY was replaced with disease-specific measures guided QALY but also a shifting towards the direction of the respective disease-specific measures.

Conclusion

To date, EUS has been found to be not available at every hospital in the UK despite the widespread adoption of EUS, and the economic evidence in this area is scant. This thesis offers various novel contributions to the economic evidence and evaluation of EUS in GOC staging, and the insights into the current practice of the utilisation of EUS in GOC staging in the UK and the advancement of health economics assessment in clinical trials. These novel

contributions from this PhD study not only would facilitate policy makers and commissioners in evidence-based decision making but also provide unique insights for future research and policy in this area and other health technology assessment areas.

Dedication

This thesis is dedicated to the two very important people in my life – My husband and my late mother. I would like to express my heartfelt gratitude to both my husband and my late mother for their unconditional love, care, and support in every way they've provided throughout the journey of my life. They are simply the rock of my life.

To my dear husband, Stephen, thank you very much for your understanding, kindness, patience and support during this crazy PhD journey. You not only have kindly taken over most of our house chores but also have acted as my great in-house PhD supervisor. Thank you very much for everything – your invaluable advice and support and for being my great listener ('patience punch bag') throughout the journey of this thesis, and for being the great chief proof-reader of this thesis. All you've done are very much appreciated.

To my late mother, thank you very much for giving me life, nurturing me, guiding me, educating me, providing me with all you had and always being the first one I called for advice and for a chat. Mom, I wouldn't be where I am today without you. You were indeed a very caring and selfless mother who had placed all your children before you in your life. I miss you dearly mom.

Acknowledgements

To my first supervisor, Professor Rhiannon Tudor Edwards (Professor in Health Economics, Bangor University), I can't thank you enough for the opportunity you've given me to embark on this PhD studentship and to grow into an independent researcher in health economics. And thank you very much for providing support, guidance and feedback throughout the journey of this thesis. To my second supervisor, Dr Zoë Hoare (Senior Statistician, Bangor University), thank you very much for the statistical advice, support, guidance and feedback you've provided throughout the journey of this thesis. To my third supervisor, Dr Hasan Haboubi (Consultant Gastroenterologist in Cardiff and Vale University Health Board: CVUHB, and Honorary Senior Clinical Lecturer at Cancer Biomarkers Groups in Swansea University), thank you very much for the clinical advice, support, guidance and feedback you've provided throughout the journey of this thesis. To the Chair of my PhD supervisory committee, Dr Sion Williams (Reader in Health Research, Bangor University), thank you very much for your advice and support during my PhD journey. Thank you, Yasmin Noorani, the information scientist at Bangor University, for your broad knowledge of library databases and your guidance on the use of RefWorks reference management tool. Thank you, Professor Stephen Attwood, for being the external reviewer to review and provide comments on the pre-launched survey questionnaire and thank you for your support in the dissemination of the survey questionnaire. Thank you, Professor Sobha Sivaprasad, for permitting the use of the health economic data from the CLARITY trial for further analysis in the exploratory study chapter in this thesis. My heartfelt thanks to all of you. I would like to express my sincerest thanks and appreciation to the Tenovus Cancer Care Charity for funding this research. I would also like to extend my sincerest thanks and appreciation to the British Society of Gastroenterology (BSG) and the Welsh Association for Gastroenterology and Endoscopy (WAGE) for their support with the survey study. To my family, relatives and friends, who has believed in me, encouraged me, showed empathy and understanding in this challenging PhD journey, thank you very much to all of you. This has given me comfort and strength towards the end of this journey. To all the patients and clinicians who participated in the research study contained within this thesis, thank you very much for your participations and feedback. All your responses are invaluable for my research. I hope the results will help shape future of cancer care service provision, particularly in the field of gastro-oesophageal cancer in the UK.

About the Tenovus Cancer Care Charity

Tenovus Cancer Care is a Wales' leading Cancer Charity, supporting people living with cancer and their loved ones in communities across the country. They not only provide cancer support services to anyone affected by cancer and their loved ones, but also work alongside Welsh and UK Governments to help represent cancer patients and their loved ones. They respond to Government's consultations by letting the Government know what they think and help shape cancer policy for the future. Furthermore, they fund all types of cancer research that are different, innovative, and have real benefits for cancer patients and their loved ones. Their aims are "to find new ways to prevent, diagnose, and treat it, as well as striving to find innovative ways to improve the lives of people living with cancer today" and "to gain insight into the issues affecting people with cancer and to help make sure cancer research positively affects cancer outcomes and experiences" (Tenovus Cancer Care, https://www.tenovuscancercare.org.uk/about-us/who-we-are/our-aims)

Tenovus Cancer Care's Mission……

> "To give hope, help and a voice to anyone affected by cancer, in and around the community. We empower people through our support and services. We champion their needs by campaigning for better treatments, outcomes and health across the nation. And we bring hope through influencing and working for advances in cancer research."

(Tenovus Cancer Care, https://www.tenovuscancercare.org.uk/about-us/who-we-are/our-aims)

Declaration and Consent

I hereby declare that this thesis is the results of my own investigations, except where otherwise stated. All other sources are acknowledged by bibliographic references. This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree unless, as agreed by the University, for approved dual awards.

I confirm that I am submitting the work with the agreement of my Supervisors.

19th December 2023

Chapter 1: Introduction

1.1 Chapter summary

This chapter begins with a personal statement and epistemological background to provide a snapshot of why this original research was undertaken. Following that, this chapter provides:

- An overview of endoscopic ultrasound (EUS) in the management of patients with gastro-oesophageal cancers (GOCs).
- Economic evidence on EUS staging for GOCs
- The conceptual framework underpinning methods of economic evaluation
- Economic evaluation of health care
- Cost-effectiveness analysis
- Cost-utility analysis.
- Economic evaluation alongside clinical trials
- Challenges in cost-utility analysis of health care interventions
- Development of methods to overcome challenges in cost-utility analysis of health care interventions
- An overview of the method first tried in the MORTISE trial of strengthening cost-utility analysis results
- Cost-effectiveness of EUS in GOC staging alongside COGNATE trial

And finally, this chapter introduces (1) the purpose and overarching aims of this thesis, (2) thesis novelty and contribution to knowledge, and (3) the structure of this thesis with a flow chart outlining the research questions, operational plans, and chapters of this thesis, and (4) dissemination of the works from this thesis.

1.2 Personal statement and epistemological background

I graduated with a Bachelor of Science degree (honours) in Food Science and Nutrition from the National University of Malaysia in August 2002. Not long after my graduation, I was awarded a 'Postgraduate Bursary for International Student' by the Bangor University. In September 2002, I flew to the UK to pursue my study in Master of Science (MSc) degree in Public Health and Health Promotion at Bangor University. For my Master's degree dissertation, I designed and undertook a health promotion project integrated with health economics component, namely "A costing study of encouraging fruit consumption in primary school children in North Wales". The project was funded by the Centre for Health Economics and Medicines Evaluation (CHEME) and supervised by Professor Rhiannon Tudor Edwards and Dr Anne Krayer. Upon completion of the project in 2004, the outcome of the study was presented to Anglesey Local Council. In 2005, I graduated with a MSc degree in Public Health and Health Promotion from Bangor University, UK.

From the experience I gained during the research project for my Master's degree dissertation, I found myself enjoy doing research and my interest in health economics has grown ever since. Then, in 2004, I joined CHEME as a research assistant in health economics, working alongside Professor Rhiannon Tudor Edwards, at Bangor University. I was trained to becoming a health economist. During my early career research journey, my interest in involving in health economics of cancer care research arises from my experience with my late mother who was diagnosed with endometrial cancer in 2005 and sadly passed away from the horrible illness four years later (29th July 2009) at the age of 58-year-old.

As I was getting interested in the field of cancer, economic evaluation alongside clinical trials and statistical methods used in QALY estimations, I decided to pursue my PhD study in health economics of cancer care on a part-time basis with Professor Rhiannon Tudor Edwards at CHEME at Bangor University, UK. I was awarded a PhD studentship by the Tenovus Cancer Care charity (https://www.tenovuscancercare.org.uk/) to complete my research. At CHEME, I was very privileged to have had the opportunity to work on the COGNATE trial as part of

my PhD study. COGNATE, a HTA-funded UK study, was a large multi-centred randomised controlled trial investigating the effectiveness and cost-effectiveness of endoscopic ultrasound (EUS) for gastro-oesophageal cancers staging (i.e. staging for oesophageal, gastro-oesophageal junction and stomach cancer).

Gastro-oesophageal cancers (GOCs) are one of the most common cancers in the UK, accounting for at least 4% of all new cancer cases (Cancer Research UK, 2021a; Cancer Research UK, 2021b). The prognosis of GOCs remains poor because these tumours are usually detected in late stage (Valero and Robles-Medranda, 2017). The UK Cancer statistics reported that the overall five-year survival is low, at approximately 20% (Cancer Research UK, 2021a; Cancer Research UK, 2021b). Thorough and accurate staging is, therefore, very important to help clinicians to choose the best treatment plan that will maximise patients' chances of survival. EUS is known to be superior compared to CT and PET for locoregional staging of gastro-oesophageal tumours, although the complementary nature of these imaging modalities must be recognised (Thakkar and Kaul, 2020; Valero and Robles-Medranda, 2017; HQIP, 2016; Allum et al., 2011; Puli et al., 2008; Lowe et al., 2005). Studies reported that it has high sensitivity and specificity in accurately diagnosing tumour invasion (T) and locoregional nodal (N) cancer stages (Thakkar and Kaul, 2020; Valero and Robles-Medranda, 2017; Puli et al., 2008; Lowe et al., 2005). The sensitivity of EUS in GOC staging has been widely studied; however, there was no rigorous evaluation of its effectiveness and cost-effectiveness in GOC staging, particularly in the form of randomised controlled trial (RCT). Hence, I undertook a cost-effectiveness analysis study of EUS in GOC staging alongside the randomised controlled trial, namely 'COGNATE' trial. Further details about this study are discussed in Chapter 2 of this thesis.

In cost-effectiveness analysis, Quality-Adjusted Life Years (QALYs) are usually used as the outcome measure of effectiveness. However, many researchers have raised their concerns that standard quality of life measurement tools, such as the EQ-5D, does not incorporate condition-/disease-specific measures to reflect the real quality of life status of a patient (Pennington et al., 2020; Wichmann et al., 2017; Pettitt et al., 2016; Tosh et al., 2012; Whitehead and Ali, 2010). In other words, EQ-5D lack essential

dimensions related to the condition-/disease-specific areas of a study. In addition to the EQ-5D, Wichmann and colleagues (2017) recommended to make use of condition-/disease-specific quality of life instruments in measuring quality of life for economic evaluations in health care. The authors stated that the QALY might be of value in informing resource allocation decisions among health care interventions if specific issues are taken into account. Therefore, I was getting interested to undertake an exploratory study of incorporating condition-/disease-specific measures into QALY estimations for cost-effectiveness analysis, using the available accessible data from both the COGNATE and CLARITY trials. Further details about this exploratory study are discussed in Chapter 5 of this thesis.

Together with the COGNATE trial team members, I led the health economics components of the COGNATE trial with guidance provided by Professor Rhiannon Tudor Edwards (first supervisor). As part of my PhD study, I was responsible for managing, collecting, cleaning, costing, analysing and interpretating the COGNATE health economics data, and writing up COGNATE health economics report for submission to the Health Technology Assessment (HTA), the funder of the COGNATE trial (see the published COGNATE HTA report (Russell et al., 2013) for further details). In the next chapter (Chapter 2), the research work relating to the economic evaluation of the COGNATE trial are discussed in detail.

During the COGNATE trial, a literature review of economic evidence of endoscopic ultrasound for GOCs staging was performed given that conducting a systematic review was not possible at that time due to resource constraints. For this reason, I decided, with agreement from supervisors, to undertake a systematic review of economic evidence in that area retrospectively for my PhD study (see Chapter 3 for further details relating to systematic review work). This not only helps to fill the gap in the literature but also give a more complete picture of the available existing economic evidence in this field.

Besides conducting an economic evaluation alongside the COGNATE trial and a retrospective systematic review, I was also interested in carrying out a survey to

obtain a more in-depth insight into the overall picture of the use of EUS in GOCs staging in the UK and its current clinical practice. Given that no such survey in this area in the UK was found until the inception of my survey in mid-October 2017, it was believed that my survey would be the first study of its kind in this area in the UK up to January 2021 when a similar UK survey was published in February 2021 by Jones et al. (2021). By conducting the survey, it would help to add new knowledge to fill the gap in the literature in this field. Further details relating to the survey are described in Chapter 4 of this thesis.

Additionally, as mentioned earlier, I was interested in improving methods of cost-utility analysis that used conventional QALY. This is because (1) there is a general concern about the insensitivity and unresponsiveness of the generic health-related quality of life measure tool – the EQ-5D, which is recommended to be used to generate QALY (NICE, 2013) – in picking up changes in patients' disease-/condition-specific related quality of life (Pennington et al., 2020; Pettitt et al., 2016; Payakachat et al., 2015; Yang et al., 2015; Mulhern et al., 2013; Tosh et al., 2012; Whitehead and Ali, 2010); and (2) both EQ-5D and disease-/condition-specific measures data are normally collected independently in a study but only EQ-5D data is used to generate QALY for cost-utility analysis of a study. In the MORTISE trial where my colleagues and I first tried the method of incorporating the disease-specific measure used in the trial (i.e. Foot Health Thermometer, FHT) into the conventional QALY estimations for use in cost-utility analysis. I have further explored this idea in other disease areas (i.e. cancer and ophthalmology) in this thesis, as discussed in Chapter 5.

We believed that these four independently related studies would collectively provide a rich narrative about the cost-effectiveness and utilisation of EUS in GOCs staging in the UK, as well as the further exploration of the novel methodology in QALY estimations in other disease areas like cancer and ophthalmology.

1.3 Overview of endoscopic ultrasound (EUS) in the management of patients with gastro-oesophageal cancers

1.3.1 Gastro-oesophageal cancers (oesophageal, gastro-oesophageal junction and gastric cancers) – incidence, mortality, and 5- and 10-year survival data

Recent data published by the Cancer Research UK (CRUK, 2021a & 2021b) showed that gastro-oesophageal (oesophageal, gastric or both) cancers (GOCs) is one of the most common cancers in the UK, with approximately 16,000 new diagnoses each year despite the incidence rates having decreased over the last decade. Of the 16,000 new cases in the UK, about 9,000 are oesophageal cancer new cases and 7,000 are stomach cancer new cases. Incidence rates for both oesophageal and stomach cancers in the UK are highest in people aged 85 to 89. Of all new oesophageal and stomach cancer cases in the UK, each year around 40% and 50% respectively are diagnosed in people aged 75 and above. Seeing the aging population in the UK has been rising steadily through the late 20th century and that this trend is projected to increase to more than a quarter of the UK population by 2066 (ONS, 2018), this is placing a huge financial pressure on the UK NHS.

CRUK (2021a & 2021b) also reported that, there are around 8,000 and 4,300 deaths from oesophageal cancer and stomach cancer, respectively, in the UK every year. Oesophageal and stomach cancers are one of the most common causes of cancer death in the UK, accounting for 3% - 5% of all cancer deaths; however, over the last decade, the mortality rates of oesophageal and stomach cancers have declined by about 9% and 30%, respectively, in the UK. Furthermore, data from CRUK (2021a & 2021b) also showed that approximately 17% of people diagnosed with oesophageal cancer survive for five years or more, and 12% for 10 years or more in the UK; for people diagnosed with stomach cancer, 22% survive for five years or more and 17% for 10 years or more in the UK.

1.3.2 Diagnosis, assessment after diagnosis and treatment for gastro-oesophageal cancers (GOCs)

Diagnosis, assessment after diagnosis and treatment for gastro-oesophageal cancers are all well-documented (Allum et al., 2011; SIGN, 2006). Clearly, patients with different types and stage of gastro-oesophageal cancers are assessed and treated differently, that is on a case-by-case basis (Jones et al., 2021). As noted in the two

published sources, during the assessment phase following diagnosis, endoscopic ultrasound (EUS), a diagnostic technique, is recommended to offer to people with gastro-oesophageal cancers to help guide ongoing management/treatment plans. Use of EUS in GOCs staging is described further in the section 1.3.3 below.

Following the necessary investigative assessments, Upper Gastro-Intestinal Multi-Disciplinary Team (Upper GI MDT) will meet and discuss to create a patient's overall treatment plan that combines different types of treatments (Allum et al., 2011). Treatment options and recommendations made by the MDT depend on several factors, including the patient's co-morbidities and overall health, nutritional status, patient's preferences, possible side effects and staging information. Treatment recommendations made by the MDT will then be discussed with patients within the context of a shared decision-making consultation to choose treatments that fit their needs and care (Allum et al., 2011).

The treatment options for gastro-oesophageal cancer are endoscopic mucosal resection (EMR), surgery, chemotherapy, radiotherapy, and palliative treatments. Depending on the type and stage of gastro-oesophageal cancer, a treatment plan that combines different types of treatments will be created and recommended by MDT (Allum et al., 2011). For early gastro-oesophageal mucosal cancer (that is where tumours were found to be mucosa), patients will be assigned to undergo EMR and endoscopic submucosal dissection (ESD). For tumours that were found to be resectable, patients will be assigned to undergo surgery with or without neo-adjuvant therapy. Neo-adjuvant therapy refers to treatment such as chemotherapy, radiotherapy, or both, is given to patients before their main treatment such as surgery with the aim to shrink tumour before the main treatment. For locally advanced tumour, for which complete resection was deemed to be impossible, patients will be assigned to receive multimodal treatment combining different types of treatments, for example, combinations of chemotherapy and radiotherapy. For metastatic cancer, where tumour spreads to another part in the body from where it started, patients will be considered for palliative treatment to help relieve symptoms and side effects with the goal of improving patient's quality of life and lengthening the patient's life.

Palliative treatment not only using medicines such as painkillers and anti-sickness drugs to control symptoms but also using cancer treatments to reduce or eliminate symptoms such as pain. Cancer treatments used as palliative treatment include chemotherapy, radiotherapy, brachytherapy, targeted cancer drugs, surgery, and radiofrequency ablation which uses thermal to kill cancer cells.

Staging investigations following a diagnosis of gastro-oesophageal cancer are therefore important to be carried out appropriately and comprehensively in achieving better staging so that a more realistic prognosis (chance of recovery or likelihood the cancer will come back) can be given, and a treatment plan can be tailored accordingly. Better staging helps reduce the likelihood of unnecessary treatment, and therefore costs and outcomes to gastro-oesophageal cancer patients in terms of QALY gain.

1.3.3 EUS for gastro-oesophageal cancers (GOCs) staging

Accurate staging of GOCs is of utmost importance in determining the stage-appropriate treatment strategy (Thakkar and Kaul, 2020; Yeo et al., 2019; Kim et al., 2018; Valero and Robles-Medranda, 2017; Allum et al., 2011; Allum et al., 2002). Staging enables selection of the most appropriate treatment plan, and thus improves survival and quality of life. Staging investigations for GOCs are usually coordinated within an agreed pathway led by an Upper Gastro-Intestinal Multi-Disciplinary Team (Upper GI MDT) (Allum et al., 2011).

Accurate staging of GOCs is usually achieved through performing a combination of investigative diagnostic imaging tests which these include computer tomography (CT), endoscopic ultrasound (EUS), positron emission tomography (PET)-CT and adjuncts to staging include magnetic resonance imaging (MRI), bronchoscopy, laparoscopy and trans-abdominal ultrasound (Allum et al., 2011; Allum et al., 2002). The guidelines for the management of oesophageal and gastric cancer stated that a CT scan should be performed at initial staging to determine whether metastatic disease is present (Allum et al., 2011; Allum et al., 2002). Metastatic disease is a condition where cancer cells have spread from the primary site of its origin to other new areas of the body. If

metastatic disease is absent, further staging with EUS in GOCs is recommended to assess and predict operability (Thakkar and Kaul, 2020; DaVee et al., 2017; Valero and Robles-Medranda, 2017; Allum et al., 2002). This is because EUS is known to be more superior to CT and PET in terms of its sensitivity for locoregional staging of gastro-oesophageal tumours (Thakkar and Kaul, 2020; Valero and Robles-Medranda, 2017; Smyth et al., 2017; Puli et al., 2008; Lowe et al., 2005). Studies reported that EUS has superior tumour invasion (T) and locoregional nodal (N) staging ability over CT and PET (Thakkar and Kaul, 2020; Smyth et al., 2017; Valero and Robles-Medranda, 2017; Findlay et al. 2015; Puli et al., 2008; Shimpi et al., 2007; Lowe et al., 2005). Specifically, EUS is sensitive for the detection of regional lymph node metastases (Thakkar and Kaul, 2020; Smyth et al., 2017; Valero and Robles-Medranda, 2017; Findlay et al. 2015; Puli et al., 2008; van Vliet et al., 2008; Lowe et al., 2005; Botet et al., 1991; Vilgrain et al., 1990).

Accurate pre-treatment staging is crucial to avoid inadequate and unnecessary treatment. EUS is recommended to use as a complementary imaging technique to other imaging techniques such as CT and PET scanning for staging GOC which this can help to minimise unnecessary treatments (NICE, 2018; Smyth et al., 2017; Findlay et al. 2015; HQIP, 2016; Dittler et al., 1995; Rösch et al., 1992); and thus potentially could save costs and offer greater health benefits to patients in terms of QALY gains.

1.4 Economic evidence on EUS staging for GOCs

Evidence on economic evaluations of EUS staging for GOCs is scarce. Only three American decision-analytic economic modelling studies (Harewood and Wiersema, 2002; Wallace et al., 2002; Hadzijahic et al., 2000) were found and no guidance on this topic from the National Institute for Health and Care Excellence (NICE) were found prior to the completion of the COGNATE trial (see below for further details about the COGNATE trial). Of the three studies, Wallace et al. (2002) was the most comprehensive, inferring that the combination of CT, EUS and fine-needle aspiration (FNA) was on average the least costly and more effective in gaining quality-adjusted life years (QALYs) compared to all other staging strategies except the combination of positron emission tomography (PET), EUS and FNA. The latter was slightly more

effective but it was more expensive and less cost-effective when judged by cost per QALY, the usual measure in both US and UK. Furthermore, there has been no rigorous evaluation of EUS staging of GOCs in the form of randomised controlled trial (RCT) until the establishment of the trial 'Cancer of the Oesophagus or Gastricus: New Assessment of the Technology of Endosonography' (COGNATE, Registration Number: ISRCTN1444215). The COGNATE trial was a pragmatic multi-centred randomised controlled trial, commissioned by the NIHR Health Technology Assessment Programme to evaluate whether EUS is effective and cost-effective in managing GOCs, not the accuracy of EUS, but its effect on patient management and outcome. The cost-effectiveness of EUS staging for GOCs of the COGNATE trial that formed part of this PhD thesis, is discussed in detail in the next chapter (Chapter 2).

1.5 Conceptual framework underpinning methods of economic evaluation

1.5.1 Health Economics – What is health economics and why it is important?

Health economics is a field of economic sciences that involves the development and application of economic theory, models and empirical methods for studying the production and consumption of health and health care (Morris, et al., 2012; Maynard and Kanavos, 2000). Health economics is sturdily based in economic theory; however, it furthers understanding of the behaviour of individuals (e.g. patients, health care professionals) and health care providers, and comprises analytical techniques developed to facilitate resource allocation decisions in health care.

Health economics has been regarded as a specialised field, providing economics ideas and methods of analysis to policy makers, governments, health care providers and society as a whole as to how best to use scarce resources to meet human wants in terms of health and health care. Health, in overall, is regarded as a "fundamental commodity" to all people and unlike other goods or services, it is non-transferrable from one person to another and non-tradeable. As health cannot be traded, it therefore cannot be analysed in the context of a market. Instead, health economists/analysts focus on the production and consumption of health and health care; thereby, improvements in health would be possible to be purchased indirectly.

Scarcity is an unavoidable issue in health care because human beings have unlimited wants for health care and resources that are used to produce health care services such as human resources, capital and raw materials, are limited (Hubbard and O'Brien, 2006; Parkin, 2000); this means, in the health care economy as a whole, there are scarce resources to meet all of the people's wants for health care (Morris et al., 2012). The issue of resources scarcity and the potential uses of those limited resources in an efficient manner in the production of health care become the two general observations underpinning health economic analysis (Rudmik and Drummond, 2013; Shiell et al., 2002). The focus of health economic analysis, viewing health as a "fundamental commodity" and "health care being an economic good", is to make choices about the production of health care and consumption of health care resources that are scarce relative to society's wants for them (Rudmik and Drummond, 2013; Morris et al., 2012; Shiell et al., 2002; Gold et al., 1996). Due to resource scarcity, choices have to be made about different ways of using resources with respect to what quantity and mix of health care to produce, how to produce it, who pays for it and how it is distributed (Rudmik and Drummond, 2013; Morris et al., 2012; Shiell et al., 2002; Maynard and Kanavos, 2000). Making these choices within limited resources is undeniably difficult. This encounters inevitable trade-offs, meaning that one would need to choose one good or service from another. In other words, one would need to sacrifice a good or service in return for another good or service they choose. Making choices and trade-offs are both the most fundamental notion in economics in which they usually refer to as opportunity costs. Opportunity costs are the costs of resources consumed measured in terms of the value of the next best alternative forgone (a good, service or programme that is not chosen) (Shiell et al., 2002; Palmer and Raftery, 1999; Russell, 1992). Basically, this means the opportunity cost of undertaking an activity is the benefit that one must forego by being unable to allocate resources to the next best activity/alternative that is not chosen.

For a country's health care system that is predominantly funded by general taxation such as the U.K., where the government fixes a proportion of their total budget to be spent on health care at the start of each period, these choices are a 'top-down' hierarchy of decisions (Morris et al., 2012). In this environment, health economists,

therefore, play an important role to the U.K. government in providing evidence to inform decisions. This enables the government to make informed choices and decisions in health care; for example, what size of the budget for health care should be, how to allocate health care resources efficiently so that the greatest output can be obtained for a given set of resources (efficiency approach) and how to distribute and finance health and health care fairly amongst the people (equity approach), based on health economics theories and empirical findings that encompass both positive and normative economics aspects (Morris et al., 2012). Arroyos-Calvera and colleagues (2019) acknowledged that efficiency is the most important single factor determining preferences between policy options, but decisions were influenced almost as much by equity as by efficiency.

Positive economics deals with objective explanation and testing and rejection of theories, for example, the rising price of crude oil on world markets will lead to an increase in cycling to work (Friedman, 1953). Normative economics carries normative statements which are subjective statements of opinion rather than objective statements that can be tested by assessing the available evidence, for example, the government should enforce minimum prices for beers and lagers sold in supermarkets and off-licences in a bid to control alcohol consumption (Culyer, 1989). In other words, positive economics is concerned with the explanation and prediction of economic phenomena, while normative economics is concerned with evaluating economic policies, practices, and states of affairs from a moral stand-point (Hausman and McPherson, 2006). In summary, health economics has both positive and normative aspects; it is one of the areas of economics concerned with issues relating to efficiency, equity, value (i.e. the role of values in making judgements), behaviour as well as effectiveness and ethics in the production and consumption of health and health care in which these issues, collectively, are the important economic concepts in facilitating decision-making in health care.

1.5.2 Economic efficiency

Faced with the problem of resource scarcity, making the most efficient use of available resources is important to achieve economic efficiency. Palmer and Torgerson (1999)

elucidated that economic efficiency can be obtained through maximising the benefit from given resources. Goodacre and McCabe (2002) categorised economic efficiency into two types - technical efficiency and allocative efficiency.

Technical efficiency is about how best to achieve a given objective with the least possible expenditure. Gravelle and Rees (1992) explained that technical efficiency is about maximising the output obtained from given quantities of input. In their published articles in 2002, Goodacre & McCabe and Shiell et al., further elucidated that, if there are two possible ways to achieve a given objective, the most technically efficient option will be that which of the two possible options should be chosen to meet the specified objective at the lowest cost.

Allocative efficiency is about deciding which objective to meet between all the objectives that compete with one another for implementation. Precisely, allocative efficiency is about deciding which competing activity (i.e. service, programme or health technology) should resources be allocated to. Goodacre and McCabe (2002) explained that, allocative efficiency can be determined by making a value judgement about relative merits of different objectives; decision must be made on which objective is most worthwhile meeting rather than which intervention will best meet our objective alone.

Deciding which objectives are worthwhile meeting will require some sort of value judgement. Because of a value judgement is required, economic evaluation will play a vital role in decision-making process to inform and illuminate issues of allocative efficiency (Goodacre and McCabe, 2002). Health economic data provide information about how much one will need to pay to achieve every objective and what health benefits one might expect from achieving certain objectives, such as how many lives will be saved by carrying out a specific activity (Goodacre and McCabe, 2002). Further details about economic evaluation of health care are discussed in section 1.6.

1.5.3 Welfarism

Economic evaluation with respect to health care and health policy is an evaluation employing a broad set of analytical techniques to describe and compare the benefits and costs of competing uses of health care resources to improve the health of individuals, patient groups and populations. Drummond and colleagues (2015) described that economic evaluation is regarded as normative as value judgements are unavoidable for economists when concluding that one treatment option is better value for money compared to another after weighing up both the benefits and costs based on ethical and ideological values held by an individual or society. More importantly, it is concerned with how benefits are to be measured and valued, and how conclusions about the desirability of any given option are to be made i.e. whether such trade-offs made are acceptable and desirable within the context of value judgements.

Boadway and Bruce (1984) defined welfare economics or welfarism as a systematic methodological approach of evaluating economic implications of alternative resource allocations i.e. assessing the social desirability of one state of the world or resources allocation, purely in terms of the utility obtained by individuals, variously defined as the happiness, satisfaction, well-being or preference. Mishan (1981) and Boadway & Bruce (1984) described welfare economics as a normative allocation economics and is concerned with formulating and justifying propositions by which alternative social states, for example how scarce health care resources are allocated, may be ranked in a logical and consistent manner. The main objective of welfare economics according to Boadway and Bruce (1984) is to provide a coherent ethical framework for making meaningful statements in a normative way of thinking about the relative social desirability of different states of the world. A dominant ethical framework in welfare economics according to Hurley (2000) has been developed for assessing whether some states of the world are as better than or socially preferable to others and is built on four key tenets (Hurley, 2000) as explained below:

(1) The utility principle- Individuals behave in a manner that is utility maximising. This means that individuals are assumed to maximise their welfare in terms of

utility in a rational way by ordering the options open to them and subsequently choosing the preferred option.

(2) Individual sovereignty (Individualism)- Individuals are themselves the best judges of what is good for them, that is, what maximises their welfare in terms of utility they derive from different states of the world (i.e. what contributes most to their utility and how much that contribution is). In welfare economics, social welfare, a measure of the well-being of society as a whole, is therefore the sum of the utilities of all individual members of society.

(3) Consequentialism- This indicates that individuals derive utility from the outcomes of their choices, as opposed to, for instance, the processes themselves. In other words, utility is derived only from the outcomes of individuals' behaviour and processes in terms of their consumption of specific types and quantities of goods and services rather than the processes themselves.

(4) Welfarism- The term Welfarism here implies that the goodness or desirability of a state of the world is judged solely by the levels of utility attained by individuals in that state of the world. This means that this tenet only takes account the individual preferences in social rankings of states of the world. This effectively diminishes the information to be considered in any individual or social ranking of states of the world (i.e. the evaluative space) to only utility. On the whole, this tenet restricts what can be considered in an evaluation (i.e. the evaluative space) to individual utility only, ignoring all other possible outcomes (or are often generically labelled as non-utility information) which may be diverse.

Together, the above four tenets build the dominant welfare economic framework with the distinct feature that only individual utilities are allowed to determine the desirability of different states of the world. Given the objective of welfare economics is to devise a decision rule that allows the social desirability of all states of the world to be ranked in a logical and consistent manner which is to be based strictly on each individual's utility for that state, it is inevitable that the relative desirability of different states of the world and the appropriate trade-offs between individuals' utility must be

made. To do this within the welfare economics context that defines feature as a normative framework for social choice, the basic judgement to be used for this is the Pareto principle.

Pareto principle according to Brouwer et al. (2008) and Morris et al. (2012) consists of two type of improvement: a 'weak' Pareto improvement and a 'strong' Pareto improvement. A 'weak' Pareto improvement is a change in the state of the world that increases the utility of all affected individuals, whereas, a 'strong' Pareto improvement is a change in the state of the world that increases the utility of at least one individual and does not involve utility loss of any other individual. When a state of the world has no further Pareto improvements can be made, this means a Pareto efficient or Pareto Optimal is reached. Brouwer et al. (2008) defined Pareto Optimal as a state of the world where an increase in one person's utility can only be achieved by reducing the utility of at least one other person. Pareto principle according to Tsuchiya and Williams (2001), however has its limitations where it does not offer any methods of ranking different Pareto optimal states. Pareto principle does not concern with 'who is better off, or about the relative size of people's gains, but only that there are no losers' in which Morris and colleagues (2012) argued that this is a rather fundamental problem.

The potential solution to this is to go beyond the Pareto principle which is called Kaldor-Hicks criterion (also known as a potential Pareto improvement criterion) that uses the compensation test introduced by Kaldor (1939) and Hicks (1939) to evaluate people's gains and losses in terms of their monetary and utility. Kaldor's compensation test (Kaldor, 1939) suggests the gainers compensate those who lose, so that the losers are no worse off and the gainers can still be better off than before the policy. Similarly, Hicks's compensation tests (Hicks, 1939) proposes benefit (loss) is the maximum (minimum) amount of money that must be taken away from (given to) an individual so that they are as well off after the change as before it, meaning that they return to the same level of utility they had before the policy. A Kaldor-Hicks improvement is achieved when the gainers are still better off in terms of utility after compensating those who lose with the intention that the whole society would end up no worse off than before (Kaldor, 1939; Hicks, 1939). However, there are a number of limitations

to compensation tests. While compensation tests allow an overall efficiency in the improvement in society's well-being (in utility terms) to be achieved, the tests may not in practice be accompanied by compensation (in monetary terms) for the loser (Morris et al., 2012; McIntosh et al., 2011). Morris and colleagues (2012) argued that this could result in a redistribution of wealth or well-being in favour of the gainer at the loser's expense in which case this would violate the normative aspects of welfare economics. Further, they explained that compensation may be costly to negotiate and organise, and the use of money as a metric to represent the corresponding changes in utility for gainers and losers may be difficult in practice as different individuals value the utility they obtained from the compensation money differently. For example, given a situation where a low-income person is the loser (loss in utility) and a high-income person is the gainer (gain in utility), it is assumed that a low-income person is more likely to value the additional utility obtained from each additional £1 of income higher than a high-income person (Morris et al., 2012). Using compensation tests, this means that the amount of money required to fully compensate the loser (the low-income person) for their utility loss would be lower than the amount of money required to compensate the gainer (the high-income person) for foregoing the change (Morris et al., 2012). This Kaldor-Hicks criterion (or potential Pareto improvement criterion) forms an underlying rationale for cost-benefit analysis.

Macintosh et al. (2011) described cost-benefit analysis as the applied side of modern welfare economics and uses a variant of the Pareto criterion by placing monetary values on the gains and losses to those affected by a change in the provision of a good e.g. health care. Macintosh et al. (2011) further reiterated that the net gain or loss from a policy change can be calculated and whether the change is potentially Pareto-improving can also be determined. As a summary, economic evaluation of health care interventions allows analysts to evaluate whether certain programmes/treatments/medical technologies are worthwhile after weighing its benefits (in monetary terms and in the same unit as its costs) against its costs.

Culyer (1990) explained that measuring social welfare within welfare economic framework, which consists Welfarism to individual utility only e.g. welfare,

satisfaction, happiness or preference, is difficult because individuals' valuation of their utility is affected by their characteristics, behaviour and circumstances. This in turn has raised the questions on the usefulness and appropriateness of welfarism in the context of health economics and medical decision making, hence, the attention for an alternative stream, called Extra-Welfarism was introduced (Seixas, 2017; Birch and Donaldson, 2003; Hurley, 2000; Hurley, 1998; Culyer, 1990).

1.5.4 Extra-Welfarism

Extra-Welfarism, like welfare economics seeks to make meaningful normative statements about the relative desirability of different states of the world. However, it is distinct from welfare economics in such a way that it rejects the tenet of Welfarism; one of the key characteristics of welfare economics that only uses individual utilities in a welfare evaluation (Boadway and Bruce, 1984). Extra-Welfarism is described as a development or an extension of the welfare economics framework being a more general framework of analysis that allows the inclusion of a range of possible outcomes (e.g. other aspects and characteristics of people such as their health, capabilities or potential achievements in society) as well as or instead of individual utilities in welfare judgements (Brouwer et al., 2008; Hurley, 1998; Culyer, 1990). Extra-welfarists argued that there is more to welfare than just individual preference-based utilities that are important in deciding whether one state of the world is preferable to another (Brouwer et al., 2008); it is considered unsatisfactory as a normative underpinning of welfare judgements for the notion of welfarist that restricts what can be considered in the evaluation space to only individual's experienced or anticipated utility (Hurley, 1998; Culyer, 1990). Hence, this explains the development of the extra-welfarist framework, a type of normative welfare economics but it is not welfarist according to Boadway and Bruce (1984), for deciding on the relative desirability of different states of the world.

Extra-Welfarism was introduced and advocated especially by Anthony Culyer, a founding father of health economics, in health economics and medical decision making where health (or capability) of individuals has been the focus in the evaluative space (i.e. the evaluations of health economics and medical decision making) as an

important human characteristic (Culyer, 1990). Health, which may be considered as an appropriate maximand in the context of healthcare decisions, is measured in QALYs (Morris et al., 2012; Brouwer et al., 2008; Culyer, 1990). Quality Adjusted Life Years (QALYs), a preference-based measure entailing and combining the main characteristics of human beings such as physical, mental and emotional health, has been proposed as an appropriate yet meaningful measure of health (Culyer, 1990). QALYs are used in the context of cost-utility analysis in economic evaluation of health care and medical decision making (see sections 'Economic evaluation of health care' and 'Cost-utility analysis' for further information about cost-utility analysis and QALY). Also, other relevant outcomes, for example, relief of burden to carers or improvement in disease-specific quality-of-life score, may be selected and included in the evaluation, next to relevant health measures; these other relevant health outcomes are used in the context of cost-effectiveness analysis in economic evaluation of health care and medical decision making (see sections 'Economic evaluation of health care' and 'Cost-effectiveness analysis' for further information about cost-effectiveness analysis).

In summary, Extra-Welfarism allows outcomes such as health, capabilities or attainments of individuals other than individual utilities (e.g. welfare, happiness, life satisfaction or preferences) to be considered in a welfare economic evaluation. Thus, the extra-welfarism framework comprising the notion of welfare economics forms the theoretical concepts underpinning economic evaluations of health care and medical decision making (Coast et al., 2008a).

1.6 Economic evaluation of health care

Economic evaluation has become as an increasingly important assessment tool for the appraisal of health care (Drummond et al., 2015; Williams et al., 2008; Cunningham, 2001), predominantly as a decision-making tool to help address resource allocation issue for a wide range of very different interventions in health care settings (Drummond et al., 2015; Drummond et al., 1996). In particular, for a publicly-funded health care system such as in the U.K., economic evaluation plays as an important structured approach to help decision makers choose between alternative ways of

using resources. Decision makers (e.g. policy makers and health care commissioners) increasingly use the evidence of value for money of a programme along with its clinical outcome in their decision-making process to determine on how to allocate scarce resources in health care settings (Drummond et al., 2015). In short, economic evaluation is concerned with combining and evaluating costs and benefits together within an evaluative framework to provide meaningful information on the worthwhileness of particular resource allocation decisions in health care (Cunningham, 2001). There are five types of economic evaluation (Robinson, 1993a-e; Gray et al., 2011; McIntosh et al., 2011; Drummond et al., 2015) – (1) cost-effectiveness analysis (CEA); (2) cost-utility analysis (CUA); (3) cost-consequences analysis (CCA); (4) cost-benefit analysis (CBA); and (5) cost-minimisation analysis (CMA). Table 1.1 below presents an overview of the five types of economic evaluation (Yeo et al., 2019). Different types of economic evaluation are required in different health care studies, depending on what is being evaluated or targeted. Cost-effectiveness and cost-utility analyses are the two types of economic evaluation that most commonly employed in health care studies and clinical research as actual health outcomes are usually of health policy makers' and health care commissioners'/providers' interests (Drummond et al., 2015). Both the cost-effectiveness and cost-utility analyses are further described in Section 1.7 and 1.8 below.

Table 1.1: Overview of types of economic evaluation (Yeo et al., 2019)

Types	**Outcome measure**	**Number of health outcomes**	**Unit of health outcome**	**Unit of measure**
Cost-minimization analysis	Evidence proves that the outcome for competing alternatives are equivalent	None	None	Cost (£)
Cost-effectiveness analysis	Health outcomes are measured in natural units: e.g. life years gained, life saved, the number of cases detected, representing a main common goal for competing alternatives	One	Natural units: e.g. life years gained, life saved, the number of cases detected, etc.	Cost (£) per additional unit of the main common goal
Cost–utility analysis	Health outcome is measured in quality-adjusted life years (QALYs), a generic preference-based measure weighing the survival or the number of additional life years (quantity of life) by the quality of life (value) of the health state a person experienced in each year	One	QALYs	Cost (£) per QALY gained
Cost–benefit analysis	Both costs and benefits (e.g. lost productivity averted, disability-adjusted life year averted, case averted) are measured in monetary terms with the financial value of the costs compared with the financial value of the benefits	Many	Pound Sterling (£)	Cost (£)
Cost–consequence analysis	A disaggregated range of relevant outcomes and costs	Many	Varied	Not a ratio

1.7 Cost-effectiveness analysis

Cost-effectiveness analysis (CEA) is a form of economic analysis that compares the relative costs and outcomes (health effects) of two alternatives of a given health intervention; it is used to assess the extent to which alternatives is more effective and less costly (Drummond et al., 2015; Glick et al., 2014). Cost-effectiveness analysis is different from cost-benefit analysis which places a monetary value to the measure of effect. In the case of cost-effectiveness analysis, the outcome of interest to both alternatives is expressed in natural units so that comparisons among alternatives can be made; the natural units used in cost-effectiveness analysis are, for example, life-years gained, improvement in disease-specific quality-of-life score, average blood pressure improvement in mm Hg, lives saved, death avoided, or cases of disease averted (Phillips and Thompson, 2001). Cost-effectiveness analysis is expressed in the form of a ratio which is also known as Incremental Cost-Effectiveness Ratio (ICER)- ICER is calculated as the difference in costs (incremental costs) divided by the difference in effects (incremental effects) between two alternatives (Drummond et al., 2015; Glick et al., 2014; Gray et al., 2011; Phillips and Thompson, 2001). For a study that compares two alternatives say alternative 1 and alternative 2, the ICER can be calculated using the formula below:

$$\text{Incremental Cost-Effectiveness Ratio, ICER} = \frac{C_2 - C_1}{E_2 - E_1}$$

Where,

C_2 = Mean cost of alternative 2

C_1 = Mean cost of alternative 1

E_2 = Mean effect of alternative 2

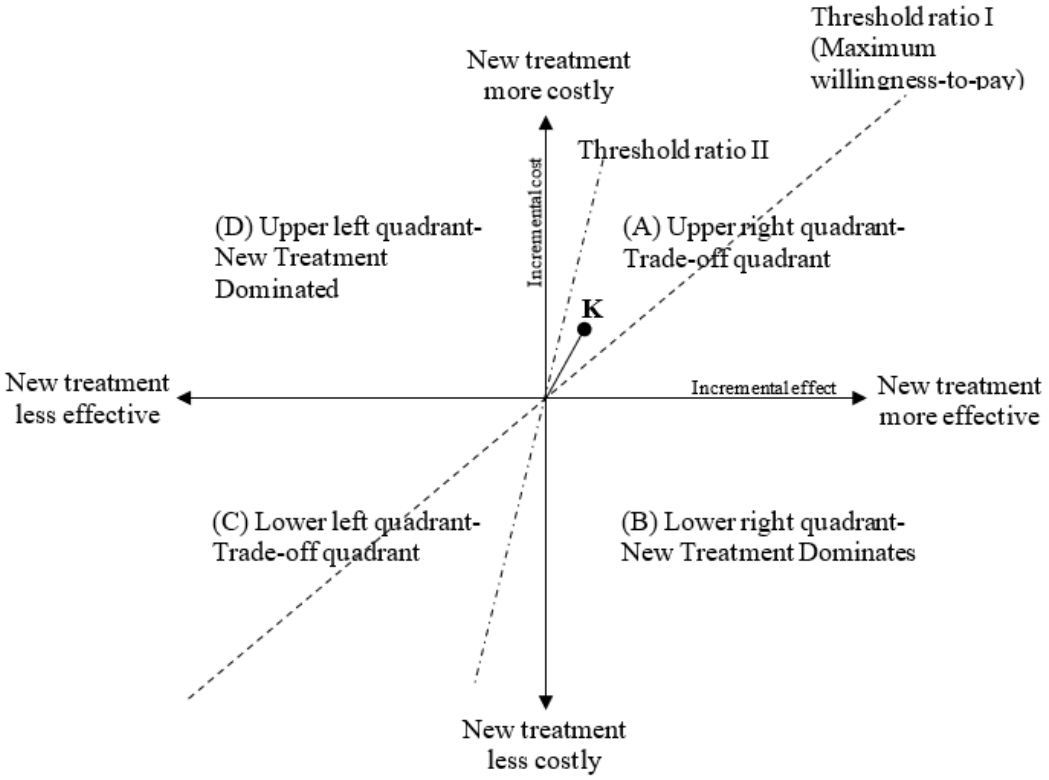
E_1 = Mean effect of alternative 1

The results of such comparisons (i.e. ICER) is expressed in the unit of cost per unit of health outcome or health effect gained (Drummond et al., 2015; Phillips and Thompson, 2001; Weinstein and Stason, 1977). The cost-effectiveness analysis result can be interpreted using cost-effectiveness plane diagram which consists of four quadrants. Figure 1.1 below shows the diagram of cost-effectiveness plane and

threshold incremental cost-effectiveness ratios (Drummond et al., 2015; Glick et al., 2014; Gray et al., 2011).

Difference in costs and effects between two alternatives (say new treatment versus current treatment) can fall into one of four quadrants. Plots lie in (A) upper right quadrant (trade-off quadrant) shows that the new treatment is more effective and more costly, (B) lower right quadrant shows that the new treatment is more effective and less costly (new treatment dominates), (C) lower left quadrant (trade-off quadrant) shows that the new treatment is less effective and less costly, and (D) upper left quadrant shows that the new treatment is less effective and more costly (new treatment dominated) (Glick et al., 2014; Gray et al., 2011). The proportion of plots below the threshold ratio I (the dashed line) represents the probability of the new treatment or programme being cost-effective. In other words, the area to the right of the dashed line (the threshold ratio I) is the region of cost-effectiveness. This means the region of cost-effectiveness is determined by the threshold willingness-to-pay ratio. Figure 1.1 shows that the incremental ratio for programme K falls outside the acceptable range of the threshold ratio I and the programme might therefore be deemed 'not cost-effective' if threshold ratio I is set as the maximum willingness-to-pay; however, if the maximum willingness-to-pay is increased to threshold ratio II (the dotted line), then programme K would be deemed cost-effective (Drummond et al., 2015).

Figure 1.1: Cost-effectiveness plane and threshold incremental cost-effectiveness ratios.



1.8 Cost-utility analysis

Cost-utility analysis (CUA) is a form of cost-effectiveness analysis that compares the incremental cost of a programme from a particular perspective to the incremental health improvement attributable to the programme (Drummond et al., 2015). In cost-utility analysis, the incremental health effect is expressed in the unit of quality-adjusted life-years (QALYs) (Drummond et al., 2015; Phillips and Thompson, 2001). Cost-utility analysis uses the similar ratio formula as for the cost-effectiveness analysis to calculate incremental cost-utility ratio (ICUR).

ICUR is expressed as cost per QALY gained (Drummond et al., 2015; Phillips and Thompson, 2001). As shown in the ICUR formula below, ICUR is calculated as the difference in costs divided by the difference in QALYs between two alternatives (e.g. services, treatments or health care technologies) (Drummond et al., 2015; Glick et al., 2014; Gray et al., 2011; Phillips and Thompson, 2001). For comparison between two

treatment groups say group 1 (Control group) and group 2 (Intervention group), the ICUR is calculated using the formula below:

$$\text{Incremental Cost-Utility Ratio, ICUR} = \frac{C_2 - C_1}{U_2 - U_1}$$

Where,

C_2 = Mean cost of group 2 (Intervention group)

C_1 = Mean cost of group 1 (Control group)

U_2 = Mean QALY of group 2 (Intervention group)

U_1 = Mean QALY of group 1 (Control group)

Comparing ICUR equation with ICER equation, they clearly show similarity, the only different is their denominator (i.e. incremental health effect). For ICUR equation in cost-utility analysis, its denominator is always measured in the unit of QALYs, a generic measure of health status that account for benefits on both quantity of life (survival) and quality of life (Drummond et al., 2015; Glick et al., 2014). Whereas for ICER equation in cost-effectiveness analysis, its denominator can be expressed in different natural units (such as life-years saved or cases averted), depending on what is the outcome of interest being evaluated in a study. Having health effects measured in different units in cost-effectiveness analysis limits its potential use in health care decision-making for policy makers and health care commissioners as it is difficult for comparisons of different studies for different disease areas, or different population groups. In contrast, cost-utility analysis uses QALY as a common unit of effect has the advantage of facilitating comparisons of health interventions across different conditions or disease areas.

Undoubtedly, the broad applicability of cost-utility analysis has made it more useful to decision-makers than is cost-effectiveness analysis that uses a measure of outcome specific to the programme under study (Drummond et al., 2015). The National Institute for Health and Clinical Excellence (NICE), the governing body for health in the UK serving the National Health Service (NHS) both in England and Wales, has recommended the use of QALYs as the preferred measure of health effects delivered

by various programmes or treatment regimens (NICE, 2013). This is because QALYs have the advantages that they combine multiple dimensions of outcome, i.e. both survival and quality of life, into a single measure that allows comparisons to be made across studies in different areas (Drummond et al., 2015; Glick et al., 2014; NICE, 2013). With health care programmes or interventions always aim to impact on individuals' length of life (survival) and health-related quality of life, QALY that seeks to reflect these two aspects in a single measure is therefore recognised as the generic measure of outcome and has been used in a wide range of clinical studies (Drummond et al., 2015). The results of cost-utility analysis can be interpreted using the cost-effectiveness plane in a similar way as in cost-effectiveness analysis (for further details about cost-effectiveness plane, see Figure 1.1 in the 'Cost-effectiveness analysis' section above). In cost-utility analysis, to determine whether a given programme/intervention or health care technology gives good value for money, the NICE in England and Wales has suggested to use a threshold range of £20,000 to £30,000 per QALY gained as a benchmark of willingness-to-pay for a unit of health gained (NICE, 2013).

1.9 Economic evaluation alongside clinical trials

Economic evaluation of health care interventions is frequently based upon prospective randomised clinical trials, but such evaluations are being made within the context of randomised clinical trials rather than being pursued as a stand-alone exercise (Adams et al., 1992; Drummond and Davies, 1991). O'Sullivan and colleagues (2005) argued that there are several potential advantages for economic analysis being piggybacked onto prospective randomised clinical trials compared to modelling studies that use retrospective data for its analysis and other types of economic assessments; the advantages are: (1) it is practical to collect patient-level economic data on costs (e.g. health resource use data) and outcomes (e.g. quality of life data) alongside a trial as personnel and processes for data collection are already in place for study's clinical measures; and (2) in terms of costs, it is often more viable to integrate economic components into a clinical trial rather than to fund a separate stand-alone economic evaluation study at a later date.

For a trial-based economic evaluation study, it is therefore important for health economists to be on board at the pre-funding stage (e.g. study design stage) through to the post-funding stages (e.g. study execution stage) so they can advise on how, when and what economic data on outcome and cost should be collected alongside the trial for use in economic evaluation (Glick et al., 2014; William et al., 2009). This is also to ensure that appropriate tools be employed to collect economic data on outcome and cost alongside clinical trials.

Among the five forms of the economic evaluation measures, NICE in the UK recommends undertaking cost-utility analysis (CUA), whether alongside a clinical trial or in an economic decision model, to help inform resource allocation decisions on which interventions/treatments/programmes to fund (NICE, 2013; Phillips and Thompson, 2001). In cost-utility analysis, benefits are measured in quality-adjusted life years (QALYs). It is widely recognised that using QALY as the measure of effect for assessing cost-effectiveness of new health care interventions/treatments/programs provides a standardised framework that allows comparison of value of interventions across different diseases and population groups (Drummond et al., 2015).

1.9.1 Quality-Adjusted Life Year (QALY)

QALY is a metric generated by combining two components – (i) quantity of life (the length of life i.e. the average number of years individuals spend in each particular health state) and (ii) quality of life (the health-related quality of life utility score of individuals i.e. a preference or utility weight associated with that health state) – into a single index (Brazier et al., 2007; Glick et al., 2014). To obtain the quality of life component to calculate the QALY in health economics to evaluate health care interventions in terms of their cost-effectiveness, generic preference-based measures (GPBMs) have been developed to include in clinical trials to collect individuals' health state data. The unique health state data collected would then be converted to health state utility value/score using a value set of preference weights, for each completion of the measure (Brazier et al., 2017; Brazier et al., 2007).

1.9.2 Generic Preference-Based Measures (GPBMs)

Generic preference-based measures (GPBMs) are also known as multi-attribute utility scales measures. Examples of GPBMs include the Quality of Well-being Self-administered (QWB-SA) scale (Pyne et al., 2003), the Health Utilities Index version 3 (HUI3) (Feeny et al., 2002), the EuroQoL-5 Dimension-3 Levels (EQ-5D-3L) (Dolan, 1997), the EuroQoL-5 Dimension-5 Levels (EQ-5D-5L) (Herdman et al., 2011), the Short Form-6 Dimensions (SF-6D) (Brazier et al., 2002), 15-Dimensional (15-D) (Sintonen, 2001) and the Assessment of Quality of Life-8 Dimensions (AQoL-8D) (Richardson et al., 2014). These preference-based measures vary greatly in terms of the content and size of their descriptive systems, the methods of valuation, and the population used to value the health states (most use a general population sample), though they all are claimed to be generic which they can be used in any population groups across different diseases/conditions (Brazier et al., 2017). A summary of the main characteristics of these six preference-based measures is presented in two separate Tables below - Table 1.2 and Table 1.3 (adapted from the review work of GPBMs by Brazier and colleagues (2017)). Table 1.2 summarises the descriptive systems of these measures, including their dimensions, severity levels and total number of health state combinations; Whereas, Table 1.3 summarises their valuation methods in terms of their valuation technique and type of model used.

Table 1.2: Descriptive systems of GPBMs (Brazier et al., 2017)

Measure	Total no. of dimensions	Dimensions	Total no. of severity levels	Total no. of health state combinations
15D	15	Breathing, depression, discomfort/symptoms, distress, eating, elimination, hearing, mental function, mobility, sexual activity, sleeping, speech, usual activities, vision vitality	4-5	31 billion
AQoL-8D	8	Coping (n = 3 items), happiness (4), independent living (4), mental health (8), pain (3), relationship (7), self-worth (3), senses (3)	4-6	2.37×10^{23}

EQ-5D-3L/ EQ-5D-5L	5/ 5	Anxiety/depression, mobility, pain/discomfort, self-care, usual activities	3/ 5	243/ 3125
HUI3	8	Ambulation, cognition, dexterity, emotion, hearing, pain, speech, vision	5-6	972,000
SF-6D	6	Energy, mental health, pain, physical functioning, role limitation, social functioning	4-6	18,000 (SF-36 v1), 18,750 (SF-36 v2) and 7500 (SF-12)
QWB-SA	3 (+68)	Mobility, physical activity, social functioning; 68 symptoms/problems	2	945

Table 1.3: Valuation methods of GPBMs (Brazier et al., 2017)

Measure	**Country**	**Valuation technique**	**Type of model**	**Health State Utility Value range[#] (min, max)**
15D	Finland	VAS	MAUT additive	0.11, 1
AQoL-8D	Australia	VAS transformed into TTO	MAUT multiplicative and statistical	-0.04, 1
EQ-5D-3L/ EQ-5D-5L	3L: UK, USA and 16 others/ 5L: UK and others	3L: Ranking, TTO, VAS/ 5L: TTO, DCE	Statistical additive / Statistical additive	3L UK: -0.59, 1 / 5L UK: -0.208, 1
HUI3	Canada, France	VAS transformed into SG	MAUT multiplicative	-0.36, 1
SF-6D	UK and 5 others	SG, ranking V2: DCE with duration	Statistical additive with interaction term	0.301, 1
QWB-SA	USA	VAS	Statistical additive, except for symptom/problem complexes	0.08, 1

[#] Health state utility values anchored at 1 representing full health, 0 (death) and values below 0 (worse than death)
15D: 15-dimensional; 3L: 3 levels; 5L: 5 levels; AQoL-8D: Assessment of Quality of Life 8 dimensions; DCE: discrete choice experiment; EQ-5D: EuroQol-5 Dimensions; GPBM: generic preference-based measure; HSUV: health state utility value; HUI3: Health Utility Index version 3; MAUT: multi-attribute utility theory; QWB-SA: Quality of Well-being Self-administered; SF-6D: Short Form 6 dimensions; SG: standard gamble; TTO: time trade-off; VAS: visual analogue scale; V2 version 2.

Of the six generic preference-based measures, EQ-5D, HUI3 and SF-6D are the three most frequently used measures (Longworth et al., 2014; Richardson et al., 2014; Brazier et al., 2017). Amongst these three frequently used measures, EQ-5D is the most widely used instrument by far and it is UK NICE's preference in 2008 (NICE, 2008) and then in 2013 (NICE, 2013) for assessing individuals' health-related quality of life for cost-utility evaluation in health care technology assessments. EQ-5D is a self-complete questionnaire that can be easily included in studies such as clinical trials or routine data collection systems with little respondent burden, and the existing scoring algorithms are easily used to generate the health state utility values for each completion of the measure (Brazier et al., 2017).

EQ-5D is the shortest generic preference-based measure with the smallest number of descriptive systems compared to all other above mentioned generic measures (15D, SF-6D, HUI3, QWB-SA and AQoL). Among the three most frequently used measures (EQ-5D, HUI3 and SF-6D), the 3-Level version of the EQ-5D (EQ-5D-3L) has the smallest descriptive system that comprises five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) with three levels of severity in each of the five dimensions (no problems, some problems, extreme problems) (EuroQoL Group, 2021). The combination of the dimensions and levels of the EQ-5D-3L defines a total of 243 health states. Then, a new 5-Level version of EQ-5D (namely EQ-5D-5L) was later (recently) introduced to further improve the sensitivity of the EQ-5D-3L and reduce ceiling effects by increasing the number of severity levels from three to five (Devlin and Brooks, 2017; Herdman et al., 2011). The EQ-5D-5L descriptive system comprises the same five dimensions as the EQ-5D-3L (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression), but it has five severity levels (no problems, slight problems, moderate problems, severe problems, unable to/extreme problems) in each of the five dimensions (EuroQoL Group, 2021). This means the combination of the dimensions and levels of EQ-5D-5L generates a higher number of health states than EQ-5D-3L's, that is a total of 3125 states (Brazier et al., 2017). Table 1.4 below shows the details of the classification of the 3-Level and 5-Level versions of the EQ-5D.

Table 1.4: EQ-5D-3L and EQ-5D-5L classification (EuroQoL Group, 2021).

Dimension	**EQ-5D version**	**Level**	**Description**
Mobility	3-Level	1	I have **no problems** in walking about
		2	I have **some problems** in walking about
		3	I am **confined to** bed
	5-Level	1	I have **no problems** in walking about
		2	I have **slight problems** in walking about
		3	I have **moderate problems** in walking about
		4	I have **severe problems** in walking about
		5	I am **unable to** walk about
Self-care	3-Level	1	I have **no problems** with self-care
		2	I have **some problems** washing or dressing myself
		3	I am **unable to** wash or dress myself
	5-Level	1	I have **no problems** washing or dressing myself
		2	I have **slight problems** washing or dressing myself
		3	I have **moderate problems** washing or dressing myself
		4	I have **severe problems** washing or dressing myself
		5	I am **unable to** wash or dress myself
Usual activities (e.g. work, study, housework, family or leisure activities)	3-Level	1	I have **no problems** with performing my usual activities
		2	I have **some problems** with performing my usual activities
		3	I am **unable to** perform my usual activities
	5-Level	1	I have **no problems** doing my usual activities
		2	I have **slight problems** doing my usual activities
		3	I have **moderate problems** doing my usual activities
		4	I have **severe problems** doing my usual activities
		5	I am **unable to** do my usual activities
Pain/discomfort	3-Level	1	I have **no** pain or discomfort
		2	I have **moderate** pain or discomfort
		3	I have **extreme** pain or discomfort
	5-Level	1	I have **no** pain or discomfort
		2	I have **slight** pain or discomfort
		3	I have **moderate** pain or discomfort
		4	I have **severe** pain or discomfort
		5	I have **extreme** pain or discomfort
Anxiety/depression	3-Level	1	I am **not** anxious or depressed
		2	I am **moderately** anxious or depressed

		3	I am **extremely** anxious or depressed
	5-Level	1	I am **not** anxious or depressed
		2	I am **slightly** anxious or depressed
		3	I am **moderately** anxious or depressed
		4	I am **severely** anxious or depressed
		5	I am **extremely** anxious or depressed

Subjects' responses to the five dimensions in the 3-Level or 5-Level version of EQ-5D would give a unique health state data. The health state data enables a health state utility value/score to be generated using a value set of preference weights, for each completion of the measure (Brazier et al., 2017; Brazier et al., 2007). EQ-5D-3L and EQ-5D-5L utility values/scores range from -0.594 to 1 and -0.208 to 1, respectively, where negative values (i.e. values less than 0) define a state as 'worse than death', 0 is the value of a health state equivalent to death, and 1 is the value of a health state representing full health (EuroQol Group, 2021).

To collect data on health state valuations to provide preference weights for all health states defined within the descriptive system of GPBMs, several techniques, as summarises in Table 1.3 above, are used for valuing the health states. The two most common techniques of valuation used across the six GPBMs have been Time Trade-Off (TTO) and Standard Gamble (SG) techniques (Brazier et al., 2017). Both of these techniques are choice-based methods which most economists have advocated for. TTO is a valuation method where subjects are asked to trade life years in the given ill-health state for a better health state in full health; Whereas, SG is a method that determines the risk of death that subjects are willing to take to avoid death in order to be in full health (Brazier et al., 2017). The valuation technique used for EQ-5D-3L is TTO as well as ranking and visual analogue scale (VAS); whereas, for EQ-5D-5L, discrete choice experiment (DCE) technique has been used as a hybrid model with TTO to value health states. DCE is a method used to elicit subjects' preferences where subjects are asked to make a choice over different hypothetical alternatives (Ryan et al., 2008). In the development of both EQ-5D-3L and EQ-5D-5L, a representative sample of the UK general population were asked to undertake a national survey to elicit health state

valuations for use in converting any health state into a single summary health index score (Dolan, 1997; Kind et al., 1999; Devlin et al., 2016).

1.10 Challenges in cost-utility analysis of health care interventions

While EQ-5D is a generic preference-based health-related quality of life measure and is NICE's preference for measuring health-related quality of life in cost-effectiveness analysis, there are concerns about the EQ-5D instrument in terms of its content validity and responsiveness in the presence of some conditions. Content validity describes the extent to which the classification comprehensively covers the different dimensions of a health condition and whether any important ones are missed. For instance, the absence of cognition in EQ-5D making the measures less valid in dementia (Brazier et al., 2017). Furthermore, it is important for analysts to examine whether the specific content of the item of a measure seems relevant to the concept they wish to measure. For example, the 'mobility' domain in the EQ-5D-5L, that is concerned with 'walking', may not be relevant for use in wheelchair user populations. In terms of responsiveness, in general, GPBMs have been shown to be unresponsive in some conditions (Brazier et al., 2017). Responsiveness of a generic instrument is related to the construct validity of the instrument that has the capacity to discriminate and the ability to respond to change over time; When an instrument is not able to reflect changes or differences in a population or seems to be less responsive to change compared to another measure (e.g. small standardised differences or changes), it is then described as being insensitive (Brazier et al., 2017; Whynes et al., 2013).

It is important to consider the potential insensitivity of the EQ-5D instrument in certain conditions (as opposed to the condition-specific measures which are potentially 'overly sensitive' and measuring differences that are not clinically meaningful – for example small difference in quality-of-life that people consider not important enough to trade for life expectancy in the utility valuation exercise). A summary critique of the literature below gives greater breadth and depth of the potential insensitivity of the EQ-5D for some medical conditions. From a review of GPBMs by Brazier and colleagues (2017d), evidence on EQ-5D suggests that EQ-5D is able to detect differences in many conditions including rheumatoid arthritis, many cancers such as lymphoma, and

depression (Herdman et al., 2020; Longworth et al., 2014; Brazier et al., 2014), musculoskeletal disease (Conner-Spady and Suarez-Almazor, 2001) and liver disease (Longworth and Bryan, 2003), but not able to reflect important differences/changes in the health outcomes associated with certain conditions (e.g. poor performance of EQ-5D in vision and hearing (Gandhi et al., 2020; Brazier et al., 2017a; Luo et al., 2015; Longworth et al., 2014), dementia (Hounsome et al., 2011), complex and severe mental health such as schizophrenia and bipolar disorder (Brazier et al., 2014), chronic obstructive airways disease (Harper et al., 1997)). In line with this, other studies have also discussed about the EQ-5D being insensitive to capture clinically important aspects of certain conditions, such as functional and symptomatic gains from a health care intervention (Pennington et al., 2020; Wichmann et al., 2017; Pettitt et al., 2016; Chevreul et al., 2015; Payakachat et al., 2015; Yang et al., 2015; Mulhern et al., 2013; Tosh et al., 2012; Whitehead and Ali, 2010; Willems et al., 2009).

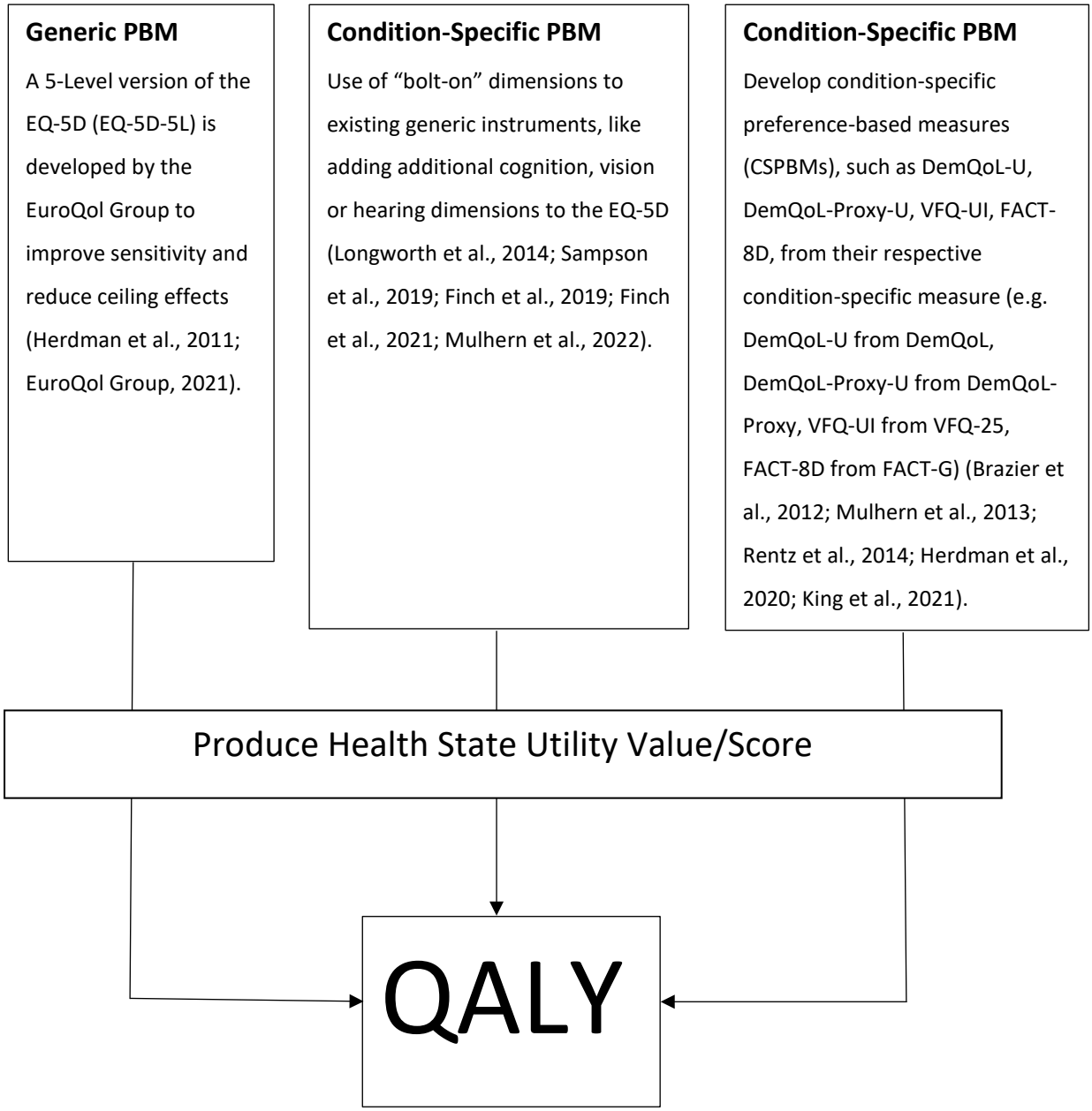
1.11 Development of methods to overcome challenges in cost-utility analysis of health care interventions

As the challenges in cost utility analysis (CUA) of health care interventions remains an issue in certain disease specific areas, researchers have developed multiple approaches in an effort to overcome the challenges in CUA of health care interventions. The most widely discussed among all the approaches is the development of a new EQ-5D, namely EQ-5D-5L. It is developed as an alternative to the EQ-5D-3L (EuroQoL Group, 2021) to further improve the EQ-5D's sensitivity and reduce ceiling effects (Jassen et al., 2013). EQ-5D-5L describes health in the five EQ-5D dimensions in greater detail as the number of descriptive levels for each dimension is increased from three to five. Studies reported that the EQ-5D-5L has superior psychometric properties compared to the EQ-5D-3L (Janssen et al., 2013; Agborsangaya et al., 2014). In terms of dimensions, EQ-5D-5L covers the same 5 core dimensions of health as in the EQ-5D-3L: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Keeping these five core dimensions the same are useful in assessing the health-related quality of life of individuals with a wide range of health conditions. However, they might not adequately capture clinically

important aspects of some conditions such as conditions affecting sensory functions like vision or hearing (Longworth et al., 2014) and mental health conditions affecting cognition (Hounsome et al., 2011).

To address this issue, preference-based health-related quality of life measures specially developed for specific conditions have been rising fast to improve EQ-5D sensitivity. The condition-specific preference-based health-related quality of life measures have been developed to extend the EQ-5D instrument to cover more dimensions – This is done by adding additional items ('bolt-on' dimensions) to the existing EQ-5D (e.g. EQ-5D with Vision 'Bolt-On' (Longworth et al., 2014)). Recent vision studies reported that a vision 'bolt-on' EQ-5D (Longworth et al., 2014; Luo et al., 2015; Gandhi et al., 2020), an extension of the EQ-5D instrument to cover vision dimension for patients with visions problems, appears to be more sensitive than the original EQ-5D (without 'bolt-on') in age-related macular degeneration (AMD) (Peixoto et al., 2021) and cataract (Gandhi et al., 2020). Further, another method has been proposed which is by transforming a condition-specific health-related quality of life instrument into a condition-specific preference-based health-related quality of life instrument (e.g. development of DemQoL-U and DemQoL-Proxy-U instruments from DemQoL and DemQoL-Proxy, respectively (Mulhern et al., 2013), and VFQ-UI from VFQ-25 (Brazier et al., 2017a, Rentz et al., 2014)). Figure 1.2 below shows a chart depicting the different methods of improving preference-based measure's (PBM's) sensitivity for certain conditions for cost-utility analysis.

Figure 1.2: A chart depicting the different methods of improving preference-based measure's (PBM's) sensitivity for certain conditions for cost-utility analysis.



Each method discussed above has both strengths and weaknesses. Development of condition-specific preference-based measures (e.g. DemQoL-U, DemQoL-Proxy-U) from their respective condition-specific measure (e.g. DemQoL, DemQoL-Proxy) is a complex development process that involves four phases: phase 1 – derivation of the health-state classification system; phase 2 – general population valuation survey and modelling to produce values for every health state; phase 3- patient/carer valuation survey; and phase 4 – application of measures to trial data. This shows that the development process for this method of transforming a condition-specific measure into a condition-specific preference-based measure is not only complex but also resource intensive. An external data set is also required to validate the classification systems of the developed condition-specific preference-based measures. Similarly, for the method of adding additional items ('bolt-on' dimensions) to multi-attribute utility instruments such as the existing EQ-5D-3L and EQ-5D-5L instruments, its development, assessment, and selection process of candidate 'bolt-on' descriptors is also a multiphases process and resource intensive. The use of 'bolt-on' dimensions of health to the EQ-5D has been proposed as a method to improve the content validity, its coverage, sensitivity, and responsiveness of the descriptive system in certain settings and health conditions in which the existing EQ-5D descriptive system may not be sensitive to the health impacts of certain conditions. While both of the above-mentioned methods have the same aim to improve the EQ-5D's sensitivity to change in certain conditions, their development process is undeniably complex, time-consuming and resource intensive. Furthermore, one other alternative to developing new measure is to undertaking mapping exercise which is described well in Wailoo et al. (2017).

Besides the different methods shown in Figure 1.2 above, in this thesis, I would like to further explore the potential use of disease-specific measure in cost-utility analysis using the method first tried in the MORTISE trial, by my colleagues and I, in other disease areas – gastro-oesophageal cancer and ophthalmology where the trials' datasets were available and accessible for use for the exploratory study for this PhD study. To fit with the purpose of the exploratory study in this thesis, minor modification to the hybrid method first tried in the MORTISE trial was performed to

allow assessment of generalisability of the approach (see Chapter 5 for further details).

In most trials, disease-specific measures are commonly collected independently alongside generic preference-based measure like the EQ-5D. However, in terms of QALY calculation for cost-utility analysis, disease-specific health-related quality of life measures are not always utilised in QALY calculation for cost-utility analysis. In other words, QALY is usually calculated from the health index score generated solely from the generic preference-based HRQoL instrument (e.g. the EQ-5D) despite a disease-/condition-specific HRQoL instrument is also used in trials.

Given the issue of lack of sensitivity of the existing EQ-5D instrument in certain disease areas, this poses the question as to whether the CUA results of health care interventions, that solely used EQ-5D, are robust and comprehensive enough to facilitate decision-making in health care settings. To ensure that generated QALY estimates cover all aspects of individual's conditions, incorporating disease-/condition-specific measure into QALY calculation might be an approach worth exploring to further improve QALY sensitivity in certain disease areas, and hence the quality of CUA results. The potential method that was first tried in the MORTISE trial allows the disease-/condition-specific measure collected in a trial be incorporated into the calculation of QALY using regression method.

The potential method first tried in the MORTISE trial was novel; my colleagues and I, as the lead trial health economist (Edwards et al., 2015), found that incorporating the condition-specific measure of the MORTISE trial (that is the Foot Health Thermometer, FHT) into QALY calculations did enhance the cost-utility analysis results of the trial. To obtain a more precise CUA results of health care interventions, this seems to be a worth exploring approach in studies/trials that have collected both EQ-5D and condition-specific measure independently in one single trial/study but only EQ-5D is utilised in QALY calculations.

To our knowledge, to date, the potential method first tried as a sensitivity analysis in the MORTISE trial that seems to make the QALY more sensitive and disease-specific has not been undertaken in other studies in other disease areas where both the EQ-5D and disease-specific health-related quality of life data are independently collected in one single study.

For this reason, with the available limited PhD resources and the accessible trials' dataset, I would like to conduct an exploratory study in this thesis to further explore the transferability/generalisability of the hybrid approach first tried in the MORTISE trial (Morton's neuroma study) in other disease areas e.g. gastro-oesophageal cancer (COGNATE trial) and ophthalmology (CLARITY trial) in cost-utility analysis for this PhD study (see Chapter 5 in this thesis for further discussions about this exploratory study).

1.12 An overview of the method first tried in the MORTISE trial of strengthening cost-utility analysis results

As there are concerns over the insensitivity of the EQ-5D-3L to change and the MORTISE trial was conceived before the EQ-5D-5L was developed, it stimulated the MORTISE team members to find a method of strengthening cost-utility analysis results in the MORTISE trial, which the method produced QALY estimates that incorporated with condition-specific measure of the MORTISE trial (i.e. Foot Health Thermometer, FHT). This was done by regressing the EQ-5D-3L data collected from the MORTISE participants as the dependent variable on their condition-specific measure data (i.e. FHT) as independent variable, with allocated treatment as the covariate. The resulting regression equation provides two complementary functions: First, it converts participants' responses to the FHT into utilities on the original EQ-5D-3L scale; Second, it uses the greater discrimination achieved by the FHT to fill gaps in the simplistic three-point scales that characterised the original EQ-5D-3L (Edwards et al., 2015).

In this thesis, the available accessible datasets from the two large clinical trials – COGNATE trial and CLARITY trial – were used for conducting an exploratory study to explore the potential use of disease-specific measure in cost-utility analysis in other disease areas using the simple linear regression method first tried in the MORTISE trial,

as explained above, with minor modification to fit the purpose of the exploratory study. COGNATE trial (Trial registration: ISRCTN1444215) (Russell et al., 2013) was a cancer study, funded by the Health Technology Assessment (HTA), that assessed the effectiveness and cost-effectiveness of endoscopic ultrasound in the management of patients with gastro-oesophageal cancer. In the COGNATE trial, EQ-5D-3L was used as the generic preference-based health-related quality of life measure, and FACT-G and FACT-AC were used as the condition-specific health-related quality of life measures. Whereas in the CLARITY trial (Trial registration: ISRCTN32207582) (Sivaprasad et al., 2018), the EQ-5D-3L questionnaire was used to collect patients' health utility data, and BCVA and VFQ-25 questionnaire were used to collect patients' condition-specific health-related quality of life data.

1.13 Cost-effectiveness of EUS in gastro-oesophageal cancer staging alongside COGNATE trial

The sensitivity of endoscopic ultrasound (EUS) in gastro-oesophageal cancer (GOC – oesophageal, gastro-oesophageal junction and gastric cancer) staging has been assessed; however, the effectiveness and cost-effectiveness of EUS staging in GOCs have not been evaluated until the inception of the HTA-funded COGNATE trial. Before COGNATE trial, there were no randomised controlled trials to evaluate the effectiveness and cost-effectiveness of EUS staging in the management of GOC patients.

Up to the inception of the COGNATE trial, although there are guidelines (Allum et al., 2002; SIGN, 2006) but no NICE guidance on this topic. The link between better staging and better management is not proven, and the benefit of EUS is not clear. This leads to the initiation of the randomised controlled trial in this area namely 'Cancer of the Oesophagus or Gastricus: New Assessment of the Technology of Endosonography' (COGNATE). This trial was designed to evaluate, not the accuracy of EUS, but the effect it had on patient management and outcome. COGNATE is an UK clinical trial commissioned by the National Co-ordinating Centre for Health Technology Assessment (NCCHTA) and led by Bangor University and Aberdeen Royal Infirmary in the UK.

The COGNATE trial (Trial registration number: ISRCTN1444215) provides a unique opportunity for me to undertake a meaningful economic study of EUS staging in the management of GOC patients. I was a health economist working on the trial alongside my PhD study. The trial ran until 2013 after which a publication of the COGNATE trial results was published in the NIHR-HTA; I co-authored the published NIHR-HTA COGNATE report. Following the completion of the COGNATE trial and the publication of the COGNATE trial NIHR-HTA report, NICE published a guidance in 2018 on the topic 'Oesophago-gastric cancer: assessment and management in adults' (NICE, 2018a; NICE, 2018b). The prospective primary health economics data collected alongside the COGNATE trial were used as part of my PhD study to evaluate how cost-effective is EUS in GOC staging as well as to explore the potential use of disease-specific measures in cost-utility analysis.

1.14 The purpose and over-arching aim of this thesis

As there has been no economic evaluation study based upon prospective primary economic data found in this area, economic evaluation study alongside the COGNATE trial was designed and conducted as part of my PhD study. The purpose of the economic evaluation study was to evaluate whether adding EUS to the standard staging algorithm in managing patients with GOC is cost-effective compared with its absence.

Further, given the scarcity of economic studies of EUS staging in GOCs it has to date including the absence of a systematic review of economic evidence and a survey study in this area, it would be very meaningful to fill these knowledge gaps by conducting a systematic review and a survey in this area as part of this thesis. The systematic review in this thesis provides the economic evidence of EUS staging in GOCs through undertaking a systematic yet comprehensive literature searches and review. Considering a holistic approach for this thesis, a health care professional survey study was designed and conducted to explore the utilisation of EUS in GOCs staging and its current practice in the UK. Finally, an exploratory study was conducted to explore the potential use of disease-specific measures in cost-utility analysis. This aims to explore

whether incorporating disease-specific measures into QALY estimations would have any potential benefits to cost-utility analysis results. All the stated points above are the purposes of this thesis.

In summary, the over-arching aims of this PhD study are to –

- Evaluate the cost-effectiveness of EUS in the management of patients with GOCs
- Offer a systematic review of economic evidence of EUS in GOCs staging
- Survey the utilisation and current practice of EUS in GOC staging in the UK, and
- Explore the potential use of disease-specific measures in cost-utility analysis in clinical trial.

1.15 Thesis novelty and contribution to knowledge

A holistic economic study of EUS in the management of GOC patients has not previously been conducted. This is the first study of its kind in this area that offers a novel opportunity to inform policy and facilitate decision-making in this clinical area. To achieve the over-arching aims of this PhD study, the following research questions are addressed, and the novel contributions are identified:

Thesis Chapter 2

Research Question 1:

Is adding EUS to standard staging algorithm cost-effective in the management of patients with GOC?

Novel Contribution:

This is the first economic evaluation study alongside clinical trial evaluating the cost-effectiveness of EUS in GOC staging. This study offers a novel opportunity to facilitate GOC multidisciplinary team and health care commissioners in their decision-making process with the evidence-based findings of the economic evaluation of EUS in GOC staging.

Thesis Chapter 3

Research Question 2:

What is the economic evidence on the use of EUS staging in patients with GOC?

Novel Contribution:

This is the first systematic review of economic evidence of EUS in GOC staging. This review offers a novel opportunity to increase the economic evidence of EUS in GOC staging and fill the knowledge gap in existing literature given the absence of comprehensive review of economic evidence of EUS in GOC staging. This systematic review provides not only the fundamental but also meaningful evidence-based information for informing health policymakers and commissioners in this clinical area.

Thesis Chapter 4

Research Question 3:

How is EUS in GOC staging used in current practice in the UK?

Novel Contribution:

Up to January 2021, this is the first UK survey on GOC Multi-Disciplinary Team (MDT) members' (GOC surgeons, gastroenterologists, oncologists) views about EUS staging in GOCs. This survey study offers a novel opportunity to provide an insight into the utilisation of EUS in GOC staging and the current practice in the UK.

Thesis Chapter 5

Research Question 4:

Has disease-specific measure got potential use in QALY estimations for cost-utility analysis?

Novel Contribution:

This exploratory study is the expansion of the concept of the "hybrid QALY technique" first tried in the MORTISE trial (on which I was the Research Officer working with the Trial Statistician, Dr Daphne Russell) (Edwards et al., 2015), to other disease areas. Data from two large clinical trials, namely COGNATE (a cancer trial) and CLARITY (an ophthalmology trial), were available for use for this exploratory study in this thesis. This offers an opportunity to inform the potential use of disease-/condition-specific

measures in QALY calculation for cost-utility analysis in clinical trials. Besides the MORTISE trial (Edwards et al., 2015), to our knowledge, there are no other studies exploring similar/the same innovative methodology of incorporating disease-/condition-specific measures into QALY estimations, as first tried in the MORTISE trial, for cost-utility analysis in other disease areas.

The disease-specific measures described below are used in the exploratory study in this thesis:

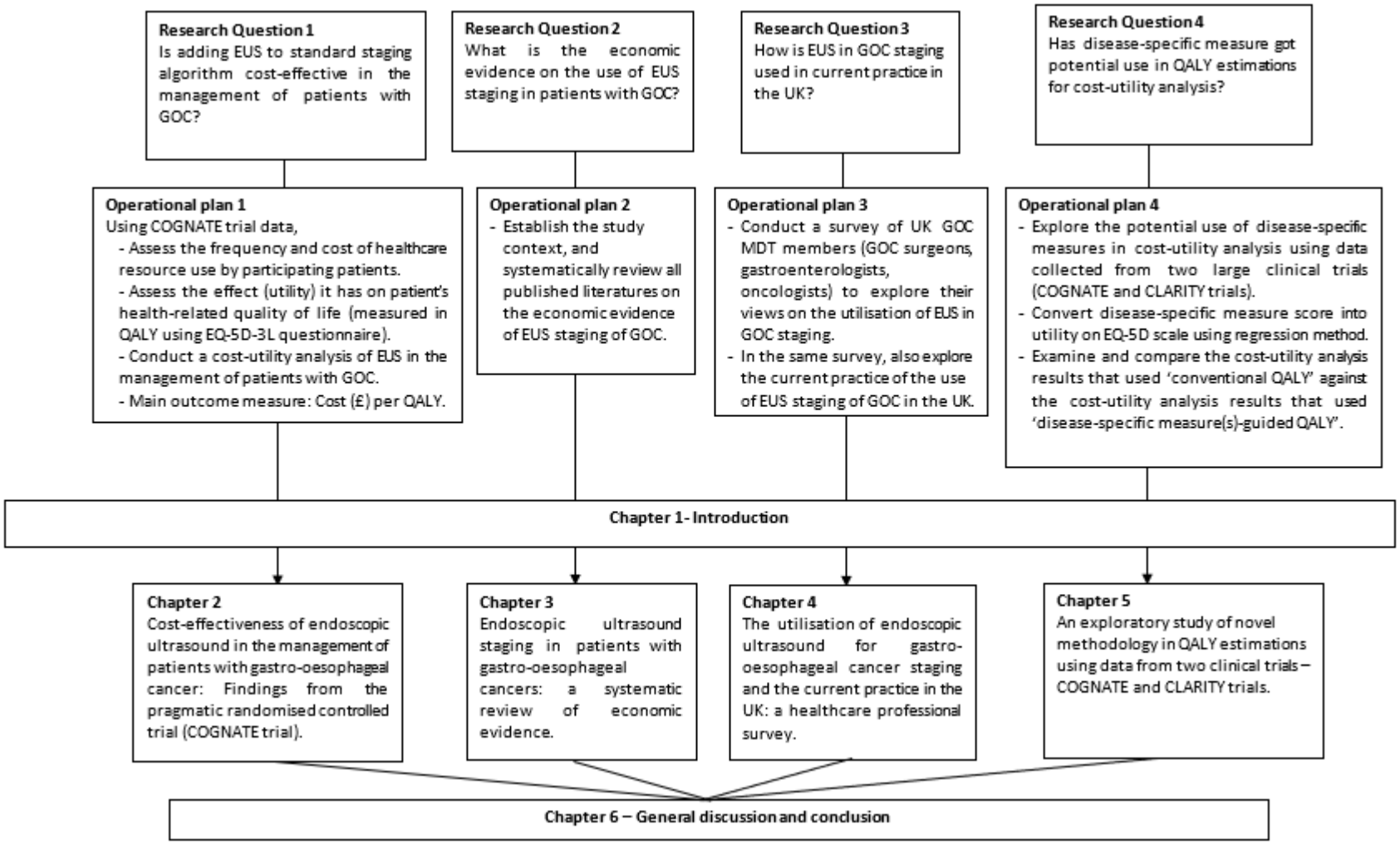
a) Functional Assessment of Cancer Therapy-General (FACT-G) – A 27-item questionnaire designed to measure health-related quality of life in cancer patients (FACIT.org, 2021a).
b) Functional Assessment of Cancer Therapy-Additional Concerns (FACT-AC) – A single integrated gastro-oesophageal cancer specific quality of life questionnaire merged from two separate FACT questionnaires – FACT-Oesophageal (FACT-E) and FACT-Gastric (FACT-Ga) (Russell et al., 2013; FACIT.org, 2021b; FACIT.org, 2021c), by the COGNATE team for ease of use for gastro-oesophageal cancer patients in the trial (Russell et al., 2013).
c) National Eye Institute 25-Item Visual Function Questionnaire (NEI VFQ-25) – A 25-item questionnaire designed to measure self-reported vision-targeted health status of people who have chronic eye diseases (NEI, 2021; Sivaprasad et al., 2018).
d) Best Corrected Visual Acuity (BCVA) – Measurement of the best vision correction that can be achieved with corrective lenses (i.e. glasses), as measured on the standard Snellen eye chart (Sivaprasad et al., 2018).

A similar approach may potentially be used in other disease areas where both the disease-/condition-specific measures data and EQ-5D data are collected independently in clinical trials. Examples of other disease-/condition-specific measures are the Asthma Quality of Life Questionnaire (Haldar et al., 2009), the Inflammatory Bowel Disease Questionnaire (Dudley-Brown et al., 2009), the Dementia Quality of Life (DemQoL) Questionnaire (Smith et al., 2005), and others, depending on what is the outcome of interests of a study.

1.16 Thesis structure

To achieve the aims of this PhD study, a flow chart of structure of this thesis showing four research questions and the corresponding operational plan for each of the four research questions is developed, as shown in Figure 1.3. Firstly, a cost-effectiveness evaluation study of the COGNATE trial is carried out to evaluate whether adding EUS to standard staging algorithm is cost-effective in managing patients with GOCs as compared with the absence of EUS. Secondly, a systematic review is carried out to systematically review economic evidence on EUS staging of gastro-oesophageal cancers (GOCs). Next, the third research objective of this thesis is to conduct a UK survey of GOC-interested Multidisciplinary Team (MDT) members (i.e. GOC surgeons, gastroenterologists, oncologists) about their views on the utilisation of EUS in GOC staging and the current clinical practice in the UK. Lastly, the fourth research objective of this thesis is to undertake an exploratory study of novel methodology in QALY estimations in other disease areas (i.e. cancer and ophthalmology) for cost-utility analysis in clinical trials. The following flow chart outlines the structure and lay out of this thesis (see Figure 1.3).

Figure 1.3: Flow chart of the structure and lay out of this thesis



1.17 Dissemination

Along with the four empirical chapters (Chapter 2-5) were written up as part of this PhD, the cost-effectiveness study of endoscopic ultrasound in the management of patients with gastro-oesophageal cancer of the COGNATE trial (Chapter 2) was published in the Health Technology Assessment (HTA) by the National Institute for Health Research (NIHR) Journals Library in 2013. Following the completion and publication of the COGNATE trial in HTA in 2013, NICE guideline (NG83) was published on the assessment and management in adults with oesophago-gastric cancer (NICE 2018a, NICE 2018b). The preliminary work of the exploratory study (Chapter 5) was presented at the Tenovus Cancer Care Charity 70th Birthday Conference at the SWALEC Stadium Conference Centre in Cardiff in 2013. Alongside, an EQ-5D 3D Model was invented by me and exhibited at the Tenovus 70th Birthday Conference in Cardiff in 2013. The systematic review protocol (Chapter 3) was published in the PROSPERO (International Prospective Register of Systematic Reviews; CRD42016043700) by the NIHR in 2016. The full systematic review (Chapter 3) was published in the BMC Cancer journal, a peer-reviewed journal, in 2019. The thesis protocol was presented at the Postgraduate Conference at Bangor University in 2016 and at the Combined Symposium of the Welsh Association for Gastroenterology and Endoscopy (WAGE) and Association of Upper GI Surgeons of Great Britain and Ireland (AUGIS) in the Life Sciences Hub Wales in Cardiff in 2016.

Thesis as a whole

Yeo, S.T., Hoare, Z, Haboubi, H, Edwards, R.T. (2016, May). *The health economics of EUS staging in patients with gastro-oesophageal cancer (GOC) and current clinical practice in the UK.* Poster presentation at the Postgraduate Conference, Bangor University, Bangor, UK.

Yeo, S.T., Hoare, Z, Haboubi, H, Edwards, R.T. (2016, May). *The health economics of EUS staging in patients with gastro-oesophageal cancer (GOC) and current clinical practice in the UK.* Poster presentation at the Combined Symposium of Welsh Association for Gastroenterology and Endoscopy (WAGE) and Association of Upper GI

Surgeons of Great Britain and Ireland (AUGIS) at the Life Sciences Hub Wales, 3 Assembly Square, Cardiff Bay, Cardiff, CF10 4PL, UK.

Cost-effectiveness analysis of endoscopic ultrasound in the management of patients with gastro-oesophageal cancer

Russell, I. T., Edwards, R. T., Gliddon, A. E., Ingledew, D. K., Russell, D., Whitaker, R., Yeo, S. T., Attwood, S. E., Barr, H., Nanthakumaran, S. and Park, K. G. M. (2013). Cancer of Oesophagus or Gastricus – New Assessment of Technology of Endosonography (COGNATE): report of pragmatic randomised trial. *Health Technology Assessment,* 17(39). https://doi.org/10.3310/hta17390

NICE Guideline NG83 was then published on this topic following the completion and publication of the COGNATE trial –

NICE (2018a). NICE Guideline: Oesophago-gastric cancer: assessment and management in adults (NG83). Retrieved from https://www.nice.org.uk/guidance/ng83/resources/oesophagogastric-cancer-assessment-and-management-in-adults-pdf-1837693014469

NICE (2018b). NICE Guideline NG83: Oesophago-gastric cancer: assessment and management in adults: Appendix L: Cost-effectiveness analyses. Retrieved from https://www.nice.org.uk/guidance/ng83/evidence/appendix-l-pdf-170036297751

Systematic review

Yeo, S. T., Bray, N., Haboubi, H., Hoare, Z., Edwards, R. T. (2016). Economic evidence for EUS staging in patients with gastro-oesophageal cancer (GOC): protocol for a systematic review. PROSPERO 2016: CRD42016043700. Retrieved from http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42016043700.

Yeo, S. T., Bray, N., Haboubi, H., Hoare, Z. and Edwards, R. T. (2019). Endoscopic ultrasound staging in patients with gastro-oesophageal cancers: a systematic review of economic evidence. *BMC Cancer,* 19, 900. https://doi.org/10.1186/s12885-019-6116-0

Preliminary work of the exploratory study

Yeo, S.T., Edwards, R.T. (2013, October). *Exploring health-related quality of life of patients with oesophageal and stomach cancer.* Poster Presentation at the Tenovus Cancer Care Charity 70th Birthday Conference, SWALEC Stadium Conference Centre, Cardiff, UK.

An *EQ-5D 3D Model* (as shown in the pictures below) was invented by Yeo, S.T. (2013, October) and exhibited at the Tenovus Cancer Care Charity 70th Birthday Conference, SWALEC Stadium Conference Centre, Cardiff, UK.

Chapter 2: Cost-effectiveness of EUS in the management of patients with gastro-oesophageal cancer: Findings from the COGNATE trial

2.1 Chapter Summary

Evidence on the cost-effectiveness of endoscopic ultrasound (EUS) staging in managing patients with gastro-oesophageal cancer is scarce; there had been no rigorous evaluation until the inception of the HTA-funded COGNATE trial (Registration Number: ISRCTN1444215). This study aimed to evaluate the cost-effectiveness of EUS from an NHS perspective, alongside the COGNATE trial, in managing patients with gastro-oesophageal cancer (GOC) compared to usual management.

A prospective cost-effectiveness analysis alongside pragmatic randomised controlled trial (COGNATE) was conducted at 48-month follow-up. A total of 223 patients with GOC who met the trial eligibility criteria were recruited from eight UK participating hospitals, consented and randomised into either intervention (standard staging algorithm plus EUS) or control (standard staging algorithm alone) group. Of these 223, 213 yielded enough data for the trial's primary analysis. Economic study sample had 107 intervention participants and 106 controls, the same as for the effectiveness analysis. The main outcome measures for economic analysis were Quality adjusted life years (QALYs), healthcare resource costs and cost per QALY over 48 months. Participants' quality of life scores data were collected using the European Quality of life-5 Dimensions-3 Levels (EQ-5D-3L) questionnaire at baseline and 1-, 3-, 6-, 12-, 18-, 24-, 36- and 48-month follow-up. Healthcare resource use data were collected electronically using a bespoke Microsoft Access database designed specifically for the trial. Healthcare resource use by participants were costed using published sources. As participants were in the trial for different lengths of time, estimates of costs and QALYs were adjusted to allow for censoring. Missing cost and QALY data were imputed as appropriate, and to allow for the censoring caused by variable follow-up in the trial. All costs and QALYs beyond 12 months were discounted at 3.5% per year as NICE recommends (NICE, 2013). All costs were in 2008 price year; for consistency throughout the thesis, the main costs were inflated to 2019 prices using the Hospital and Community Health Services (HCHS) pay and price inflation indices and the NHS Cost Inflation Index (NHSCII). A non-parametric bootstrapping with 5000 iterations

was performed to derive confidence intervals around the point estimates of incremental cost-effectiveness ratios (ICERs) when appropriate, and also to draw scatter plots of the joint distribution of the cost and effect pairs on incremental cost-effectiveness plane. Then, the corresponding cost-effectiveness acceptability curves (CEACs) was generated to quantify uncertainty and convey to policymakers the probability that EUS is cost-effective at different thresholds.

Primary analysis at 48 months, with costs and QALYs discounted at 3.5%, showed that, on average, intervention patients gained 0.1969 more QALYs than controls (95% bootstrapped CI from -0.0640 to 0.4575); and cost £2,860 less per patient (95% bootstrapped CI from -£7,987 to £2,192) (2019 prices: £3,490 less per patient, 95% bootstrapped CI from -£9,746 to £2,675), explained by fewer bed-days as inpatients. At the NICE thresholds of £20,000 and £30,000 per QALY, there is 97% probability that EUS is cost effective. Sensitivity analyses using two alternative estimates of the unit cost of EUS (£1,477 as outpatient or £3,781 as inpatient rather than £551 as day patient in the primary analysis) suggest that EUS remains cost effective with a probability of 95% and 86%, respectively. Subgroup analysis prompted by effectiveness findings showed that EUS was more cost-effective for participants reporting poorer health (below median EQ-5D score at baseline) despite saving less. Even for initially healthier participants, the intervention had at least 73% chance of being cost effective at the NICE threshold if EUS costs £551.

EUS has potential to save costs and is cost-effective with 97% probability. These economic findings provide strong evidence in favour of EUS scans for all gastro-oesophageal cancer patients with the potential to benefit. A systematic review of economic evidence of EUS in GOC is required to fill the gap in the literature review world and an exploratory methodological study on the modification of QALY for use in cost-effectiveness/cost-utility analysis is recommended.
Source of funding; National Institute of Health Research – Health Technology Assessment Programme. (***Trial registration:*** ISRCTN 1444215).

2.2 Introduction

Gastro-oesophageal cancer (GOC) is among the most common cancers worldwide with estimated annual totals exceeding 1.4 million new diagnoses and 1.1 million deaths, despite declining incidence over the past few decades (Jemal et al., 2011; Cook et al., 2009; Bosetti et al., 2008). In the UK, GOC was the seventh most frequent cancer with about 16,000 people diagnosed each year; and the fourth most common cause of cancer death with more than 12,000 deaths annually (Cancer Research UK, 2021a; Cancer Research UK, 2021b). GOC occurs mainly in men over 55 years of age (Allum et al., 2011).

Thorough and accurate staging for all patients is essential in managing GOC, especially to make the most of advances in treatment (Allum et al., 2011). Staging aims to plan appropriate treatment, and thus improve survival and quality of life. Computed Tomography (CT) was once standard practice in staging GOC, but its sensitivity is poor compared with endoscopic ultrasound (EUS), a diagnostic tool that combines endoscopy and ultrasonography (Thakkar and Kaur, 2020; Valero and Robles-Medranda, 2017; Puli et al., 2008). Though few studies reported the superiority of EUS over CT for staging GOC rather than diagnostic accuracy (Russell et al., 2013), EUS became common practice for staging following recommendations from three national bodies (Allum et al., 2011). However there had been no randomised trial to evaluate the effectiveness and cost-effectiveness of EUS in managing GOC. Thus, the link between better staging and better management was unproven, and the benefit of EUS was not clear. Therefore, a group of four co-lead investigators: (1) Professor Ian Russell (ITR), trialist and the methodological chief investigator of the COGNATE trial, (2) Professor Rhiannon Tudor Edwards (RTE), the health economist, (3) Dr David Ingledew (DI), senior lecturer in psychology and (4) Professor Kenneth Park (KP), consultant surgeon and clinical chief investigator, designed the trial 'Cancer of the Oesophagus or Gastricus: New Assessment of the Technology of Endosonography' (COGNATE) to evaluate, not the accuracy of EUS, but its effect on patient management and outcome (Russell et al., 2013). COGNATE was a pragmatic randomised controlled trial, commissioned by the NIHR Health Technology Assessment Programme to evaluate whether EUS is effective and cost-effective in managing GOC.

A literature review conducted during the COGNATE trial found only five relevant economic studies on EUS staging of GOC – two review articles (Lennon and Penman, 2007; Harris et al., 1998) and three American decision-analytic studies (Wallace et al., 2002; Harewood and Wiersema, 2002; Hadzijahic et al., 2000). Wallace et al. (2002) was the most comprehensive, inferring that the combination of CT, EUS and fine needle aspiration (FNA) was on average less costly and more effective in increasing quality-adjusted life years (QALYs) than all other staging strategies except the combination of positron emission tomography (PET), EUS and FNA. Though the latter was slightly more effective, it was more expensive and less cost-effective when judged by cost per QALY, the usual criterion in both US and UK. In the absence of guidance on this topic from the National Institute of Health & Care Excellence (NICE) prior to the inception of the COGNATE trial, we conducted a full cost-effectiveness analysis of EUS for GOC alongside the COGNATE trial.

2.2.1 My role and contributions to the COGNATE trial as a PhD student

I undertook the economic evaluation study alongside the COGNATE trial as part of my PhD studies. I managed, cleaned, costed, analysed and interpreted the economic evaluation data of the COGNATE trial. Professor Rhiannon Tudor Edwards (RTE) designed and led the economic evaluation within the COGNATE trial, and acts as guarantor. Dr Daphne Russell (DR) led the statistical component of the COGNATE trial. Professor Ian T Russell (ITR) was methodological lead of the COGNATE trial. I, together with RTE and DR, undertook the collection, analysis, and interpretation of data. Guided by RTE, I prepared the economic evaluation report of the COGNATE trial for submission to HTA (the funder of the COGNATE trial) with statistical advice/support provided by DR; ITR edited the overall COGNATE HTA report prior to submission. Following the submission to HTA, I prepared successive drafts of the economic evaluation manuscript; RTE, DR and ITR commented on the drafts. Then, I revised the drafts critically and have written up as a chapter for my PhD thesis. My PhD is supervised by (1) Professor Rhiannon Tudor Edwards (RTE; health economics lead of the COGNATE trial and health economics advisor/PhD supervisor), (2) Dr Zoë Hoare (ZH; statistical advisor/PhD supervisor), and (3) Dr Hasan Haboubi (HH; clinical

advisor/PhD supervisor). All my PhD supervisors (RTE, ZH and HH) commented on the final draft of this chapter.

2.3 Methods

2.3.1 Study setting and population

The COGNATE trial was conducted in eight UK hospitals (Russell et al., 2013). After standard staging algorithms, 567 potentially eligible patients with GOC were invited to participate, of whom 453 consented. However, 230 patients were then excluded, almost all after multi-disciplinary team (MDT) meetings to agree management plans. The remaining participants were randomised – 111 to intervention group (EUS) and 112 to control group (usual care). After further MDT meetings to review management plans in the light of supplementary investigations, notably EUS scans in the intervention group, participants were followed till death or the end of data collection on 31 July 2009, that is between 12 and 54 months from recruitment. Of these trial participants, 213 (96%) yielded enough data for effectiveness and cost-effectiveness analyses.

2.3.2 Measurement of outcomes

Researchers collected data on clinical outcomes and service use from hospital records, and on measures of effectiveness directly from participants – at baseline, discharge from hospital after initial treatment, and follow-up clinics after one, three, six, 12, 18, 24, 36 and 48 months. As this process yielded full economic data, imputed by the Trial Statistician (Dr Daphne Russell (DR)), if necessary (see further details in Section 2.3.4 and 2.3.5), for 107 intervention participants and 106 controls, the study sample for the economic evaluation was the same as for the effectiveness analysis.

Guided by Brazier et al (2007) and NICE (2013), the COGNATE trial economic team led by Professor Rhiannon Tudor Edwards (RTE) adopted QALYs as the measure of effect for economic analysis. I, with the guidance from RTE and DR, estimated QALYs by the area under the curve of participants' health-related quality of life measured by the EQ-5D-3L, a generic preference-based instrument developed and validated by the

European Quality of Life group (EuroQoL) to yield a single utility for each participant's health status at each timepoint (EuroQoL, 2021; Brooks, 1996).

2.3.3 Measurement of costs

I, guided by RTE, estimated costs from the perspective of the NHS at prices for 2008, these were inflated to 2019 in the final analysis. I, guided by RTE, recorded the type and frequency of participants' contacts with NHS secondary care, which account for almost all direct NHS costs for these patients, including investigation, treatment and palliation. The COGNATE health economics team (RTE and I) and the COGNATE trialist and methodological lead (Professor Ian Russell, (IR)) focused on EUS, surgery, chemotherapy, radiotherapy, other drugs, outpatient visits, day care and inpatient stays. The COGNATE health economics team (RTE and I) designed an electronic version of the Client Service Receipt Inventory (CSRI) – a trial-specific structured form that enables research staff to report the type and frequency of participants' contacts with care (Ridyard and Hughes, 2010; Knapp and Beecham, 1990) – and incorporated it into the COGNATE database with assistance from the COGNATE IT specialist (Mr Kevin Mawdsley (KW)).

Guided by RTE, I derived national unit costs for most of these resources from published sources including Unit Costs of Health and Social Care 2008 (Curtis, 2008), NHS Reference Costs for 2008 (Department of Health, 2009) and Prescription Cost Analysis (PCA) 2008 (NHS Information Centre, 2009). I, guided by RTE, estimated the cost of stents from Shenfine et al. (2005) and converted that to 2008 prices using the health services pay and price index (Curtis, 2008). As the unit cost of a PET scan was not available from the NHS, RTE and I estimated it in collaboration with the Aberdeen Royal Infirmary Nuclear Medicine and Finance Departments. I, guided by RTE, undertook detailed costing with surgeons, radiologists, oncologists and others at COGNATE trial sites of expensive procedures like chemotherapy, multimodal treatment, neo-adjuvant therapy, PET scanning, radiotherapy and surgery. IR, RTE and I used 'time and motion' studies in collaboration with trial sites to cost endoscopic ultrasound procedures. I, guided by RTE and IR, addressed variation between these national and local estimates through sensitivity analysis.

For consistency throughout this thesis, the main costs in this chapter were inflated to 2019 prices using the Hospital and Community Health Services (HCHS) pay and price inflation indices and the NHS cost Inflation Index (NHSCII) from the published sources (Curtis and Burns, 2019; Curtis and Burns, 2016).

2.3.4 Analysis strategy

Guided by RTE, I estimated incremental cost-effectiveness ratios (ICERs) as the difference in the mean cost of participants' care between intervention and control groups divided by the corresponding difference in mean QALYs gained. I used Stata 10.1 (Stata Inc., 2008) to bootstrap with 5000 replicates, thus ameliorating skewed cost data, and to derive 95% confidence intervals around estimated ICERs when findings require trade-off between increased benefits and increased costs (or, more rarely, between reduced costs and reduced benefits). I also used these replicates to construct incremental cost-effectiveness planes plotting incremental costs against incremental effects and the corresponding cost-effectiveness acceptability curves (CEACs) plotting the probability that EUS is cost-effective at different thresholds (Fenwick et al., 2004). These enable policy makers to compare the probability that an intervention is cost effective at different thresholds, particularly at both the NICE thresholds of £20,000 and £30,000 per QALY (NICE, 2013), and to compare ICERs (in cost per QALY gained) across other diseases and patient groups.

Guided by RTE and DR, I discounted all costs and QALYs beyond 12 months at 3.5% per year, as recommended by NICE (NICE, 2013). All analyses took account of censoring resulting from the termination of data collection at the end of the trial and the resulting variation in length of follow up between those recruited early and late (Glick et al., 2014).

DR performed the imputation and censoring to the COGNATE economic data. First, DR used the SPSS MVA procedure (SPSS Inc., 2007) to impute missing or partially missing QALY data at assessments before the end of the trial (Briggs et al., 2003; Lin et al., 1997). DR then projected cost and QALY data for each participant beyond the end of

the trial to 48 months, by multiplying the probability of being alive in each time interval by the imputed costs and QALYs of survivors. The imputation of missing data and the projection of data to 48 months performed by the Trial Statistician (DR) are described in detail in Section 2.3.5 below. With discussion and agreement of the Methodological Lead of COGNATE (IR), DR discarded the final interval between 48 and 60 months because only ten participants were still alive at 48 months, but none reached 60 months (Russell et al., 2013).

I bootstrapped analyses (with 5000 replicates) at 48 months to explore whether the estimates of the costs and benefits of EUS relative to conventional staging were sensitive to key assumptions. As the cost of EUS varied between national tariffs, local tariffs and the literature, I, with the guidance from RTE and IR, investigated sensitivity to the choice between: £551 for day patients as used in our primary cost-effectiveness analysis; £1,477 for outpatients; and £3,781 for inpatients. Finally, two exploratory subgroup analyses at 48 months: (1) classifying participants as below or above the median baseline EQ-5D score of the whole sample, investigated whether cost-effectiveness of EUS varied with self-reported EQ-5D scores at baseline, and (2) classifying participants as below or above median age of 65 years, examined whether cost-effectiveness of EUS varied with age.

2.3.5 Missing values and imputation

In the COGNATE trial, the Methodological Lead (IR) and Senior Trial Statistician of COGNATE (DR) were responsible for dealing with missing data and performing imputation. They explained that participants for whom some outcome data are missing (Glick et al., 2007; Carpenter and Kenward, 2007) were not excluded, whenever possible, to avoid bias in analysis by treatment allocated; therefore, these missing data were imputed from known data about these participants and other participants whose outcome data are known (Briggs et al., 2003; Lin et al., 1997).

The main trial recruited for 3.5 years. To maximise statistical power, all participants were followed for as long as possible, between 1 and 4.5 years. The main aim of conducting imputation for the outcome data was to achieve complete data within the

design rather than to extend data beyond 31 July 2009, the end of the trial and thus the censoring date. For health economic analysis, the costs and benefits, that were estimated by STY with guidance from RTE, were also imputed for all participants for the same period by DR for used in the bootstrap methods to assess cost-effectiveness. Two follow-up times were chosen for investigation: 12 months, the minimum unless a participant withdrew from the trial, and 48 months, which took account of all information on both survival (as there were no subsequent deaths) and quality of life (as the last quality of life questionnaire to participants was at 36 months). This allows analyses of both effectiveness and cost-effectiveness providing explicit links between the effectiveness and cost-effectiveness sections of the published COGNATE report (Russell et al., 2013).

For all participants who died during follow-up, potentially there was complete information as dead participants did not use any more resources and it was agreed to set their subsequent quality of life to zero, rather than missing. Survival data were also received on all participants up to complete withdrawal or the end of the trial. DR explained that survival data does not need imputation as survival analysis allows for variable follow-up.

There were two types of data received in the COGNATE trial:
(a) Clinical data, including demographic and resource use, extracted by research professionals from hospital notes and entered retrospectively onto the electronic database – In general, these data did not need to be imputed, because research professionals were asked to collect complete data except for pre-randomisation tumour stage, for which 'missing' was permitted. However, DR imputed resource use in secondary care from the end of the trial to 48 months.

(b) Patient-reported outcome measures at baseline and follow-up – DR imputed the few missing data for the main effectiveness analysis. For the cost-effectiveness analysis and effectiveness sensitivity analysis, DR needed to impute quality of life to 48 months.

The COGNATE Trial Statistician (DR) carried out the imputation of missing quality of life data in three phases where each phase used SPSS (SPSS Inc., Chicago, IL, USA) Missing Values Analysis (MVA) procedure (Briggs et al., 2003). Table 2.1 below provides further information about the different phases of the imputation of missing quality of life data and cost data. DR explained that the SPSS MVA procedure performed by her uses single imputation for missing data, it simultaneously estimates all missing values in a data set, on one or more data sets (Russell et al., 2013). DR also performed the final phase of imputation on missing cost data where it used estimated survival probabilities from the Cox Regression model.

DR used Cox Regression to model the simultaneous effect on survival of several characteristics, for the main quality-adjusted survival (i.e. QALY) comparisons. Both primary quality-adjusted survival analysis and secondary survival analyses in the COGNATE trial using Cox Regression models considered several baseline characteristics, including EQ-5D and FACT baseline scores, for inclusion as covariates. DR did this, not only to take account of any baseline imbalance between groups despite stratification, but also to improve the precision and generalisability of the model. DR and IR always included centre, condensed to three groups of similar size: Aberdeen, Gloucester and the rest. DR and IR always used the baseline score of a given measure to predict a later score of that measure. Other characteristics considered in step-wise model building included: age and gender; site, stage and type of tumour; the initial management plan agreed before randomisation; but not WHO status, as most participants had a WHO status of 1. To get the best from 'initial management plan' in predicting outcomes, DR created a binary variable to distinguish between conservative prior plans (namely chemotherapy, radiotherapy, both or neither) and therapeutic prior plans (namely endoscopic resection or surgery in some form). As DR and IR had expected, this later proved very good at predicting outcomes. As conservative plans choose between all possible combinations of chemotherapy and radiotherapy, DR followed the example of many MDTs by describing this and the resulting binary variable as 'multimodal'.

Table 2.1: The imputation procedure phases of missing quality-of-life data and cost data, performed by the Trial Statistician (DR). The text in the Table 2.1 below is quoted directly from the published COGNATE report (Russell et al., 2013).

Imputation Phase	Description	Reference
Phase 1	**Imputing quality of life: phase 1 – psychometric and effectiveness analyses** Phase 1 used all quality-of-life items (27 in FACT-G covering four subscales; 33 in FACT-AC; five in EQ-5D, one for each domain; and EQ-VAS, a single item) answered in interviews at the same time to estimate the missing items in those interviews. This yielded a complete set of responses for existing interviews. Initial psychometric analyses, using responses at 0, 1 and 3 months, reduced the number of items in FACT-AC by two, after which we repeated phase 1. We used the resulting data to calculate scores for EQ-5D and FACT scales and subscales for all existing interviews.	Russell et al., 2013, p22.
Phase 2	**Imputing quality of life: phase 2 – until 12 months** DR then discarded item scores, except EQ-VAS, in favour of scale scores across time. DR created a single data set comprising all 213 participants and set scale scores to zero after death. To this data set DR added the allocated treatment, and baseline characteristics to improve estimates. Phase 2 used only time points up to 12 months, the minimum period of follow-up in the trial. DR used MVA to impute scale scores at times without interviews for those who were still alive at 12 months, and then for those who had died by 12 months. This yielded a complete imputed data set with all quality-of-life scores at all times up to 12 months. Three participants withdrew before 12 months. While phase 2 included them among survivors, phase 3 adjusted their estimated quality of life at times after withdrawal to take account of the probability of death.	Russell et al., 2013, p22
Phase 3	**Imputing quality of life: phase 3**	Russell et al., 2013, p22

No more participants withdrew completely after 12 months. Beyond 12 months, however, the survival status of progressively more participants is unknown because of censoring at the end of the trial. Phase 3 therefore estimated both the probability of being alive at each of the three remaining times and the quality of life of the participant if alive. Multiplying these two estimates yields the expected quality of life. This procedure adapts to quality-of-life data the process for imputing censored cost data described by Lin et al. (1997).

In the first part of phase 3, DR derived the probability of censored participants being alive at 18 months from the Cox regression model for survival. DR calculated similar probabilities at 24 and 36 months, and also at 48 months for use in cost-effectiveness analysis and effectiveness sensitivity analysis. In the second part of phase 3, DR used three separate MVA imputations to extend the data set from phase 2 to 18 months, 24 months and 36 months for those not known to be dead at those times. DR multiplied each imputation by the probability that each participant would have been alive at this time. Finally DR set quality of life scores for people known to be dead to zero, or the equivalent for FACT-AC, for which 0 is the best possible score.

By the end of phase 3, we had complete quality-of-life information for all 213 participants at all time points before the end of the trial, and survival status at that date, enabling us to estimate QALYs for primary analysis. We also had expected quality of life scores, but not survival status, for all 213 participants at all time points up to 36 months, for cost-effectiveness analysis and fully imputed sensitivity analysis for effectiveness.

Phase 4	**Imputing: phase 4 – secondary care costs** STY and DR combined data on resource use in secondary care into six periods – up to 12 months, 12–18 months, 18–24 months, 24–36 months, 36–48 months and 48–60 months. STY costed and summed unimputed frequency data to give the total cost in each period for each participant. However, DR and IR were able to discard the final period (48–60 months) because by 48 months only 10 participants were still in the trial, none of whom reached 60 months. For each of the first five periods, DR imputed the expected cost for unobserved participants by adapting the method of Lin et al. (1997), although not exactly as DR had adapted it to quality-of-life data in phase 3 above. Of these four phases of data imputation, phase 1, which imputes missing answers to questions within a scale from answers to related questions in the same scale, is not appropriate to costs because costs have no 'related questions within a scale' in the psychometric sense. Phase 2 was not necessary because we observed costs until censoring at 12 months or later. As with Lin et al. (1997), our costs were spread over intervals, while we had collected and imputed (by DR) quality-of-life scores at exact time points, which included the ends of the cost intervals. Hence, although the survival probabilities were exactly the same in Phases 3 and 4, DR used two for each cost interval – those of being alive at the start and the end of the interval. Unlike quality-of-life scores, however, costs are highly skewed and unsuitable for the MVA procedure (Glick et al., 2007). In general, therefore, DR estimated costs in unobserved intervals from the mean cost among people in the same allocated treatment group who were alive and observed throughout the interval. Nevertheless, DR used separate estimates for the cost of the year before death, because they are consistently and considerably higher than all years other than the first.	Russell et al., 2013, p23

2.3.6 Ethics approval and conduct

Before starting the trial, approval was gained from the Scotland A Multicentre Research Ethics Committee (reference 04/MRE10/10) and ten Local Research Ethics Committees and associated research governance units (Russell et al., 2013). Trial participants were asked for informed consent before allocating them to either intervention (EUS) or control (usual staging) group.

2.4 Results

2.4.1 Characteristics of study participants

Randomisation duly ensured that intervention and control groups had similar demographic characteristics and baseline quality of life scores [see COGNATE HTA report (Russell et al., 2013)].

2.4.2 Quality adjusted life-years (QALYs)

At 48 months, the intervention group had gained a mean of 1.362 discounted QALYs (SD 0.999) and the control group a mean of 1.165 (SD 0.976). Thus, EUS group gained on average 0.197 more QALYs than the control group (95% bootstrapped confidence interval (CI) from -0.0640 to +0.4575) (Table 2.2a).

Table 2.2a: Fully imputed costs* and QALYs by allocated treatment group over 48 months[a]

Type of cost	Intervention (n=107); mean (SD)	Control (n=106); mean (SD)	Intervention minus Control (Bootstrapped 95% CI)
Over 48 months			
Total cost / patient (£)	29,190 (14,902)	32,049 (22,019)	-2,860 (-7,987, 2,192)
Total cost / patient (£)¥	35,619 (18,184)	39,108 (26,869)	-3,490 (-9,746, 2,675)
Discounted QALYs (yrs)	1.3616 (0.9989)	1.1647 (0.9756)	0.1969 (–0.0640 to 0.4575)

* All costs were in 2008 prices unless stated otherwise.

a These estimates use fully imputed data; Table 2.4 uses unimputed data to elaborate these costs over 5 years from randomisation.

¥ 2019 prices – For consistency throughout this thesis, the total cost per patient (£) was inflated to 2019 prices, using the Hospital and Community Health Services (HCHS) pay and price inflation indices and the NHS cost Inflation Index (NHSCII) from the published sources (Curtis and Burns, 2016; Curtis and Burns, 2019).

Table 2.2b shows the mean (SD) of the EQ-5D scores in survivors at each time point, stratified by trial arm. It appears that participants are getting worst in their EQ-5D where their mean EQ-5D utility scores are declining over the study period. Similarly, over the study period, fewer participants remained in the trial. Between the two groups, participants in the EUS group had higher mean EQ-5D utility score compared to participants in the non-EUS group; however, their differences in mean EQ-5D utility scores are not statistically significant.

Table 2.2b: EQ-5D utilities by allocated group: unadjusted means (including deaths and imputed survivors). This table is quoted directly from the published COGNATE report.

Time of EQ-5D	**Non-EUS group**		**EUS group**			**EUS minus non-EUS**	
	n	**Mean (SD)**	**N**	**Mean (SD)**	**Total n**	**Difference (95% CI)[a]**	
Survivors (imputed to end of trial if necessary) and deaths before end of trial							
Baseline	106	0.801 (0.164)	107	0.807 (0.198)	213	0.007	(–0.042 to 0.056)
1 month	106	0.733 (0.285)	107	0.729 (0.265)	213	-0.005	(–0.079 to 0.070)
3 months	106	0.615 (0.305)	107	0.658 (0.289)	213	0.043	(–0.037 to 0.123)
6 months	106	0.535 (0.323)	107	0.550 (0.347)	213	0.015	(–0.076 to 0.106)
12 months	106	0.449 (0.391)	107	0.509 (0.376)	213	0.061	(–0.043 to 0.164)
18 months	87	0.377 (0.400)	90	0.394 (0.399)	177	0.017	(–0.102 to 0.135)
24 months	73	0.251 (0.347)	77	0.330 (0.392)	150	0.079	(–0.041 to 0.198)
36 months	44	0.189 (0.312)	47	0.226 (0.373)	91	0.037	(–0.106 to 0.181)
All participants (fully imputed beyond end of trial if necessary)[b]							
18 months	106	0.353 (0.387)	107	0.400 (0.382)	213	0.047	(–0.057 to 0.151)
24 months	106	0.260 (0.341)	107	0.328 (0.364)	213	0.068	(–0.027 to 0.163)
36 months	106	0.152 (0.260)	107	0.211 (0.306)	213	0.060	(–0.017 to 0.136)

SD, standard deviation.

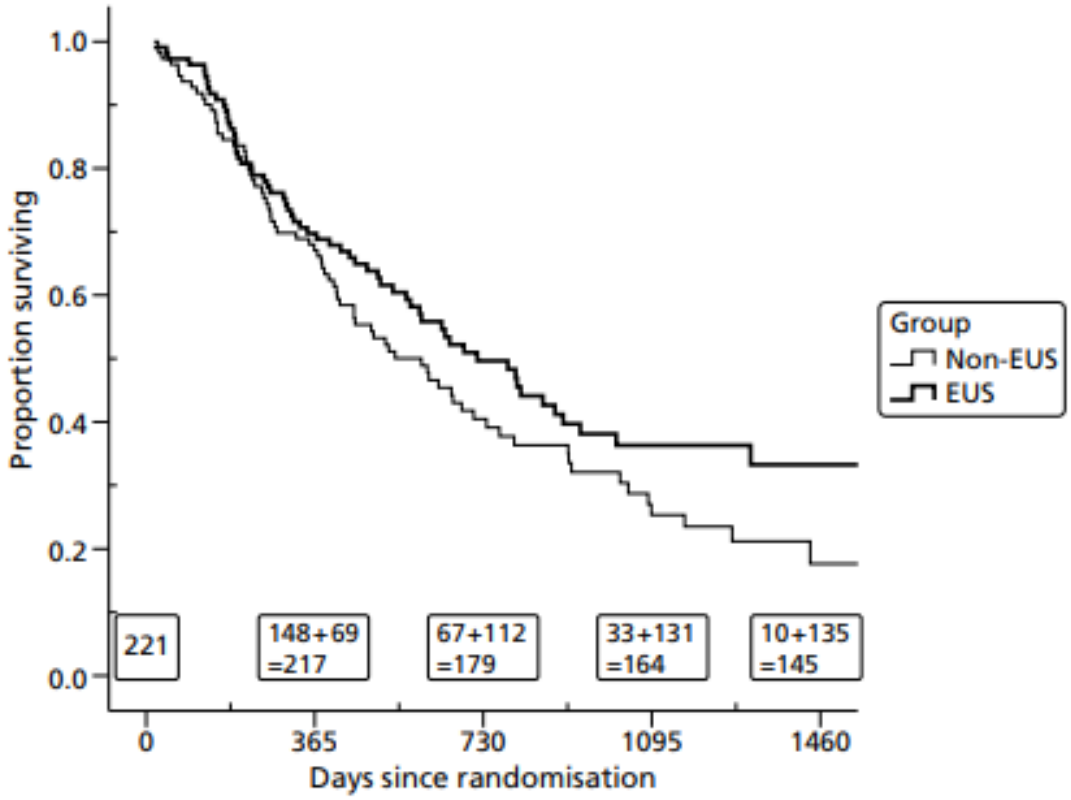
a Confidence intervals for comparison only; to avoid multiple testing, analysis focuses on 12-month follow-up.

b Confidence intervals in this subtable, especially for 36 months, are narrower than in corresponding rows of previous subtable owing to increase in sample size

Figure 2.0 below shows the observed survival curves stratified by allocated trial arm (EUS group vs non-EUS group). The survival curve of the intervention group (EUS) lies above the curve of the control group (non-EUS). The proportion of surviving is similar in the two groups up to about 12 months, after which the curves diverge. The COGNATE Trial Statistician, DR, explained that these observed survival curves have taken account of withdrawals and different lengths of survival time in the trial (between 1 and some 4 years) to provide unbiased estimates of the proportion of the original sample still alive at each time. DR further explained that because of censoring, the lower portion of the curve suffers from more random variation. For example, a single death represents a change in survival of about 0.01 in the first year of the trial and about 0.02 at three years.

Figure 2.0: The observed survival curves by allocated trial arm (EUS group vs non-EUS group).

(Note: rectangular boxes above the x-axis record the number of participants available to estimate survival curve, subdivided into participants still alive + cumulative deaths; thus these boxes exclude participants censored when the trial ended or they withdrew completely. This figure is drawn directly from the published COGNATE report).



2.4.3 Use of healthcare resources

Table 2.3 shows the mean unimputed frequencies of resource use in secondary care for all causes including hospital-prescribed drugs, unimputed but adjusted to avoid bias from censoring, over six time intervals for the 213 participants in the primary cost-effectiveness analysis (see full table in Appendix 2.1). The sparsity of these data, especially the presence of many zeroes and thus high skewness, meant that imputation was not possible for individual frequencies (Russell et al., 2013). Instead Table 2.3 includes three columns estimating mean frequencies for complete cases from the unimputed data (see full table in Appendix 2.1).

Table 2.3: Estimated mean frequency of contacts with secondary healthcare including hospital-prescribed drugs by 213 participants over 60 months[a].

	Intervention group (Estimated mean frequency in each interval)						Estimated mean frequency for complete case	Control group (Estimated mean frequency in each interval)						Estimated mean frequency for complete case	Total sample (Estimated mean frequency in each interval)						Estimated mean frequency for complete case
Time period (months)[b]	**0-12**	**12-18**	**18-24**	**24-36**	**36-48**	**48-60**		**0-12**	**12-18**	**18-24**	**24-36**	**36-48**	**48-60**		**0-12**	**12-18**	**18-24**	**24-36**	**36-48**	**48-60**	
Effective sample size, n[c]	**106.5**	**97.5**	**83**	**63**	**33.5**	**10**		**105**	**96**	**80.5**	**60**	**31**	**10**		**211.5**	**193.5**	**163.5**	**123**	**64.5**	**20**	
Outpatient visits	11.31	2.47	2.90	2.97	1.28	0.80	**21.73**	11.29	2.06	1.45	1.75	1.16	0.00	**17.71**	11.30	2.27	2.19	2.37	1.22	0.40	**19.75**
Inpatient stay for any cause (no of bed days)	13.28	2.91	1.80	3.05	0.57	0.00	**21.60**	18.00	2.02	3.23	3.47	3.16	0.00	**29.88**	15.62	2.47	2.50	3.25	1.81	0.00	**25.66**
EUS as day case (no. of day cases)	0.90	0.00	0.00	0.00	0.00	0.00	**0.90**	0.03	0.00	0.00	0.00	0.00	0.00	**0.03**	0.47	0.00	0.00	0.00	0.00	0.00	**0.47**
Surgery (count, no. of bed days)	0.61, 13.68	0.00	0.00	0.00	0.00	0.00	**0.61, 13.68**	0.62, 15.74	0.00	0.00	0.00	0.00	0.00	**0.62, 15.74**	0.61, 14.70	0.00	0.00	0.00	0.00	0.00	**0.61, 14.70**
EMR as day case	0.08	0.02	0.01	0.00	0.00	0.00	**0.11**	0.08	0.01	0.00	0.00	0.00	0.00	**0.09**	0.08	0.02	0.01	0.00	0.00	0.00	**0.10**
EMR as inpatient (count, no. of bed days)	0.03, 0.06	0.00	0.00	0.00	0.00	0.00	**0.03, 0.06**	0.01, 0.03	0.00	0.00	0.00	0.00	0.00	**0.01, 0.03**	0.02, 0.04	0.00	0.00	0.00	0.00	0.00	**0.02, 0.04**
Chemotherapy (no. of cycles)	2.24	0.25	0.35	0.33	0.00	0.00	**3.17**	2.82	0.14	0.26	0.13	0.13	0.00	**3.48**	2.53	0.19	0.31	0.24	0.06	0.00	**3.32**
Radiotherapy (no. of fractions)	5.44	0.35	0.48	0.17	0.00	0.00	**6.44**	5.35	0.54	0.17	0.43	0.00	0.00	**6.50**	5.39	0.44	0.33	0.30	0.00	0.00	**6.47**
Tests and investigations (count)	66.03	8.69	7.41	7.63	2.90	4.60	**97.26**	78.00	7.34	7.48	11.32	9.32	0.00	**113.46**	71.97	8.02	7.44	9.43	5.98	2.30	**105.15**
PAMS	11.65	0.56	0.13	0.56	0.15	0.00	**13.05**	17.45	0.32	0.84	0.28	0.00	0.00	**18.90**	14.53	0.44	0.48	0.42	0.08	0.00	**15.96**
Hospital prescribed drugs (no. of items) [d]	28.93	5.26	4.67	5.40	3.01	5.90	**53.18**	31.79	4.31	4.09	3.70	4.29	2.40	**50.58**	30.35	4.79	4.39	4.57	3.63	4.15	**51.87**

a Full table is in Appendix 2.1.

b 0-12 months is the first year after randomisation; 12-18 months is the interval between 12 and 18months after randomisation; etc.

c Sample sizes include only participants still in the trial, namely survivors observed throughout the interval, and those who died before the end of the interval but had been randomised early enough for that interval to end before 31 July 2009. We included participants (dead or alive) for whom the trial ended during the interval, but gave them half the weight of those in the trial throughout the interval.

d Excluding drugs prescribed for surgery or chemotherapy.

These show that total inpatient stays were considerably shorter in the EUS group (21.6 days) than in the control group (29.9 days). Almost all these stays were cancer-related, which accounted for 20.8 days (96.3% of 21.6 days) and 29.5 days (98.7% of 29.9 days) respectively. In contrast, outpatient visits over 48 months were more frequent in the EUS group than the control group (21.7 visits versus 17.7 visits), perhaps because of longer survival.

2.4.4 Costs over 48 months

Guided by RTE, I estimated participants' costs in each time interval and resource category by multiplying individual resource use in Table 2.3 by the corresponding unit cost from published national sources and discounting costs beyond 12 months at 3.5% per annum. Table 2.4 shows mean values of the costs of selected types of service use (full table is presented in Appendix 2.2). The sparsity of these data meant that, like frequencies, DR could not impute them beyond the end of the trial.

However, costs summed over all resource types within each interval were less skewed, and could therefore be imputed for individual participants. So estimates were summed across categories to yield the total cost of each interval when participants remained in the trial; imputed costs for later intervals; summed observed and imputed costs over 48 months to yield a total cost for each participant; and calculated mean total cost per participant for intervention and control groups. Table 2.2a summarises the fully imputed total cost per participant by allocated group over 48 months for the primary cost-effectiveness analysis. Over 48 months, the mean cost of secondary care including hospital prescribed drugs was £29,190 (SD £14,902) (2019 prices: £35,619, SD £18,184) in the intervention group and £32,049 (SD £22,019) (2019 prices: £39,108, SD £26,869) for controls, including the cost of endoscopic ultrasound scans. Thus, the EUS group cost £2,860 less (95% bootstrapped CI from –£7,987 to £2,192) (2019 prices: £3,490 less, 95% bootstrapped CI from -£9,746 to £2,675) on average than the control group, mainly because participants in the EUS group spent fewer days as inpatients. Although not statistically significant, this difference is 9% of the mean total cost for controls over 48 months.

Table 2.4: Estimated mean cost of secondary healthcare including hospital-prescribed drugs by 213 participants over 60 months[a].

	Intervention group (Estimated mean cost (£) in each interval)						Estimated mean cost (£) for complete case	Control group (Estimated mean cost (£) in each interval)						Estimated mean cost (£) for complete case	Total sample (Estimated mean cost (£) in each interval)						Estimated mean cost (£) for complete case
Time period (months)[b]	**0-12**	**12-18**	**18-24**	**24-36**	**36-48**	**48-60**		**0-12**	**12-18**	**18-24**	**24-36**	**36-48**	**48-60**		**0-12**	**12-18**	**18-24**	**24-36**	**36-48**	**48-60**	
Effective sample size, n[c]	**106.5**	**97.5**	**83**	**63**	**33.5**	**10**		**105**	**96**	**80.5**	**60**	**31**	**10**		**211.5**	**193.5**	**163.5**	**123**	**64.5**	**20**	
Outpatient visits	1290	262	324	318	133	70	**2397**	1317	226	165	187	113	0	**2007**	1303	244	246	254	123	35	**2205**
Inpatient stay for any cause	6179	1060	741	1117	239	0	**9336**	8228	791	1397	1296	1556	0	**13268**	7196	927	1064	1205	872	0	**11263**
EUS as day case	497	0	0	0	0	0	**497**	16	0	0	0	0	0	**16**	258	0	0	0	0	0	**258**
Surgery	8226	0	0	0	0	0	**8226**	8976	0	0	0	0	0	**8976**	8598	0	0	0	0	0	**8598**
EMR as day case	39	10	6	0	0	0	**56**	40	5	0	0	0	0	**45**	40	8	3	0	0	0	**51**
EMR as inpatient	25	0	0	0	0	0	**25**	13	0	0	0	0	0	**13**	19	0	0	0	0	0	**19**
Chemotherapy	1278	172	250	230	0	0	**1929**	1613	60	170	82	76	0	**2001**	1445	116	211	158	36	0	**1965**
Radiotherapy	712	40	62	22	0	0	**836**	660	69	24	64	0	0	**816**	686	55	43	42	0	0	**826**
Other procedures	1189	306	166	208	0	0	1869	1105	103	121	368	215	0	1912	1147	205	143	287	104	0	1885
Tests and investigations	1003	103	124	97	28	44	**1399**	1041	109	126	132	77	0	**1485**	1022	106	125	114	52	22	**1440**
PAMS	274	13	10	46	2	0	**345**	434	28	27	13	0	0	**502**	354	20	19	30	1	0	**423**
Total 2ndary care cost	**20714**	**1966**	**1683**	**2037**	**402**	**114**	**26916**	**23441**	**1392**	**2030**	**2141**	**2037**	**0**	**31041**	**22068**	**1681**	**1854**	**2088**	**1188**	**57**	**28935**
Hospital prescribed drugs[d]	997	178	162	379	269	227	**2212**	1816	201	165	236	310	28	**2756**	1403	189	163	310	289	128	**2482**
Total cost (secondary care+ hospital-prescribed drugs)	**21710**	**2144**	**1844**	**2417**	**671**	**341**	**29128**	**25257**	**1592**	**2195**	**2378**	**2346**	**28**	**33797**	**23471**	**1870**	**2017**	**2398**	**1476**	**185**	**31417**

EMR, endometrial mucosal resection; PAMs, Professions Allied to Medicine.

a Full table is in Appendix 2.2.

b 0-12 months is the first year after randomisation; 12-18 months is the period between 12 and 18months after randomisation; etc.

c Sample sizes include only participants still in the trial: survivors observed throughout the interval; and those who died before the end of the interval but had been randomised early enough for that interval to end before 31 July 2009. Participants (dead or alive) for whom the trial ended during the interval were included, but gave them half the weight of those in the trial throughout the interval.

d Excluding drugs prescribed for surgery or chemotherapy.

2.4.5 Primary cost-effectiveness analysis at 48 months

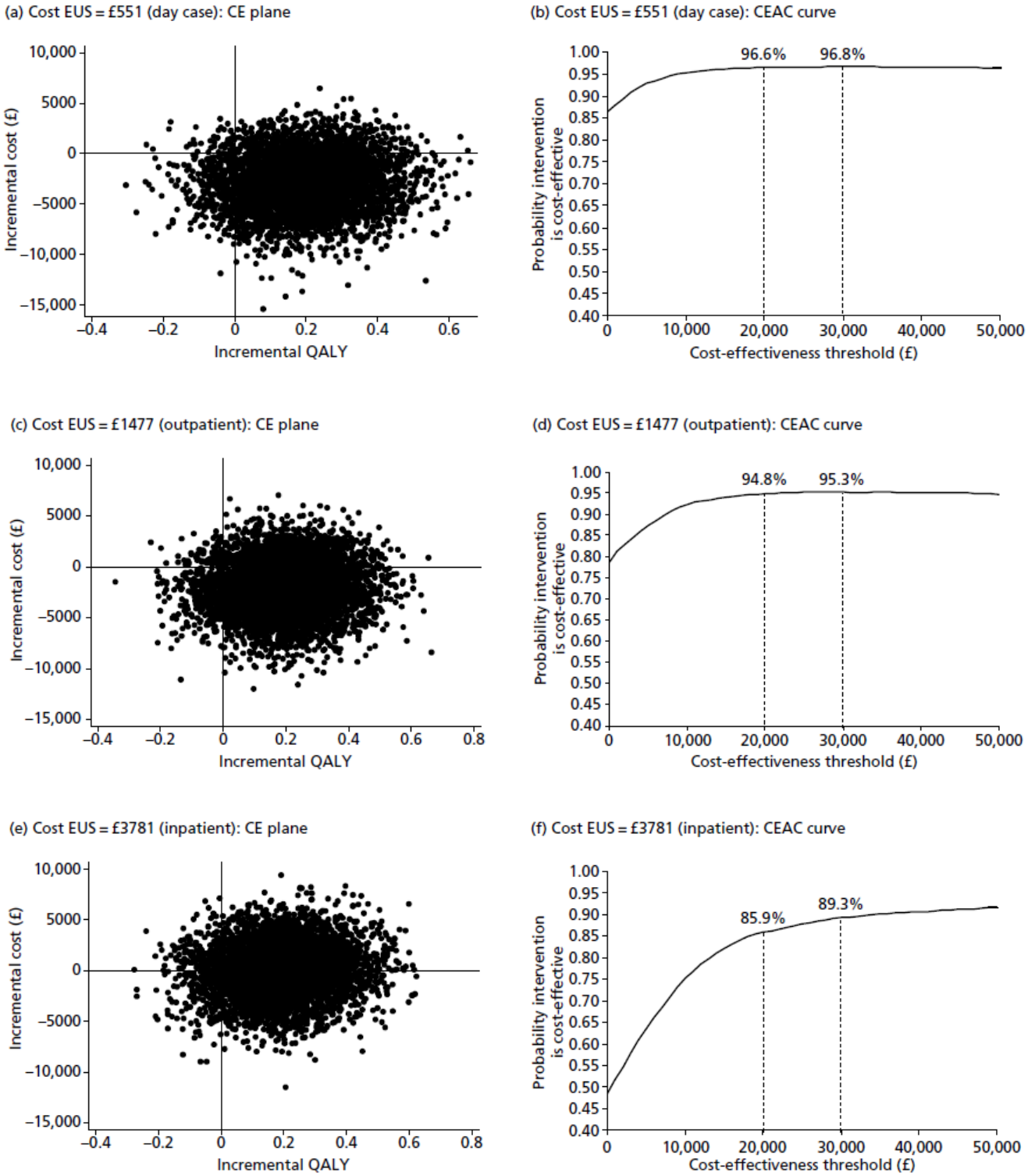
In summary, we estimated that by 48 months EUS participants had on average gained 0.197 more QALYs and cost £2,860 less (2019 prices: £3,490 less) than controls. Although neither difference alone is significant (Table 2.2a), cost-effectiveness analysis considers both in combination. Figure 2.1a is the cost-effectiveness plane showing the joint distribution of costs and effects (in QALYs) arising from 5000 replicates each yielding one point. There are 3988 (79.8%) points in the South-East quadrant ('win-win'); 326 (6.5%) in the South-West (net costs and QALYs both negative); 640 (12.8%) in the North-East (net costs and QALYs both positive); and 46 (0.9%) in the North-West ('lose-lose'). Figure 2.1b is the corresponding CEAC showing the probability that EUS is more cost-effective than usual care (equivalent to the proportion of points that achieve cost-effectiveness) for a range of cost-effectiveness thresholds that decision-makers are willing to pay to gain one QALY. The curve intercepts the y-axis in Figure 2.1b at 0.863, meaning that, even if decision-makers put no value on an extra QALY, endoscopic ultrasound has 86.3% probability of being cost-effective (the proportion of points below the x-axis in Figure 2.1a). When the threshold rises to the lower and upper NICE criteria of £20,000 and £30,000 per QALY, this probability rises to 96.6% and 96.8% respectively (see Figure 2.1b). And when the threshold is infinite, it falls to 92.6% (the proportion of points to the right of the y-axis in Figure 2.1a). Hence, we can be reasonably confident that EUS is cost-effective in managing GOC. However, the saving of nearly £3,000 (2019 prices: saving of nearly £3,500) depends on the cost of EUS. So, we need sensitivity analysis to assess how dependent it is.

2.4.6 Sensitivity analysis with different costs of EUS

The Department of Health offers three different unit costs for EUS – £551 for day patients, £1,477 for outpatients and £3,781 for inpatients (Department of Health, 2009). To find out how that choice affects our cost-effectiveness findings, I repeated our analysis using all three estimates. Our primary analysis used the day cost of £551 for two reasons: firstly, most trial participants received their EUS scans as day patients; and secondly our confirmatory costing, based on detailed analysis of the staffing and time needed to deliver EUS to trial participants as day patients, was close to £500.

Figures 2.1c and 2.1d use the unit cost of £1,477 for receiving EUS scans as outpatients, reducing the mean cost saving at 48 months to £2,055 (2019 prices: £2,508). Figure 2.1d shows that this slightly reduces the probability that EUS is cost-effective – to 78.7% at a threshold of zero, 94.8% at £20,000 and 95.3% at £30,000; however, it is still 92.6% at a threshold of infinity.

Figure 2.1: Cost effectiveness planes and acceptability curves for discounted QALYs at 48 months by unit cost of endoscopic ultrasound.



Figures 2.1e and 2.1f use the unit cost of £3,781 for receiving EUS scans as inpatients, reducing the mean cost saving to £53 (2019 price: £65). Figure 2.1f shows that this further reduces the probability that EUS is cost-effective – to 48.5% at a threshold of zero, 85.9% at £20,000 and 89.3% at £30,000;

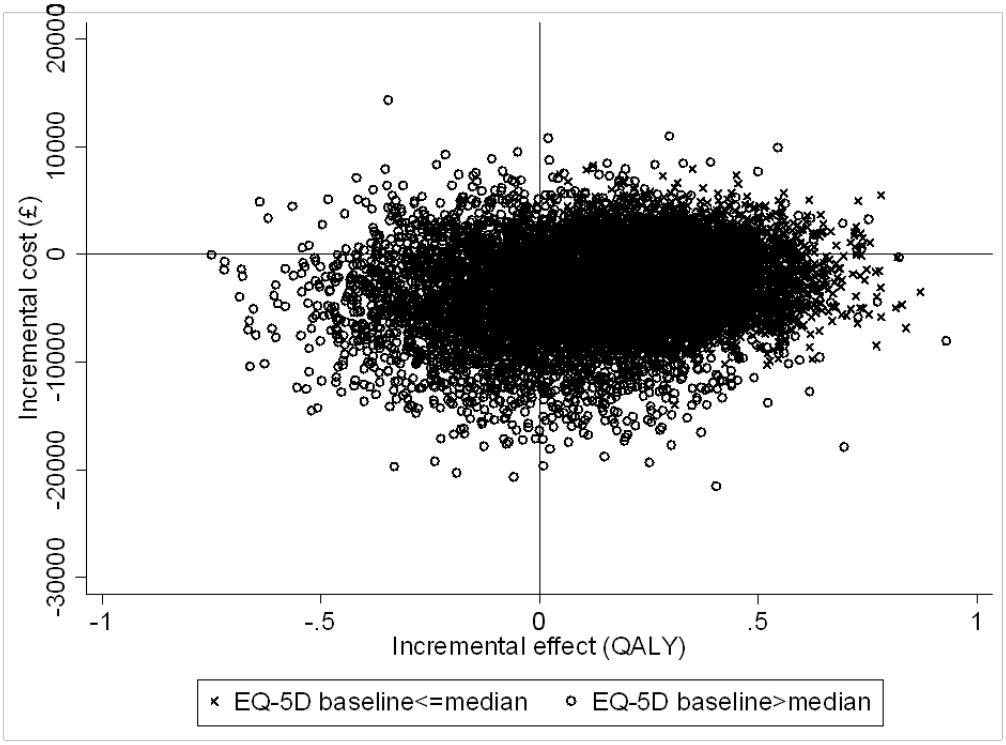
however, it is still 92.6% at a threshold of infinity. Not surprisingly, as the unit cost increases, the probability that EUS is cost-effective at 48 months decreases. As the threshold increases to infinity, where QALYs not costs affect the probability, all three probabilities converge on 92.6%. Thus, our sensitivity analysis confirms that EUS is probably cost-effective at 48 months; for the NICE thresholds range of £20,000 to £30,000 per QALY, that probability ranges from 86% when EUS costs nearly £4,000 to 97% when it costs slightly more than £500.

2.4.7 Subgroup analysis

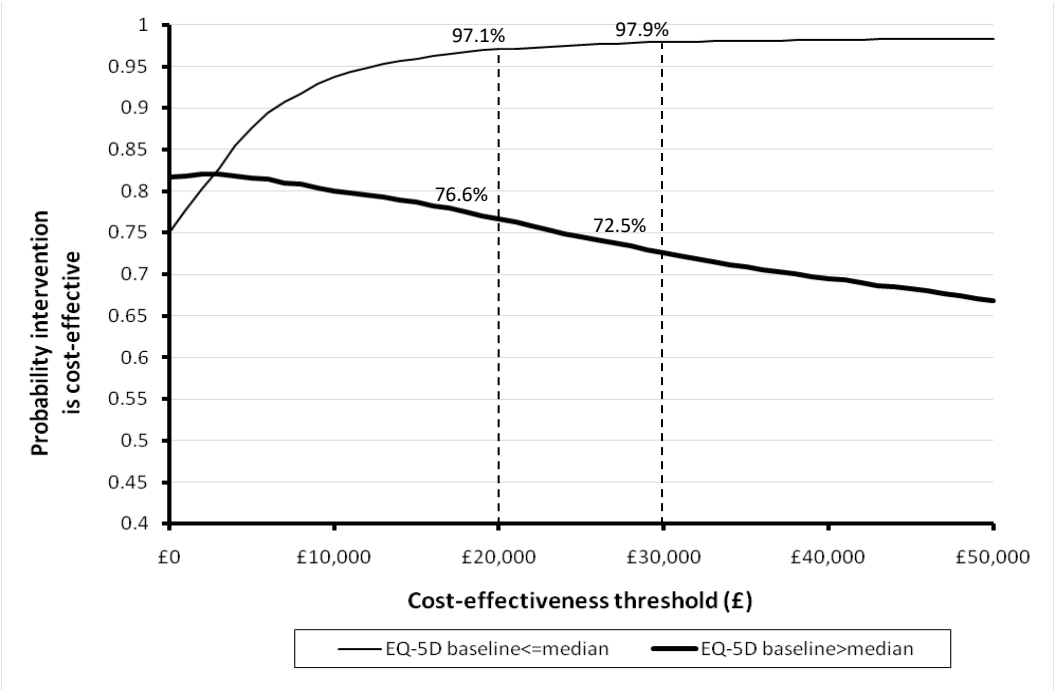
Though subgroup analysis reduces sample sizes and statistical power, we responded to the trial finding that EUS is more effective for participants with lower self-reported health related quality of life (HRQoL), those who reported worse initial HRQoL at baseline (Russell et al., 2013). To maximise statistical power, it was agreed to split the effective trial population of 213 at the median baseline EQ-5D utility score of 0.796 into two equally sized groups (below and above that median baseline EQ-5D utility score). This divided the intervention group into 54 above that median and 53 below, and the control group into 47 above that median and 59 below. Participants with lower self-reported HRQoL at baseline had a much higher mean QALY gain (0.307) than participants with higher self-reported HRQoL at baseline (0.0183) but a smaller saving of £1,918 (2019 prices: saving of £2,340) compared with £4,259 (2019 prices: saving of £5,197). Combining these effectiveness and cost findings by bootstrapping, Figure 2.2b shows that, despite the tendency of smaller samples to reduce the probabilities of the previous analyses, EUS in participants with lower self-reported HRQoL at baseline has 97.1% probability of being cost-effective at a threshold of £20,000 per QALY, 97.9% at £30,000 and 97.9% at infinity; whereas in participants with higher self-reported HRQoL at baseline, these probabilities fall to 76.6%, 72.5% and 53.3% respectively. For participants with lower self-reported HRQoL at baseline, therefore, EUS is both significantly more effective (Russell et al., 2013) and substantially more cost-effective.

Figure 2.2: Cost-effectiveness plane and acceptability curve for discounted QALYs at 48 months by whether baseline EQ-5D score below or above median.

(a) Cost-effectiveness plane



(b) CEAC curve



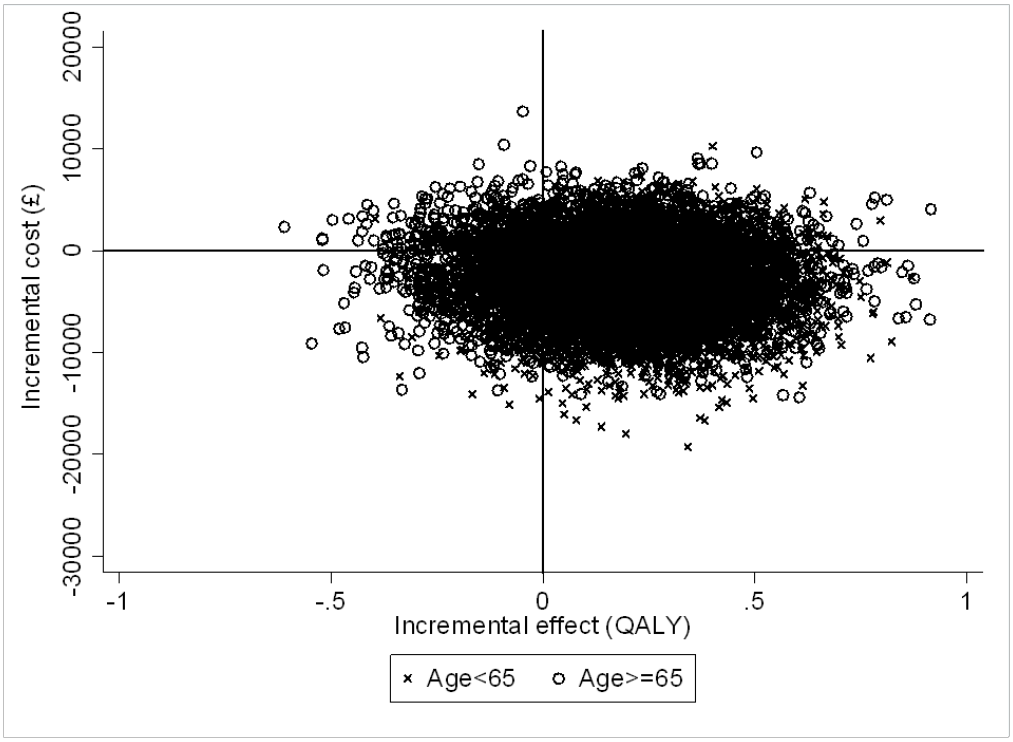
In addition, with guidance from DR, RTE and IR, I conducted a second subgroup analysis, by splitting the trial sample into subgroups at median age of 65 years; this divided the intervention group into 50 participants aged 65 years and above, and 57 participants aged below 65 years, and the control group into 50 participants aged 65 years and above, and 56 participants aged below 65 years. Younger participants had a higher mean QALY gain (0.2350) than older participants (0.1528) and a larger mean cost saving (£3,454 vs £2,246) (2019 prices: £4,215 vs £2,741).

Bootstrapping shows that EUS in younger participants has 94.5% and 94.9% probability of being cost-effective at the NICE thresholds of £20,000 and £30,000 per QALY, respectively; in older participants these probabilities are 82.3% and 81.6%, respectively (Figure 2.3b).

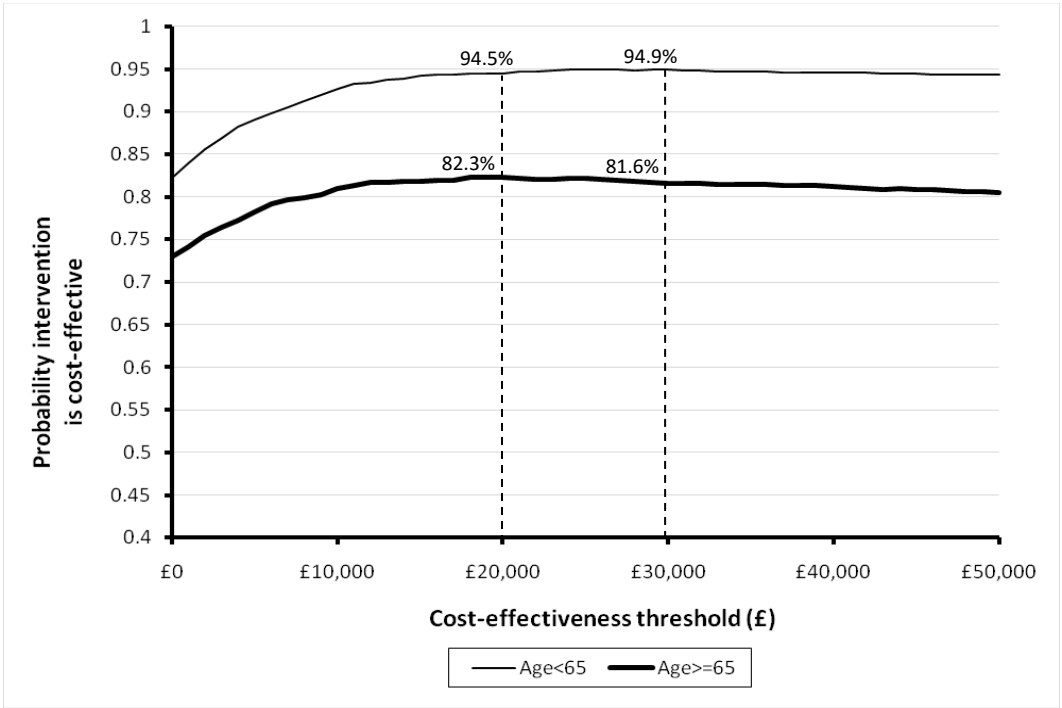
Comparing these two subgroup analyses, we see that the CEACs in Figure 2.2b are farther apart than those in Figure 2.3b over thresholds between £20,000 and £30,000. Even with a reduced sample size, the probability of EUS being cost-effectiveness between £20,000 and £30,000 is much higher in patients with lower self-reported HRQoL at baseline than in the whole sample. To explain why the differential effect of sickness does not translate into a differential effect of age, we observe that the proportions of patients with lower self-reported HRQoL at baseline among those under and over 65 years were surprisingly similar – 52% and 53% respectively. In short, baseline health status appears to be a better predictor than age of whether EUS is beneficial in gastro-oesophageal cancer.

Figure 2.3: Cost-effectiveness plane and acceptability curve for discounted QALYs at 48 months by whether age below or above median of 65 years

(a) Cost-effectiveness plane



(b) CEAC curve



2.5 Discussion

2.5.1 Principal findings

As described in the published COGNATE trial report, Table 28 (see Appendix 2.3) and Table 32 (see Appendix 2.4) (Russell et al., 2013), EUS group had higher surgery avoided cases (44.4%, 8 out of 18 surgery avoided cases) compared to non EUS group (10.0%, 1 out of 10 surgery avoided case). Additionally, there were considerably shorter hospital inpatient stays for the participants in the EUS group (21.6 days) compared to the non-EUS group (29.9 days), contributing to a lower mean total cost per participant in the EUS group compared to the non EUS group (EUS £29,190 vs non EUS £32,049). This suggests, at 48 months after randomisation, with costs discounted at 3.5%, EUS had saved costs (£2,860, but not significantly – see Table 2.2a) and gained QALYs (0.1969, but not significantly – see Table 2.2a). Combining these findings, it can be concluded that EUS is significantly cost-effective, in the sense that the probability of EUS being cost-effective was reached and exceeded 95% at the NICE threshold of £20,000 to £30,000 per QALY when the cost of a scan is set at the national unit cost of £551 for day-case and £1,477 for outpatient. However, the probability of cost-effectiveness falls below 95% at the highest national unit cost of £3,781 for inpatient scans. It was judged, that a cost close to £500 is much more plausible because most trial participants received their scans as day cases and, more importantly, our own detailed analysis of the staffing and time needed to deliver EUS to trial participants as day cases was close to £500.

Though the COGNATE clinical and economic analysis (see Figures 2.1(b)) provide strong evidence that EUS is likely to be cost-effective in gastro-oesophageal cancer, we ask whether some participants benefit more than others. The effectiveness analysis addressed this question by adding covariates to primary analysis. Table 24 in the published COGNATE report (Russell et al., 2013, page 53), used Cox regression, showing that participants reporting poorer health at baseline gained significantly more QALYs over a range of follow-up periods averaging 24 months than those reporting better health at baseline. We found a similar pattern, although less significant, in fully imputed 48-month QALYs (see Effectiveness, Quality-adjusted survival: sensitivity analyses for fully imputed data, and Appendix 6.18 presented in

the published COGNATE report (Russell et al., 2013, page 150)). Whereas, the economic analysis addressed this question by performing two exploratory subgroup analyses - (1). To maximise statistical power, it was agreed to split the effective trial population of 213 at the median baseline EQ-5D utility score of 0.796, which this divided the trial population into two groups (below and above that median baseline EQ-5D utility score); and (2). We split the trial population of 213 at 65 years, close to the median, which this divided the trial population into two groups (below and 65 years and above). Exploratory subgroup analysis at 48 months suggests that the participants reporting poorer health at baseline saved fewer costs than the participants reporting better health at baseline (but far from significantly – see Figure 2.2(b)), and had a much higher QALY gain. Combining these findings suggests that EUS is probably more cost-effective for patients reporting poorer health at baseline. Exploratory subgroup analysis also hints that EUS could be slightly better for younger patients. However these weak subgroup analyses do not detract from the unequivocal finding that EUS is almost certainly cost-effective for gastro-oesophageal cancer. In short, baseline health status appears to be a better predictor than age of whether EUS is beneficial in gastro-oesophageal cancer.

In summary, the COGNATE trial and economic evaluation showed several substantial benefits of EUS:

A. EUS participants gained an average of 0.197 more QALYs than controls;
B. EUS participants cost the NHS an average of £2,860 less (2019prices: £3,490 less) than controls as a result of fewer or shorter hospital stays;
C. Though neither benefit A nor B was statistically significant on its own, economic analysis showed that together they conclude that EUS has 96.6% probability of being cost-effective at the lower NICE threshold of £20,000 per QALY;
D. Sensitivity analysis showed that for the NICE thresholds range of £20,000 to £30,000 per QALY, probability of EUS being cost-effective ranges from 86% when EUS costs nearly £4,000 to 97% when it costs slightly more than £500; and
E. EUS participants with initial quality of life worse than average gained an average of 0.307 more QALYs than their controls – significantly greater than the average net gain of 0.0183 QALYs by those EUS participants with initial quality of life

better than average. Despite the tendency of smaller samples to reduce the probabilities of the previous analyses, EUS in participants with lower self-reported HRQoL at baseline had 97.1% and 97.9% probability of being cost-effective at the NICE thresholds of £20,000 and £30,000 per QALY, respectively; whereas in participants with higher self-reported HRQoL at baseline, these probabilities fell to 76.6% and 72.5%, respectively.

F. Younger participants had a higher mean QALY gain (0.2350) than older participants (0.1528) and a larger mean cost saving (£3,454 vs £2,246) (2019 prices: £4,215 vs £2,741). EUS in younger participants had 94.5% and 94.9% probability of being cost-effective at the NICE thresholds of £20,000 and £30,000 per QALY, respectively; whereas in older participants these probabilities were 82.3% and 81.6%, respectively.

2.5.2 Comparison with other studies

No direct comparisons could be made as, to date (as of 2021), this is the first study of its kind in this area in the evaluation of the economic values of EUS in managing patients with GOC. Furthermore, this economic evaluation of EUS in GOC was carried out using primary economic data collected prospectively alongside a pragmatic randomised controlled trial (namely the COGNATE trial); hence this economic study offers novel contributions to the economic evidence of EUS in the management of GOC patients as well as to the health technology assessments in the UK.

2.5.3 Strengths and limitations of the study

The economic analysis was conducted and reported in accordance with the published standard operating procedure for economic evaluation alongside randomised controlled trials (Edwards et al., 2008), published guidance (Ramsey et al., 2005; Glick et al., 2014; Drummond et al., 2015) and published CHEERS checklist (Husereau et al., 2013).

The main weakness of the COGNATE trial is that it was some 10 years too late (Russell et al., 2013), committed to the practice of evidence-based medicine, however, we achieved high standards of recruitment, retention and data collection, and created a

data set of high quality in a population with low self-reported HRQoL. That overall COGNATE trial, including its economic evaluation study, achieved largely unequivocal findings despite recruiting only a minimal sample is a tribute to its rigour. The economic evaluation study of the COGNATE trial had the same sample size as for the effectiveness study (n=213).

The SPSS MVA procedure performed by the COGNATE Trial Statistician (DR) uses single imputation for missing data. As mentioned earlier, the SPSS MVA procedure used by DR simultaneously estimates all missing values in a data set, on one or more data sets. The analysis of the trial was undertaken in 2010 when multiple imputation was becoming more common place (White et al., 2010). It is likely that if this analysis was undertaken months later, then it is likely multiple imputation would have been used as it has become the most widely used method of imputation for missing data in trial-based cost-effectiveness analysis as reported in the review undertaken by Leurent et al. (2018). It is acknowledged that single imputations are likely to underestimate the uncertainty of the estimates and produce overly precise results (Kang, 2013).

The COGNATE trial was designed in response to a 2001 HTA commissioning brief which referenced a 1998 systematic review by Harris et al. (Harris et al., 1998). The case for EUS as an adjunct to the usual staging test for gastro-oesophageal cancer was not proven. However, due to a variety of circumstances beyond the control of the HTA and the COGNATE team recruitment to the trial did not begin until 7 years after the publication of the systematic review. By which point, treatment pathways had progressed and EUS was being used by many clinicians who did not wish to forgo EUS for any patients. This lack of equipoise for many reduced the potential sample of sites that could be used for recruitment. Two centres – Aberdeen and Gloucester did remain in equipoise and recruited effectively over a period of 3.5 years while a remaining six centres allowed individual clinicians in equipoise to recruit where possible resulting in effectively three centres delivering the recruitment intended to be completed across ten centres.

Given that many centres had already adopted EUS, the primary outcome of survival was replaced by quality-adjusted survival, often called quality-adjusted life-years, for two reasons: (1) this reduced the target sample size from 700 to 400; and (2) QALYs is the criterion preferred by NICE. As the number of active recruiting centres fell, the target effect size for both survival and quality of life was increased from 0.3 to 0.4 enabling a reduction in the target sample size from 400 to 220 and the recruitment period extended by 6 months to achieve this.

The main weakness of the COGNATE trial was that it was 10 years too late, however, committed to the practice of evidence-based medicine, the trial was delivered in the most pragmatic way in difficult circumstances to undertake a rigorous evaluation.

Conducting a systematic review of economic evidence of EUS in the management of GOC would be beneficial in designing the COGNATE trial; however, due to time constraints prior to the trial design, we could only conduct a literature review.

NICE (2013) recommended the use of QALY, measured by EQ-5D, the preferred generic preference-based health-related quality of life instrument, for cost-effectiveness analysis. Complying with the NICE's recommendations, in the COGNATE trial, we conducted cost-effectiveness analysis using QALY, solely measured by EQ-5D, as the measure of effect though studies have argued that EQ-5D may not be sensitive enough to pick up changes in patients' disease-/condition-specific related quality of life (Pennington et al., 2020; Wichmann et al., 2017; Pettitt et al., 2016; Tosh et al., 2012; Whitehead and Ali, 2010).

2.5.4 Future research

Upon completion of the COGNATE trial in 2013, comprehensive economic-specific systematic review of EUS in GOC staging remains lacking, with the exception of studies by Harris and colleagues (Harris et al. 1998). However, the systematic review by Harris et al. (1998) was almost exclusively covered the clinical evidence of EUS in GOC staging. Therefore, future work on conducting a systematic review of economic evidence of EUS in GOC staging is recommended to fill the gap in the economic

literature in this area and offer policy makers, commissioners and researchers an insight into the up-to-date findings of economic evidence in this field. Furthermore, given the lack of sensitivity and responsiveness of EQ-5D in picking up changes in patients' disease-/condition-specific related quality of life, further study on methodological advancement of cost-utility analysis is recommended to explore the modification of QALY that would help to better reflect the 'real' health status of patients, for use in cost-effectiveness/cost-utility analysis.

2.6 Conclusion

The economic evaluation findings of the COGNATE trial provide convincing evidence that the use of endoscopic ultrasound is cost-effective in managing patients with GOC. The improvement in quality-adjusted life years (especially for participants initially less healthy); and the consistent reductions in resource use, especially the duration of hospital stays – all resulting in 97% probability that EUS is cost-effective – is conclusive. Hence, we believe that policy makers and commissioners of NHS cancer services can be confident that EUS contributes much to the effective and efficient management of gastro-oesophageal cancer. We therefore recommend the provision of EUS scans for all patients with the capacity to benefit.

2.7 Novel contributions

This is the first economic evaluation study of EUS in GOC staging conducted alongside a randomised controlled trial (COGNATE trial) in the UK in the field of diagnosis and treatment planning of GOC. The economic findings of the COGNATE trial add to the literatures of EUS staging in GOC. This chapter offers a novel opportunity to inform evidence-based health care policy decision-making on the use of EUS in GOC staging within the finite public resources.

2.8 How does the COGNATE trial develop the next few chapters of this thesis (Chapter 3, 4, and 5)?

There is a need to fill the literature gap of the economic evidence of endoscopic ultrasound in gastro-oesophageal cancer staging given the only published systematic review (by Harris et al. (1998)) was more than two decades ago and it is not an

economic-specific systematic review. Therefore, following the COGNATE trial, a systematic review in this field was conducted retrospectively to identify, appraise, and explore the existing economic literatures on endoscopic ultrasound in managing patients with gastro-oesophageal cancer. The systematic review is presented and discussed in detail in Chapter 3.

As the outcomes of the COGNATE trial have shown that EUS appears to be cost-effective, it is of great interests to study the utilisation of EUS in gastro-oesophageal cancer staging across the NHS hospitals in the UK. Therefore, a survey on the use of EUS in GOC staging in the UK was conducted. The findings of the survey are presented and discussed in Chapter 4 which give a comprehensive understanding on the utilisation of EUS in GOC staging and the current practice in the UK.

In the cost-effectiveness analysis of EUS performed in the COGNATE trial, QALY was measured solely using EQ-5D tool. It is widely perceived and argued that EQ-5D is not a faultless tool in measuring QALY (Pettitt et al., 2016; Tosh et al., 2012; Whitehead and Ali, 2010). In Chapter 5, a novel method of generating regression models between conventional QALYs (measured purely by EQ-5D) and disease-specific measures to convert disease-specific measure scores into utilities (i.e. disease specific measures-guided QALY or hybrid QALY) on the original EQ-5D scale was explored. Datasets from two large clinical randomised controlled trials – COGNATE trial and CLARITY trial – were used to demonstrate the novel methodological approaches utilised in this exploratory study.

Chapter 3: Endoscopic ultrasound staging in patients with gastro-oesophageal cancers: a systematic review of economic evidence

An edited version of this chapter is published in BMC Cancer Journal (see Appendix 3.23):

Yeo, S. T., Bray, N, Haboubi, H., Hoare, Z. and Edwards, R. T. (2019). Endoscopic ultrasound staging in patients with gastro-oesophageal cancers: a systematic review of economic evidence. *BMC Cancer,* 19, 900. doi: https://doi.org/10.1186/s12885-019-6116-0

3.1 Chapter Summary

The sensitivity of endoscopic ultrasound (EUS) in staging gastro-oesophageal cancers (GOCs) has been widely studied. However, the economic evidence of EUS staging in the management of patients with GOCs is scarce. This review aimed to examine all economic evidence (not limited to randomised controlled trials) of the use of EUS staging in the management of patients with GOCs, and to offer a review of economic evidence on the costs, benefits (in terms of GOC patients' health-related quality of life), and economic implications of the use of EUS in staging GOC patients.

The protocol was registered prospectively with PROSPERO (CRD42016043700; http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42016043700). MEDLINE (ovid), EMBASE (ovid), The Cochrane Collaboration Register and Library (including the British National Health Service Economic Evaluation Database), CINAHL (EBSCOhost) and Web of Science (Core Collection) as well as reference lists were systematically searched for studies conducted between 1996-2018 (search update 28/04/2018). All the searches were again updated to cover studies between 2018-2021 (search update 13/12/2021). Two authors independently screened the identified articles, assessed study quality, and extracted data. Study characteristics of the included articles, including incremental cost-effectiveness ratios, when available, were summarised narratively.

Of the 197 articles retrieved, 6 studies met the inclusion criteria: 3 were economic studies and another 3 were economic modelling studies. Of the 3 economic studies, one was an incremental cost-utility analysis and two were cost analyses. Of the 3 economic modelling studies, one was an incremental cost-utility analysis and two were cost-minimisation

analyses. Both of the incremental cost-utility analyses reported that use of EUS as an additional staging technique provided, on average, more QALYs (0.0019-0.1969 more QALYs) and saved costs (by £1,969-£3,364 per patient, 2017 price year) compared to staging strategy without EUS.

The data available suggest use of EUS as a complementary staging technique to other staging techniques for GOCs appears to be cost saving and offers greater QALYs. Nevertheless, future studies are necessary because the economic evidence around this EUS staging intervention for GOCs is far from robust. More health economic research and good quality data are needed to judge the economic benefits of EUS staging for GOCs.

3.2 Introduction

Gastro-oesophageal (oesophageal or gastric, or both) cancers (GOCs) are one of the most common cancers in the UK with approximately 16,000 new diagnoses each year (Cancer Research United Kingdom [CRUK], 2021a & 2021b). Oesophageal and gastric cancers were the seventh and fourteenth most common cause of cancer death respectively in the UK in 2016, as shown from the latest available statistics reported by the CRUK (CRUK, 2021a & 2021b). It is estimated that a total of around 12,330 people died from these cancers – that is 34 deaths per day (CRUK, 2021a & 2021b). Accurate staging of GOCs is vital for determining prognosis and planning appropriate treatment. Accurate staging in the management of GOCs will not only help avoid unnecessary surgical interventions but also will ultimately help reduce the financial pressure on the NHS, which is particularly important given the limited resources available to cancer services and the growing incidence of GOCs (Dubecz et al., 2013).

Accurate staging of GOCs can be achieved by a combination of investigative techniques. The techniques used for staging GOC include computer tomography (CT), endoscopic ultrasound (EUS), positron emission tomography (PET) and adjuncts to staging include magnetic resonance imaging (MRI), bronchoscopy, laparoscopy and trans-abdominal ultrasound (Thakkar and Kaul, 2020; Valero and Robles-Medranda, 2017; Allum et al., 2011; Allum et al., 2002). CT has been recommended for use at initial staging assessment to determine whether the cancer cells have spread from the primary site of its origin into new areas of the body (i.e. metastasis), but in the absence of metastatic disease, endoscopic ultrasound (EUS) has been advocated as the preferred technique for the assessment and prediction of operability (Thakkar and Kaul, 2020; DaVee et al., 2017; Valero and Robles-Medranda, 2017; Allum et al., 2011; Allum et al., 2002). This is due to the fact that EUS is superior to CT for locoregional staging of oesophageal and gastric tumours (Thakkar and Kaul, 2020; Valero and Robles-Medranda, 2017; Allum et al., 2011; Allum et al., 2002).

Studies and guidelines for the management of oesophageal and gastric cancer have reported that EUS has superior tumour invasion (T) and locoregional nodal (N) staging ability over CT and PET given its sensitivity, particularly for detection of regional lymph node metastases, although the complementary nature of these investigative techniques must be recognised (Thakkar and Kaul, 2020; Valero and Robles-Medranda, 2017; HQIP, 2016; Allum et al., 2011;

Puli et al., 2008; van Vliet et al., 2008; Lowe et al., 2005; Botet et al., 1991a; Vilgrain et al., 1990). The sensitivity of EUS for staging of GOC has been widely evaluated; however, the economic evidence of EUS staging in the management of GOC patients is scarce. Furthermore, the effectiveness and cost-effectiveness of EUS staging of GOC had not been assessed, particularly in the form of randomised controlled trials (RCT), until the establishment of "COGNATE" trial - a HTA-funded RCT UK study (Russell et al., 2013).

Given that the economic evidence of EUS for staging of GOC is scant, conducting a systematic review of the economic evidence on EUS staging in patients with GOC is therefore important. It not only gives a meaningful evidence-based insight, from an economic perspective, for researchers and clinical experts in this field but also health care commissioners. In view of that, this systematic review aimed to examine all economic evidence (not just from randomised controlled trials) of the use of EUS staging in the management of patients with GOC. The protocol of this systematic review was registered on PROSPERO, an international prospective register of systematic reviews [PROSPERO registration number: PROSPERO 2016:CRD42016043700; http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42016043700] (Yeo et al., 2016, see Appendix 3.1). This paper offers a review of economic evidence on the costs, benefits (in terms of GOC patients' health-related quality of life), and economic implications of the use of EUS for staging GOC patients.

3.3 Methods

This review was carried out and reported in accordance with the updated Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009; Liberati et al., 2009) (see Appendix 3.2).

3.3.1 Searches and Study Selection

Searches for this systematic review were conducted using a range of electronic databases: MEDLINE (ovid), EMBASE (ovid), The Cochrane Collaboration Register and Library (including Cochrane Central Register of Controlled Trials (CCRCT), Cochrane Reviews, Database of Abstracts of Reviews of Effects (DARE), Health Technology Assessment (HTA), NHS Economic Evaluation Database (NHS EED), Cochrane Methodology Register (CMR)), CINAHL

(EBSCOhost), Web of Science (Core Collection). Searches were restricted to publications from the last 20 years (1996-2016) as per the registered protocol on PROSPERO [Registration number: PROSPERO 2016:CRD42016043700; http://www.crd.york.ac.uk/PROSPERO_REBRANDING/display_record.asp?ID=CRD42016043700] (Yeo et al., 2016). To ensure that the review was as up-to-date as possible, the searches were re-run on all databases to cover 2016-2018 (search update on 28/04/2018). All the searches were again updated to cover studies between 2018-2021 (search update on 13/12/2021).

In order to ensure a comprehensive search was achieved and any relevant research had not been missed, online searches were also conducted through the following internet search engines and appropriate websites to identify grey literature, reports, ongoing and unpublished studies from conference papers and abstracts: Google, Google Scholar, Department of Health (DoH), National Institute for Health and Clinical Excellence (NICE), National Institute for Health Research (NIHR) Journals Library, NIHR UK Clinical Trials Gateway, The National Cancer Research Institute (NCRI), Cancer Research Wales (CRW), Wales Cancer Research Centre (WCRC), Welsh Government (WG), Health and Care Research Wales (HCRW), Cancer Research UK (CRUK) and other relevant charitable organisation websites.

The reference lists of papers that were included in the review were searched for further publications that had not been identified in the electronic searches. Contacts with study authors were made to locate further relevant literature and publications.

Guided by the review objectives, the search terms as shown in Table 3.1 were developed using the PICO framework (Lang, 2004; Schardt et al., 2007). The PICO framework was utilised to help shape, design and construct the search process to identify all relevant published and unpublished materials from various sources. Titles, abstracts and full-text papers were searched for using these search terms.

Table 3.1: Search terms by category, guided by PICO framework, for the systematic review.

No.	Search Term Category	Search Terms
1.	Disease	neoplas* OR
		cancer*OR
		carcin* OR
		tumo* OR
		adenocarcinoma* OR
		squamous cell carcinoma* OR
		malig* OR
		metasta*
		AND
2.	Type of disease	gastro* OR
		oesophag* OR
		esophag* OR
		gastro-oesophag* OR
		gastro-esophag* OR
		gastroesophag* junction* OR
		gastro-esophag* junction* OR
		gastrooesophag* junction* OR
		gastro-oesophag* junction* OR
		esophagogastric junction* OR
		esophago-gastric junction* OR
		oesophagogastric junction* OR
		oesophago-gastric junction* OR
		oesophageal squamous cell carcinoma* OR
		esophageal squamous cell carcinoma* OR
		gut* OR
		gullet* OR
		food pipe OR
		stomach OR

		upper GI OR
		upper-GI OR
		upper gastrointestin* OR
		upper-gastrointestin* OR
		upper digestive tract* OR
		upper-digestive tract* OR
		intraepithelial OR
		intramucosal OR
		node* OR
		nodal
		AND
3.	Intervention	endosono* OR
		EUS OR
		endoscopic ultraso* OR
		endoscopic-ultraso* OR
		EUS-FNA OR
		EUS-fine needle aspiration OR
		EUS fine-needle aspiration OR
		Endosonography-guided FNA OR
		Endoscopic ultrasound-fine needle aspiration OR
		Endoscopic ultrasound-guided fine needle aspiration OR
		Endoscopic ultrasound-guided fine-needle aspiration OR
		Endoscopic-ultrasound-guided fine-needle aspiration OR
		Endoscopic ultrasound guided fine needle aspiration OR
		Echoendoscop* OR
		*Echo-endoscop**
		AND
		Staging OR
		Preoperative staging OR
		Pre-operative staging
		AND

4.	Outcome	econom* OR health economics OR economic evaluation OR cost-effective* OR cost effect* OR cost utility OR cost-utility OR cost-conseq* OR cost conseq* OR cost-benefit OR cost benefit OR cost-minimisation OR cost minimisation OR cost-minimization OR cost minimization OR cost* OR cost* analys* OR unit cost OR unit-cost OR unit-costs OR unit costs OR drug cost OR drug costs OR hospital costs OR health-care costs OR health care cost OR medical cost OR medical costs OR cost* efficacy* OR cost* analys* OR cost* allocation* OR

	cost* control* OR
	cost* illness* OR
	cost* affordable* OR
	cost* fee* OR
	cost* charge*
	economic model* OR
	markov* OR
	budget* OR
	healthcare economics OR
	health care economics OR
	cost analys* OR
	health-care cost* OR
	health care cost* OR
	hrqol OR
	Health related quality of life OR
	health-related quality of life OR
	quality-adjusted life year* OR
	quality adjusted life year* OR
	qaly OR
	Quality of life OR
	quality-of-life OR
	QoL

The search strategy for each of the five electronic databases was developed, checked and tested by an information specialist before finalising the search terms; this process was informed by the search strategy of a wider evidence synthesis that includes a systematic review of non-economic studies of treatments for resectable adenocarcinoma of the stomach, gastro-oesophageal junction and lower oesophagus (Ronellenfitsch et al., 2013). An example of search strategy used in the Medline Ovid database is shown as below in Table 3.2. Full details of the search strategy for each of the five databases are presented in Appendix 3.3-3.7.

Table 3.2: An example of search strategy used in the Medline Ovid database.

1	exp Endosonography/
2	endosono$.tw.
3	endoscopic ultraso$.tw.
4	endoscopic-ultraso$.tw.
5	EUS.tw.
6	(echoendoscop$ or echo-endoscop$).tw.
7	((endosono$ or endoscopic ultraso$ or endoscopic-ultraso$ or EUS) adj6 aspiration).tw.
8	1 or 2 or 3 or 4 or 5 or 6 or 7
9	staging.tw.
10	((Preoperative or pre-operative) adj6 staging).tw.
11	9 or 10
12	8 and 11
13	exp Adenocarcinoma/
14	adenocarcinoma$.tw.
15	13 or 14
16	exp Esophagus/
17	exp Esophagogastric Junction/
18	(gastroesophag$ adj3 junction$).tw.
19	(gastro-esophag$ adj3 junction$).tw.
20	(gastrooesophag$ adj3 junction$).tw.
21	(gastro-oesophag$ adj3 junction$).tw.
22	esophagogastric junction$.tw.
23	esophago-gastric junction$.tw.
24	oesophagogastric junction$.tw.

25 oesophago-gastric junction$.tw.

26 exp Stomach/

27 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26

28 15 and 27

29 exp Esophageal Neoplasms/

30 exp Stomach Neoplasms/

31 (esophag$ adj5 neoplas$).tw.

32 (oesophag$ adj5 neoplas$).tw.

33 (esophag$ adj5 cancer$).tw.

34 (oesophag$ adj5 cancer$).tw.

35 (esophag$ adj5 carcin$).tw.

36 (oesophag$ adj5 carcin$).tw.

37 (esophag$ adj5 tumo$).tw.

38 (oesophag$ adj5 tumo$).tw.

39 (esophag$ adj5 metasta$).tw.

40 (oesophag$ adj5 metasta$).tw.

41 (esophag$ adj5 malig$).tw.

42 (oesophag$ adj5 malig$).tw.

43 (esophag$ adj5 adenocarcinoma$).tw.

44 (oesophag$ adj5 adenocarcinoma$).tw.

45 (stomach adj5 neoplas$).tw.

46 (stomach adj5 cancer$).tw.

47 (stomach adj5 carcin$).tw.

48 (stomach adj5 tumo$).tw.

49 (stomach adj5 metasta$).tw.

50 (stomach adj5 malig$).tw.

51	(stomach adj5 adenocarcinoma$).tw.
52	(gastric adj5 neoplas$).tw.
53	(gastric adj5 cancer$).tw.
54	(gastric adj5 carcin$).tw.
55	(gastric adj5 tumo$).tw.
56	(gastric adj5 metasta$).tw.
57	(gastric adj5 malig$).tw.
58	(gastric adj5 adenocarcinoma$).tw.
59	29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58
60	28 or 59
61	(gut$ adj5 (neoplas$ or cancer$ or carcin$ or tumo$ or adenocarcinoma$ or metasta$ or malig$)).tw.
62	(gullet$ adj5 (neoplas$ or cancer$ or carcin$ or tumo$ or adenocarcinoma$ or metasta$ or malig$)).tw.
63	(food pipe adj5 (neoplas$ or cancer$ or carcin$ or tumo$ or adenocarcinoma$ or metasta$ or malig$)).tw.
64	(("upper GI" or "upper-GI") adj5 (neoplas$ or cancer$ or carcin$ or tumo$ or adenocarcinoma$ or metasta$ or malig$)).tw.
65	(("upper gastrointestin$" or "upper-gastrointestin$") adj5 (neoplas$ or cancer$ or carcin$ or tumo$ or adenocarcinoma$ or metasta$ or malig$)).tw.
66	((upper digestive tract$ or upper-digestive tract$) adj5 (neoplas$ or cancer$ or carcin$ or tumo$ or adenocarcinoma$ or metasta$ or malig$)).tw.
67	61 or 62 or 63 or 64 or 65 or 66
68	60 or 67
69	12 and 68
70	exp Economics/

71	health economics.mp.
72	Economic evaluation.mp.
73	exp Cost-Benefit Analysis/
74	(cost$ adj2 (effective$ or utilit$ or consequence$ or benefit$ or minimi$)).tw.
75	Cost effectiveness analysis.mp.
76	cost utility analysis.mp.
77	cost consequences analysis.mp.
78	cost minimisation analysis.mp.
79	cost minimization analysis.mp.
80	exp "Costs and Cost Analysis"/
81	(unit cost or unit-cost or unit-costs or unit costs or drug cost or drug costs or hospital costs or health-care costs or health care cost or medical cost or medical costs).tw.
82	(cost$ adj2 (efficac$ or analys$ or allocation$ or control$ or illness$ or affordable$ or fee$ or charge$)).tw.
83	exp Models, Economic/
84	(decision adj1 (tree$ or analys$ or model$)).tw.
85	Markov$.tw.
86	exp Economics, Pharmaceutical/ or exp Economics, Medical/ or exp Economics, Hospital/
87	(econom$ or cost$ or price$ or pricing or pharmacoeconomic$ or pharmaeconomic$ or pharmaco-economic$).tw.
88	exp "Fees and Charges"/
89	exp Budgets/
90	(financ$ or fee$).tw.
91	((value or values or valuation) adj2 (money or monetary or life or lives or costs or cost)).tw.
92	exp Health Expenditures/

93	(low adj cost).mp.
94	(high adj cost).mp.
95	(health?care adj cost$).mp.
96	(cost adj estimate$).mp.
97	exp Hospital Costs/
98	exp "Cost Savings"/
99	exp "Quality of Life"/
100	*"Quality of Life"/
101	70 or 71 or 72 or 73 or 74 or 75 or 76 or 77 or 78 or 79 or 80 or 81 or 82 or 83 or 84 or 85 or 86 or 87 or 88 or 89 or 90 or 91 or 92 or 93 or 94 or 95 or 96 or 97 or 98 or 99 or 100
102	69 and 101

3.3.2 Inclusion and Exclusion Criteria

Table 3.3 presents the inclusion and exclusion criteria, using the economic evidence review design framework outlined in the University of York Centre for Reviews and Dissemination (2009): Population, Interventions, Comparators, Outcomes, and Type of Evidence. Due to resources constraints, only studies written in English were included. This includes international studies that have been translated or written in English.

Table 3.3: Inclusion and exclusion criteria for the systematic review.

	Inclusion Criteria	Exclusion Criteria
Population	All adults (>18 years: aged 19 and above (Bray et al., 2014; Franssen et al., 2020)) who had cancer (i.e. localised tumour) of the oesophagus, stomach or gastro-	Population aged below 19 years and had metastatic oesophageal, gastro-oesophageal or gastric cancer.

	oesophageal junction; free of metastatic disease.	
Interventions	Use of endoscopic ultrasound (EUS) (also known as endosonography, echoendoscopy) staging in patient with oesophagus, gastro-oesophageal and gastric cancer.	Use of endoscopy only or ultrasound only, and use of EUS for non-cancer staging purposes e.g. treatment of cancer
Comparators	Standard staging algorithm e.g. trans-abdominal ultrasound scan, computed tomography (CT) scan. Partial economic evaluations, when no formal comparator was used, were included.	
Outcomes	All relevant full economic evaluation studies outcomes including (but not be restricted to) cost per QALY and cost per life-year gained; All other relevant economic outcomes including (but not be restricted to) resource use, direct and indirect costs, incremental benefits e.g. quality-adjusted survival or quality-adjusted life years (QALYs), health-related quality of life, cancer-specific quality of life and utility gained –	All outcomes unrelated to economic evidence of EUS staging of the oesophagus, gastro-oesophageal junction or gastric cancer.

	this includes partial economic evaluation studies outcomes, which costs or consequences alone of a single intervention (e.g. EUS staging of GOC) were described, were included.	
Type of Evidence	Full economic evaluation evidence (i.e. cost-effectiveness, cost-utility and cost-benefit analyses) related to EUS staging of oesophageal, gastro-oesophageal junction and gastric cancer were considered. Other economic studies that contain partial economic evaluation or no evaluation context (e.g. cost analyses, cost-description studies, cost-outcome descriptions, budgetary studies, outcome-description studies in terms of utility gained, health-related quality of life and cancer-specific quality of life measures such as QALYs and FACT-G score) were also considered. Economic evaluation studies conducted alongside RCTs, non-RCTs, quasi-experimental trials, epidemiological research, cohort studies, and modelling studies were considered.	Non-research studies such as editorials, case reports or other descriptive studies.

General	Language – English. Years – 1996-2016, 2016-2018 and 2018-2021.	Language – Not written or translated into English. Years – Before 1996.

3.3.3 Data Extraction

Titles and abstracts of all studies identified were screened and assessed for relevance against the inclusion criteria by two independent reviewers (STY and NB). The inclusion or exclusion of each study was checked and confirmed. All potentially relevant full-text papers were then obtained and screened against the inclusion criteria, with disagreements resolved through discussion.

Following screening, relevant data from all full-text papers included in the review were extracted by the primary reviewer (STY) using an adapted standardised form (University of York Centre for Reviews and Dissemination, 2009), and checked by the second reviewer (NB). Two adapted standardised forms were developed and used for data extraction – one for economic studies (see Appendix 3.8 or 3.9) and another for economic modelling studies (see Appendix 3.10).

3.3.4 Quality Assessment

The quality of all full-text papers included in the review were assessed and rated independently by the two reviewers using the Critical Appraisal Skills Programme (CASP) economic evaluation checklist (CASP, 2017) tool for economic studies and the Philips et al's economic modelling checklist (Philips et al., 2004) tool for economic modelling studies. The papers were critically appraised to assess to what extent the content of these papers complied with the criteria of good practice in economic evaluation and if there was any obvious bias. Disagreements between the reviewers were resolved through discussion.

3.3.5 Data Synthesis

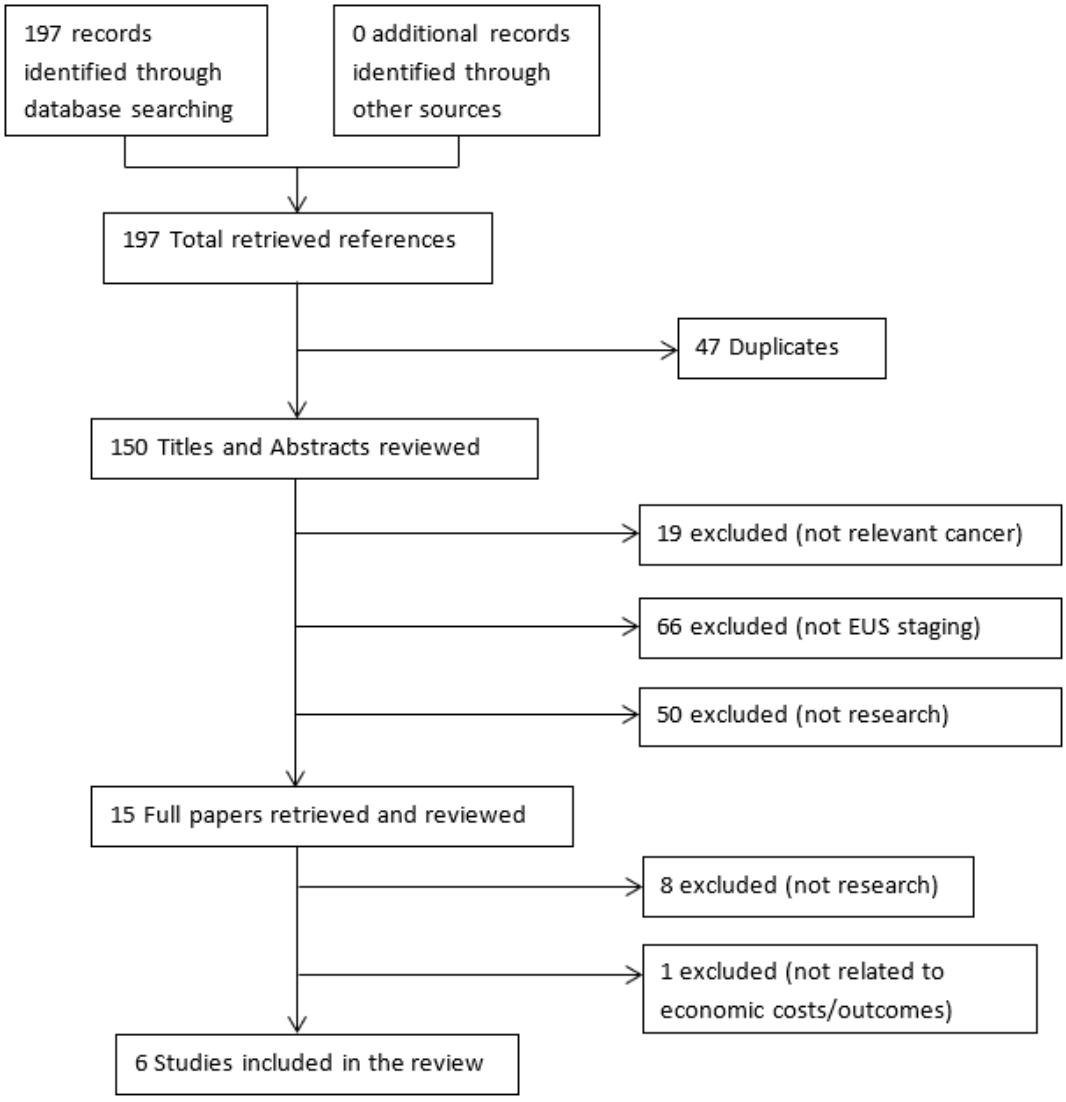
All studies included in the review were summarised and compared across studies in a narrative form to answer the review objectives. The aims, methods, and results of the studies reviewed were synthesised narratively. This demonstrates the heterogeneity of the studies in terms of characteristics (University of York Centre for Reviews and Dissemination, 2009). Due to the heterogeneity of the studies in terms of the study type and outcomes across the studies, meta-analysis was not appropriate (University of York Centre for Reviews and Dissemination, 2009). Costs were converted into British pounds sterling, £, using the appropriate exchange rate published in the International Monetary Fund (International Monetary Fund, 2017) and inflated to 2017 price year using the hospital and community health services (HCHS) index (Curtis and Netten, 2005; Curtis and Burns, 2015; Curtis and Burns, 2016; Curtis and Burns, 2017) for the studies included in the review.

3.4 Results

3.4.1 Literature Search: Identification of Studies

Overall, the search from 1996 to 2016 identified 197 potentially relevant studies, six of which fulfilled the inclusion criteria and were included in the review (Figure 3.1). Of the six studies included, three were economic analysis studies and 3 were economic modelling studies.

Figure 3.1: Flowchart of the study selection process.



To ensure that the review was as up-to-date as possible, the searches were re-run on all databases to cover 2016-2018 (search update on 28/04/2018) and again to cover 2018-2021 (search update on 13/12/2021); 30 potentially relevant papers were identified in the 2018 search update and 24 in the 2021 search update but none met the inclusion criteria. In such case, the final number of studies included in the review remained at six.

3.4.2 Study Descriptions

Table 3.4 and 3.5 summarises the characteristics of the six studies included in the review. There were three economic analysis studies (Table 3.4) and three economic modelling studies (Table 3.5). Further details of study descriptions for each of the six studies reviewed can be found in Appendix 3.11-3.16. Five of the studies included in the review were US studies, and one was a UK study. Of the three economic analysis studies, two were cost analyses (Chang et al., 2003; Shumaker et al., 2002) and one was a cost-effectiveness analysis (Russell et al., 2013). Of the three economic modelling studies, two were cost-minimisation analyses (Harewood et al., 2002; Hadzijahic et al., 2000) and one was a cost-effectiveness analysis (Wallace et al., 2002). All of the three economic modelling studies used decision tree modelling techniques to explore staging strategies.

The six studies included in the review differed quite markedly in terms of their design. Only one study used primary cost and outcome data collected in prospective evaluation (Russell et al., 2013), one study used data collected in prospective case series (Chang et al., 2003), one study used retrospective data (Shumaker et al., 2002), and the remaining three studies synthesised data from secondary sources in a decision tree model (Wallace et al., 2002; Harewood et al., 2002; Hadzijahic et al., 2000). Of the six studies, only one (Russell et al., 2013) was a randomised controlled trial and included participants diagnosed with gastro-oesophageal cancer (i.e. oesophageal, gastro-oesophageal junction or gastric cancer); the other five were non-trial studies and included participants diagnosed with oesophageal cancer. Amongst the six studies, Russell et al. (2013) was again the only study which evaluated costs of health care resource use covering secondary care contacts and hospital prescribed drugs in addition to cost of EUS, collected prospectively in the trial.

In terms of health outcome measures, two studies (Russell et al., 2013; Wallace et al., 2002) included quality-adjusted life year (QALY) as the measure of effect and conducted a cost-effectiveness analysis to assess the gain in QALYs relative to the costs of different staging strategies. The remaining four studies (Chang et al., 2003; Shumaker et al., 2002; Harewood et al., 2002; Hadzijahic et al., 2000) did not explore QALY or other quality of life measures but only cost.

Table 3.4: Summary table of the structure and results of the three economics papers included in the systematic review.

Authors, year, country	Aims of the study	Type of participants (*n*)	Type of study, methodology	Study perspective	Price year, currency (unit)	Type of intervention / staging technique	Method of delivery	Length of follow-up	Cost of intervention / staging technique	Type of economic analysis conducted	Outcomes / results / conclusions*
Shumaker et al. (2002), USA.	To determine (1) the relative proportions of each oesophageal cancer stage in a group of patients referred for preoperative staging with EUS, (2) the proportion of patients with EUS stage 1 and 4 tumours	Patients with oesophageal cancer receiving preoperative staging with EUS (n=180, 82% men and mean age 66.5 years).	Cost analysis using a retrospective review of a large multicentre national computerised endoscopic database. Data between February 1998 and October 2000 were extracted, reviewed and analysed.	Not stated specifically, the authors described US Medicare data	Price year: 2000 Currency: US dollars (USD$)	NA: retrospective review of a large national endoscopic database.	NA	NA	The cost of EUS for preoperative staging of oesophageal cancer was estimated at $634 per patient (£697 per patient, 2017 price year)	Cost analysis study: the potential cost savings of performing preoperative EUS in oesophageal cancer patients.	Preoperative staging of oesophageal cancer with EUS can facilitate cost savings by reducing the need for additional treatments in stage 1 and 4 oesophageal cancer (a significant proportion of patients – 26% in this series).

	that would not be treated with combined modality therapy, and (3) to estimate the potential cost savings of performing preoperative EUS in oesophageal cancer patients.										
Chang et al. (2003), USA	To determine the impact of EUS combined with FNA on patients'	Patients diagnosed with oesophageal cancer (squamous-cell or adenocarcinoma) who were	Cost analysis alongside prospective case series.	Not stated specifically, the study was undertaken in California, USA.	Not stated specifically, the authors described their cost analyses	NA: cost analysis study alongside prospective case series.	NA	Based on the data used in the cost analyses, the length of follow-	The cost of EUS-FNA biopsy based on the published direct costs of endosonograp	Cost analysis study: the cost of care for these patients was calculated to explore whether or	Patients' decisions on whether to undergo medical or surgical treatment

	choice of therapy and on the cost of care.	referred to the University of California's Irvine Medical Center for preoperative EUS staging between August 1993 and August 1997 (n=60, 39 men, 21 women and mean age 68±10 years). These patients were all being considered for surgical resection and had undergone standard evaluation including CT which showed no evidence of distant metastases.			were based on the published direct costs of endosonography-guided aspiration biopsy and thoracotomy published in 1997 (Gress et al., 1997). Currency: US dollars (USD$)			up was, on average, 17 months (range 1-51 months).	hy-guided aspiration biopsy (Gress et al., 1997) was estimated at $1,975 per patient (outpatient) (£3,528 per patient, 2017 price year).	not the use of EUS decreases the cost of managing patients with oesophageal cancer.	correlated significantly with their overall tumour staging, suggesting that the information provided by EUS played a significant role in patients' decision-making. EUS-guided therapy potentially reduces the cost of managing patients with oesophageal cancer by USD$12,340 per patient

											(£10,510 per patient, 2017 price year) due to reduced number of thoracotomies undertaken (patient choice).
Russell et al. (2013), UK	To examine whether the addition of EUS to usual staging uses resources cost-effectively.	Patients with proven cancer of the oesophagus, stomach or gastro-oesophageal junction; medically fit for both surgery (even if not planned) and chemotherapy, free of metastatic disease and had	Cost-effectiveness analysis alongside a multi-centre randomised controlled trial (RCT) namely 'COGNATE trial'. The study explored whether giving EUS scan in addition to standard staging algorithms would be more cost-	NHS perspective, focusing on health-care resources used by participants including investigation, treatment and palliation, and other elements of secondary and	Price year 2008 Currency: Pounds Sterling (£)	Cancer staging with EUS vs. without EUS	Patients randomised to intervention group received EUS scan in addition to standard staging algorithms. Patients randomised to control	Study follow-up period was 54 months or until death, whichever came first. Main analyses of the study (including health economic	The cost of EUS scan was £551 (day case) (£648, 2017 price year), £1,477 (outpatient) (£1,737, 2017 price year) and £3,781 (inpatient) (£4,447, 2017 price year).	Cost-effectiveness analysis using QALY as a measure of effect – The difference in cost and QALY between intervention and control groups was calculated; the probabilities	EUS reduced net use of health-care resources by £2,860 (£3,364, 2017 price year) and had an increase of 0.1969 in estimated mean QALYs. Combining these

		not started treatment. Both their ASA (America Society of Anesthesiologists) grading and their WHO performance status had to be 1 or 2 (n=213, 165 male; mean age 64.4 years; EUS group (n=107); No EUS group (n=106)).	effective compared to standard staging algorithms.	pharmaceutical care.			group received standard staging algorithms.	analysis) used 48 months.		of the EUS intervention being cost-effective at different willingness-to-pay thresholds were estimated.	estimated benefits and savings yields probability of 96.6% that EUS is cost-effective in the sense of achieving the NICE criterion of costing less than £20,000 to gain a QALY.

NA, Not applicable; ICER, incremental cost-effectiveness ratio; EUS, endoscopic ultrasound; EUS-FNA, endoscopic ultrasound-fine needle aspiration; NHS, national health service; QALY, quality-adjusted life year; NICE, National Institute for Health and Care Excellence

*Converted to pound sterling (£) at 2017 prices.

Table 3.5: Summary table of the structure and results of the three economic modelling papers included in the systematic review.

Authors, year, country	Aims of the study	Type of participants	Type of study, methodology	Perspective of the model	Price year, currency (unit)	Type of intervention / staging technique	Analysis	Time horizon	Outcome measures	Outcomes / results / conclusions*
Hadzijahic et al. (2000), USA	To determine whether it is less costly to request CT or EUS first to identify advanced oesophageal cancer; to determine which variables most affect the overall cost of identifying advanced disease.	Oesophageal cancer patients who underwent both CT and EUS between July 1995 and April 1999 (n=124, mean age = 62.7 years, 98 (79%) men, and 72 (58%) white).	Cost-minimisation study using decision tree model to compare which of the two initial staging strategies (EUS first or CT first strategy) would cost less to detect advanced disease in patients diagnosed endoscopically with	Not stated specifically, the study took local referral centre perspective.	Price year: 1999 Currency: US dollars (USD$).	CT first strategy vs. EUS first strategy.	Decision analysis using decision tree model.	Not stated specifically .	Overall cost of identifying advanced disease of the two strategies: EUS first and CT first strategies.	Initial CT is the least costly strategy if the probability of finding advanced disease by initial CT is greater than 20%, if the probability of finding advanced disease by initial EUS is less than 30%, or if the cost of EUS is greater than 3.5 times the cost of CT. EUS found advanced disease more frequently than CT (44% vs. 13%; $p <$ 0.0001) and initial EUS was the least costly strategy (Initial EUS strategy expected cost was US$804 (£824, 2017 price year) vs.

			oesophageal cancer.							initial CT strategy expected cost $844 (£867, 2017 price year)).
Harewood et al. (2002), USA	To examine which staging/management technique was the least costly: EUS FNA, CT-guided FNA or surgical management of oesophageal tumours.	Patients with apparently "resectable" oesophageal cancer on CT (i.e. patients with non-metastatic oesophageal cancer).	Cost-minimisation study using decision tree model to determine which strategy is least costly among the different alternatives: CT-FNA, EUS-FNA and 'proceed straight to surgery' options.	Third party payer perspective.	Price year: 2001 Currency: US dollars (USD$).	CT-FNA vs. EUS-FNA vs. 'proceed directly to surgery'.	Decision analysis using decision tree model.	Not stated specifically.	Least costly staging strategy among the three strategies (CT-FNA vs. EUS-FNA vs Surgery)	EUS FNA was the least costly strategy at $13,811 (£14,578, 2017 price year), followed by surgery at $13,992 (£14,768, 2017 price year) and CT-FNA at $14,350 (£15,147, 2017 price year). EUS FNA remained the least costly option, provided that the prevalence of celiac lymph node (CLN) involvement was greater than 16%. Below this value, surgery became the least costly strategy. The final outcome of the model was also sensitive

										to variation in the sensitivity of EUS FNA. Provided that the sensitivity of EUS-FNA was greater than 66%, EUS-FNA remained the least costly staging option in the management of oesophageal tumours. Despite changing the values of two or three variables simultaneously in the two- and three-way sensitivity analyses, the result still showed that EUS FNA remained the least costly strategy.
Wallace et al. (2002), USA	To compare the health care costs and effectiveness of multiple	All Medicare-eligible patients whose invasive oesophageal cancer was	Cost-effectiveness study using decision tree model to compare the	Third-party payer perspective	Price year: 2000 Currency: US dollars (USD$).	The costs and effectiveness of the six strategies were compared – CT alone vs.	Decision analysis using decision tree model.	Not stated specifically	Cost, QALYs and cost per QALY of the six strategies	Under baseline assumptions, CT+EUS-FNA was the least costly strategy and offered more QALYs, on average, than all other strategies with

	staging options for patients with oesophageal cancer.	diagnosed between January 1991 and December 1996. Data were obtained retrospectively from the SEER–Medicare databases.	costs and effectiveness of six strategies (CT alone vs. CT+EUS vs. CT+TL vs. CT+EUS+TL vs CT+PET+EUS vs. PET+EUS).			CT+EUS vs. CT+TL vs. CT+EUS+TL vs CT+PET+EUS vs. PET+EUS.				the exception of PET+EUS-FNA. The latter was slightly more effective but also more costly. The marginal cost-effectiveness ratio comparing PET+EUS-FNA with CT+EUS-FNA was $60,544 per QALY (£66,588 per QALY, 2017 price year). These findings were robust and changed very little in all of the sensitivity analyses.

ICER, incremental cost-effectiveness ratio; EUS, endoscopic ultrasound; EUS-FNA, endoscopic ultrasound-fine needle aspiration; CT, computed tomography; PET, positron emission tomography; TL, thoracoscopy and laparoscopy; QALY, quality-adjusted life year.
*Converted to pound sterling (£) at 2017 prices.

3.4.3 Quality Assessment

Each of the six studies included in the review were critically appraised against the appropriate source of quality appraisal checklist: the CASP economic evaluation checklist (CASP, 2017) was used for the three economic studies, and Philips et al's economic modelling checklist (Philips et al., 2004) was used for the remaining three economic modelling studies. Table 3.6 and 3.7 summarised the quality assessment of the three economic studies and three economic modelling studies, respectively. Further details of the quality assessment results for each of the six studies reviewed can be found in Appendix 3.17-3.22.

Table 3.6 shows the study quality of the three economic studies was generally good, although only one study (Russell et al., 2013) met all quality criteria on the CASP economic evaluation checklist. The study by Shumaker et al. (2002) scored the second highest, followed by Chang et al. (2003). Of these three economic studies, two had missing key information: Chang et al. (2003) reported neither cost perspective, cost inflation, discounting nor price year, and sensitivity analysis was not undertaken; likewise, Shumaker et al. (2002) did not state whether their reported costs were discounted or inflated as appropriate.

Table 3.7 shows the study quality of the three economic modelling studies included in the review was satisfactory, scoring moderately well on the Philips et al's economic modelling checklist. In descending order of quality, the study by Wallace et al. (2002) scored the highest followed by Harewood et al. (2002) and Hadzijahic et al. (2000). One study (Hadzijahic et al., 2000) did not state the perspective of the model and all three (Hadzijahic et al., 2000; Harewood et al., 2002; Wallace et al., 2002) did not specify the time horizon of the decision tree model. There was insufficient detail of how parameters in the model were identified (Wallace et al., 2002) and how data were modelled (Harewood et al., 2002). There was also a lack of clarity with regards to the source of probabilities and cost data used in the decision tree model (Hadzijahic et al., 2000).

Table 3.6: Quality assessment results of economic studies included in the systematic review.

Question no.	CASP economic evaluation checklist questions*¥	Response (√, x, NC or NA)		
		Studies (author and year)		
		Shumaker et al. (2002)	Chang et al. (2003)	Russell et al. (2013)
1	Was a well-defined question posed?	√	√	√
2	Was a comprehensive description of the competing alternatives given?	NA	NA	√
3	Does the paper provide evidence that the programme would be effective (i.e. would the programme do more good than harm)?	√	√	√
4	Were the effects of the intervention identified, measured and valued appropriately?	NA	NA	√
5a	Were all important and relevant resources required and health outcome costs for each alternative identified?	NC	NC	√

5b	Were all important and relevant resources required and health outcome costs for each alternative measured in appropriate units?	√	√	√
5c	Were all important and relevant resources required and health outcome costs for each alternative valued credibly?	√	NC	√
6	Were costs and consequences adjusted for different times at which they occurred (discounting)?	x	x	√
7	What were the results of the evaluation?	√	√	√
8	Was an incremental analysis of the consequences and cost of alternatives performed?	NA	NA	√
9	Was an adequate sensitivity analysis performed?	√	x	√
10	Is the programme likely to be equally effective in your context or setting?	√	√	√
11	Are the costs translatable to your setting?	x	x	√

12	Is it worth doing in your setting?	√	√	√
Score, ratio™ (%)		8/11 (73%)	6/11 (55%)	14/14 (100%)

NA: Not Applicable; NC: Not Clear

*Available from: http://www.casp-uk.net/casp-tools-checklists

¥Adapted from: Drummond MF, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press, 1987

™Ratio = b/a, where b = sum of tick; a = sum of items (excluding 'NA' items)

Table 3.7: Quality assessment results of economic modelling studies included in the systematic review.

Quality Criterion	Philips et al' economic modelling checklist questions*	Response (√, x, NC or NA)		
		Studies (author and year)		
		Hadzijahic et al. (2000)	Harewood et al. (2002)	Wallace et al. (2002)
S1	Is there a clear statement of the decision problem?	√	√	√
	Is the objective of the evaluation and model specified and consistent with the stated decision problem?	√	√	√
	Is the primary decision-maker specified?	NC	√	√
S2	Is the perspective of the model stated clearly?	X	√	√
	Are the model inputs consistent with the stated perspective?	NC	√	√
	Has the scope of the model been stated and justified?	√	√	√
	Are the outcomes of the model consistent with the perspective, scope and overall objective of the model?	√	√	√

S3	Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	√	√	√
	Are the sources of data used to develop the structure of the model specified?	√	√	√
	Are the causal relationships described by the model structure justified appropriately?	NA	NA	NA
S4	Are the structural assumptions transparent and justified?	√	√	√
	Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	√	√	√
S5	Is there a clear definition of the options under evaluation?	√	√	√
	Have all feasible and practical options been evaluated?	√	√	√
	Is there justification for the exclusion of feasible options?	NA	NA	NA

S6	Is the chosen model type appropriate given the decision problem and specified causal relationships within the model?	√	√	√
S7	Is the time horizon of the model sufficient to reflect all important differences between options?	X	x	x
	Are the time horizon of the model, the duration of treatment and the duration of treatment effect described and justified?	X	x	x
S8	Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions?	√	√	√
S9	Is the cycle length defined and justified in terms of the natural history of disease?	NA	NA	NA
D1	Are the data identification methods transparent and appropriate given the objectives of the model?	√	NC	√

	Where choices have been made between data sources, are these justified appropriately?	NA	√	√
	Has particular attention been paid to identifying data for the important parameters in the model?	√	√	x
	Has the quality of the data been assessed appropriately?	X	x	x
	Where expert opinion has been used, are the methods described and justified?	NA	NA	x
D2	Is the data modelling methodology based on justifiable statistical and epidemiological techniques?	√	NC	√
D2a	Is the choice of baseline data described and justified?	√	√	√
	Are transition probabilities calculated appropriately?	NA	NA	NA
	Has a half-cycle correction been applied to both cost and outcome?	NA	NA	NA
	If not, has this omission been justified?	NA	NA	NA

D2b	If relative treatment effects have been derived from trial data, have they been synthesised using appropriate techniques?	NA	NA	NA
	Have the methods and assumptions used to extrapolate short-term results to final outcomes been documented and justified?	NA	NA	NA
	Have alternative assumptions been explored through sensitivity analysis?	√	√	√
	Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?	NA	NA	NA
	Have alternative assumptions regarding the continuing effect of treatment been explored through sensitivity analysis?	NA	NA	NA
D2c	Are the costs incorporated into the model justified?	√	√	√
	Has the source for all costs been described?	√	√	√
	Have discount rates been described and justified given the target decision-maker?	NC	NA	√

D2d	Are the utilities incorporated into the model appropriate?	NA	NA	√
	Is the source for the utility weights referenced?	NA	NA	x
	Are the methods of derivation for the utility weights justified?	NA	NA	x
D3	Have all data incorporated into the model been described and referenced in sufficient detail?	NC	√	√
	Has the use of mutually inconsistent data been justified (i.e. are assumptions and choices appropriate)?	NC	NC	√
	Is the process of data incorporation transparent?	√	x	x
	If data have been incorporated as distributions, has the choice of distribution for each parameter been described and justified?	NA	NA	NA
	If data have been incorporated as distributions, is it clear that second order uncertainty is reflected?	NA	NA	NA
D4	Have the four principal types of uncertainty been addressed?	x	x	x
	If not, has the omission of particular forms of uncertainty been justified?	x	x	x

D4a	Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?	x	x	x
D4b	Is there evidence that structural uncertainties have been addressed via sensitivity analysis?	x	x	x
D4c	Has heterogeneity been dealt with by running the model separately for different subgroups?	X	x	x
D4d	Are the methods of assessment of parameter uncertainty appropriate?	√	√	√
	If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	NC	√	√
C1	Is there evidence that the mathematical logic of the model has been tested thoroughly before use?	X	x	x

C2	Are any counterintuitive results from the model explained and justified?	NA	NA	NA
	If the model has been calibrated against independent data, have any differences been explained and justified?	NA	NA	NA
	Have the results of the model been compared with those of previous models and any differences in results explained?	X	x	x
Score, ratio™ (%)		21/38 (55%)	24/38 (63%)	28/43 (65%)

NA: Not Applicable; NC: Not Clear

*****Available from: Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, Woolacott N and Glanville J. (2004) Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess,* 8(36)

™Ratio = b/a, where b = sum of tick; a = sum of items (excluding 'NA' items)

3.4.4 Data Synthesis Results

All of the six studies included in the review exhibit EUS as a complementary imaging technique to other imaging modalities such as CT and PET scanning for staging gastro-oesophageal cancer. This is in agreement with a previously published meta-analysis study of diagnostic test characteristics for EUS, CT, and PET scanning (van Vliet et al., 2008), concluding that the three approaches were complementary.

Results from three of the economic studies (Russell et al., 2013; Chang et al., 2003; Shumaker et al., 2002) show staging of oesophageal or gastro-oesophageal cancer with EUS could potentially save costs. Similarly, results from two of the modelling studies (Harewood et al., 2002; Hadzijahic et al., 2000) show that EUS or EUS-fine-needle aspiration biopsy (FNA) is the least costly staging technique for oesophageal cancer. The study by Wallace et al. (2002) shows that EUS-FNA in addition to CT scan is the least costly strategy than all other strategies i.e. CT alone, CT+ thoracoscopy and laparoscopy (TL), CT+EUS-FNA+TL, CT+PET+EUS-FNA and PET+EUS-FNA.

Results from the two studies (Russell et al., 2013; Wallace et al., 2002) in which quality-adjusted life year (QALY) and cost data were available demonstrate the use of EUS (Russell et al., 2013) or EUS-FNA (Wallace et al., 2002) as an additional staging technique for gastro-oesophageal cancer offered more QALYs and costed less, on average, compared to staging techniques without EUS. Russell et al. (2013) reported that EUS resulted in a QALY gain of 0.1969 QALYs and saved costs by £2,860, on average, per patient (£3,364 per patient, 2017 price year); combining these benefits and savings demonstrates that EUS is likely to be cost-effective with a probability of 96% at the UK NICE's threshold of £20,000-£30,000 per QALY (NICE, 2013).

Similarly, Wallace et al. (2002)'s modelling study showed that using EUS-FNA as an additional staging technique offered greater QALYs and saved more costs, on average, than staging strategy without EUS. For example, the combination of CT and EUS-FNA (CT+EUS-FNA) provided 0.0019 more QALYs and saved US$1,790, on average, per patient (£1,969 per patient, 2017 price year) compared to CT alone strategy. The authors argued that, among all the six staging strategies evaluated (i.e. CT alone, CT+EUS-FNA, CT+TL, CT+EUS-FNA+TL, CT+PET+EUS-FNA and PET+EUS-FNA), CT+EUS-FNA was the least costly strategy (US$40,363) (£44,392, 2017 price year) and offered higher QALYs on average (0.9649) than all other strategies with the exception of PET+EUS-FNA (US$44,521 for 1.0336 QALYs) (£48,965, 2017 price year). The latter was slightly

more effective (by 0.0687 QALYs on average) but more costly (by US$4,158 on average [£4,573, 2017 price year]) compared with CT+EUS-FNA, yielding a marginal cost-effectiveness ratio of US$60,544 per QALY (£66,588 per QALY, 2017 price year), a ratio that is less than that of other medical treatments but above accepted thresholds in the USA and UK.

3.5 Discussion

3.5.1 Main Findings

This systematic review of economic evidence of EUS staging in patients with GOC, conducted up till 2016, updated in 2018 and again in 2021, overall, it revealed a considerably small number of relevant studies. Studies varied in quality, study design and method. Study quality was generally satisfactory across all the studies included in the review, but only one of these studies (Russell et al., 2013) met all reporting and quality criteria. Differences in study design make it difficult to draw definitive conclusions as to whether the use of EUS as an additional staging technique could be considered cost-effective. For example, a head-to-head comparison of the results couldn't be made from the Russell et al. (2013) and Wallace et al. (2002) studies to draw definitive conclusions. Although both of these studies had evaluated both costs and QALYs, their respective study designs were too different to allow direct comparison; one was an economic evaluation study using primary data (Russell et al., 2013) and the other an economic modelling study using secondary data (Wallace et al., 2002). Nevertheless, the economic evidence identified in this review, especially the better quality studies, provided useful findings on the value of EUS staging in the management of GOC patients, which could be of importance to policymakers and healthcare commissioners.

Among the six studies included in the review, two studies (Russell et al., 2013; Wallace et al., 2002) are the most robust in terms of including and comparing the relative costs and QALYs of different staging strategies, for example GOC staging with and without EUS. Findings from both of these two studies demonstrated that use of EUS as an additional imaging technique could save costs and offer greater QALY gains. This could be due to the fact that EUS has been known to be beneficial in terms of sensitivity for locoregional staging of GOC, as it allows a more detailed evaluation of locoregional disease extent (T and N stage) (Thakkar and Kaul, 2020; Valero and Robles-Medranda, 2017; Allum et al., 2011; Takizawa et al., 2009; Puli et al., 2008; van Vliet et al., 2008; Allum et al., 2002; Grimm et al., 1993; Botet et al., 1991a; Botet et al., 1991b). For that reason, using EUS as a complementary imaging technique to other imaging techniques such as

CT and PET scanning for staging GOC could undoubtedly help minimise unnecessary treatments (Valero and Robles-Medranda, 2017; Allum et al., 2011; Allum et al., 2002; Dittler et al., 1995; Rösch et al., 1992); and thus potentially could save costs and offer greater health benefits to patients in terms of QALY gains. The EUS cost saving evidence was also supported by the remaining four studies (Chang et al., 2003; Shumaker et al., 2002; Harewood et al., 2002; Hadzijahic et al., 2000) evaluating only the cost of EUS e.g. whether EUS is a cost saving strategy or the least costly staging strategy. Russell et al. (2013) further argued that EUS has a considerably high probability of being cost-effective under current recommended UK NICE's threshold of £20,000 to £30,000 per QALY (NICE, 2013). Thus, despite the scarcity of economic evidence in this field, from these studies identified in the review, there is some positive economic evidence relating to the cost-effectiveness of EUS in the management of patients with GOC.

3.5.2 Strengths and Limitations of Review Methods

This review adds to the literature by providing critical evaluation of the health economics evidence of EUS staging in gastro-oesophageal cancers (GOCs), for which there is a lack of well-conducted economic studies. Though a systematic review in this field was published 20 years ago (Harris et al., 1998), this systematic review is the most up-to-date collection of economic literature in this area. Twenty years on since the review by Harris et al. (1998), still only six papers were found in the area of health economics of EUS staging in GOC. This shows that there is a lack of prioritisation of research in this area.

Broad search terms were used to develop a comprehensive search strategy for each of the databases used in this systematic review. The resultant retrieved studies were quality appraised, using both the Critical Appraisal Skills Programme (CASP) economic evaluation checklist (CASP, 2017) and the Philips et al's economic modelling checklist (Philips et al., 2004) for the retrieved economic studies and economic modelling studies, respectively. The narrative summary of the review not only described the economic evidence of EUS staging in GOC but also served as a platform for providing a holistic insight into the health economics research available to date in this area. The latter is particularly helpful for commissioners, clinicians and researchers to elicit information and potentially to facilitate the development of further research in this area.

This review has several limitations. Heterogeneity of the included studies in the review in terms of study designs and methods meant that a meta-analysis of studies was not possible and a

narrative summary was used. In terms of impact that EUS has on patients' quality of life and its costs, the lack of the availability of health economics research in this area means that it is considerably difficult, particularly for commissioners and clinicians, to guide evidence-based practice from an economic perspective.

3.5.3 Further Research

By 2021, this series of systematic reviews show that the economic evidence available to date in this area is still scarce. There was a lack of health economic research collecting data, especially primary data, on both costs and effects (such as utility values to construct QALYs) of EUS staging in GOC. To improve this, there is a need for more primary health economic research in this area, particularly integrated clinical and economic trials of EUS staging in GOC that can offer robust evidence of costs and effects.

3.6 Conclusions

There is still not a great deal of evidence on costs and benefits of EUS staging for GOC, the data available from this review suggest use of EUS as a complementary staging technique to other staging techniques for GOC appears to be cost saving and offers greater QALYs. Based on the randomised controlled trial conducted in the UK identified in this review, EUS seems to have high probability of being cost-effective at the UK NICE's threshold of £20,000-£30,000 per QALY. Nevertheless, future studies are necessary because the economic evidence around EUS staging interventions for GOC is far from robust. More health economic research and good quality data are needed to judge the economic benefits of EUS staging for GOC, particularly primary health economic research that collects primary data on the costs and effects (such as QALYs) of EUS staging in GOC.

3.7 Novel contributions

The conduct of a systematic review of the economic evidence of EUS staging in GOC is novel in the field of health economics, endoscopic ultrasound and gastro-oesophageal cancer.

The detailed, comprehensive, and transparent reporting of the methodology of this systematic review aids future work on systematic reviews of EUS staging in GOC, particularly on an update of this systematic review.

The extensive search terms designed for this systematic review as well as the detailed and comprehensive search strategy developed for each search database used in this systematic review aid the design of future systematic reviews of EUS staging within the context of health economics.

Up to 2021, this is the first economic-specific systematic review of EUS in GOC staging. Thus, it is believed that the systematic identification, screening, extraction, appraisal, and exploration of the available existing literatures can provide robust evidence-based information for policy makers in their decision making on the economic value of the use of EUS staging in GOC setting.

Chapter 4: The utilisation of endoscopic ultrasound for gastro-oesophageal cancer staging and the current practice in the UK: a healthcare professional survey

4.1 Chapter Summary

As well as a lack of cost-effectiveness analysis studies of endoscopic ultrasound (EUS) in gastro-oesophageal cancer (GOC) staging having been conducted alongside a randomised controlled trial until the inception of the HTA-funded COGNATE trial (Chapter 2), the utilisation of EUS in GOC in clinical practice has also not been widely surveyed. Prior to the launch of our survey in mid-October 2017, only one US survey study exploring the utilisation of EUS in oesophageal cancer staging has been completed but no equivalent UK study was found. As of January 2021, this is the first UK EUS survey study of its kind in the context of GOC that aimed to explore the views of UK GOC clinicians on the utilisation of EUS in GOC staging, its current clinical practice in the UK, factors associated with referral for EUS, and factors that are considered to have influence on the utilisation of EUS in GOC staging.

An online survey of health care professionals on the utilisation of EUS and the current clinical practice in the UK was conducted. Bristol Online Survey tool was used to design the online survey questionnaire. Questions covering clinician's knowledge and feeling about the use of EUS, clinician's experience in the field of EUS, current use of EUS, reasons where EUS may be limited, the usefulness and availability of EUS and clinician's clinical management style, were explored. Health care professionals were invited through two UK professional bodies – British Society of Gastroenterology (BSG) and Welsh Association of Gastroenterologists and Endoscopists (WAGE) – via email to take part in the online survey. Ethical approval was granted from the School of Healthcare Sciences Ethics and Research Committee at Bangor University. Completed online survey responses were collected and transferred on to the University's encrypted computer for analysis. Descriptive statistics and exploratory logistic regression analyses were performed using the IBM SPSS Statistics 25 software package.

Ninety-eight respondents completed the online survey questionnaire. Most of the respondents were from England (n=69; 70.4%), male (n=76; 77.6%) with a mean age of 46.6 years (SD = 9.4 years) and gastroenterologists (n=79; 80.6%). The majority support the use of EUS for staging GOC (n=89; 90.8%) as they believed that EUS offers the best local regional

staging of GOC, offers opportunity to biopsy during staging and the cost-effectiveness of using EUS to stage GOC. Of the 98 respondents, 81 (82.7%) reported that they have experience in the field of EUS either by requesting, performing or both. These 81 respondents were asked to complete further questions including (but not limited to): the perceived usefulness of EUS, availability of EUS and clinical management style. In terms of perceived usefulness of EUS, over 90% of the 81 respondents felt that EUS is more useful for staging oesophageal cancer (n=78; 96.3%) and gastro-oesophageal junction cancer (n=78; 96.3%) than gastric cancer (n=58; 71.6%). About three-quarters of the 81 respondents reported that EUS is available within their hospital. With regards to EUS referral, these 81 respondents estimated that they referred, on average, more oesophageal and gastro-oesophageal junction cancer patients for EUS staging (57.2% and 60.1%, respectively) compared to gastric cancer patients (12.7%). Clinical vignette results showed that clinicians' recommendations in treatment decisions varied depending upon tumour stage. When potential factors associated with referral for EUS were explored, two factors – 'attend Multidisciplinary Team (MDT) meeting' and 'clinicians' age' – demonstrated significant association ($p<0.05$) with referral of GOC patients for EUS, with 'attend MDT meeting' was the most important factor.

Attendance at MDT meetings is likely to increase EUS referral; hence, it may be beneficial to encourage GOC clinicians to attend MDT meetings as this may help to avoid unnecessary treatment in GOC patients. More research into the field of EUS for GOC staging are required to increase the use of EUS in clinical practice. This would aid clinical decision making in treatment recommendations for GOCs. Benefits may be gained if EUS could be offered in more clinical settings to avoid delays in the diagnosis and treatment planning of GOCs. However, it is acknowledged that there are logistical, cost and training challenges in providing EUS at every hospital in the UK.

4.2 Introduction

Endoscopic ultrasound (EUS) has been advocated as the useful staging modality to assess the extent of local disease spread and prediction of operability for gastro-oesophageal cancer (GOC) patients with non-metastatic disease (Thakkar and Kaul, 2020; DaVee et al., 2017; Valero and Robles-Medranda, 2017; Allum et al., 2002). There is evidence that EUS has superior local regional staging ability in assessing the depth of tumour invasion (T) and the presence of locoregional lymph node involvement (N) for GOC compared to computed tomography (CT) and positron emission tomography (PET), particularly for detection of regional lymph node metastases although the complementary nature of these staging techniques must be given emphasis on (Thakkar and Kaul, 2020; DaVee et al., 2017; Valero and Robles-Medranda, 2017; RCSE, 2013; Allum et al., 2011; Puli et al., 2008; van Vliet et al., 2008; Pfau et al., 2007; Lowe et al., 2005; Botet et al., 1991a; Botet et al., 1991b; Vilgrain et al., 1990). In addition, from an economic perspective, EUS has been shown to be cost saving and offers greater quality-adjusted life years (QALYs) (Yeo et al., 2019, Russell et al., 2013).

Despite the sensitivity, specificity, effectiveness, and cost-effectiveness of EUS in GOC staging have been evaluated in the setting of research, the utilisation of EUS in GOC in clinical practice has not been widely surveyed. To date (up until the launch of our survey in Autumn 2017), only one US survey study by Kim and Koch (1999) exploring the utilisation of EUS in oesophageal cancer staging has been completed but no equivalent UK study was found.

Given that no survey study in this field has been conducted in the UK before the launch of our survey in Autumn 2017, it is therefore crucial to fill the gap of knowledge missing in this field by conducting one in the UK. This survey study aimed to explore the views of UK GOC clinicians (i.e. gastroenterologists, oncologists) on the utilisation of EUS in GOC staging, its current clinical practice in the UK and potential factors associated with referral for EUS. The factors that are considered to have influence on the utilisation of EUS in GOC staging in clinical practice including (but not limited to) the following were also explored: clinicians' knowledge of EUS for staging GOC, individuals' perceptions about clinical utility of EUS, availability of EUS and individuals' clinical management style.

4.3 Methods

4.3.1 Study design

An online survey design was used involving closed and open questions. The online survey was designed and carried out in accordance with the published guidance for online questionnaire surveys (Regmi et al., 2016).

4.3.2 Study sample

Study samples were invited through two UK professional bodies – British Society of Gastroenterology (BSG) and Welsh Association for *Gastroenterology* and Endoscopy (WAGE) – via email to take part in the online survey of this study. Based on the estimate of an approximately 300 GOC clinicians in the UK, the margin of error of 5% - 8% and the confidence level of 95%, it was calculated that 100 – 169 responses (i.e. people who have completed and submitted online the survey questionnaire) were aimed. Higher number of respondents would be favourable as it would give higher precision of any estimates made. Given that there is no national database of clinicians with an interest in GOC in the UK, we have made an estimate on the number of GOC clinicians in the UK based on the data available from the UK NHS Digital on number of hospitals in the UK (Healthcare Quality Improvement Partnership, 2016), and hence an estimate of number of Multidisciplinary Teams (MDTs) and GOC-interested clinicians (personal communications with Dr Hasan Haboubi, Consultant Gastroenterologist and Clinical Advisor to this thesis). Our estimate is further supported by the data from AUGIS (2020) where it estimated that there are approximately 150 resectional upper GI surgeons performing surgery on gastro-oesophageal cancer in the UK. Accounting for Gastroenterologists/Radiologists/Oncologists that may support the surgeons/consultants within the MDTs (but who may not all be involved in EUS), this number (150) were doubled. Therefore, the combination of both of these datasets suggests that there are approximately 300 individuals in the UK who have an interest in GOC and insight into the potential use of EUS, and thus this was taken as a figure by which to calculate the study sample needed for the survey.

4.3.3 Procedure

This survey study was conducted through two UK professional bodies – British Society of Gastroenterology (BSG) and Welsh Association of Gastroenterologists and Endoscopists

(WAGE). An email invite was sent out through BSG and WAGE to their registered members (i.e. clinicians – gastroenterologists, surgeons, radiologists and oncologists) to complete the online survey. The email invite (see Appendix 4.1) consists of a support note from the Chairman of the Oesophageal Section of the BSG, the information about the survey study, and a web-link to the online survey questionnaire. The online survey was open running for six months, from mid-October 2017 to mid-April 2018. To help increase the survey responses, two email reminders were sent out through BSG and WAGE in the final two months of the six-month opening period of the online survey.

Clinicians' consent to take part in the study was determined by the submission of the completed online survey – A debrief statement, that includes the researcher's contact information and a statement to restate that respondents have the right to withdraw, but that by submitting they are agreeing to participate, was included at the end of the online survey (see Appendix 4.2). The online survey contains no identifiable information, so confidentiality was maintained.

In recognition of respondents' time given to complete the online survey, they were offered to enter into a prize draw at the end of their completed online survey. If they wish to enter into a prize draw to win a £100 Amazon voucher, they were asked to provide their email address at the end of their completed survey.

Completed online survey responses were collected and transferred into Excel and SPSS format on the University's encrypted computer for analysis. All data provided by respondents were handled in confidence and in accordance with the Data Protection Act 1998. Subsequent statistical analyses were then carried out using SPSS statistical software package, namely IBM SPSS Statistics 25 (IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp., 2017).

4.3.4 The Online survey questionnaire

An online survey questionnaire (see Appendix 4.2) was designed using the Bristol Online Survey (BOS) tool. The questionnaire consists of five sections: (1) demographic characteristics of clinicians; (2) clinician's knowledge and feelings about the use of EUS; (3) clinician's

experience in the field of EUS; (4) current use of EUS, reasons where EUS may be limited, the usefulness and availability of EUS; and (5) clinician's clinical management.

In the first section of the online survey questionnaire, demographic details included (but not limited to) age, gender, years in practice since primary medical qualification, and hospital at which respondents are practicing. In the second section, respondents were asked about their knowledge and feelings about the use of EUS. In this section, respondents were also asked if they support the use of EUS for staging GOC and the rationale of their response. In the third section, they were asked about their experience in the field of EUS. For those respondents who responded that they have experience in the field of EUS, they were required to complete the next sections of the online survey filling with questions relating to their experience in the field of EUS. Questions about whether the respondents have experience in requesting and/or performing EUS were asked, and if they do, they were asked to estimate the number of EUS requested and/or performed per 3-month period. In the fourth section, questions relating to their current use of EUS, reasons where EUS may be limited, the perceived usefulness and availability of EUS, were asked. The perceived usefulness of staging with EUS was assessed by asking the respondents to rate on a five-level Likert scale. The availability of EUS within respondents' hospital was assessed; if EUS is unavailable within their hospital, they were asked as to which hospital GOC patients were usually referred to for EUS scans, and whether it was a teaching hospital. In this section, they were also asked to estimate the total number of gastro-oesophageal malignancies (oesophageal, gastro-oesophageal junction and gastric cancers) cases seen, and the number referred for EUS. Finally, in the fifth section, a series of pre-validated clinical vignettes (Kim and Koch, 1999) was presented to assess how do different clinicians manage different situations of GOC patients based upon small changes in tumour stage.

In accordance with the good practice for the design and application of online questionnaire surveys described in Regmi et al. (2016), a pilot study was carried out with potential participants (MDT GOC-interested clinicians) to check that the questions and instructions were clear, adequate and in the right sequence, as well as the contents were comprehensive and clear. By piloting the survey questionnaire, it allows experts in the field to review the survey and provide their feedback (Regmi et al., 2016; van Teijlingen and Hundley, 2002). Our

pilot study showed that the online survey questionnaire took approximately 10 minutes to complete.

4.3.5 Ethics

Ethical approval for this survey study was granted from the Bangor University's School of Healthcare Sciences Ethics and Research Committee (study's reference number: 2017-16023, 20 June 2017, see Appendix 4.3). National Health Service Research Ethics Committee (NHS REC) approval and Research and Development (R&D) permission were not required as this survey study was classified as a service evaluation study. However, this survey study was registered with five out of the seven Welsh Health Boards' Clinical Audit Office, as required: Betsi Cadwaladr University Health Board (BCUHB), Hywel Dda University Health Board (HDUHB), Cardiff and Vale University Health Board (CVUHB), Powys Teaching Health Board (PTHB) and Cwm Taf University Health Board (CTUHB). Whereas, the remaining two Welsh Health Boards did not require the need to register the survey study with their Clinical Audit Office: Abertawe Bro Morgannwg University Health Board (ABMUHB) (this has now changed its name to "Swansea Bay University Health Board") and Aneurin Bevan University Health Board (ABUHB).

4.3.6 Statistical analyses

Statistical analyses were performed using the IBM SPSS Statistics 25 software package (IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp., 2017). Descriptive statistics analyses were performed whereby number and percentage, or mean, standard deviation, median, interquartile range and range were reported, as appropriate. In addition, an exploratory logistic regression analysis was undertaken to determine factors associated with referral for EUS for GOC staging. Based on the evidence from the published US survey study by Kim and Koch (1999) and guided by clinical inputs from an UK-based GOC expert clinician (also the clinical advisor on my supervisory team), several factors were proposed as potential predictors of referral for EUS for GOC staging in clinical practice. Factors considered for inclusion in the logistic regression model were: (1) clinical management style (aggressive vs. conservative), (2) usefulness of EUS (somewhat/very useful/essential vs. slightly/not useful), (3) availability of EUS (yes vs. no), (4) attend MDT meeting (yes vs. no), (5) age, and (6) years in practice. Clinical management style was derived from respondents' responses to the pre-

validated clinical vignette describing the different EUS staging results of a patient with heartburn and is found to have a 2cm diameter distal oesophageal adenocarcinoma. As guided by Kim and Koch (1999), respondents who chose curative surgery as a method of management for stage III (T3N1) oesophageal cancer (i.e. mass penetrates all wall layers with two round 1cm peri-oesophageal node) were considered to have an aggressive management style. For the variable 'usefulness of EUS', it was dichotomised into the two groups (i.e. somewhat/very useful/essential vs. slightly/not useful) as guided by Kim and Koch (1999).

Independent samples T-test and Chi-square test were used for continuous variables and categorical variables, respectively, as appropriate. For the two continuous variables i.e. 'Age' and 'Years in Practice', firstly, normality test was carried out to examine whether an independent samples T-test (a parametric test) or a Mann-Whitney U test (a non-parametric equivalent to the independent samples T-test) should be employed. Results showed that both of the continuous variables ('Age' and 'Years in Practice') are normally distributed ($p>0.05$); thus, an independent sample T-test (a parametric test) was undertaken to examine whether there is a statistically significant difference in means of 'Age' and 'Years in Practice' between groups (somewhat/very useful/essential vs. slightly/not useful). For the four categorical variables (i.e. 'clinical management', 'perceived usefulness of EUS', 'availability of EUS' and 'attend MDT meeting'), Chi-square test was performed to examine association between groups (i.e. 'perceived usefulness of EUS' vs. other categorical variables).

For the exploratory logistic regression analysis, both univariate and multivariate logistic regression analyses were performed. Univariate logistic regression analysis was undertaken to assess whether or not there was a significant association between each of the six factors (independent variables) and the factor 'referral for EUS' (dependent variable). Taking account of confounding factors, multivariate logistic regression analysis was performed using a backward stepwise method to derive a multivariate model. To make sure that none of the confounding factors was erroneously removed from the final model, removed variables were added back individually into the final model and the odds ratios for statistically significant variables were assessed. If the odds ratios for statistically significant variables changed by more than 10%, the removed variables were added back and forced into the final model. The significance level at $p=0.05$ was used for all statistical comparisons.

4.4 Results

4.4.1 Survey responses

Online survey questionnaires were completed and submitted by 98 respondents who were invited through two UK's Gastroenterology professional bodies – BSG and WAGE. The minimum sample suggested by the calculations was 100 – this would achieve a margin of error of approximately 8%, 98 responses were achieved.

4.4.2 Descriptive statistics

Descriptive statistics analyses results were reported for all five sections of the survey questionnaire:

(A) Clinicians' demographic details
(B) Clinicians' knowledge and feelings about the use of EUS
(C) Clinicians' experience in the field of EUS i.e. request and/or perform EUS
(D) Clinicians' current use of EUS and reasons where this may be limited, the usefulness and availability of EUS
(E) Clinicians' clinical management

4.4.2.1 Section A: Clinicians' demographic details (n=98)

In terms of geographical spread of the survey, most (n = 69; 70.4%) were from England, 17 (17.3%) from Wales, 6 (6.1%) from Scotland, 3 (3.1%) from Northern Ireland and 3 (3.1%) had missing information. The majority (n = 66; 67.3%) heard about this survey through BSG email, with the next highest through a WAGE email (n=11; 11.2%).

The study sample was predominantly male (n = 76; 77.6%) with a mean age of 46.6 years (SD = 9.4), 17 (17.3%) were female with a mean age of 43.6 years (SD=7.4) (Table 4.1). Of the 98 respondents, the majority (n=79; 80.6%) were gastroenterologists and the remaining were GI surgeons (n = 13; 13.3%), oncologists (n = 1; 1.0%), and radiologists (n = 3; 3.1%). More than 85% of the study sample (n=85; 86.7 %) were members of the British Society of Gastroenterology (BSG) with 11 having more than one membership of other medical societies including WAGE, or Association of Upper Gastrointestinal Surgeons (AUGIS), British Society of Gastrointestinal and Abdominal Radiology (BSGAR), UKEUS, Royal College of Radiologists (RCR) and others. Fourteen (14.2%) were a member of WAGE with 10 of these also a member of other professional bodies such as BSG, Association of Surgeons of Great Britain and Ireland

(ASGBI), United European Gastroenterology (UEG) and British Association of Cancer Research (BACR).

The survey respondents have practiced in their field for an average of 21.7 (SD = 9.2) years. Of the 98 respondents, the majority (n=72; 73.5%) have regular (i.e. weekly) exposure and contact with gastro-oesophageal cancer (GOC) patients; slightly more than half (n=54; 55.1%) are based at Teaching hospitals and the remaining 42 respondents (42.9%) at District General hospitals. Over 90% (n=89; 90.8%) reported that their hospital does run an Upper GI cancer Multi-Disciplinary Team (MDT) meeting. Of the 98 study sample, the majority (n=73; 74.5%) attended an Upper GI MDT meeting.

Table 4.1: Demographic characteristics of the survey sample (n=98).

Characteristics	Number (%)
Responses by region	
England	69 (70.4)
Wales	17 (17.3)
Scotland	6 (6.1)
Northern Ireland	3 (3.1)
Missing	3 (3.1)
Heard about this survey via….	
BSG email	66 (67.3)
BSG Newsletter	7 (7.1)
BSG website	4 (4.1)
BSG website + BSG email	2 (2.0)
WAGE email	11 (11.2)
Other (such as colleague and Twitter)	7 (7.1)
Missing	1 (1.0)
Gender	
Male	76 (77.6)
Female	17 (17.3)

Missing	5 (5.1)
Age (years): mean (SD); range	
Overall (n = 94, n = 4 missing)	46.2 (9.1); 32–65
Male (n = 74, n = 2 missing)	46.6 (9.4); 32–65
Female (n = 16, n = 1 missing)	43.6 (7.4); 32–57
Which of the following best describes you?	
Gastroenterologist	79 (80.6)
GI Surgeon	13 (13.3)
Oncologist	1 (1.0)
Radiologist	3 (3.1)
Missing	2 (2.0)
You are a member of...	
BSG	74 (75.5)
WAGE	4 (4.1)
BSG + WAGE	7 (7.1)
BSG + other	3 (3.1)
WAGE + other	2 (2.0)
All – BSG+WAGE+Other	1 (1.0)
Other (e.g. AUGIS, UK EUS User etc)	7 (7.1)
Years in practice since medical qualification (n=97): mean (SD); range	21.7 (9.2); 7–42
Which of the following best describes your exposure and contact with gastro-oesophageal cancer patient?	
Yes – Regularly (i.e. weekly)	72 (73.5)
Yes – Sometimes (i.e. monthly)	20 (20.4)
Yes – Rarely (i.e. yearly)	3 (3.1)
Yes – But hardly ever (Longer than yearly)	1 (1.0)
No – Never at all	0 (0.0)

Missing	2 (2.0)
Is your current hospital a.....	
Teaching Hospital	54 (55.1)
District General Hospital	42 (42.9)
Missing	2 (2.0)
Does your hospital run an Upper GI cancer MDT meeting?	
Yes	89 (90.8)
No	5 (5.1)
Missing	4 (4.1)
Do you attend an Upper GI cancer MDT meeting?	
Yes	73 (74.5)
No	23 (23.5)
Missing	2 (2.0)

Figures are numbers (percentages) unless stated otherwise.
SD, standard deviation; MDT, Multi-Disciplinary Team

4.4.2.2 Section B: Clinicians' knowledge and feelings about the use of EUS (n=98)

Table 4.2 shows the results of the questions asking the respondents about their knowledge and feelings about the use of EUS. Eighty-six respondents (87.8%) reported that they have experience in the staging of GOC. Oesophageal cancer was the type of cancer that the majority (n=91; 92.9%) think EUS can accurately stage, followed by gastro-oesophageal junction cancer (n=84; 85.7%), gastric cancer (n=36; 36.7%) and other lesions (n=10; 10.2%).

Of the 98 survey sample, 70 (71.4%) respondents agreed that EUS helps to reduce unnecessary surgery for GOC patients, 12 (12.2%) did not agree and 16 (16.3%) were not sure. Of those (n=70) that agreed EUS helps, just over 40% in total (n=29; 41.4%) estimated that EUS helps somewhat a lot (n=25; 35.7%) or a lot (n=4; 5.7%) in reducing unnecessary surgery, however more than half in total (n=41; 58.6%) estimated that EUS helps somewhat little (n=34; 48.6%) or very little (n=7; 10.0%).

Of the 98 respondents, most (n=42; 42.9%) reported that, if EUS is not currently an option used for staging GOC patients and if it was to become available to them, it is likely that they will choose to use EUS for staging GOC patients in the future. Interestingly, almost all of the 98 respondents (n=89; 90.8%) reported that they support the use of EUS for staging GOC, mainly for the reason that EUS offers the best local regional staging of GOC (n=77; 86.5%).

When asked about the published COGNATE HTA report, just about one-third of the 98 respondents (n=28; 28.6%) reported that they are aware of the report and most of them (n=22; 78.6%) have read it. For all of those who are aware but have not read the report (n=6; 21.4%) noted that it is likely that they will read the report. And for those who are not aware of the published COGNATE HTA report (n=70, 71.4%), most of these respondents (n=53; 75.7%) reported that it is likely that they will read the report.

Lastly, when the 98 respondents were asked if they have any experience in the field of EUS i.e. request and/or perform EUS, the majority (n=81; 82.7%) responded that they have. These 81 respondents were then asked to estimate the total number of EUS they requested and/or performed per 3-month period, questions related to their current use of EUS and reasons where this may be limited, the usefulness and availability of EUS, and also questions related to their clinical management.

Table 4.2: Clinician's knowledge and feelings about the use of EUS (n=98).

Characteristics	Number (%)
Do you have experience in the staging of GOC?	
Yes	86 (87.8)
No	12 (12.2)
Which of the following can EUS accurately stage? (Tick all that apply)	
Oesophageal cancer	91 (92.9)
Gastro-oesophageal junction cancer	84 (85.7)
Gastric cancer	36 (36.7)
Other lesion (e.g. not specify)	10 (10.2)
Does EUS help to reduce unnecessary surgery for GOC patients?	
Yes	70 (71.4)
No	12 (12.2)
I am not sure	16 (16.3)
If yes, by how much would you estimate EUS reduces unnecessary surgery?	
Very little	7 (10.0)
Somewhat little	34 (48.6)
No difference	0 (0.0)
Somewhat a lot	25 (35.7)
A lot	4 (5.7)
If EUS is not currently an option used for staging GOC patients and if it was to become available to you, how likely will you choose to use EUS for staging GOC patients in the future?	
Not applicable	45 (45.9)
Never	0 (0.0)
Not likely	5 (5.1)

Somewhat likely	18 (18.4)
Very likely	13 (13.3)
Always	11 (11.2)
Missing	6 (6.1)
Do you support the use of EUS for staging GOC?	
Yes	89 (90.8)
No	9 (9.2)
If yes, why? (Tick all that apply)	
EUS offers the best local regional staging of GOC	77 (86.5)
EUS offers opportunity to biopsy during staging	57 (64.0)
Cost-effectiveness of using EUS to stage GOC	29 (32.6)
Other (e.g. superficial cancers can be removed by Endoscopic techniques)	5 (5.6)
If no, why? (Tick all that apply)	
Lack of expertise in performing EUS	2 (22.2)
Inter-operator variability in interpreting images	5 (55.6)
Cost-effectiveness of using other staging modalities to stage GOC	2 (22.2)
Other (e.g. Failure to obtain the required information due to structuring disease in the oesophagus)	5 (55.6)
Are you aware of the published UK COGNATE trial HTA report?	
Yes	28 (28.6)
If yes, have you read the report?	
Yes	22 (78.6)
No	6 (21.4)
If no, how likely is it that you will read the report?	
Not likely	0 (0.0)
Somewhat likely	1 (16.7)
Very likely	5 (83.3)

No	70 (71.4)
If no, how likely is it that you will read the report?	
Not likely	17 (24.3)
Somewhat likely	28 (40.0)
Very likely	25 (35.7)
Do you have any experience in the field of EUS (i.e. request and/or perform EUS)?	
Yes	81 (82.7)
No	17 (17.3)

Figures are numbers (percentages) unless stated otherwise.

4.4.2.3 Section C: Clinicians' experience in the field of EUS i.e. by requesting and/or performing EUS (n=81)

Table 4.3 shows that 71 out of 81 respondents who have experience in the field of EUS are EUS requesters; these 71 respondents estimated that, on average, they request 12.2 (SD 17.2) EUS per 3-month period. Of the 81 respondents, 38 are EUS performers; these 38 respondents estimated that, on average, they perform 40.6 (SD 30.1) EUS every 3 months.

Table 4.3: Clinician's experience in the field of EUS i.e. by requesting and/or performing EUS (n=81).

Clinician's experience in the field of EUS (i.e. by requesting and/or performing EUS)	Number (%)
What best describes your experience in the field of EUS?	
I request EUS	40 (49.4)
I perform EUS	9 (11.1)
I both request and perform EUS	28 (34.6)
I do not regularly request EUS and do not perform EUS	3 (3.7)
I do not request EUS and do not regularly perform EUS	1 (1.2)
Estimate how many EUS are requested per 3-month period (n=71): mean (SD); IQR; range	12.2 (17.2); 2.5-13.5; 0-90
Estimate how many EUS are performed per 3-month period (n=38): mean (SD); IQR; range	40.6 (30.1); 17.5-55.0; 0-110

Figures are numbers (percentages) unless stated otherwise.
SD, standard deviation; IQR, interquartile range

4.4.2.4 Section D: Clinicians' current use of EUS and reasons where this may be limited, the usefulness and availability of EUS (n=81)

Table 4.4 shows the descriptive statistics results of the questions asking the respondents about their current use of EUS and reasons where this may be limited, the usefulness and availability of EUS.

The majority (n=47; 58.0%) reported that they do not routinely request EUS for staging all newly diagnosed GOC patients. Likewise, the majority (n=36; 44.4%) are not likely to request EUS for suspected relapse of GOC cancer. Two-thirds (n=54; 66.7%) reported that they do have a standard hospital protocol regarding the use of EUS for staging GOC to follow. In terms of what proportion of EUS requests are made through the Upper GI Cancer MDT in or outside their practice, most of the respondents reported that majority or all EUS requests are made through MDT in their practice (total n (%) = 74 (91.4%)) and outside their practice (total n (%) = 53 (45.5%)). About 80% (n=65) reported that there are specific situations where they don't use EUS. Advanced metastatic disease (n=57; 87.7%) was reported to be the main specific situation followed by patient morbidity (n=39; 60.0%), tight oesophageal stricture (n=38; 58.5%), other situations e.g. advanced local disease (n=10; 15.4%) and availability of EUS (n=5; 7.7%). When the respondents were asked what the reasons behind request and utilisation of other imaging modalities are compared to EUS in GOC assessment, availability was the reason mostly chosen by the respondents (n=29; 44.6%) (see Table 4.4).

In terms of respondents' thoughts on how useful staging is with EUS in assisting their clinical management in patients newly diagnosed with GOCs (i.e. oesophageal cancer, gastro-oesophageal junction cancer, gastric cancer), gastric cancer had the highest number of respondents who reported that EUS was not considered useful (n=23; 28.4%) compared to oesophageal cancer (n=3; 3.7%) and gastro-oesophageal junction (GOJ) cancer (n=3, 3.7%). A total number of 48 (59.2%), 36 (44.4%) and 34 (41.9%) respondents thought that EUS is slightly/somewhat useful for gastric cancer, GOJ cancer and oesophageal cancer, respectively. Oesophageal cancer had the majority of responses (n=44, 54.4%) reporting that EUS is very useful/essential, followed by GOJ cancer (n=42, 51.8%) and gastric cancer (n=10, 12.3%).

Regarding the availability of EUS, the majority (59 out of 81, 72.8%) reported that EUS is available within their hospital. Of these 59 respondents, 39 (66.1%) are practicing at a

Teaching Hospital and 19 (32.2%) at a District General Hospital. Whereas, 21 out of 81 (25.9%) respondents reported that EUS is not available within their hospital and remaining 1 respondent (1.2%) had missing data. Of these 21 respondents, 15 (71.4%) are practicing at a District General Hospital and 5 (23.8%) at a Teaching Hospital. These 21 respondents estimated that the average distance between their hospital and the hospital to which they usually refer their patients for EUS is 21.4 miles (SD 21.7 miles). Of these 21 respondents, almost all (n=19, 90.5%) reported that the hospital for which their patients are usually referred to for EUS screening is a teaching hospital.

In addition, all of the 81 respondents were asked to estimate the total number of patients seen and approximately how many of these patients seen underwent EUS in the past 12 months for each type of the GOCs. On average, most of the oesophageal and GOJ cancer patients seen had EUS: Out of the estimated average of 30.6 oesophageal cancer patients seen, 17.5 (57.2%) had EUS; whereas for GOJ cancer patients, 11.6 out of 19.3 (60.1%) patients seen underwent EUS. For gastric cancer patients, only a few had EUS (2.53 out of 20.0 (12.7%) patients seen had EUS).

Table 4.4: Clinicians' current use of EUS and reasons where this may be limited, the usefulness and availability of EUS (n=81).

Clinicians' current use of EUS and reasons where this may be limited, the usefulness and availability of EUS	Number (%)
(A) <u>Clinician's current use of EUS and reasons where this may be limited</u>	
Do you routinely request EUS for staging all newly diagnosed GOC patients?	
Not applicable (I am unable to request EUS)	2 (2.5)
Yes	32 (39.5)
No	47 (58.0)
How likely are you to request EUS for suspected relapse of GOC cancer?	
Not applicable (I am unable to request EUS)	2 (2.5)
Never	4 (4.9)
Not likely	36 (44.4)

Somewhat likely	24 (29.6)
Very likely	9 (11.1)
Always	0 (0.0)
Missing	6 (7.4)

Do you have a standard hospital protocol regarding the use of EUS for staging GOC to follow?

Yes	54 (66.7)
No	26 (32.1)
Missing	1 (1.2)

What proportion of EUS requests are made through the Upper GI Cancer MDT in your practice or outside?

In your practice

All in MDT	37 (45.7)
Majority in MDT	37 (45.7)
Approximately equal requests through MDT and outside MDT	4 (4.9)
Majority outside MDT	2 (2.5)
All outside MDT	1 (1.2)

Outside your practice

All in MDT	22 (7.2)
Majority in MDT	31 (38.3)
Approximately equal requests through MDT and outside MDT	4 (4.9)
Majority outside MDT	3 (3.7)
All outside MDT	0 (0.0)
Missing	21 (25.9)

Are there specific situations where you don’t use EUS?

Yes	65 (80.2)
No	16 (19.8)

If yes…

a) What are the specific situations where you don't use EUS? (n=65)

Availability	5 (7.7)
Tight oesophageal stricture	38 (58.5)
Patient morbidity	39 (60.0)
Advanced metastatic disease	57 (87.7)
Other (e.g. advanced local disease etc.)	10 (15.4)

b) What are the reasons behind request and utilisation of other imaging modalities compared to EUS in GOC assessment? (n=65)

Availability	29 (44.6)
Cost	8 (12.3)
Expertise	17 (26.2)
Resources e.g. staffing constraints and etc.	15 (23.1)
Time taken to perform EUS procedure	10 (15.4)
Patient comfort	10 (15.4)
Other (e.g. when EUS cannot technically be performed etc.)	27 (41.5)

(B) <u>Usefulness of EUS</u>

In patients newly diagnosed with following gastro-oesophageal malignancies, how useful is staging with endoscopic ultrasound in assisting your clinical management?

a) Oesophageal cancer

Not useful	3 (3.7)
Slightly useful	10 (12.3)
Somewhat useful	24 (29.6)
Very useful	22 (27.2)
Essential	22 (27.2)

b) Gastro-oesophageal junction (GOJ) cancer

Not useful	3 (3.7)
Slightly useful	7 (8.6)

Somewhat useful	29 (35.8)
Very useful	24 (29.6)
Essential	18 (22.2)

c) Gastric cancer

Not useful	23 (28.4)
Slightly useful	30 (37.0)
Somewhat useful	18 (22.2)
Very useful	6 (7.4)
Essential	4 (4.9)

(C) <u>Availability of EUS</u>

Is EUS available within your hospital?

Yes	59 (72.8)
No	21 (25.9)
Missing	1 (1.2)

If no…

a) Please estimate how far (in miles) the hospital is situated from your hospital for which you usually refer your patients to for EUS screening (n=21): Mean (SD); IQR; range	21.4 (21.7); 6–30; 1–80

b) For the hospital that patients are usually referred to for EUS screening, is this a teaching hospital? (n=21)

Yes	19 (90.5)
No	2 (9.5)

In the past 12 months, approximately how many patients with each of the following gastro-oesophageal malignancies did you see in total, and approximately how many of these patients underwent endoscopic ultrasound?

a) Oesophageal cancer –

Total patients seen: Mean (SD); IQR; range	30.6 (35.8); 10-35; 0-200
Had EUS: Mean (SD); IQR; range	17.5 (19.7); 4-22.5; 0-89

b) GOJ cancer –

Total patients seen: Mean (SD); IQR; range	19.3 (22.5); 4-30; 0-100
Had EUS: Mean (SD); IQR; range	11.6 (12.6); 2-20; 0-50

c) Gastric cancer –

Total patients seen: Mean (SD); IQR; range	20.0 (31.8); 4-20; 0-200
Had EUS: Mean (SD); IQR; range	2.53 (3.9); 0-4.3; 0-20

Figures are numbers (percentages) unless stated otherwise.
SD, standard deviation; IQR, inter-quartile range; MDT, multi-disciplinary team.

4.4.2.5 Section E: Clinicians' patient management (n=81)

In the pre-validated clinical vignettes, adapted from Kim and Koch (1999), we assessed how clinicians would treat the simulated patient following EUS staging i.e. whether they would consider very radical treatment like surgery versus other treatment options. Table 4.5 shows that the majority chose to refer the simulated patient for oesophagectomy either with or without chemotherapy/radiotherapy for all of the three pre-validated clinical vignettes presented to them: (1) mass penetrates all wall layers (T3; CRUK, 2019) with two round 1cm peri-oesophageal node (N1; CRUK, 2019) (n=56, 69.1%); (2) mass penetrates all wall layers (T3; CRUK, 2019) but no nodes are seen (N0; CRUK, 2019) (n=66, 81.5%); and (3) mass is confined to muscularis (T2; CRUK, 2019), no nodes are seen (N0; CRUK, 2019) (n=75, 92.6%).

For the hypothetical scenario showing the simulated patient had a stage III (T3N1) oesophageal cancer (i.e. mass penetrates all wall layers with two round 1cm peri-oesophageal node), respondents who chose curative surgery as a method of management were considered to have an aggressive management style as guided by Kim and Koch (1999) and we evaluated this factor 'aggressive management style' in the following exploratory regression analyses section to assess whether it was also a predictor of preference for EUS referral.

Table 4.5: Clinicians' patient management (n=81).

Clinician's patient management based on a series of pre-validated clinical vignette	Number (%)
A 58 year-old homemaker presents with heartburn and is found to have a 2cm diameter distal oesophageal adenocarcinoma. Staging endoscopic ultrasound is performed. What treatment would you recommend for each of the following EUS results?	
a) Mass penetrates all wall layers with two round 1cm peri-oesophageal node:	
• Referral for oesophagectomy (with or without chemo/XRT)	56 (69.1)
• Referral for chemotherapy and/or radiotherapy without surgery	24 (29.6)
• No treatment, with endoscopic palliation for future symptoms	1 (1.2)

b) Mass penetrates all wall layers but no nodes are seen:

• Referral for oesophagectomy (with or without chemo/XRT)	66 (81.5)
• Referral for chemotherapy and/or radiotherapy without surgery	15 (18.5)
• No treatment, with endoscopic palliation for future symptoms	0 (0.0)

c) Mass is confined to muscularis, no nodes are seen:

• Referral for oesophagectomy (with or without chemo/XRT)	75 (92.6)
• Referral for chemotherapy and/or radiotherapy without surgery	5 (6.2)
• No treatment, with endoscopic palliation for future symptoms	0 (0.0)
Missing	1 (1.2)

Figures are numbers (percentages) unless stated otherwise.

4.4.3 Exploratory logistic regression analyses

In this section, firstly, demographic characteristics of the study sample (n=81) used in the exploratory logistic regression analyses was reported. Next, independent samples t-tests and Chi-square tests results for all the factors considered for inclusion in the model were reported. Finally, the exploratory logistic regression analyses results with respect to potential factors associated with referral for EUS were presented for each type of the gastro-oesophageal cancers: oesophageal cancer, gastro-oesophageal junction cancer, gastric cancer and gastro-oesophageal cancers as a whole.

4.4.3.1 Demographic characteristics of the logistic regression analyses study sample (n=81)

Table 4.6 shows the demographic characteristics of the 81 respondents who reported that they have had experience in requesting and/or performing EUS. In terms of geographical spread of the survey, these respondents were mostly from England (n = 58, 71.6%), followed by 13 (16.0%) from Wales, 5 (6.2%) from Scotland, 2 (2.5%) from Northern Ireland and 3 (3.7%) did not provide an answer. The majority heard about the survey through BSG email, and the next most was through WAGE email (n = 9, 11.1%).

The study sample was predominantly male (n = 66, 81.5%) with a mean age of 46.8 years (SD = 9.2), 12 (14.8%) were female with a mean age of 44.6 years (SD=7.4) (Table 4.6). Of the 81 respondents, the majority (n=64, 79.0%) were gastroenterologists and the remaining were GI surgeons (n = 12, 14.8%), oncologist (n = 1, 1.2%), and radiologists (n = 3, 3.7%). More than 80% (n=69, 85.2%) were a member of British Society of Gastroenterology (BSG) with 7 of these also a member of other professional bodies such as Welsh Association for Gastroenterology and Endoscopy (WAGE) etc; Ten (12.3%) were a member of WAGE with 6 of these also a member of other professional bodies such as BSG etc.

These 81 survey respondents have practiced in their field for an average of 22.5 (SD = 9.0) years. The majority (n=64, 79.0%) have regular (i.e. weekly) exposure and contact with gastro-oesophageal cancer (GOC) patients; slightly more than half (n=45, 55.6%) are based at Teaching hospitals and the remaining 34 respondents (42.0%) at District General hospitals. Nearly 90% (n=72, 88.9%) reported that their hospital does run an Upper GI cancer Multi-Disciplinary Team (MDT) meeting; with the majority (n=66, 81.5%) reporting that they attend an Upper GI MDT meeting.

Table 4.6: Demographic characteristics of the survey sample used in exploratory logistic regression analyses (n=81).

Characteristic	Number (%)
Responses by region	
England	58 (71.6)
Wales	13 (16.0)
Scotland	5 (6.2)
Northern Ireland	2 (2.5)
Missing	3 (3.7)
Heard about this survey via….	
BSG email	54 (66.7)
BSG Newsletter	6 (7.4)
BSG website	4 (4.9)

BSG website + BSG email	2 (2.5)
WAGE email	9 (11.1)
Other (such as colleague and Twitter)	5 (6.2)
Missing	1 (1.2)
Gender	
Male	66 (81.5)
Female	12 (14.8)
Missing	3 (3.7)
Age (years): mean (SD); range	
Overall (n = 78, n = 3 missing)	46.7 (8.9); 32-65
Male (n = 65, n = 1 missing)	46.8 (9.2); 32-65
Female (n = 11, n = 1 missing)	44.6 (7.4); 33-57
Which of the following best describes you?	
Gastroenterologist	64 (79.0)
GI Surgeon	12 (14.8)
Oncologist	1 (1.2)
Radiologist	3 (3.7)
Missing	1 (1.2)
You are a member of...	
BSG	62 (76.5)
WAGE	4 (4.9)
BSG + WAGE	3 (3.7)
BSG + other	3 (3.7)
WAGE + other	2 (2.5)
All – BSG+WAGE+Other	1 (1.2)
Other (e.g. AUGIS, UK EUS User etc)	6 (7.4)
Years in practice since medical qualification (n=81): mean (SD); range	22.5 (9.0); 8-42

Which of the following best describes your exposure and contact with gastro-oesophageal cancer patient?	
Yes – Regularly (i.e. weekly)	64 (79.0)
Yes – Sometimes (i.e. monthly)	12 (14.8)
Yes – Rarely (i.e. yearly)	2 (2.5)
Yes – But hardly ever (Longer than yearly)	1 (1.2)
No – Never at all	0 (0.0)
Missing	2 (2.5)
Is your current hospital a.....	
Teaching Hospital	45 (55.6)
District General Hospital	34 (42.0)
Missing	2 (2.5)
Does your hospital run an Upper GI cancer MDT meeting?	
Yes	72 (88.9)
No	5 (6.2)
Missing	4 (4.9)
Do you attend an Upper GI cancer MDT meeting?	
Yes	66 (81.5)
No	13 (16.0)
Missing	2 (2.5)

Figures are numbers (percentages) unless stated otherwise.
SD, standard deviation; MDT, Multi-Disciplinary Team

4.4.3.2 Independent Sample T-Test Results

For gastro-oesophageal junction and gastro-oesophageal cancers, independent sample T-test results (see Table 4.7) showed that there was no statistically significant difference ($p>0.05$) in the mean of 'age' and 'Years in Practice' between 'usefulness of EUS' groups (i.e.

'somewhat/very useful/essential' vs 'slightly/not useful' groups). However, for oesophageal cancer, results showed that there was statistically significant different ($p<0.05$) in the mean of 'Years in Practice' between 'usefulness of EUS' groups, but no significance difference ($p>0.05$) in the mean of 'age' between groups (see Table 4.7). For gastric cancer, both the mean of 'Age' and 'Years in Practice' were statistically significant different ($p<0.05$) between the two groups.

4.4.3.3 Chi-Square Test Results

Chi-square test results showed that there were no significant associations ($p>0.05$) between 'perceived usefulness of EUS' and the other categorical variables (i.e. 'clinical management style', 'availability of EUS' and 'attend MDT meeting') for all types of gastro-oesophageal cancers (i.e. oesophageal cancer, gastro-oesophageal junction cancer, gastric cancer and gastro-oesophageal cancers) (see Table 4.8 to Table 4.19).

4.4.3.4 Exploratory Logistic Regression Analyses Results

To determine as to what factors associated with 'referral for EUS', for oesophageal cancer, results showed that 'attend MDT meeting' was the only factor that demonstrated significant ($p<0.05$) association with 'referral for EUS' for both the univariate ($p = 0.017$) and multivariate logistic regression ($p=0.013$). From the multivariate logistic regression, this significant factor 'attend MDT meeting' had an odds ratio of 25.214 (see Table 4.20) – This implied that for those clinicians who attend MDT meeting, the odds of their oesophageal cancer patients being referred for EUS is 25.214 times higher than the odds of those not attending MDT meeting.

For gastro-oesophageal junction cancer, four out of six factors demonstrated significant association with 'referral for EUS' in the univariate logistic regression (see Table 4.21). These four factors are: 'attend MDT meeting' ($p=0.009$), 'perceived usefulness of EUS' ($p=0.040$), 'EUS available' ($p=0.041$) and 'aggressive management style' ($p=0.025$). Whereas in the multivariate logistic regression, results showed that 'attend MDT meeting' and 'aggressive management style' are the only two factors that demonstrated significant association with 'referral for EUS' with a p-value of 0.014 and 0.037, respectively. Both of these factors, 'attend MDT meeting' and 'aggressive management style', had an odds ratio of 14.593 and 11.866, respectively – This implied that for clinicians who 'attend MDT meeting', the odds of their

gastro-oesophageal junction cancer patients being referred for EUS is 14.593 times higher than the odds of those not attending MDT meeting; For clinicians who have 'aggressive management style', the odds of their gastro-oesophageal junction cancer patients being referred for EUS is 11.866 times higher than the odds of those without an aggressive management style.

For gastric cancer, univariate logistic regression results showed that none of the six examined factors had significant association with 'referral for EUS'. However, in the multivariate regression, 'age' appeared to show as the only factor that had association with 'referral for EUS' but this association was not significant ($p=0.063 > 0.05$) (see Table 4.22).

For gastro-oesophageal cancer (GOC), univariate logistic regression results showed that 'attend MDT meeting' is the only factor that had significant association with 'referral for EUS' ($p=0.005$) (see Table 4.23). Whereas, in the multivariate logistic regression, 'age' and 'attend MDT meeting' are the two factors that showed significant association with 'referral for EUS' with a p-value of 0.041 and 0.033, respectively. These factors, 'age' and 'attend MDT meeting', had an odds ratio of 0.606 and 990.733, respectively (see Table 4.23). This indicated that for younger clinicians, the odds of their gastro-oesophageal cancer patients being referred for EUS is 0.606 times lower than the odds of older clinicians referring their gastro-oesophageal cancer patients for EUS; While for those clinicians who 'attend MDT meeting', the odds of their gastro-oesophageal cancer patients being referred for EUS is 990.733 times higher than the odds of those not attending MDT meeting. Furthermore, in relation to the factor 'hospital type' (Teaching Hospital vs. District General Hospital (DGH)), both the multivariate logistic regression and univariate logistic regression were re-run to include the factor 'hospital type', this additional analysis result showed that the factor 'hospital type' had no significant association ($p>0.05$) with referral for EUS for GOC staging with a p-value of 0.998 and 0.728, respectively.

Table 4.7: Independent sample T-test results for oesophageal, gastro-oesophageal junction, gastric and gastro-oesophageal cancers study sample (N=81).

	Perceived Usefulness of EUS groups		
Oesophageal Cancer	Group 1 (n=13) Mean (SD)	Group 2 (n≤68) Mean (SD)	p-value*
Age (years), n=78	42.85 (8.51)	47.48 (8.86)	0.087
Years in Practice, n=81	17.38 (8.12)	23.49 (8.86)	0.024*
Gastro-Oesophageal Junction Cancer	Group 1 (n≤10) Mean (SD)	Group 2 (n≤71) Mean (SD)	p-value*
Age (years), n=78	45.89 (8.05)	46.81 (9.07)	0.772
Years in Practice, n=81	19.90 (8.10)	22.87 (9.10)	0.330
Gastric Cancer	Group 1 (n≤53) Mean (SD)	Group 2 (n≤28) Mean (SD)	p-value*
Age (years), n=78	48.22 (8.13)	43.85 (9.77)	0.039*
Years in Practice, n=81	24.08 (8.59)	19.54 (9.11)	0.030*
Gastro-Oesophageal Cancers (Oesophageal, Gastro-Oesophageal Junction, and Gastric Cancers)	Group 1 (n≤19) Mean (SD)	Group 2 (n≤62) Mean (SD)	p-value*
Age (years), n=78	46.78 (8.02)	46.68 (9.23)	0.969
Years in Practice, n=81	21.68 (8.42)	22.76 (9.20)	0.651

*Significance level at p=0.05

Group 1= Slightly/Not useful; Group 2= Somewhat/Very useful/Essential

Table 4.8: Association between management style and perceived usefulness of EUS for Oesophageal Cancer (N=81).

Oesophageal Cancer			
Management	Perceived usefulness of EUS		
Style	Slightly/Not Useful	Somewhat/Very Useful/Essential	Total
Conservative	4 (16.0%)	21 (84.0%)	25
Aggressive	9 (16.1%)	47 (83.9%)	56
Total	13 (16.0%)	68 (84.0%)	81

Management style Vs. Perceived usefulness of EUS, Pearson Chi-Square = 0.000 (p=0.994)

Table 4.9: Association between management style and perceived usefulness of EUS for Gastro-Oesophageal Junction Cancer (N=81).

Gastro-Oesophageal Junction Cancer			
Management	Perceived usefulness of EUS		
Style	Slightly/Not Useful	Somewhat/Very Useful/Essential	Total
Conservative	5 (20.0%)	20 (80.0%)	25
Aggressive	5 (8.9%)	51 (91.1%)	56
Total	10 (12.3%)	71 (87.7%)	81

Management style Vs. Perceived usefulness of EUS, Pearson Chi-Square = 1.958 (p=0.162)

Table 4.10: Association between management style and perceived usefulness of EUS for Gastric Cancer (N=81).

Gastric Cancer			
Management	Perceived usefulness of EUS		
Style	Slightly/Not Useful	Somewhat/Very Useful/Essential	Total
Conservative	15 (60.0%)	10 (40.0%)	25
Aggressive	38 (67.9%)	18 (32.1%)	56
Total	53 (65.4%)	28 (34.6%)	81

Management style Vs. Perceived usefulness of EUS, Pearson Chi-Square = 0.472 (p=0.492)

Table 4.11: Association between management style and perceived usefulness of EUS for Gastro-Oesophageal Cancers (Oesophageal, Gastro-Oesophageal Junction and Gastric Cancers) (N=81).

Gastro-Oesophageal Cancers (Oesophageal, Gastro-Oesophageal Junction and Gastric Cancers)			
Management	Perceived usefulness of EUS		
Style	Slightly/Not Useful	Somewhat/Very Useful/Essential	Total
Conservative	8 (32.0%)	17 (68.0%)	25
Aggressive	11 (19.6%)	45 (80.4%)	56
Total	19 (23.5%)	62 (76.5%)	81

Management style Vs. Perceived usefulness of EUS, Pearson Chi-Square = 1.470 (p=0.225)

Table 4.12: Association between availability and perceived usefulness of EUS for Oesophageal Cancer (N=81).

Oesophageal Cancer			
Availability of EUS	Perceived usefulness of EUS		
	Slightly/Not Useful	Somewhat/Very Useful/Essential	Total
Available	8 (13.6%)	51 (86.4%)	59
Not available	5 (23.8%)	16 (76.2%)	21
Total	13 (16.3%)	67 (83.8%)	80[¥]

Availability of EUS Vs. Perceived usefulness of EUS, Pearson Chi-Square = 1.196 (p=0.274)
[¥]Out of 81 cases, 1 missing case

Table 4.13: Association between availability and perceived usefulness of EUS for Gastro-Oesophageal Junction Cancer (N=81).

Gastro-Oesophageal Junction Cancer			
Availability of EUS	Perceived usefulness of EUS		
	Slightly/Not Useful	Somewhat/Very Useful/Essential	Total
Available	6 (10.2%)	53 (89.8%)	59
Not available	4 (19.0%)	17 (81.0%)	21
Total	10 (12.5%)	70 (87.5%)	80[¥]

Availability of EUS Vs. Perceived usefulness of EUS, Pearson Chi-Square = 1.116 (p=0.291)
[¥]Out of 81 cases, 1 missing case

Table 4.14: Association between availability and perceived usefulness of EUS for Gastric Cancer (N=81).

Gastric Cancer			
Availability of EUS	Perceived usefulness of EUS		
	Slightly/Not Useful	Somewhat/Very Useful/Essential	Total
Available	41 (69.5%)	18 (30.5%)	59
Not available	12 (57.1%)	9 (42.9%)	21
Total	53 (66.3%)	27 (33.8%)	80[¥]

Availability of EUS Vs. Perceived usefulness of EUS, Pearson Chi-Square = 1.056 (p=0.304)
[¥]Out of 81 cases, 1 missing case

Table 4.15: Association between availability and perceived usefulness of EUS for Gastro-Oesophageal Cancers (Oesophageal, Gastro-Oesophageal Junction and Gastric Cancers) (N=81).

Gastro-Oesophageal Cancers (Oesophageal, Gastro-Oesophageal Junction and Gastric Cancers)			
Availability of EUS	Perceived usefulness of EUS		
	Slightly/Not Useful	Somewhat/Very Useful/Essential	Total
Available	13 (22.0%)	46 (78.0%)	59
Not available	6 (28.6%)	15 (71.4%)	21
Total	19 (23.8%)	61 (76.3%)	80[¥]

Availability of EUS Vs. Perceived usefulness of EUS, Pearson Chi-Square = 0.366 (p=0.545)
[¥]Out of 81 cases, 1 missing case

Table 4.16: Association between attend MDT meeting and perceived usefulness of EUS for Oesophageal Cancer (N=81).

Oesophageal Cancer			
Attend MDT	Perceived usefulness of EUS		
meeting	Slightly/Not Useful	Somewhat/Very Useful/Essential	Total
No	3 (23.1%)	10 (76.9%)	13
Yes	9 (13.6%)	57 (86.4%)	66
Total	12 (15.2%)	67 (84.8%)	79[¥]

Attend MDT Meeting Vs. Perceived usefulness of EUS, Pearson Chi-Square = 0.751 (p=0.386)
[¥]Out of 81 cases, 2 missing cases

Table 4.17: Association between attend MDT meeting and perceived usefulness of EUS for Gastro-Oesophageal Junction Cancer (N=81).

Gastro-Oesophageal Junction Cancer			
Attend MDT	Perceived usefulness of EUS		
meeting	Slightly/Not Useful	Somewhat/Very Useful/Essential	Total
No	2 (15.4%)	11 (84.6%)	13
Yes	7 (10.6%)	59 (89.4%)	66
Total	9 (11.4%)	70 (88.6%)	79[¥]

Attend MDT Meeting Vs. Perceived usefulness of EUS, Pearson Chi-Square = 0.246 (p=0.620)
[¥]Out of 81 cases, 2 missing cases

Table 4.18: Association between attend MDT meeting and perceived usefulness of EUS for Gastric Cancer (N=81).

Gastric Cancer			
Attend MDT meeting	Perceived usefulness of EUS		
	Slightly/Not Useful	Somewhat/Very Useful/Essential	Total
No	7 (53.8%)	6 (46.2%)	13
Yes	45 (68.2%)	21 (31.8%)	66
Total	52 (65.8%)	27 (34.2%)	79[¥]

Attend MDT Meeting Vs. Perceived usefulness of EUS, Pearson Chi-Square = 0.992 (p=0.319)
[¥]Out of 81 cases, 2 missing cases

Table 4.19: Association between attend MDT meeting and perceived usefulness of EUS for Gastro-Oesophageal Cancers (Oesophageal, Gastro-Oesophageal Junction and Gastric Cancers) (N=81).

Gastro-Oesophageal Cancers (Oesophageal, Gastro-Oesophageal Junction and Gastric Cancers)			
Attend MDT meeting	Perceived usefulness of EUS		
	Slightly/Not Useful	Somewhat/Very Useful/Essential	Total
No	2 (15.4%)	11 (84.6%)	13
Yes	15 (22.7%)	51 (77.3%)	66
Total	17 (21.5%)	62 (78.5%)	79[¥]

Attend MDT Meeting Vs. Perceived usefulness of EUS, Pearson Chi-Square = 0.347 (p=0.556)
[¥]Out of 81 cases, 2 missing cases

Table 4.20: Predictors of referral of oesophageal cancer patients for EUS – Results from the EUS survey sample in the UK (n=81).

	Univariate logistic regression				Multivariate logistic regression			
Predictors (independent variable)	β-coefficient	p-value*	Odds Ratio	95% CI for odds ratio	β-coefficient	p-value*	Odds Ratio	95% CI for odds ratio
Age	-0.030	0.497	0.970	0.888 to 1.059	-0.192	0.212	0.825	0.610 to 1.116
Years in practice	-0.018	0.689	0.982	0.900 to 1.072	0.129	0.373	1.138	0.856 to 1.512
Attend MDT meeting	2.197	0.017	9.000	1.471 to 55.072	3.227	0.013	25.214	1.952 to 325.663
Perceived usefulness of EUS	1.372	0.101	3.943	0.767 to 20.273	0.752	0.544	2.121	0.187 to 24.116
EUS available	1.124	0.147	3.077	0.673 to 14.077				
Aggressive management style	0.239	0.761	1.271	0.272 to 5.946	-0.932	0.431	0.394	0.039 to 3.998

*Significance level at p=0.05

Table 4.21: Predictors of referral of gastro-oesophageal junction cancer patients for EUS – Results from the EUS survey sample in the UK (n=81).

	Univariate logistic regression				Multivariate logistic regression			
Predictors (independent variable)	β-coefficient	p-value*	Odds Ratio	95% CI for odds ratio	β-coefficient	p-value*	Odds Ratio	95% CI for odds ratio
Age	-0.002	0.969	0.998	0.924 to 1.079	-0.344	0.068	0.709	0.490 to 1.026
Years in practice	0.026	0.518	1.026	0.948 to 1.111	0.300	0.107	1.350	0.938 to 1.943
Attend MDT meeting	2.303	0.009	10.000	1.786 to 55.976	2.681	0.014	14.593	1.738 to 122.510
Perceived usefulness of EUS	1.615	0.040	5.029	1.080 to 23.404	0.508	0.716	1.661	0.108 to 25.545
EUS available	1.422	0.041	4.145	1.060 to 16.206	2.403	0.051	11.052	0.994 to 122.819
Aggressive management style	1.578	0.025	4.846	1.216 to 19.311	2.474	0.037	11.866	1.156 to 121.833

*Significance level at p=0.05

Table 4.22: Predictors of referral of gastric cancer patients for EUS – Results from the EUS survey sample in the UK (n=81).

	Univariate logistic regression				Multivariate logistic regression			
Predictors (independent variable)	β-coefficient	p-value*	Odds Ratio	95% CI for odds ratio	β-coefficient	p-value*	Odds Ratio	95% CI for odds ratio
Age	-0.054	0.102	0.947	0.888 to 1.011	-0.064	0.063	0.938	0.878 to 1.003
Years in practice	-0.049	0.126	0.952	0.894 to 1.014				
Attend MDT meeting	0.160	0.853	1.174	0.216 to 6.389				
Perceived usefulness of EUS	1.059	0.071	2.882	0.912 to 9.107				
EUS available	-0.511	0.396	0.600	0.185 to 1.951				
Aggressive management style	0.363	0.518	1.438	0.478 to 4.323				

*Significance level at p=0.05

Table 4.23: Predictors of referral of gastro-oesophageal cancer (oesophageal, gastro-oesophageal junction and gastric cancers) patients for EUS – Results from the EUS survey sample in the UK (n=81).

	Univariate logistic regression				Multivariate logistic regression			
Predictors (independent variable)	β-coefficient	p-value*	Odds Ratio	95% CI for odds ratio	β-coefficient	p-value*	Odds Ratio	95% CI for odds ratio
Age	-0.009	0.864	0.991	0.889 to 1.104	-0.502	0.041	0.606	0.374 to 0.980
Years in practice	0.018	0.751	1.018	0.912 to 1.137	0.331	0.104	1.392	0.935 to 2.072
Attend MDT meeting	2.931	0.005	18.750	2.393 to 146.927	6.898	0.033	990.733	1.747 to 561721.166
Perceived usefulness of EUS	-0.470	0.684	0.625	0.065 to 6.030	0.243	0.892	1.275	0.038 to 42.310
EUS available	1.504	0.119	4.500	0.681 to 29.748	0.784	0.676	2.190	0.055 to 86.608
Aggressive management style	-0.720	0.532	0.487	0.051 to 4.666	-3.138	0.266	0.043	0.000 to 10.950

*Significance level at p=0.05

4.5 Discussion

4.5.1 Main findings and comparisons with other studies

Correct staging of gastro-oesophageal cancers (GOCs) is important in determining the appropriate diagnosis, prognosis and treatment options. Endoscopic ultrasound (EUS) has emerged as a superior diagnostic tool in the assessment of GOCs in T- and N-staging. Previous studies have shown the combination use of EUS and other imaging modalities such as computed tomography (CT), positron emission tomography (PET) or integrated PET-CT provides complimentary information for locoregional staging and detection of distant metastasis (Allum et al., 2011; Takizawa et al., 2009; Pfau et al., 2007). Furthermore, utilisation of EUS in clinical practice has been shown to have impact on changes in the management of patients (Nickl et al., 1996), which inevitably associates directly with survival in patients (Gress et al., 1995).

To our knowledge, up to January 2021, this is the first survey of its kind in the field of EUS staging in gastro-oesophageal cancer (GOC) in the UK that explored clinicians' views on the utilisation of EUS in GOC staging, its current clinical practice in the UK, potential factors associated with referral for EUS, and factors that are considered to have influence on the utilisation of EUS in GOC staging. More than four-fifths of the survey respondents reported that they have had experience in the field of EUS either by requesting, performing or both, and support the use of EUS for staging GOC as they believed that EUS offers the best local regional staging of GOC and the opportunity to biopsy during staging. Furthermore, they acknowledged that EUS helps to reduce unnecessary surgery for GOC patients. These findings are in line with the findings reported in the recent survey undertaken by Jones and colleagues (Jones et al., 2021) investigating the use of EUS in the diagnosis and treatment of oesophageal cancer in the UK. Their survey study found that (1) EUS is mostly used as a complementary staging technique to PET-CT, either following PET-CT or at the same time as PET-CT, in the diagnostic pathway of oesophageal cancer; (2) for surgical planning, EUS is most commonly used to assess unresectable T4 disease, which is in line with the reported greater accuracy of EUS over PET-CT for T-staging (Thakkar and Kaul, 2020; DaVee et al., 2017; Valero and Robles-Medranda, 2017; Allum et al., 2011; Puli et al., 2008; Pfau et al., 2007); and (3) for treatment planning, EUS is most commonly used for selecting between surgery and chemoradiotherapy for more advanced but potentially resectable tumour which it helps to avoid unnecessary surgery for oesophageal patients (Jones et al., 2021).

With respect to respondents' knowledge and feelings about the use of EUS, most felt that, among the three type of GOCs, EUS can accurately stage oesophageal cancer, followed by gastro-oesophageal junction cancer and lastly gastric cancer. This sequence is in line with Lennon and Penman's (2007) review study where they summarised that the accuracy of EUS in gastric cancer staging was reported to be less than in oesophageal cancer staging (Lennon and Penman, 2007). Nevertheless, compared with other imaging modalities such as computer tomography (CT), EUS is still considered to be superior to CT for locoregional staging of gastric cancer (Polkowski et al., 2004). Whilst our study was focussed on the investigation of EUS in GOC cancers, with the gastric component implying Siewert type II or III tumours, some clinicians may have interpreted the use of EUS for other Gastric cancers (Riphaus et al., 2013a) including non-adenocarcinomas such as GIST tumours (Gao et al., 2017; Eckardt and Jenssen, 2015; Riphaus et al., 2013b).

Interestingly, although the majority of respondents supported the use of EUS for GOC staging, they would not necessarily routinely request EUS for staging all newly diagnosed GOC patients. This could be due to several reasons such as clinicians might be required to follow a standard hospital protocol regarding the use of EUS for staging GOC and/or they might be required to make EUS requests through the multi-disciplinary team (MDT) first rather than independently. Additionally, it could be due to other specific situations where EUS might not be used. For example, availability of EUS, tight oesophageal stricture, patient morbidity, advanced metastatic disease or other advanced local disease, lack of expertise in performing EUS and interpreting of EUS images, resources constraints, time taken to perform EUS procedure, or patient comfort. Supported by the similar UK survey, conducted recently by Jones et al. (2021), investigating the use of EUS for oesophageal cancer diagnosis and treatment planning, the authors found that there is a proportion of their survey respondents would request EUS on a case-specific basis dependent on PET-CT results, whereas, some would request EUS at the same time as PET-CT or routinely following PET-CT (Jones et al., 2021). Given that EUS has a good evidence base for oesophageal cancer staging, but is more difficult to interpret the images at the junction, or in cancers extending into the upper stomach, particularly when a hiatus hernia is involved, it is not surprising to see a step wise reduction in referral or use of EUS from oesophageal (Siewert I) to Oesophago-Gastric Junction (Siewert II) to Gastric (Siewert III) cancers.

In terms of availability, based on the findings of this survey study, EUS is available within the majority of hospitals at which the respondents are practicing. For those who reported EUS is not

available within their hospital, these respondents usually refer patients to another hospital situated on average about 21 miles away from their hospital for EUS screening. This could lead to delays in the diagnosis and treatment planning of GOCs, and furthermore, this could also put extra burden on patients and their families or friends in terms of time loss and out-of-pocket expenses (e.g. loss of income, childcare cost, travel costs) incurred by them for attending an EUS procedure in other hospital. Taking this into consideration, benefit may be gained if EUS were offered in more clinical settings. However, it is acknowledged that there are challenges in terms of logistics, costs and training in providing EUS at every hospital in the UK. The self-reported approximate number of patients referred for EUS staging is, on average, higher in oesophageal and gastro-oesophageal junction cancer patients compared to gastric cancer patients. This is in keeping with the review conducted by Lennon and Penman (2007) indicating that EUS is more accurate in staging oesophagus cancer than gastric cancer as EUS can be more difficult to interpret in the stomach and sometimes over-stages these cancers.

With regards to clinicians' perceptions about the usefulness of EUS staging in assisting clinical management, responses varied across the three types of GOCs (oesophageal, gastro-oesophageal junction and gastric cancers). More than four-fifths of respondents felt that EUS is somewhat useful to very useful and essential for the evaluation of oesophageal and gastro-oesophageal junction cancers. Whereas, for gastric cancer, just over one-third of respondents reported the same. This implied that EUS was believed to be of less useful in staging gastric cancer compared to oesophageal and gastro-oesophageal junction cancers – This is in line with the earlier findings of this survey study regarding clinician's knowledge about which type of GOC EUS can accurately stage. Consistently, responses to the subsequent question related to the approximate number of EUS referred out of the approximate number of GOC cases seen showed that gastric cancer patients were reported to be of less likely to be referred for EUS compared to oesophageal and gastro-oesophageal junction cancer patients.

Responses to the pre-validated clinical vignettes varied depending upon detailed staging information about the patient's oesophagus tumour conditions. Clinicians made their treatment recommendations regarding resection versus less invasive or palliative therapy based upon detailed differences in tumour stage resulting from EUS scan/staging. This involves the depth of tumour invasion (T) into the oesophageal wall and the involvement of regional lymph nodes (N). As studies demonstrated, such fine distinctions can be examined most accurately by EUS

compared to other imaging modalities (Thakkar and Kaul, 2020; DaVee et al., 2017; Valero and Robles-Medranda, 2017; Allum et al., 2011; Takizawa et al., 2009; Puli et al., 2008; Pfau et al., 2007; Botet et al., 1991a; Botet et al., 1991b).

Various factors could influence the referral of GOC patients for EUS in clinical practice. These include factors such as variations in clinical management styles, perceptions of clinical utility of EUS, availability of EUS, whether or not clinicians attend MDT meeting, clinicians' age and years in practice (Kim and Koch, 1999; personal communications with Dr Hasan Haboubi, Consultant Gastroenterologist and Clinical Advisor to this thesis). In this UK survey study, the association between these potential factors and referral of GOC patients for EUS was explored. Results showed that clinical management style varied among respondents but was not significantly associated with referral for EUS for GOCs. This is in agreement with Kim and Koch (1999)'s findings. Similarly, three other factors – 'clinicians' perception about the utility of EUS', 'availability of EUS' and 'clinicians' years in practice' – were found to have association with referral for EUS for GOC staging but were not significant. Of these three factors, the findings of the two factors – 'availability of EUS' and 'perceived utility of EUS' – are in contrast to Kim and Koch (1999)'s survey study results where their study showed that the availability of EUS and perceived utility of EUS were both significantly associated with EUS referral. This could be due to the fact that the availability of EUS has been increased since two decades ago although it varies across the hospitals in the UK; clinicians' perceptions about the clinical utility of EUS is believed to have also increased since two decades ago alongside the evolution of EUS technology in clinical contexts. There are also other cultural differences in medical practice between the US health care system which is privately funded versus the UK National Health Service which is publicly funded which may account for this difference, as well as the potential differences in access to these services and capacity to provide them in a publicly funded system. It is acknowledged that, with time, EUS has become accepted much more widely for various indications including diagnosis/staging of gastro-oesophageal cancers (Thakkar and Kaul, 2020; DaVee at al., 2017; Valero and Robles-Medranda, 2017; Sahai, 2012). For the remaining two factors ('attend MDT meeting' and 'age'), they both were significantly associated with referral for EUS for GOC staging. The odd ratios of these two factors indicated that clinicians who attend MDT meeting appear to be more likely to refer GOC patients for EUS, and clinicians who are younger in age appear to be less likely to refer patients for EUS. This could be due to the fact that MDT meeting is recognised as a platform for individual clinician to discuss and make informed

clinical decisions safely and collectively, together with a group of professionals from multi-disciplines, regarding appropriate diagnostic procedures and treatment decisions for individual patient. The clinicians that attend these meetings usually have a specific sub-speciality interest in the management of GOCs and are more likely to be aware of the potential advantages and limitations of its use. Furthermore, by being able to communicate with other experts in the field through the MDT, may also have easier access to requesting EUS. Caution may have to be made with interpreting this result as the 98 respondents may have replied to the questionnaire due to their enthusiasm for the use of EUS and may therefore not be fully representative of the national picture. This self-selected group, the majority of whom attended MDTs and also the majority of whom performed EUS may introduce some bias to the results which always needs to be considered with questionnaire data interpretation. With respect to the factor 'age' where younger clinicians were shown to be less likely to refer patients for EUS, this is probably due to the differences in training and experience between younger clinicians and older clinicians may have their preferences of EUS for GOC staging for patients. EUS is a relatively new diagnostic technique and as such has only been incorporated into the training of clinicians in more recent years. Therefore, there may be some differences seen here due to experience of using EUS, as well as more detailed understanding of the areas of its use, despite previous studies have shown that the combination use of EUS and other imaging modalities found to be superior and are complementary to one another (Allum et al., 2011; Takizawa et al., 2009; Pfau et al., 2007). With respect to the factor 'hospital type' (Teaching Hospital versus District General Hospital (DGH)), it was found to have no significant association with referral for EUS for GOC staging though Teaching Hospitals may have better outcomes than General Hospitals in cancer care (Burke et al., 2017). This means that 'hospital type' is not an important factor for EUS referral despite Teaching Hospitals potentially having more funding (and staff) and therefore more access to specialist tests, and may be more radical in their approach to patient care compared to General Hospitals which may be more conservative. Possible explanations for this findings could be (1) The respondents may have been a self-selected group of individuals who completed the survey questionnaire because they either perform or have easy access to EUS irrespective of the hospital site (Teaching versus DGH), and (2) It is also possible that the local geography of Wales with some more rural setups means that these DGHs have now adapted and now EUS is not considered to be such a specialist service anymore (perhaps historically may have been placed in larger University Teaching Hospitals but now more readily available across all areas) thereby reducing the variability in requesting/performing EUS between hospitals.

Last but not least, one of the most surprising findings of the survey was the response to the awareness of the COGNATE trial where less than 30% of the respondents indicated that they were aware of the COGNATE trial. This could be due to the interval of four years between the COGNATE report being published and the survey being disseminated. In addition, clinicians may not recognise the COGNATE acronym of the trial. It is possible that this survey question could have been phrased slightly differently and provided more information. Furthermore, the trial started 10 years late, clinicians might already have the belief that the EUS has benefit for patients and therefore they might not be interested in reading the trial report.

4.5.2 Strengths and Limitations of the study

Up to January 2021, this robust survey is the first study to explore the utilisation of EUS for GOC staging and its clinical practice in the UK. This survey offers an insight into the utilisation of EUS for GOC staging in the UK and the potential factors associated with referral for EUS. The findings from this survey not only have important implications for improving further in the use of EUS in the UK but also providing useful evidence-based information to policy makers and commissioners in NHS cancer care services to help inform policy and improve clinical practice in the UK.

This survey study has several limitations. Firstly, despite a recent review reported that 12-month recall period is commonly used in a significant proportion of health surveys (Heijink et al., 2011), the accuracy of self-reported data on the approximate number of GOC patients referred for EUS is still subject to recall bias (Althubaiti, 2016). While recall bias is not uncommon in self-reported data of health surveys, the main outcome variable used in our exploratory regression analyses was simply whether any patients had been referred for EUS in the past 12 months. Although GOC clinicians might not accurately recall the number of patients referred for EUS, it is believed that they are highly likely to have remembered correctly whether they had referred any patients at all for EUS (Kim and Koch, 1999). Secondly, the total number of responses to this survey provides an adequate level of precision despite being marginally below the minimum target number of sample size calculated for this study. This may limit the power of the study to identify factor less strongly associated with referral for EUS and may affect the generalisability of the study findings to GOC clinicians in the UK. Lastly, the questionnaire designed for the online survey was a 17-page long questionnaire that consisted of five sections. Although the online survey could be

completed in approximately less than 15 minutes, the number of sections may have been overly burdensome for participants (Appendix 4.2).

4.5.3 Further Research and Recommendations

Additional cancer research studies in the field of EUS for GOC staging are recommended as it will help to consolidate the knowledge and experiences clinicians have acquired to date. Consequently, further research in this area is required to strengthen the link between clinicians' knowledge and the clinical utility of EUS in GOC staging which undoubtedly will together positively influence the clinical practice in the UK.

In conjunction with additional research studies, increased education efforts are needed to encourage utilisation of EUS in GOC staging given that EUS has been recommended as a complementary staging technique to other staging techniques in the evaluation of GOCs (Allum et al., 2011; Pfau et al., 2007), and it has been shown to be cost saving and offers greater QALYs (Yeo et al., 2019; Russell et al., 2013). Concomitant efforts must also be made to offer more EUS-related trainings and workshops to increase clinicians' awareness and perception of clinical utility of EUS. Increased availability of EUS more widely across the hospitals in the UK would help to not only increase the accessibility and utilisation of EUS for GOC staging but also avoid delays in the diagnosis and treatment planning of GOCs.

As performing EUS requires technical proficiency, increased intensive yet comprehensive training in EUS for managing GOC should also be given further consideration to help develop technically trained and skilled experts in performing EUS in the diagnostic and therapeutic procedures for GOC as well as in evaluating and interpreting EUS images (Dietrich et al., 2019; Penman et al., 2011; Carroll and Penman, 2004; Catalano et al., 1995). This allows clinicians to rely upon the detailed information resulting from EUS imaging to guide their decisions on patient's management plan.

4.6 Conclusions

Accurate diagnosis and safe clinical practice are every clinician's responsibility in the management of patients. Most UK GOC clinicians, who took part in this survey, recognised the value of EUS for GOC staging. Despite this, there is still a proportion of clinicians felt that EUS is less useful in staging gastric cancer compared to oesophageal and gastro-oesophageal junction

cancers although EUS has been proved to be more accurate than other imaging modalities in evaluating locoregional disease. Clinicians that 'attend Upper GI cancer MDT meeting' was shown to be the most important factor associated with referral of GOC patients for EUS. Attendance at MDT meetings is likely to increase EUS referral; hence, it may be beneficial to encourage GOC clinicians to attend MDT meetings as this may help to avoid unnecessary treatment in GOC patients. However, MDT meetings are costly and there will be opportunity cost of mandating MDT meetings attendance (De Ieso et al., 2013). Therefore, in the currently challenging healthcare financial environment in the UK, caution needs to be taken in recommending attendance at MDT meetings be made mandatory. Though studies showed that there are considerable benefits could be gained from attending MDT meetings (Forrest et al., 2005; Stephens et al., 2006; Mazzaferro and Majno, 2011; De Ieso et al., 2013), the feasibility of mandating MDT meetings attendance in clinical practice would need to be further explored as there will be costs challenges involved (De Ieso et al., 2013). More research into the field of EUS for GOC staging are required to increase the use of EUS in clinical practice. This would aid clinical decision making in treatment recommendations for GOCs. Furthermore, benefits may be gained if EUS could be offered in more clinical settings to avoid delays in the diagnosis and treatment planning of GOCs. However, it is acknowledged that there are logistical, cost and training challenges in providing EUS at every hospital in the UK.

By 2021, the main thrust of this thesis remains, through a series of updated systematic reviews, that suggest use of EUS in combination with other staging techniques such as PET, CT and MRI in the management of patients with GOC appears to be cost saving and offers greater QALY benefits to patients. Whilst EUS has been shown as being cost-effective (COGNATE; Chapter 2), and useful in locoregional staging (Thakkar and Kaul, 2020; Valero and Robles-Medranda, 2017; Puli et al., 2008), its use in the clinical practice in the UK remains unchanged over time – The only UK guidelines by Allum et al. (2011) recommended EUS staging for all GOC patients with non-metastatic cancer; however, my survey in this Chapter (conducted between mid-October 2017 and mid-April 2018) found that EUS is used in non-metastatic cancer with caveat that it helps clinicians in patient's management. The survey findings are in line with the NICE Guidelines (NICE, 2018) and the recent UK survey by Jones and colleagues (2021).

Chapter 5: An exploratory study of novel methodology in QALY estimations using data from two clinical trials – COGNATE and CLARITY trials

5.1 Chapter summary

Although the Quality-Adjusted Life Year (QALY) is currently used as a standard metric to evaluate new health care interventions or technologies and optimise resource allocation in different health systems internationally, it is widely acknowledged that the QALY measure is not always sensitive enough to capture particular aspects of certain conditions, such as functional and symptomatic gains from a health care intervention, in a single index. This chapter describes the exploration of the transferability/generalisability of the concept of the "hybrid QALY technique" first tried in the MORTISE trial (on which I was a Research Officer working with the Trial Statistician Dr Daphne Russell) in other disease areas (e.g. cancer and ophthalmology).

Following the MORTISE trial, I tried this novel hybrid QALY technique using data from two large clinical trials, each with a 12-month follow-up (COGNATE, a cancer trial (Trial registration: ISRCTN1444215), and CLARITY, a non-cancer trial (Trial registration: ISRCTN32207582)) funded by the HTA and NIHR-EME, respectively. I employed regression models between disease-specific measures area-under-the-curve (AUC) and conventional QALYs for both trials following the methodology from the MORTISE trial. The disease-specific measures for the COGNATE trial were the Functional Assessment of Cancer Therapy – General (FACT-G) and the Functional Assessment of Cancer Therapy – Additional Concerns (FACT-AC); and, for the CLARITY trial were the Best Corrected Visual Acuity (BCVA) and the Visual Function Questionnaire – 25 (VFQ-25). Using these two clinical datasets, exploratory linear regression analyses were conducted in IBM SPSS Statistics 25 software package with conventional QALYs as the dependent variable and disease-specific measures as the independent variable. The resulting regression models were used to calculate the disease-specific measure guided QALY (hybrid QALY). The relationship between the disease-specific measure and conventional QALY variables based on observed data for each model were explored. Cost-utility analyses were performed using the STATA version 13 software package to compare the costs and effects of alternative interventions where incremental cost-effectiveness planes and cost-effectiveness acceptability curves (CEACs) were constructed and compared with and without the use of disease-specific measures guided QALY.

In both trials, cost-utility analysis results showed not only more certainty around the estimates of incremental cost-effectiveness ratios (ICERs) when conventional QALY was replaced with the disease-specific measures guided QALY but also a shifting towards the direction of the respective disease-specific measures. In general, the effect of the disease-specific measures guided QALY on the findings of cost-utility/cost-effectiveness is unique for each trial. For example, in this exploratory study I found that, for the COGNATE trial, there was a decrease in the mean difference in the disease-specific measure guided QALYs ($p < 0.05$) compared to the mean difference in the conventional QALY ($p < 0.05$). However, the cost-effectiveness acceptability curves for all the disease-specific measures guided QALYs showed that the probability of the new intervention being cost-effective at the NICE thresholds of £20,000 and £30,000 per QALY remained very closely the same as the probabilities for the conventional QALY.

Applying a holistic approach to modifying QALY used in economic evaluation of health care interventions and/or technologies is crucial to improving decision-making around health care expenditure. From this exploratory study, I have learnt that disease-specific measures could be potentially useful in constructing a hybrid QALY for use in cost-utility analysis in clinical trials to help inform health care resource allocation decisions more effectively and rationally, at least within specific conditions or disease areas. We believe that incorporating disease-specific measures into QALYs estimations will yield hybrid QALYs that more appropriately reflect the 'real' status of patient's quality of life compared to conventional QALYs. The methodology of QALY modification explored in this Chapter is novel and could potentially contribute meaningfully towards future research. In addition, the insights gained from this exploratory study serve to expand future research in economic evaluation methodology on modifying QALY based on the disease-specific measures used in clinical trials.

5.2 Introduction

In economic evaluation, cost-utility analysis is a well-known tool to evaluate whether new health care interventions and/or technologies are good value for money, where health benefits are expressed in the unit of QALYs. One of the most broadly used generic preference-based instruments for the assessment of Health-Related Quality of Life (HRQoL) is the EuroQol five-dimensional (EQ-5D) questionnaire in which QALYs can be generated (EuroQoL Group, 2021). QALY is a composite metric that combines both quantity (survival) and quality of life (utility) into one single index value, thereby providing a common unit to allow comparisons across different disease areas (Kind et al., 2009). QALY can be calculated using the area under the curve (AUC) method from a graph that plots EQ-5D utility values against survival time, where the AUC equates to the total QALY value.

QALY is widely recognised as standard metric of health outcomes in evaluating new health care interventions or technologies to inform health care resource allocation decisions. Despite that, concerns have been raised as to whether the QALY generated solely from EQ-5D could appropriately reflect the 'real' status of patient's HRQoL. Several authors have argued that generic instruments, such as the EQ-5D, are not sensitive enough to capture particular aspects of certain conditions (Pennington et al., 2020; Wichmann et al., 2017; Pettitt et al., 2016; Payakachat et al., 2015; Yang et al., 2015; Mulhern et al., 2013; Tosh et al., 2012; Whitehead and Ali, 2010). For example, it would not always be possible to capture all benefits of medical conditions such as functional and symptomatic gains from a new health care intervention or technology.

To pick up a more specific HRQoL for a particular condition or disease area, condition-/disease-specific measures that consist of questions that are more focused on the studied condition or disease area may be used. Examples of condition-/disease-specific HRQoL measures are the Functional Assessment of Cancer Therapy – General (FACT-G) scale (cancer-related) for HRQoL assessment in patients undergoing cancer therapy (FACIT.org, 2021a; Russell et al., 2013), Functional Assessment of Cancer Therapy – Additional Concerns (FACT-AC) scale (gastro-oesophageal cancer specific) for HRQoL assessment in gastro-oesophageal cancer patients undergoing cancer therapy (Russell et al., 2013), and Visual Function Questionnaire-25 (VFQ-25) for vision-related quality of life assessment in patients with chronic eye diseases (e.g. proliferative diabetic retinopathy (PDR), glaucoma) (National Eye

Institute, 2020; Sivaprasad et al., 2018). FACT-G is a validated 27-item questionnaire with 5-point Likert scale that measures four domains of HRQoL in cancer patients: physical, social, emotional, and functional well-being (FACIT.org, 2021a; Russell et al., 2013). FACT-AC is a single integrated gastro-oesophageal cancer specific quality of life questionnaire merged from two separate FACT questionnaires – FACT-Esophageal (FACT-E) and FACT-Gastric (FACT-Ga) (FACIT.org, 2021b; FACIT.org, 2021c; Russell et al., 2013) by the COGNATE psychometric team for ease of use for gastro-oesophageal cancer patients in the trial (Russell et al., 2013). VFQ-25 is a 25-item National Eye Institute Visual Function Questionnaire (NEI-VFQ) (NEI, 2021; Sivaprasad et al., 2018). Table 5.0a below shows the summary details of these validated disease-specific HRQoL measures (FACT-G, FACT-AC and VFQ-25) used for the exploratory analysis in this Chapter.

Table 5.0a: Summary details of the validated disease specific HRQoL measures: FACT-G, FACT-AC and VFQ-25.

Disease-specific HRQoL measure	**Definition**	**Scoring**	**Subscale**	**Interpretation**
Functional Assessment of Cancer Therapy-General (FACT-G)	FACT-G is a psychometric instrument measuring cancer-specific quality of life in patients undergoing cancer therapy. This questionnaire was administered in the COGNATE trial to measure patients' HRQoL related to cancer at baseline, 1, 3, 6 and 12 months after randomisation, and at 18, 24 and 36 months where possible.	27 items, each with a 5-point Likert scale (minimum score = 0, and maximum score = 4). FACT sums scores on the four subscales to derive FACT-G total score (score range: 0-108). Scoring FACT-G to derive FACT-G Total Score: 1. In each subscale, sum individual items scores. 2. Multiply the sum individual item scores by the number of items in the subscale, 3. Then, divide by the number of items answered in each subscale. This produces the subscale score. 4. Lastly, sum all four subscales' scores to derive FACT-G total score (score range: 0-108). Then, the average FACT-G score is calculated by dividing the FACT-G Total Score (ranging from 0-108) by the total number of the FACT-G items (that is 27 items) where this generates an average FACT-G score of 0 (minimum average FACT-G score) to 4 (maximum average FACT-G score).	Four subscales of HRQoL in cancer patients: 1. Physical Well-Being (seven items), 2. Social or Family Well-Being (seven items), 3. Emotional Well-Being (six items), and 4. Functional Well-Being (seven items).	The higher the score, the better the quality of life (QOL).

Functional Assessment of Cancer Therapy-Additional Concerns (FACT-AC)	FACT-AC is a psychometric instrument measuring gastro-oesophageal cancer (GOC) specific quality of life in GOC patients undergoing cancer therapy. As FACT-Oesophageal and FACT-Gastric modules have many similar questions, the FACT team encouraged the COGNATE psychometric team to combine them into a single 'Additional Concerns' module. Factor analysis was used by the COGNATE psychometric team to examine the structure of FACT and thereby assessed whether to aggregate these two modules into one. This questionnaire was administered in the COGNATE trial to measure patients' HRQoL related to gastro-oesophageal cancer specific at baseline, 1, 3, 6 and 12 months after randomisation, and at 18, 24 and 36 months where possible.	33 items, each with a 5-point Likert scale (minimum score = 0, and maximum score = 4). The COGNATE psychometric team derived its FACT-AC scale from Gastric Additional Concerns, comprising 19 items, and Oesophageal Additional Concerns, comprising 17 items, by removing overlapping items and psychometrically weak items using methods described by Streiner and Norman (2008). In this way, the COGNATE psychometric team effectively merged the Gastric Additional Concerns and Oesophageal Additional Concerns scales to form a single integrated Gastro-Oesophageal Concerns scale for easy use by all COGNATE trial patients. Scoring 33-item FACT-AC subscale: 1. Sum all individual items scores. 2. Multiply the sum individual item scores by the number of items in the subscale. 3. Then, divide by the number of items answered. This produces the FACT-AC subscale score (score range: 0-132). To derive FACT-AC Total Score: 4. Sum scores on all the subscales of the FACT-G (four subscales) and FACT-AC (one overall subscale of 33 items) to derive FACT-AC Total Score (score range: 0-240). Then, the average FACT-AC score is calculated by dividing the FACT-AC Total Score (ranging from 0-240) by the total number of the FACT-AC items (that is 60 items) where this generates an average FACT-AC score of 0 (minimum average FACT-AC score) to 4 (maximum average FACT-AC score).	No subscale	The higher the score, the better the QOL.

Visual Function Questionnaire-25 (VFQ-25)	VFQ-25 is a vision-related quality of life questionnaire to assess patients with chronic eye diseases (e.g. proliferative diabetic retinopathy (PDR), glaucoma) This questionnaire was administered in the CLARITY Trial to measure patients' vision-related quality of life at baseline and 52 weeks.	25 items (for each item, 100 = Best score, and 0 = Worst possible score), and an overall VFQ-25 composite score. In each subscale, items within each subscale are averaged together to create the 12-subscale scores (100 = Best score, and 0 = Worst score). Items that are left blank (missing data) are not taken into account when calculating the scale scores. Subscales with at least one item answered can be used to generate a subscale score. Hence, scores represent the average for all items in the subscale that the respondent answered. Then, the average VFQ-25 score is calculated by dividing the VFQ-25 Total Score (that is the sum of the 12 subscales score with each subscale score ranging 0-100) by the total number of the VFQ-25 subscales (that is a total of 12 subscales) where this generates an average VFQ-25 score of 0 (minimum average VFQ-25 score, worse score) to 100 (maximum average VFQ-25 score, best score). To calculate an overall VFQ-25 composite score, average the vision-targeted subscale scores (100 = Best score, and 0 = Worst score), excluding the general health rating question.	12 subscales: 1. General health (1 item). 2. General vision (1 item). 3. Ocular pain (2 items). 4. Near activities (3 items). 5. Distance activities (3 items). 6. Social functioning (2 items). 7. Mental health (4 items). 8. Role difficulties (2 items). 9. Dependency (3 items). 10. Driving (2 items). 11. Colour vision (1 item). 12. Peripheral vision (1 item)	The higher the score, the better the visual functioning QOL.

In clinical studies, objective measures (existing and measurable, independent of individual experiences in which data is derived from medical record, such as serum albumin levels) can be collected alongside subjective measures (based on individual awareness or experience in which data collected from participants' self-reported responses to surveys about HRQoL, such as vision functioning and general health perceptions) (Cleary, 1997). For example, in the CLARITY trial (a vision study), both the visual-related objective measure (i.e. Best-Corrected Visual Acuity (BCVA)) and subjective measure (i.e. VFQ-25 questionnaire) were collected for the evaluation of intravitreal aflibercept in patients with proliferative diabetic retinopathy. BCVA is used to measure objectively the best possible vision an eye can see with corrective lenses on a standardised Snellen eye chart from a specific viewing distance (Sivaprasad et al., 2018). Whereas, the VFQ-25, a visual function specific HRQoL measure, is used to assess subjectively the impact of improving visual function on patients' lives where aspects of health relating to visual functioning are measured, such as ocular pain, near activities, distance activities, role difficulties, and driving. Accommodating the way in which individuals interpret and synthesise different aspects of health based on their awareness or experience is important in assessing the impact of a new intervention on individuals' HRQoL. It is advocated that subjective measures are an inherent part of the study design (Cleary, 1997).

Given health is a multi-faceted concept needing multiple indicators to assess different aspects of health, subjective measures are undeniably a holistic measure covering wide aspects of health for assessing individuals' HRQoL as opposed to objective measures that assess only one very specific point of measure of study interest (e.g. BCVA for visual acuity measure or serum albumin levels for biological measure depending on the area of study interest). Thus, for the exploratory study in this Chapter, it would be appropriate to understand the varying implications of using both the objective measures (i.e. BCVA) and subjective measures (i.e. VFQ-25) from the CLARITY trial as part of the exploration of the hybrid method.

In clinical trials, although disease-specific data are collected alongside EQ-5D data, QALYs generated solely from the EQ-5D are used to perform cost-utility analysis to evaluate whether the new intervention or technology is cost-effective compared to standard care. However, there is a growing concern that the QALY generated solely from the EQ-5D may not appropriately reflect the 'real' status of individual's quality of life. Many authors argued that the EQ-5D is less sensitive and less responsive to changes in individual's quality of life specific

to a studied condition/disease area (Pennington et al., 2020; Wichmann et al., 2017; Pettitt et al., 2016; Payakachat et al., 2015; Yang et al., 2015; Mulhern et al., 2013; Tosh et al., 2012; Whitehead and Ali, 2010). Hence, having the opportunity to explore the hybrid method first tried in the MORTISE trial (a Morton Neuroma study on which I was a Research Officer working with the Trial Statistician Dr Daphne Russell) in other disease areas for this Chapter in this thesis, it would allow us to explore whether the innovative idea of incorporating disease-specific measures into the conventional QALY (the hybrid method) would help enhance cost-utility analysis results as demonstrated in the MORTISE trial.

The exploratory study in this Chapter was conducted to further explore the hybrid method introduced in the MORTISE trial (Edward et al., 2015). The aim of this study was to further explore the potential use of disease-specific measures in QALY calculation (i.e. hybrid method) for cost-utility analysis in other disease areas (e.g. cancer and ophthalmology). This was done by establishing a model between disease-specific measures and conventional QALY, as first tried in the MORTISE trial (Edwards et al., 2015), using the accessible data from the COGNATE trial (gastro-oesophageal cancer study) and the CLARITY trial (ophthalmology study). This novel concept of the hybrid method that takes account of all aspects of patient's HRQoL measured through both the generic preference-based HRQoL and disease-specific HRQoL instruments would reflect the effect of a new health care intervention or technology on patient's QoL more comprehensively.

5.3 Methods

5.3.1 Source of data

Twelve-month health economics data from two large clinical trials – COGNATE, a cancer trial (Trial registration: ISRCTN1444215) (Russell et al., 2013), and CLARITY, an ophthalmology trial (Trial registration: ISRCTN32207582) (Sivaprasad et al., 2018), funded by the HTA and NIHR-EME respectively – were used for the exploratory analysis in this chapter.

The COGNATE trial was a multicentre prospective randomised controlled trial (RCT) that evaluated the effectiveness and cost-effectiveness of adding endoscopic ultrasound (EUS) to the standard staging algorithm in the management of patients with gastro-oesophageal cancers compared to standard staging algorithm. As the disease-specific measures in the COGNATE trial had complete imputed data set at 12 months, the exploratory analysis in this

chapter was undertaken based on these data. The disease-specific measures employed in the COGNATE trial, Functional Assessment of Cancer Therapy – General (FACT-G) and Functional Assessment of Cancer Therapy – Additional Concerns (FACT-AC), were used for the exploratory analysis in this Chapter. To deal with missing quality of life data, imputation was performed as Phase 2 imputation by DR (the Trial Statistician) (see Section 2.3.5 in Chapter 2 for further details; Russell et al., 2013). DR explained that she used the SPSS MVA procedure to impute scale scores at times without interview for those who were still alive at 12 months, and then for those who had died by 12 months. DR performed the SPSS MVA procedure that uses single imputation for the missing quality of life data. This yielded a complete imputed data set with all quality of life scores at all times up to 12 months (see Section 2.3.5 in Chapter 2 for further details; Russell et al., 2013).

The CLARITY trial was a multicentre prospective randomised controlled trial that evaluated the clinical efficacy, mechanism and cost-effectiveness of intravitreal aflibercept therapy for patients with proliferative diabetic retinopathy (PDR) compared to standard care panretinal photocoagulation (PRP) at 12 months. The disease-specific measures employed in the CLARITY trial, Best-Corrected Visual Acuity (BCVA) and Visual Function Questionnaire-25 (VFQ-25), were used for the exploratory analysis in this Chapter. Approval for use of the CLARITY trial data as part of the exploratory analysis in this chapter had been granted from the Principal Investigator of the CLARITY trial (See E-mail communication in Appendix 5.1). The economic sample for the 12-month follow up CLARITY trial was undertaken on 202 participants (101 per arm) with complete cost and outcome data where full data were available. This economic sample of 202 participants represents 96.7% of the clinical sample included in primary outcome Intention-To-Treat analysis.

5.3.2 Principal underlying analysis

The methodological approach employed in the MORTISE trial has only been assessed in one disease area (Morton's neuroma) and has not been assessed in cancer or ophthalmology. Employing the technique there would allow us to identify the possible transferability/generalisability of the technique to other disease areas and therefore potentially further use the technique in other disease areas. For exploring the transferability/generalisability of the technique as a methodology to achieve the potential aim of using disease-specific measures in cost-utility analysis in clinical trials, minor

modification to the technique used in the MORTISE trial was undertaken where only the studied disease-specific measure(s) was/were included as the parameter(s) for the independent variable(s) in regression model. Hence, the focus of this exploratory study on the aim of understanding and expanding the approach taken in the MORTISE trial for a methodology used across disease areas is novel and it is the first study of its kind.

5.3.3 Procedure and Analysis Strategy

Twelve-month data were extracted from the COGNATE and CLARITY trials and managed using MS Excel. Cost data from the two trials were all inflated to 2019 prices using the Hospital and Community Health Services (HCHS) pay and price inflation indices and the NHS cost Inflation Index (NHSCII) from the published sources (Curtis and Burns, 2016; Curtis and Burns, 2019). In both trials, difference in effects between arms (i.e. QALY, FACT-G and FACT-AC in COGNATE trial; QALY, BCVA and VFQ-25 in CLARITY trial) were all adjusted for differences in baseline effects (Manca et al., 2005). This adjustment was undertaken by performing linear regression in IBM SPSS Statistics 25 (IBM Corporation, Armonk, NY, USA) using the corresponding baseline effect and trial arm as the only covariates for effects adjustment (Laramée et al., 2013). For example, difference in QALY between arms was adjusted for differences in baseline EQ-5D-3L index score with baseline EQ-5D-3L index score and trial arm as the covariates in the linear regression for QALY adjustment.

5.3.4 Procedure and Analysis Strategy – Exploring the potential use of disease-specific measure in cost-utility analysis

To explore the potential use of disease-specific measures in cost-utility analysis in other disease areas, I used the same methodological approach as the one first tried in the published MORTISE trial – that is using linear regression to regress conventional QALY on disease-specific measure to obtain disease-specific guided QALY (hybrid QALY). For ease of exploration of examining the impact of disease-specific measures has had on cost-utility analysis results, a minor modification to the method of linear regression first tried in the published MORTISE trial was undertaken for this exploratory study in this Chapter. In the publised MORTISE trial, the linear regression model was established with disease-specific measure and treatment group included as the independent variables and conventional QALY as the dependent variable. Whereas, for this exploratory study, a regression model was established where only the studied disease-specific measure was included as the independent

variable and conventional QALY as the dependent variable – this was to understand the broader impact of the methodology as a generalisable technique. With only the studied disease-specific measure included as independent variable in regression model, this would provide us a clearer picture of the changes to the cost-utility analysis result that are solely influenced by the studied disease-specific measure. For both trials data, using the adjusted effect data for baseline differences, regression models were establised between disease-specific measure and conventional QALY. This was achieved by conducting linear regression in IBM SPSS Statistics 25 software package (IBM Corporation, Armonk, NY, USA) with conventional QALY as dependent variable and disease-specific measure(s) as independent variable(s). Two regression models were developed for the COGNATE trial and three regression models were developed for the CLARITY trial. From the resulting regression models, hybrid QALY (i.e. disease-specific guided QALY) were able to be produced incorporating the disease-specific measure into the QALY. Furthermore, in each regression model, the relationship between disease-specific measure and conventional QALY variables, based on observed data, were explored.

Another alternative approach taken to achieve this aim is to undertake a mapping exercise. Wailoo et al. (2017) describe a good practice process for developing the regression models in mapping studies. Here in this Chapter however, the approach is to further generalise the methodology in the MORTISE trial to explore the transferability of the hybrid approach to other disease areas.

Linear regression models in this exploratory study were built in IBM SPSS Statistics 25 software package (IBM Corporation, Armonk, NY, USA) where conventional QALY was used as dependent variable and disease-specific measure(s) was used as independent variable(s). A coefficient value attached to each independent variable and a constant value were generated. A simple linear regression equation is as below:

$Y = \beta_1 X + \beta_0 + \varepsilon$, where

Y is all observed values for dependent variable

β_1 is the coefficient value for independent variable X

X is all observed values for independent variable

β_0 is the constant a.k.a bias

ε is the random error term

Thus, the simple linear regression equation used in this exploratory study is described as below:

$Y = \beta_1 X + \beta_0 + \varepsilon$, where

Y is all observed conventional QALY values

β_1 is the coefficient value for disease-specific measure area-under-the-curve value

X is all observed disease-specific measure area-under-the-curve values

β_0 is the constant a.k.a bias

ε is the random error term

The observed QALY is listed as the dependent variable in the regressions, and the area-under-the-curve of the studied disease-specific measure is listed as the independent variable in the regressions for both the COGNATE and MORTISE trials. For COGNATE trial, the independent variable used in the regressions are – FACT-G-Area-Under-The Curve (FACT-G-AUC) and FACT-AC-AUC. To derive FACT-G-AUC and FACT-AC-AUC, firstly, for scaling consistency, the average FACT-G scores and the average FACT-AC scores were scaled from 0 to 4 to a scale of 0 to 1 similar to the scale of the EQ-5D index score. This was done by dividing the FACT scores by 4. Then, the FACT-G-AUC and FACT-AC-AUC were calculated for use in the regressions as independent variable. For CLARITY trial, the independent variable used in the regressions are –BCVA-AUC and VFQ-25-AUC. To derive BCVA-AUC and VFQ-25-AUC, firstly, for scaling consistency, the BCVA scores and the average VFQ-25 scores were scaled from 0 to 100 to a scale of 0 to 1 similar to the scale of the EQ-5D index score. This was done by dividing the BCVA scores and the average VFQ-25 scores by 100. Then, the BCVA-AUC and the VFQ-25-AUC were calculated for use in the regressions as independent variable. The calculation method for FACT-G-AUC, FACT-AC-AUC, BCVA-AUC and VFQ-25-AUC is the same as for the calculation method for QALY, that is using area-under-the curve method by weighting quality of life scores (e.g. scaled average FACT-G scores, for example, for COGNATE trial) with quantity of life (e.g. survival or study time points).

Table 5.0b below shows an example of fitted regression models for the COGNATE trial and CLARITY trial, with a numeric example demonstrating how a FACT-G-AUC score (for example

for the COGNATE trial) or a VFQ-25-AUC score (for example for the CLARITY trial) is combined with their corresponding coefficient to calculate the hybrid QALY.

Table 5.0b: The fitted regression models for the COGNATE and CLARITY trials with a numeric example for each trial.

Regression models	**COGNATE trial**	**CLARITY trial**
Example model	$Y = \beta_1$FACT-G-AUC + $\beta_0 + \varepsilon$, where Y is all observed conventional QALY values β_1 is the coefficient value for FACT-G-AUC β_0 is the constant a.k.a bias ε is the random error term	$Y = \beta_1$VFQ-25-AUC + $\beta_0 + \varepsilon$, where Y is all observed conventional QALY values β_1 is the coefficient value for VFQ-25-AUC β_0 is the constant a.k.a bias ε is the random error term
Numeric example*	Below is a numeric example demonstrating how a FACT-G-AUC value, for example, is combined with the coefficient to calculate the hybrid QALY- For example: If $\beta_1 = 0.9$, $\beta_0 = 0.5$ and a FACT-G-AUC value = 0.52, Therefore, Hybrid QALY = (0.9 x 0.52) + 0.5 Hybrid QALY = (0.468) + 0.5 Hybrid QALY = 0.968	Below is a numeric example demonstrating how a VFQ-25-AUC value, for example, is combined with the coefficient to calculate the hybrid QALY- For example: If $\beta_1 = 0.75$, $\beta_0 = 0.05$ and a VFQ-25-AUC value = 0.66, Therefore, Hybrid QALY = (0.75 x 0.66) + 0.05 Hybrid QALY = (0.495) + 0.05 Hybrid QALY = 0.545

*****In this table these are examples and in results there are 'true' values quoted.

5.3.5 Exploratory cost-utility analysis

Costs and effects (both the conventional and exploratory QALYs) data from the COGNATE and CLARITY trials were used to perform cost-utility analysis in STATA version 13 software package (Stata Corp, College Station, TX, USA). In cost-utility analysis, the difference between

intervention and control groups in the mean costs of participants' health care resource use were compared to the corresponding difference in mean effects to estimate incremental cost-effectiveness ratios (ICERs) where appropriate. To overcome the skewed cost data and to quantify the uncertainty around the estimates of incremental cost-effectiveness ratios, a simulation of 5,000 non-parametric bootstrapping iterations was conducted using STATA version 13 (Stata Corp). These were used to estimate 95% confidence intervals (CIs) for incremental costs and incremental effects, and ICERs where appropriate, and then to construct cost-effectiveness planes (CE planes) – a scatter plot of the joint distribution of incremental costs and effects – and CEACs (Glick et al., 2007; Briggs and Gray, 1999). The CEACs displays the probability that an intervention is more cost-effective than the alternative across a range of willingness-to-pay thresholds for a QALY (Glick et al., 2007; Fenwick et al., 2006; Fenwick et al., 2004). Then, the constructed CE planes and CEACs of each of the disease-specific measure guided QALYs were compared to explore the changes to the results of the CE plane and CEAC of the conventional QALY.

5.4 Results

5.4.1 COGNATE trial 12-month follow-up results

Table 5.1 shows that, for the COGNATE trial, the regression coefficients for - FACT-G-AUC (β-coefficient = 0.396) and FACT-AC-AUC (β-coefficient = 0.249) were statistically significant (p-value < 0.05) when conventional QALY were regressed on FACT-G-AUC and FACT-AC-AUC individually (see Model 1 and 2, respectively, in Table 5.1). This indicates that there is relationship between (a) FACT-G-AUC and conventional QALY, and (b) FACT-AC-AUC and conventional QALY. The magnitude of FACT-G-AUC coefficient (0.396) is greater than the magnitude of FACT-AC-AUC coefficient (0.249), indicating that the relationship between FACT-G-AUC and conventional QALY is stronger than the relationship between FACT-AC-AUC and conventional QALY. This is also shown by the higher adjusted R-squared value for Model 1 (r^2 = 0.267) compared to Model 2's (r^2 = 0.107), indicating a higher strength of the relationship between the model using FACT-G as predictor and the dependent variable (i.e. the conventional QALYs). A higher adjusted R-squared value also represents a better fit for the model i.e. smaller differences between the observed data (the conventional QALYs) and the fitted values (the predicted QALY values). Among the two models explored, the adjusted R-squared value for Model 1 was the highest, followed by Model 2's. The regression coefficient magnitude of 0.396 for the FACT-G-AUC in Model 1 implies that for every 1-unit

increase in FACT-G-AUC, there is a 0.396-unit increase in conventional QALY. As a result, a hybrid QALY (i.e. disease-specific guided QALY) that takes account of the unit change in FACT-G-AUC can be generated. The fitted regression equations for the two models for COGNATE trial can be expressed as below-

Model 1:

Y = (0.396 x FACT-G-AUC) + 0.324, where

Y is all observed conventional QALY values

0.396 is the coefficient value for FACT-G-AUC

0.324 is the constant a.k.a bias

Model 2:

Y = (0.249 x FACT-AC-AUC) + 0.421, where

Y is all observed conventional QALY values

0.249 is the coefficient value for FACT-AC-AUC

0.421 is the constant a.k.a bias

A numeric example demonstrating how the hybrid QALY was calculated-

Using Model 1 for example,

If FACT-G-AUC equals to 1.5

Therefore,

Hybrid QALY = (0.396 x FACT-G-AUC) + 0.324

Hybrid QALY = (0.396 x 1.5) +0.324

Hybrid QALY = 0.918

Table 5.1: Exploratory analyses using the COGNATE trial 12-month data (N = 213): Results from the linear regression models using disease-specific measures to guide QALYs.

Model no. (n=213)	Predictor(s)[¥]	β-coefficient	Standard Error	p-value*	95% Confidence Interval	Adjusted R-squared
1.	**FACT-G-AUC**	0.396	0.045	<0.05	0.308 to 0.484	0.267
	Constant	0.324	0.027	<0.05	0.270 to 0.378	
2.	**FACT-AC-AUC**	0.249	0.048	<0.05	0.153 to 0.344	0.107
	Constant	0.421	0.028	<0.05	0.365 to 0.476	

*Significance level at p=0.05
¥All predictors were corrected for baseline

Table 5.2 below shows the means and standard deviations of the COGNATE trial's effects data for intervention (EUS) and control (non-EUS) groups, the mean difference between groups with the corresponding bootstrap standard error and bootstrap 95% confidence interval. The mean difference between groups for both the FACT-G-AUC-guided QALY and FACT-AC-AUC-guided QALY were lower than the mean difference between groups for the conventional QALY. Amongst both of the disease specific measures-guided QALYs, the mean difference in FACT-G-AUC-guided QALY (0.0092; bootstrap 95% CI = 0.0004 to 0.0181) was statistically significant different between groups, likewise the mean difference in conventional QALY (0.0304; bootstrap 95% CI = 0.0131 to 0.0470). However, the mean difference in FACT-AC-AUC-guided QALY was not statistically significant different between groups (mean difference = 0.0052; bootstrap 95% CI = –0.0002 to 0.0114). For the area-under-the curve (AUC) of the two disease-specific measures scores, FACT-G-AUC and FACT-AC-AUC, that were individually taken account into QALY calculation in the regressions to generate disease-specific measures guided QALY (hybrid QALY), results showed that there was a statistically significant difference in mean FACT-G-AUC between groups (mean difference = 0.0954; bootstrap 95% CI = 0.0063 to 0.1844), however this difference was tiny so may not be clinically meaningful. However, for FACT-AC-AUC, there was no statistically significant difference in mean between groups (mean difference = 0.0835; bootstrap 95% CI = –0.0088 to 0.1783).

Table 5.2: Effect data of COGNATE trial – Mean and standard deviation (SD) of allocated group, mean difference between groups, bootstrap standard error and bootstrap 95% confidence interval.

Effect*	**Intervention (EUS) (n=107) Mean (SD)**	**Control (Non-EUS) (n=106) Mean (SD)**	**Mean difference** between groups**	**bootstrap standard error**	**bootstrap 95% confidence interval**
FACT-G-AUC	2.460 (0.350)	2.365 (0.334)	0.0954	0.0460	0.0063 to 0.1844
FACT-AC-AUC	2.341 (0.358)	2.257 (0.344)	0.0835	0.0482	-0.0088 to 0.1783
QALY	0.578 (0.069)	0.548 (0.058)	0.0304	0.0087	0.0131 to 0.0470
FACT-G-AUC-Guided QALY	0.568 (0.035)	0.559 (0.033)	0.0092	0.0046	0.0004 to 0.0181

FACT-AC-AUC-Guided QALY	0.566 (0.022)	0.561 (0.021)	0.0052	0.0030	-0.0002 to 0.0114

*All effects were corrected for baseline

**Mean difference between groups = Mean of Intervention group minus Mean of Control group

Table 5.3 shows the mean total costs per patient (£) by allocated groups and the mean difference between groups over 12 months in the COGNATE trial. Results showed that use of EUS as an additional staging technique, on average, had a lower total costs (£26,379, SD £16,048) compared to staging strategy without EUS (£30,538, SD £21,755). This yielded, on average, a cost saving of £4,159 (bootstrap 95%CI -£9,318 to £888) in total costs.

Table 5.3: COGNATE trial – Costs by allocated group over 12 months (2019 price year, £).

	Intervention (EUS) (n=107) Mean (SD)	**Control (Non-EUS) (n=106) Mean (SD)**	**Intervention minus Control[†] (bootstrap 95% CI)**
Secondary care costs (£)	25,174 (14,891)	28,354 (18,402)	-3,180 (-7,659 to 1267)
Medication costs (£) (All hospital prescribing non chemotherapy and surgery drug costs)	1,205 (1,782)	2,184 (4,818)	-979 (-2,079 to -142)
Total costs/patient (£)	26,379 (16,048)	30,538 (21,755)	-4,159 (-9,318 to 888)

[†] Negative differences imply EUS group has lower cost than non-EUS group.

Note: Sum didn't add up due to rounding issue

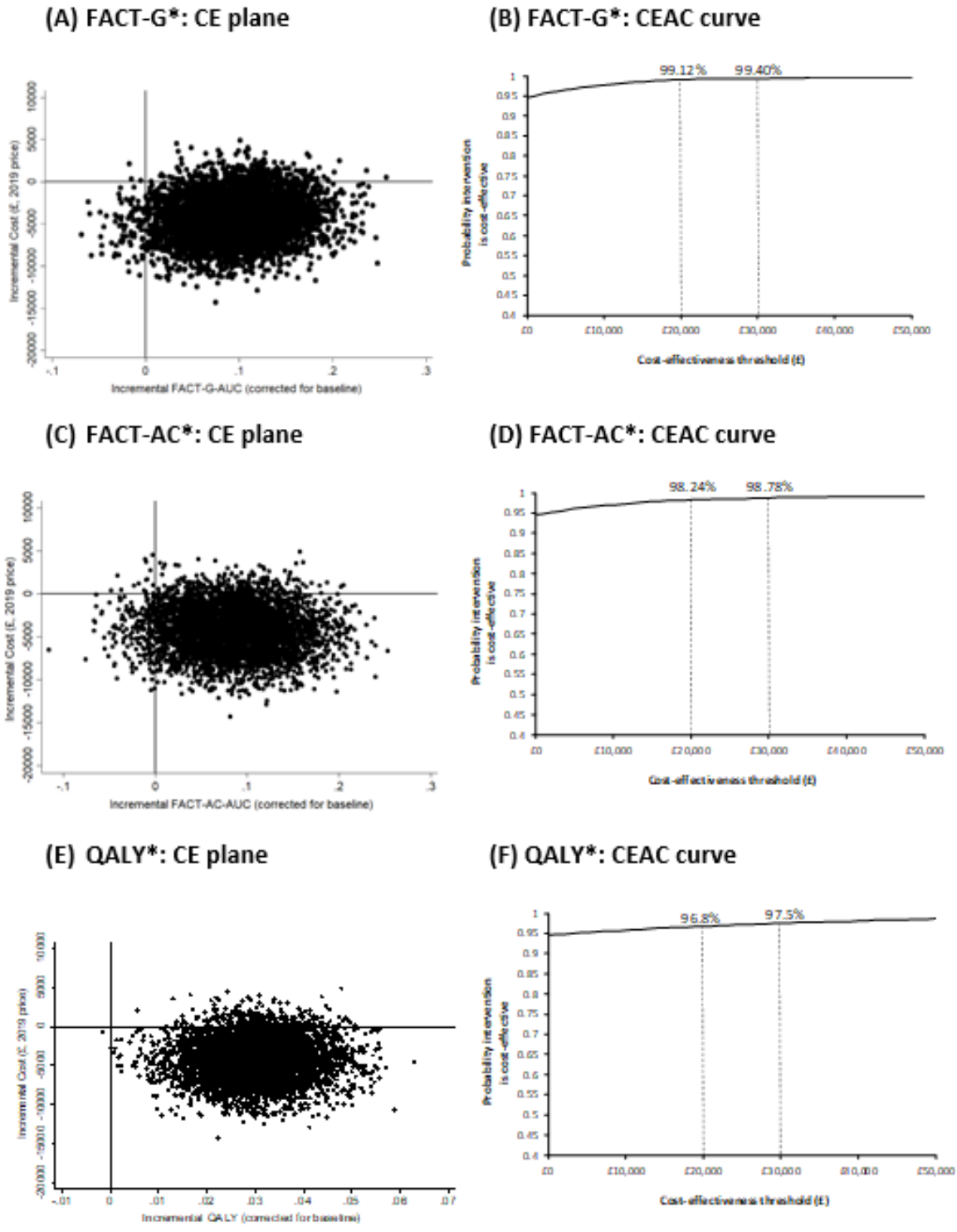
Figure 5.1 presents the cost-effectiveness plane (CE plane) and their corresponding cost-effectiveness acceptability curve (CEAC) for each of the two disease-specific measures guided QALYs generated in this exploratory study: FACT-G-AUC-guided QALY (Figure 5.1(G) and (H)), and FACT-AC-AUC-guided QALY (Figure 5.1(I) and (J)). In addition, the CE plane and the corresponding CEAC graphs for FACT-G-AUC (Figure 5.1(A) and (B)), FACT-AC-AUC (Figure 5.1(C) and (D)) and conventional QALY (Figure 5.1(E) and (F)) were also plotted.

Compared to the CE plane plotting incremental cost against incremental conventional QALY (Figure 5.1(E)), the CE planes in Figure 5.1(G) and (I) showed that the scatter plots became more concentrated around the estimates once the disease-specific measures scores were taken into account in the QALY calculation to generate disease-specific measures guided

QALY. This is in line with the results of the bootstrap standard error in Table 5.2 that showed that all the disease-specific measures guided QALYs had a smaller bootstrap standard error than the conventional QALY. The smaller the boostrap standard error indicates the lesser the variability of the data. Graphically, a visual representation of the bootstrapped conventional QALY data and bootstrapped disease-specific measure guided QALY data was demonstrated using the Box and Whisker plot in Figure 5.2. Results showed that all bootstrapped disease-specific measures guided QALYs data had shorter box length compared to the bootstrapped conventional QALY data. Shorter box lengths indicate narrower distribution, that is, less dispersed data (i.e. more concentrated data). Furthermore, the scatter plots in Figure 5.1(G) and (I) showed shifting towards disease-specific measure's direction once the disease-specific measure scores were incorporated into QALY calculation to generate disease-specific measures guided QALY.

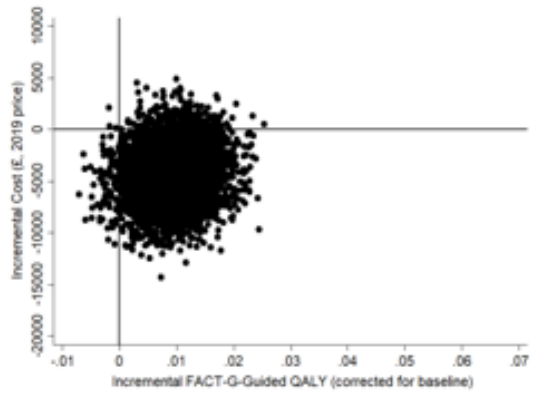
With regards to how cost-effective is EUS in GOC staging, CEAC shows that EUS is 95.44% and 95.96% being cost-effective at the UK NICE's thresholds of £20,000 and £30,000 per QALY respectively, when FACT-G-AUC-guided QALY was used as a measure of effect; and 94.96% and 95.20% when FACT-AC-AUC-guided QALY was used as a measure of effect – These probabilities findings are similar to the findings when the conventional QALY was used as a measure of effect where EUS is shown to be 96.8% and 97.5% cost-effective at £20,000 and £30,000, respectively.

Figure 5.1: Cost-effectiveness planes with 5000 bootstrapped incremental costs and effects pairs from the COGNATE trial data and the corresponding cost-effectiveness acceptability curves (CEACs): an exploratory analysis of the usability of disease-specific measures in cost-utility analysis

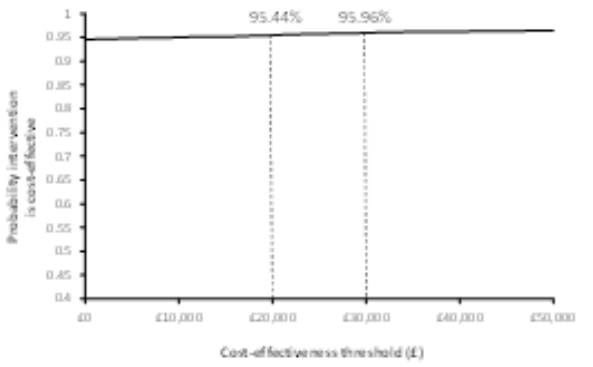


***Footnote:** FACT-G-AUC, FACT-AC-AUC and QALY were corrected for baseline.

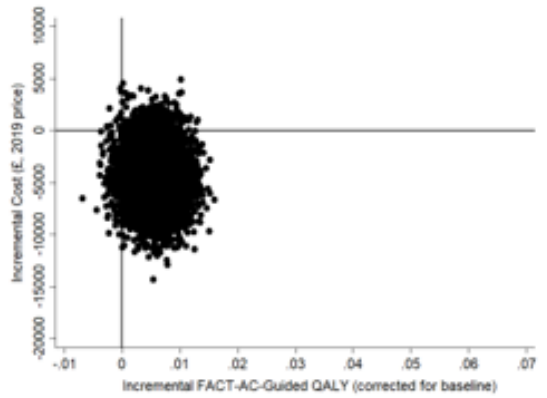
(G) FACT-G-Guided QALY*: CE plane



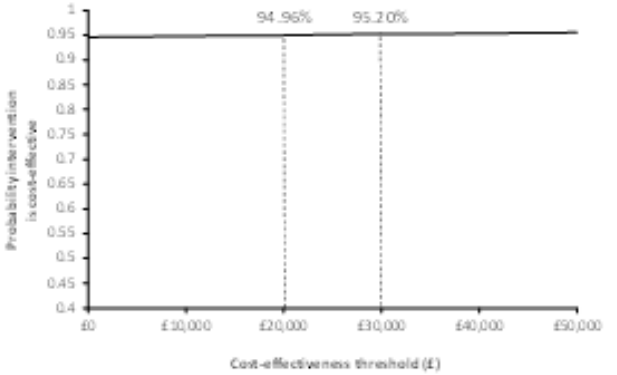
(H) FACT-G-Guided QALY*: CEAC curve



(I) FACT-AC-Guided QALY*: CE plane

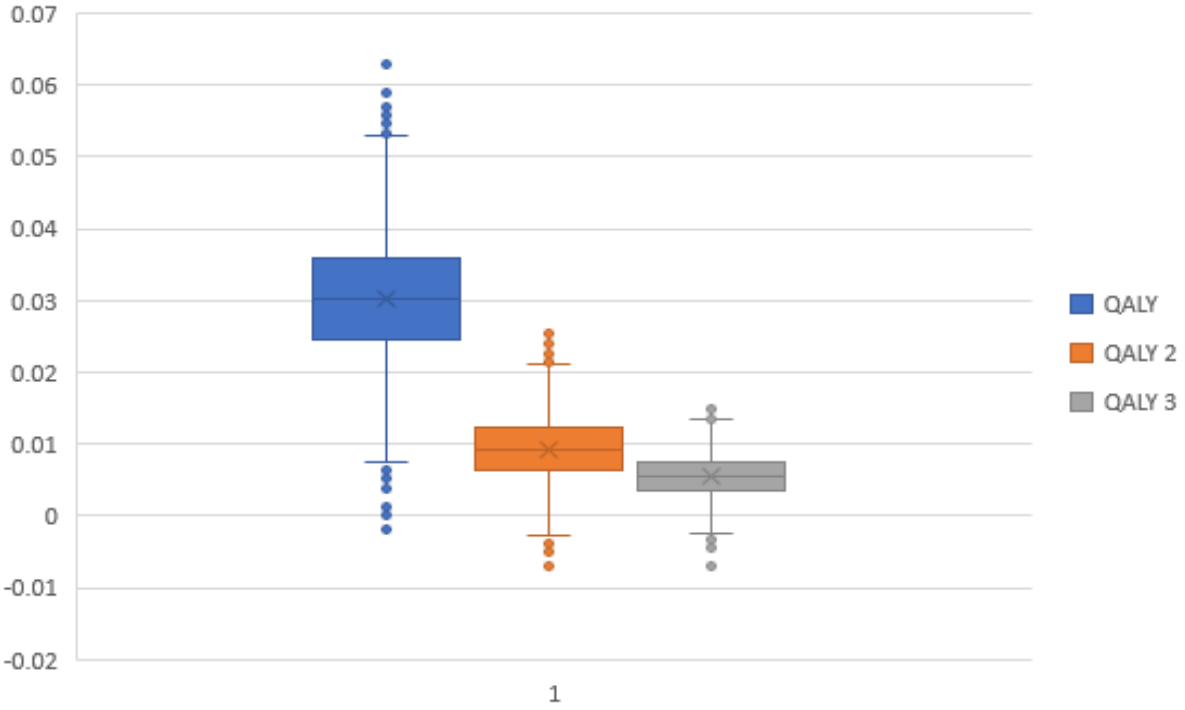


(J) FACT-AC-Guided QALY*: CEAC curve



***Footnote:** FACT-G-Guided QALY and FACT-AC-Guided-QALY were all corrected for baseline.

Figure 5.2: Box and Whisker plot of bootstrapped effects* data from COGNATE trial: QALY= Conventional QALY, QALY 2= FACT-G-guided QALY and QALY 3= FACT-AC-guided QALY



***Footnote:** Conventional QALY, FACT-G-Guided QALY and FACT-AC-Guided QALY were all corrected for baseline.

5.4.2 CLARITY trial 12-month follow-up results

Table 5.4 shows that, for the CLARITY trial, only the regression coefficient on VFQ-25-AUC was statistically significant ($p<0.05$) when conventional QALYs were regressed on VFQ-25-AUC alone (β-coefficient = 0.901) (see Model 2 in Table 5.4) and jointly with BCVA (β-coefficient = 0.911) (see Model 3 in Table 5.4). This indicates that there is relationship between VFQ-25-AUC and conventional QALY. This is also shown by the higher adjusted R-squared value for Model 2 ($r^2 = 0.258$) and Model 3 ($r^2 = 0.255$), where VFQ-25-AUC was used as independent variable in both of the models, compared to Model 1's ($r^2 = 0.006$), where VFQ-25-AUC was not used as independent variable in the model. This indicates that there was a stronger relationship in the models using VFQ-25-AUC as independent variable and conventional QALY as the dependent variable. Also, a higher adjusted R-squared value represents a better fit for the model. The adjusted R-squared value for Model 3 ($r^2 = 0.255$) was similar to Model 2's ($r^2 = 0.258$). Among the three models explored, the adjusted R-squared value for Model 2 was the highest ($r^2 = 0.258$), followed by Model 3's ($r^2 = 0.255$) and then Model 1's ($r^2 = 0.006$). The β-coefficient magnitude of VFQ-25-AUC implies that conventional QALY is expected to increase by 0.901 units for every 1-unit increase in VFQ-25-AUC when conventional QALYs were regressed on VFQ-25-AUC alone, and to increase by 0.911 units for every 1-unit increase in VFQ-25 when conventional QALYs were regressed on both VFQ-25-AUC and BCVA-AUC together. As a result, a hybrid QALY that takes account of the unit change in VFQ-25-AUC (also named as 'VFQ-25-AUC-guided QALY' in this Chapter) can be generated. The fitted regression equations for the three models for CLARITY trial can be expressed as below-

Model 1:

Y = (0.283 x BCVA-AUC) + 0.566, where

Y is all observed conventional QALY values

0.283 is the coefficient value for BCVA-AUC

0.566 is the constant a.k.a bias

Model 2:

Y = (0.901 x VFQ-25-AUC) + 0.022, where

Y is all observed conventional QALY values

0.901 is the coefficient value for VFQ-25-AUC

0.022 is the constant a.k.a bias

Model 3:

Y = (-0.064 x BCVA-AUC) +(0.911 x VFQ-25-AUC) + 0.065, where

Y is all observed conventional QALY values

-0.064 is the coefficient value for BCVA-AUC

0.911 is the coefficient value for VFQ-25-AUC

0.065 is the constant a.k.a bias

A numeric example demonstrating how the hybrid QALY was calculated-

Using Model 2 for example,

If VFQ-25-AUC equals to 0.920,

Therefore,

Hybrid QALY = (0.901 x VFQ-25-AUC) + 0.022

Hybrid QALY = (0.901 x 0.920) +0.022

Hybrid QALY = 0.851

Table 5.4: Exploratory analyses using the CLARITY trial 12-month data (N = 202): Results from the linear regression models using disease-specific measures to guide QALY.

Model no. (n=202)	Predictor(s)[¥]	β-coefficient	Standard Error	p-value*	95% Confidence Interval	Adjusted R-squared
1.	**BCVA-AUC**	0.283	0.192	0.141	-0.095 to 0.661	0.006
	Constant	0.566	0.156	0.000	0.258 to 0.874	
2.	**VFQ-25-AUC**	0.901	0.107	0.000	0.690 to 1.111	0.258
	Constant	0.022	0.093	0.813	-0.161 to 0.205	
3.	**BCVA-AUC**	-0.064	0.171	0.710	-0.401 to 0.274	
	VFQ-25-AUC	0.911	0.110	0.000	0.693 to 1.129	0.255
	Constant	0.065	0.148	0.662	-0.228 to 0.358	

*****Significance level at p=0.05; [¥]All predictors were corrected for baseline

Table 5.5 below shows the means and standard deviations of the CLARITY trial's effects data for intervention (Aflibercept) and control (PRP) groups, the mean difference between groups with the corresponding bootstrap standard error and bootstrap 95% confidence interval. The mean difference between groups for a single or both disease-specific measures guided QALY

demonstrated mixed results compared to the mean difference in the conventional QALY between groups. Amongst the three disease-specific measures guided QALYs, the mean difference between groups for VFQ-25-AUC-guided QALY (-0.022, bootstrap 95% CI -0.052 to 0.008) and VFQ-25-AUC-BCVA-AUC jointly-guided QALY (-0.023, bootstrap 95% CI -0.053 to 0.007) were both consistent with the mean difference in the conventional QALY (-0.022, bootstrap 95%CI -0.081 to 0.035), except for the BCVA-AUC-guided QALY that had a positive mean difference between groups (0.004, bootstrap 95% CI -0.002 to 0.010). However, there were no statistically significant difference in means between groups for all the QALYs – the conventional QALY and the three exploratory disease-specific measures guided QALYs – in the CLARITY trial.

Between the two disease-specific measures, BCVA and VFQ-25, used in the regressions, results showed that BCVA-AUC had a positive mean difference between groups (1.312, bootstrap 95% CI -0.870 to 3.391), whereas VFQ-25-AUC had a negative mean difference between groups (-2.456, bootstrap 95% CI -5.726 to 0.875), and both of these means difference were not statistically significant.

Table 5.5: Effect data of CLARITY trial – Mean and standard deviation (SD) of allocated group, mean difference between groups, bootstrap standard error and bootstrap 95% confidence interval.

Effect*	**Intervention (Aflibercept) (n=101) Mean (SD)**	**Control (PRP) (n=101) Mean (SD)**	**Mean difference** between groups**	**bootstrap standard error**	**bootstrap 95% confidence interval**
BCVA-AUC	81.947 (7.315)	80.635 (7.772)	1.312	1.087	-0.870 to 3.391
VFQ-25-AUC	84.736 (12.657)	87.192 (10.572)	-2.456	1.656	-5.726 to 0.875
QALY	0.785 (0.234)	0.807 (0.174)	-0.022	0.029	-0.081 to 0.035
BCVA-Guided QALY	0.798 (0.021)	0.794 (0.022)	0.004	0.003	-0.002 to 0.010
VFQ-25-Guided QALY	0.785 (0.114)	0.808 (0.095)	-0.022	0.015	-0.052 to 0.008

BCVA&VFQ-25-Guided QALY	0.784 (0.114)	0.808 (0.095)	-0.023	0.015	-0.053 to 0.007

*All effects were corrected for baseline
**Mean difference between groups = Mean of Intervention group minus Mean of Control group

Table 5.6 shows the mean total costs per patient (£) by allocated groups and the mean difference between groups over 12 months in the CLARITY trial. Results showed that Aflibercept group had a higher mean total costs per patient (£8,210, SD £8,182) compared to PRP group (£2,430, SD £3,156). This generated a positive mean difference of £5,781 (bootstrap 95%CI £4,266 to £7,712) in total costs between groups.

Table 5.6: CLARITY trial – Costs by allocated group over 12 months (2019 price year).

	Intervention (Aflibercept) (n=101) Mean (SD)	**Control (PRP) (n=101) Mean (SD)**	**Intervention minus Control[†] (bootstrap 95% CI)**
Community care cost (£)	222 (315)	188 (233)	34 (-39 to 112)
Secondary care cost (£)	2,874 (7,851)	1,325 (2,763)	1,549 (145 to 3,382)
Intervention delivery cost (£)	4,854 (1,761)	517 (204)	4,337 (3,996 to 4,687)
Medication cost (Ophthalmology-related) (£)	261 (694)	400 (1,249)	-139 (-432 to 124)
Total costs/patient (£)	8,210 (8,182)	2,430 (3,156)	5,781 (4,266 to 7,712)

[†] Negative differences imply Aflibercept group has lower cost than PRP group.
Note: Sum didn't add up due to rounding issue

Although all the effects in the CLARITY trial showed no statistically significant differences, CE planes and CEAC graphs were plotted for exploration purposes. Figure 5.3 presents the CE planes and their corresponding CEACs for all the three disease-specific measures guided QALYs generated in this exploratory study – BCVA-AUC-guided QALY (Figure 5.3(G) and (H)), VFQ-25-AUC-guided QALY (Figure 5.3(I) and (J)) and BCVA-AUC&VFQ-25-AUC-guided QALY (Figure 5.3(K) and (L)) – as well as for the BCVA-AUC (Figure 5.3(A) and (B)), VFQ-25-AUC (Figure 5.3(C) and (D)) and conventional QALY (Figure 5.3(E) and (F)).

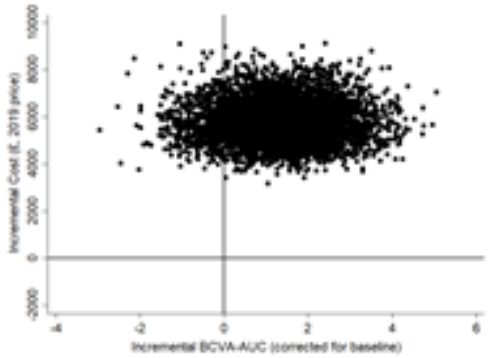
Compared to the CE plane plotting incremental cost against incremental conventional QALY (Figure 5.3(E)), the CE planes in Figure 5.3(G), (I) and (K) showed that, likewise in the COGNATE trial, the scatter plots became more concentrated around the estimates once the incremental

conventional QALY was replaced by the incremental disease-specific measures guided QALY. This is in agreement with the bootstrap standard error presented in Table 5.5 that showed that all the disease-specific measures guided QALYs had a smaller bootstrap standard error than the conventional QALY, indicating lesser variability of the data. Graphically, a visual representation of the bootstrapped conventional QALY data and disease-specific measures guided QALYs data was demonstrated using the Box and Whisker plot in Figure 5.4. Overall, all bootstrapped disease-specific measures guided QALYs data had shorter box length compared to the bootstrapped conventional QALY data – This means narrower distribution, that is, less scattered data (i.e. more concentrated data). Besides, the scatter plots in Figure 5.3 (G), (I) and (K) shows shifting towards disease-specific measure's direction once the incremental conventional QALY was replaced by the incremental disease-specific measures guided QALY in cost-utility analysis.

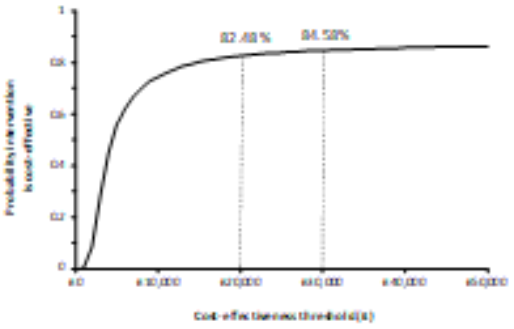
The CEACs graphs in Figure 5.3 show that, at a high willingness-to-pay threshold, for example at £500,000 threshold, Aflibercept had 12.76% probability of being cost-effective when conventional QALY was used as a measure of effect in CLARITY trial; however, this probability (12.76%) dropped to 1.24%, 1.40% and 1.12% when BCVA-AUC-guided QALY, VFQ-25-AUC-guided QALY and BCVA-AUC&VFQ-25-AUC-guided QALY were instead used as a measure of effect, respectively.

Figure 5.3: Cost-effectiveness planes with 5000 bootstrapped incremental costs and effects pairs from the CLARITY trial data and the corresponding cost-effectiveness acceptability curves (CEACs): an exploratory analysis of the usability of disease-specific measures in cost-utility analysis

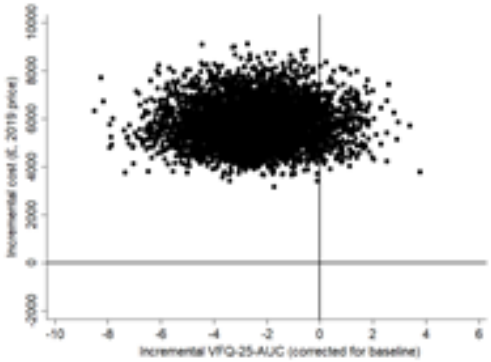
(A) Best Corrected Visual Acuity (BCVA)*: CE plane



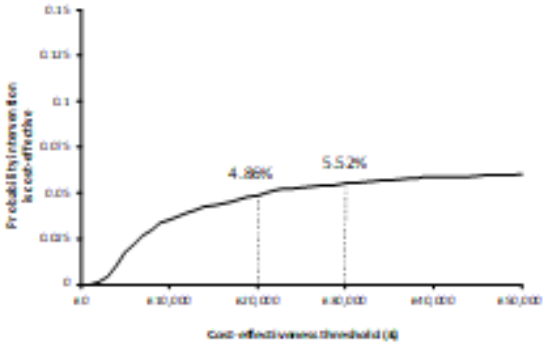
(B) Best Corrected Visual Acuity (BCVA)*: CEAC curve



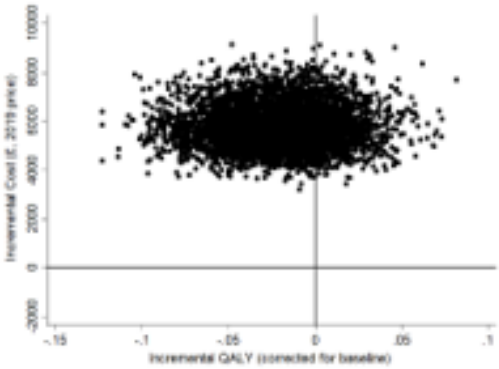
(C) Visual Function Questionnaire-25 (VFQ-25)*: CE plane



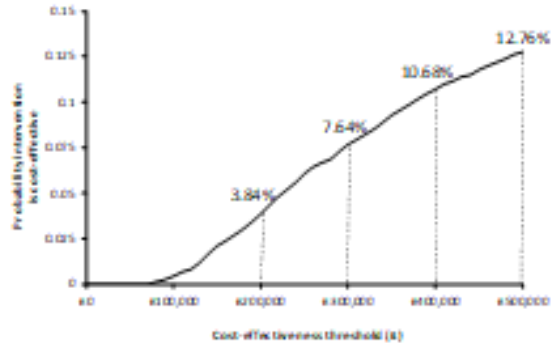
(D) Visual Function Questionnaire-25 (VFQ-25)*: CEAC curve



(E) QALY*: CE plane

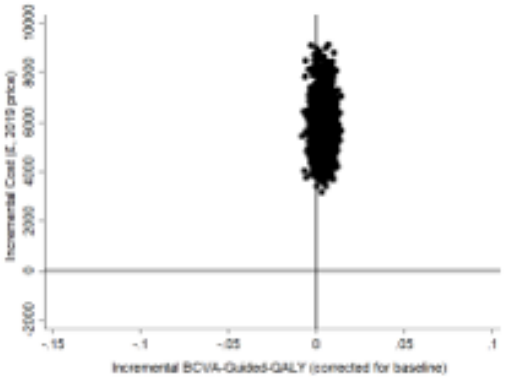


(F) QALY*: CEAC curve

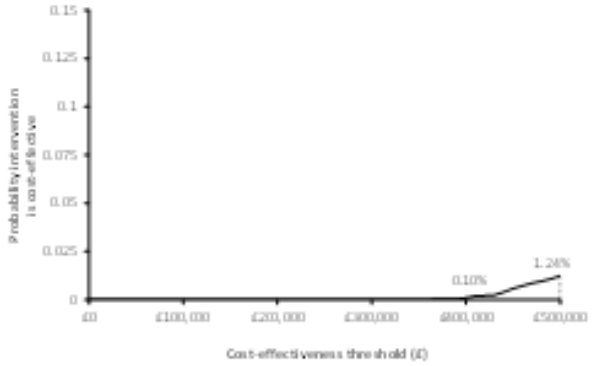


***Footnote:** BCVA-AUC, VFQ-25-AUC and QALY were corrected for baseline

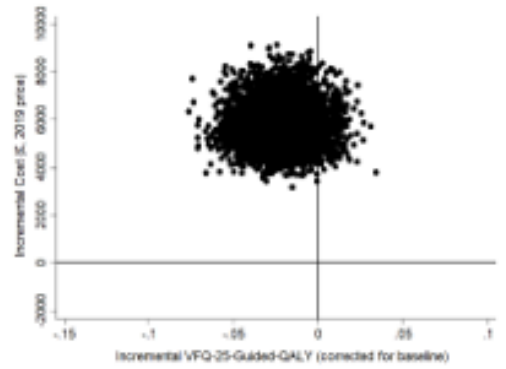
(G) BCVA-Guided QALY*: CE plane



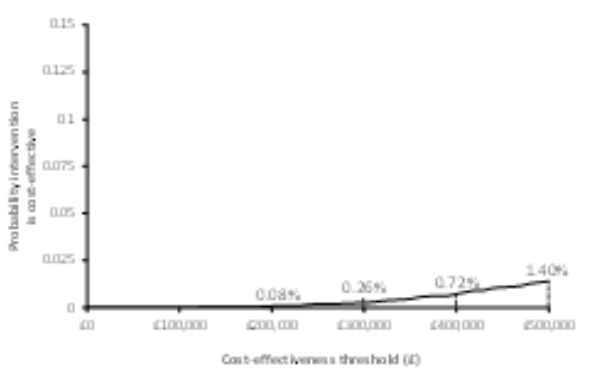
(H) BCVA-Guided QALY*: CEAC curve



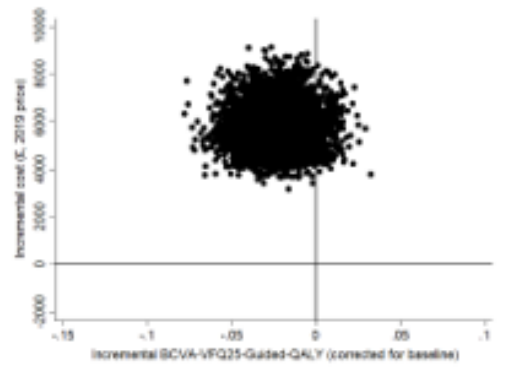
(I) VFQ-25-Guided QALY*: CE plane



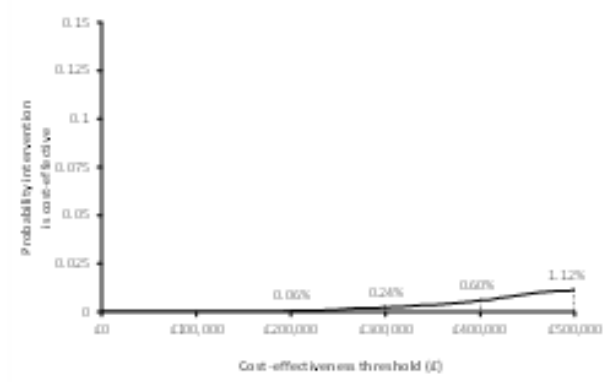
(J) VFQ-25-Guided QALY*: CEAC curve



(K) BCVA&VFQ-25-Guided QALY*: CE plane

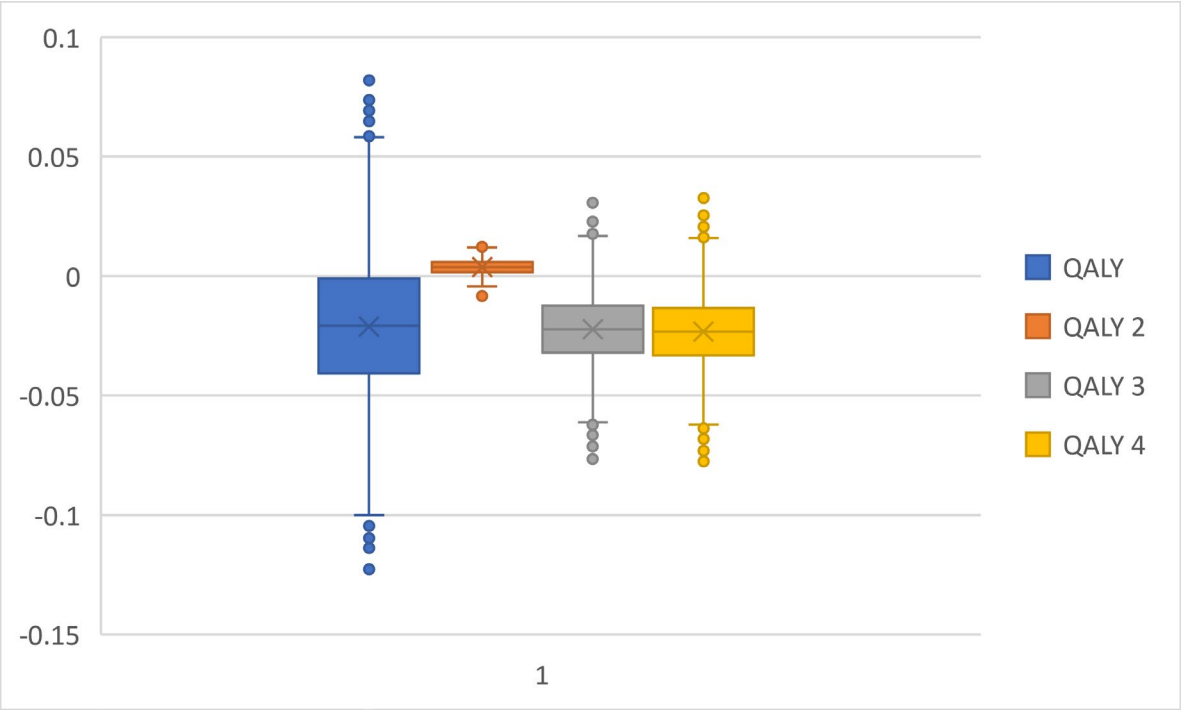


(L) BCVA&VFQ-25-Guided QALY*: CEAC curve



***Footnote:** BCVA-Guided QALY, VFQ-25-Guided QALY and BCVA&VFQ-25-Guided QALY were corrected for baseline.

Figure 5.4: Box and Whisker plot of bootstrapped effects* data from CLARITY trial: QALY= Conventional QALY, QALY 2= BCVA-guided QALY, QALY 3= VFQ-25-guided QALY and QALY 4= BCVA&VFQ-25-guided QALY.



***Footnote:** Conventional QALY, BCVA-Guided QALY, VFQ-25-Guided QALY and BCVA&VFQ-25-Guided QALY were all corrected for baseline.

5.5 Discussion

5.5.1 Main Findings

Summary of results –

From the COGNATE trial, the regression coefficient (β-coefficient) generated using FACT-G-AUC or FACT-AC-AUC showed a statistically significant relationship with the conventional QALYs (p-value < 0.05). FACT-G-AUC (β-coefficient = 0.396) showed a stronger relationship with the conventional QALYs compared to FACT-AC-AUC (β-coefficient = 0.249). From the CLARITY trial, only VFQ-25-AUC showed a statistically significant relationship (p-value < 0.05) with the conventional QALYs either it was regressed individually or jointly with BCVA.

Main findings –

Overall, among the five models explored, the models using (1) FACT-G individually as -independent variable in regression for the COGNATE trial and (2) VFQ-25 alone or together with BCVA as independent variable(s) in regression for the CLARITY trial had a higher adjusted R-squared value indicating a stronger relationship between the independent variable and the

dependent variable (i.e. the conventional QALYs). Furthermore, the statistically significant coefficients of the independent variables (FACT-G-AUC for the COGNATE trial; VFQ-25-AUC for the CLARITY trial) imply that there is a relationship between the independent variable and the dependent variable (i.e. the conventional QALY), where a one-unit change in the independent variable will give a mean change in the conventional QALYs. In such case, the hybrid QALY that takes account of disease-specific measures can be generated.

In the CLARITY trial, the independent variable 'BCVA' was found to have no statistically significant relationship with the dependent variable 'conventional QALY' when it was used either solely or jointly with VFQ-25 as independent variable in regression. This could be due to the fact that BCVA is an objective measure assessing only one very specific point of measure. And, unlike subjective measures, an objective measure is independent of individual experiences in which its data are derived from medical record. Hence, this may be a contributing factor in which BCVA was found to not even have a statistical significant relationship with the conventional QALY when it was used individually or jointly with VFQ-25. Therefore, it seems like objective measures may not be as appropriate as subjective measures for use as independent variable to better tune conventional QALYs in the hybrid method. This is the first exploratory study of its kind, thus, more studies are needed to affirm this.

When assessing the disease-specific measures guided QALY (i.e. the hybrid QALY), for the COGNATE trial, both the mean difference values of the FACT-G-guided QALY and FACT-AC-guided QALY were lower than the mean difference value of the conventional QALY. Even so, the mean difference values of the disease-specific measures guided QALYs argubly may be more appropriate for use in cost-utility analysis compared to the mean difference value of the conventional QALYs. This is because the disease-specific measures guided QALYs are better tuned to the effects experienced within the population group. By using the disease-specific measures guided QALY instead of the conventional QALYs for cost-utility analysis, it may reflect better the 'real' health-related quality of life status of the particular condition-specific population group. As a consequence, using it for cost-utility analysis may make the cost-utility analysis result of a trial more focused to the studied disease area. This could be explained by the inclusion of the studied disease-specific elements in the QALY calculation which could make the cost-utility analysis result of a trial better tuned to the effects experienced within the disease population group. However, more studies covering a broader

range of disease areas are required before this novel concept of the hybrid method could be accepted widely for producing a better tuned cost-utility analysis result to the specific studied disease area.

For the CLARITY trial, all the hybrid QALYs (BCVA-AUC, VFQ-25-AUC and BCVA-AUC&VFQ-25-AUC-guided QALYs) as well as the conventional QALY had insignificant mean difference between trial arms. Of the three hybrid QALYs, only the VFQ-25-AUC-guided QALY and the VFQ-25-AUC-BCVA-AUC jointly-guided QALY had a very similar mean difference value between trial arms and these mean difference values were closely similar to the mean difference value of the conventional QALY. The insignificant mean difference values of all the QALYs (conventional and hybrid QALYs) in the CLARITY trial could be due to the fact there is no effect between the treatment groups which would confirm the findings of the original CLARITY trial. Previous studies have expressed concerns with regards to the validity of the EQ-5D in visual disorders (Tosh et al., 2012). Brazier et al. (2017a) found that the preference-based VFQ (VFQ-UI) was more sensitive to changes in visual acuity-related quality of life than the EQ-5D in patients with diabetic macular edema. For this reason and despite the insignificant findings of all the effects in the CLARITY trial, the hybrid QALY (i.e. disease-specific measures guided QALY) would still seem to be more appropriate and meaningful than the conventional QALYs for use in cost-utility analysis for the cost-effectiveness evaluation within a specific disease context.

Of all four variables (FACT-G, FACT-AC, VFQ-25 and BCVA) tested for the hybrid method in this exploratory study, the BCVA is qualitatively different from the other three measures in that it is a measure of visual function that is measurable and independent of individual experiences (objective measure) rather than self-reported disease specific QoL measure that is based on individual awareness and experience (subjective measure) (Cleary, 1997). The scatter plots on the cost-effectiveness plane in Figure 5.3(g) that showed a very narrow looking shape with the plots concentrating around zero incremental BCVA-guided QALY could be largely explained by the fact that the BCVA is a specific measurable variable and not a self-reported disease specific QoL measure that covers different aspects of health. It seems like the objective measure (here in BCVA) combined in this way for the hybrid method may not be appropriate as it is not influenced by the individuals as subjective measures are (here in VFQ-25, FACT-G, FACT-AC and EQ-5D). This conversely demonstrated that the hybrid QALY method

might be not equally suited as a clinical measure (e.g. BCVA) which is measured objectively and is independent of individual experiences, compared to disease specific QoL measures that are measured subjectively and based on individual experience. This might explain the limited independent explanatory power of BCVA for EQ-5D scores.

5.5.2 Comparisons with Other Studies

No direct comparisons could be made as this is the first study of its kind in the area of cost-utility analysis in cancer and opthalmology trials. Although no studies (other than MORTISE) have used the same method, several studies have used alternative methods to try to achieve the goal of improving the disease specific sensitivity of preference-based measures in cancer or vision research. For example, Herdman et al. (2020) compared the validity and responsiveness of FACT-8D and ED-5D-5L and found that the EQ-5D performed as well as or better than the FACT-8D in patients with lymphoma. Luo et al. (2015) and Gandhi et al. (2020) concluded that the vision bolt on EQ-5D was more discriminative than the standard EQ-5D in vision problems. Brazier et al (2017a) found that the VFQ-UI was more sensitive than the EQ-5D in diabetic macular edema.

With regards to which approach(es) might be most appropriate in future studies where there is concern that the conventional EQ5D is not sensitive, studies could choose to use approach(es) such as using 'Preference-based condition-specific quality of life measure' (e.g. VFQ-UI, DemQoL-U), or 'EQ5D-Condition specific bolt on' (e.g. EQ5D-vision bolt on), or a mapping approach but choosing which approaches to use depends on the available research resources and budget in a study/trial. This is because the intensity of these approaches is variable. Different approaches require different amount of resources and budget.

The benefit of the hybrid approach explored here in this Chapter 5 is that analysts can perform it as a sensitivity analysis like the MORTISE trial without extra resource if disease-specific quality of life measure is collected alongside EQ-5D. In terms of being able to assess disease-specific cost-utility quickly as a sensitivity analysis, this is very resource conservative.

As disease specific sensitivity analysis (i.e. the hybrid approach) is not going to be resource intensive compared with all the other methods where they require development, validation

and verification, this is another way of looking at the data collected in trial in a more disease-specific way. It will not necessary be performed as a primary analysis but maybe as a sensitivity analysis. This would give indication directly for patients in those disease populations within the context of the trial that is being analysed in.

Compared with other methods, in terms of the impact on patient treatment, analysts would get two estimates of cost-utility analysis – One would be generic that would be able to compare across all disease areas which is advocated by the NICE guidance in the UK (i.e. QALY) and the other one would be more disease specific (i.e. hybrid QALY or disease-specific measure guided QALY) and would optimise information generated in the trial.

Another benefit of this hybrid approach is that analysts would be able to build regression algorithm from the actual data itself within the respective trial for use to generate disease-specific measure guided QALY (hybrid QALY) directly for use in cost-utility analysis for a more disease-specific sensitive result.

Furthermore, the scatter plots of bootstrapped incremental cost and effect pairs on the incremental cost-effectiveness planes of both the COGNATE and CLARITY trials illustrate the uncertainty surrounding the estimates of incremental costs and incremental effects of two competing alternatives had become less dispersed (i.e. more precise) and the cloud of the scatter plots had shifted towards the direction of the respective disease-specific measures. This is in line with the MORTISE trial's findings where the similar methodology was employed to modify its conventional QALY (Edwards et al., 2015). The MORTISE trial was a study looking at the effectiveness and cost-effectiveness of steroid (Methylprednisolone) injections versus anaesthetic alone for the treatment of Morton's neuroma (Edwards et al., 2015; Thomson et al., 2013). The authors of the MORTISE trial found that the cloud of the scatter plots of the bootstrapped incremental costs and incremental QALYs that had taken account of the Foot Health Thermometer (FHT), the disease-specific measure of the trial, had demonstrated a shift towards the direction of the FHT (Edwards et al., 2015).

5.5.3 Strengths and Limitations of the Study

Though to date, only one study – that is the MORTISE trial by Edwards et al. (2015) – had utilised the disease-specific measure of the trial in cost-utility analysis of steroid injections

versus anaesthetic alone for the treatment of Morton's neuroma (MORTISE trial), this exploratory study described in this chapter is the first study of its kind in exploring the transferability/generalisability of the hybrid method first tried in the MORTISE trial in other disease areas (here in gastro-oesophageal cancer (COGNATE trial) and ophthalmology (CLARITY trial)).

The linear regression approach may be improved upon as only disease-specific measure was included as independent variable in the regressions in this exploratory study which this was aimed to ease the identification of the effect of hybrid QALY has had on the cost-utility results for the exploratory purposes. Consideration of other types of regression modelling could be undertaken in future studies. One other benefit of the hybrid technique is that the hybrid QALYs are generated from the actual data within the trial.

Based on the findings from this exploratory study using the datasets from the COGNATE and CLARITY trials, the cost-utility analyses using the hybrid QALYs (i.e. disease-specific measure guided QALYs) seem to yield more precise findings than using the conventional QALY. Likewise, Edwards and colleagues from the MORTISE trial (Edwards et al., 2015) found that the results of their cost-utility analysis were enhanced when the FHT scores were taken into account in the QALY calculation to generate FHT-AUC-guided QALY for the cost-utility analysis.

Though research to date suggests that disease-specific measures have potential use in getting a better tuned QALY for cost-utility analysis, currently this novel concept of the hybrid method for QALYs however can only be used to make comparison within the same conditions/disease areas. More studies covering a broader range of conditions/disease areas are required to allow a wider coverage of knowledge of the effect of this concept of the hybrid method for QALYs on the cost-utility analysis results in other disease areas.

In general, the effect of hybrid QALY (i.e. disease-specific measures guided QALY) on the findings of cost-utility analysis is unique for each trial. In a trial that employed this approach, specific regression models between disease-specific measures-AUC and EQ-5D QALY (i.e. conventional QALY) will have to be developed for each trial. The established regression model will subsequently be used to calculate disease-specific measures guided QALY – A hybrid

QALY where disease-specific measures are factored into the EQ-5D scales. The Morton's neuroma study by Edwards and colleagues (2015) that employed the similar methodology in their cost-utility analysis described that the resulting regression model uses the greater discrimination achieved by the disease-specific measure of the study (i.e. the Foot Health Thermometer) to fill gaps in the simplistic three-point scales that characterised the original EQ-5D-3L that was used in their study. Therefore, using hybrid QALY (i.e. disease-specific measures guided QALY) in cost-utility analysis may provide better tuned results to the effects experienced within the disease population group. However, further research in this area is recommended to give a far-reaching insight into the effect of incorporating condition-/disease-specific measures in QALY calculations on the findings of cost-utility analysis in other condition/disease areas.

Last but not least, in this exploratory study, only disease-specific measures-AUC were used as independent variables in the regressions. Therefore, any changes in the results of cost-utility analysis using disease-specific measures guided QALY can be identified as solely due to the effect of disease-specific measures.

5.5.4 Future Research

Up to now, only one clinical trial [the MORTISE trial by Edwards et al. (2015)] had utilised disease-specific measure guided QALY (i.e. hybrid QALYs) to perform cost-utility analysis. This concept of hybrid QALY technique was further explored to assess the transferability/generalisability of the technique first tried in the MORTISE trial to other disease areas. Based on the explorative findings from these three disease areas, disease-specific measures guided QALYs (i.e. hybrid QALYs) seems to have shown effect on the findings of cost-utility analysis. However, more studies across broad range of disease areas are needed to gain a far-reaching insight into the added value of disease-specific measures in cost-utility analysis in clinical trials.

In addition, given that most if not all clinical studies normally collect both the EQ-5D and disease-specific measures data, conducting a sensitivity analysis using the hybrid QALY (i.e. disease-specific measures guided QALY) for cost-utility analysis could be recommended as part of the standard reporting procedure in health economics research alongside clinical

trials. Identifying resource conservative method (i.e. hybrid QALY method) would be a way of optimising with what we already have to provide disease specific sensitive estimates.

In terms of the impact on patient treatment, analysts would get two estimates of cost-utility analysis - One would be generic that would be able to compare across all disease areas which is advocated by the NICE guidance in the UK (i.e. QALY) and the other one would be more disease specific (i.e. hybrid QALY or disease-specific measure guided QALY) and would optimise information generated in the trial.

5.5.5 Implications for service commissioning

Based on the findings from this exploratory study, the disease-specific measures guided QALY seems to be more meaningful in reflecting the 'real' status of participants' quality of life compared to conventional QALY that measured solely from a generic health-related quality of life measure (e.g. the EQ-5D). This is because EQ-5D is not always sensitive enough to pick up changes in participants' disease-specific related quality of life (Pennington et al., 2020; Wichmann et al., 2017; Pettitt et al., 2016; Payakachat et al., 2015; Yang et al., 2015; Mulhern et al., 2013; Tosh et al., 2012; Whitehead and Ali, 2010). A holistic approach was taken in this exploratory study by incorporating the disease-specific measures area-under-the curve into QALY calculation to generate disease-specific measures guided QALY for use in cost-utility analysis. In other words, the disease-specific measures guided QALYs were derived by converting participants' responses to the disease-specific measures of the trial into utilities on the original EQ-5D scale – an innovative approach that demonstrates the adjustment of conventional QALY to the disease context of the trial. Having modified findings of cost-utility/cost-effectiveness of a new intervention or programme that takes account of studied disease area is crucial as the modified findings could help better decisions be made in an efficient way around health care resource allocations by policy makers or service commissioners to prioritise health care services within limited resources.

5.6 Conclusions

Applying a holistic approach to modifying QALY used in economic evaluation of health care interventions and/or technologies may improve decision-making around health care expenditure. From this exploratory study, we have learnt that disease-specific measures could be potentially useful in constructing a hybrid QALY for use in cost-utility analysis in

clinical trials to help inform health care resource allocation decisions more effectively and rationally, at least within specific conditions or disease areas, which use the same disease-specific guided QALY. It is possible that QALYs that take account of disease-specific measures are likely to more appropriately reflect the 'real' status of patient's quality of life compared to conventional QALYs. The concept of the hybrid QALY technique explored in this chapter is novel and could potentially contribute meaningfully towards future research. In addition, the insights gained from this exploratory study serve to expand future research in economic evaluation methodology on modifying QALY based on disease-specific measures used in clinical trials.

Chapter 6: Discussion

6.1 Chapter summary

The main findings of this thesis along with their corresponding research questions are presented in this final chapter.

In addition, comparisons of the main findings of this thesis with other studies are presented, as well as the implications for service commissioning, the strengths and limitations, and recommendations for future research. In this chapter, I also review the novel contribution of this thesis in a changing field of economic evaluation methodology aside from the novel contribution leading from other chapters in this thesis.

Finally, conclusions are drawn about the value of a holistic or universal "one size fits all" economic study of EUS in GOC staging to decision-makers in UK healthcare organisations, as well as the value of the range of novel methodologies in QALY estimations in clinical trials.

6.2 Research questions and main findings

To date (up to 2021), (1) with the exception of the published COGNATE trial HTA report by my COGNATE team colleagues and I, there are no other published studies evaluating the cost-effectiveness of EUS in GOC staging alongside randomised trials until the inception of the HTA-funded COGNATE trial that formed part of this thesis; (2) with the exception of my published systematic review paper in 2019 (Yeo et al., 2019) as a result of this PhD, there are no other published studies undertaking a rigorous economic-specific systematic review of EUS in GOC staging; (3) there are no published studies surveying the utilisation of EUS in GOC staging and the current practice in the UK up until January 2021 when a similar, but not the same, UK survey by Jones and colleagues was published in February 2021, and (4) there are no further published studies exploring the novel hybrid methodology in QALY estimations for cost-utility analysis in clinical trials, except in the MORTISE trial study where my colleagues and I first tried the hybrid QALY methodology idea (Edwards et al., 2015). To fill this gap, four independently related studies on EUS staging in GOCs were developed to collectively provide a holistic economic study on the cost-effectiveness, economic evidence, and utilisation of EUS in GOC staging in the UK, as well as the further exploration of the novel hybrid methodology in QALY estimations in other disease areas i.e. cancer (COGNATE trial) and ophthalmology (CLARITY trial) for cost-utility analysis in clinical trials. In achieving these milestones, four

research questions were developed and presented in earlier Chapters (see "Chapter 1: Introduction"). In this final chapter, these research questions are listed below with a summary of the main findings.

Thesis Chapter 2

Research Question 1: Is adding EUS to the standard staging algorithm cost-effective in the management of patients with GOC?

Main Findings: From an NHS perspective, economic evaluation of the COGNATE trial (a HTA-funded, pragmatic, multi-centred, randomised controlled trial) showed that, on average, participants in the EUS group (intervention group) gained 0.197 more QALYs compared to participants in the non-EUS group (control group); and cost an average of £2,860 less per patient (2019 price: £3,490 less per patient) than controls as a result of fewer or shorter hospital stays. Combining these effectiveness and cost findings by bootstrapping, EUS has 97% probability of being cost-effective at the NICE thresholds of £20,000 and £30,000 per QALY.

Sensitivity analyses using two alternative estimates of the unit cost of EUS (£1,477 as outpatient or £3,781 as inpatient rather than £551 as day patient in the primary analysis) suggested that EUS remains cost-effective with a probability of 95% and 86% respectively at the lower NICE threshold of £20,000 per QALY.

Sub-group analysis by initial quality of life showed that EUS was more cost-effective for participants reporting poorer initial health (below median EQ-5D score at baseline) despite saving less. Even for initially healthier participants (above median EQ-5D score at baseline), the intervention had at least 73% probability of being cost-effective at the NICE threshold if EUS costs £551.

Additional sub-group analysis by age group showed that EUS was more cost-effective for younger participants (age<65) with a probability of 95% at the NICE thresholds. Even for older participants (age>=65), the intervention had a considerably high probability (greater than 80%) of being cost-effective at the NICE thresholds of £20,000 and £30,000 per QALY.

Thesis Chapter 3

Research Question 2: What is the economic evidence on the use of EUS staging in patients with GOC?

Main Findings: This is the first robust systematic review of economic evidence of EUS staging in GOCs. This review was carried out and reported in accordance with the updated PRISMA guidelines (Moher et al., 2009; Liberati et al., 2009). Six studies, conducted between 1996-2018, met the inclusion criteria: three were economic studies and another three were economic modelling studies. The searches were updated to 2021 and no relevant paper was found. Study quality was generally satisfactory across all the six studies, but only one of these studies met all reporting and quality criteria – that was the study by Russell et al. (2013).

Of the three economic studies, one was an incremental cost-utility analysis and two were cost analyses. Of the three economic modelling studies, one was an incremental cost-utility analysis and two were cost-minimisation analyses. Both of the incremental cost-utility analyses reported that use of EUS as an additional staging technique offered, on average, more QALYs (0.0019-0.1969 more QALYs) and saved costs (by £1,969-£3,364 per patient, 2017 price year; by £2,037-£3,481 per patient, 2019 price year) compared to staging strategy without EUS.

Thesis Chapter 4

Research Question 3: How is EUS in GOC staging used in current practice in the UK?

Main Findings: Up to January 2021, this is the first survey study of its kind in the field of EUS staging in GOC in the UK that explored clinicians' views on the utilisation of EUS in GOC staging, its current clinical practice in the UK, potential factors associated with referral for EUS, and factors that are considered to have influence on the utilisation of EUS in GOC staging.

The majority (greater than 90%, n = 89 out of 98) support the use of EUS for staging GOC as they believed that EUS offers the best local regional staging of GOC and the opportunity to biopsy during staging. Furthermore, they acknowledged that EUS helps to reduce unnecessary surgery for GOC patients.

More than four-fifths of the survey respondents (n = 81 out of 98) reported that they have had experience in the field of EUS either by requesting, performing or both. In terms of

clinicians' perceptions about the clinical utility of EUS in GOC staging, these respondents perceived EUS is more useful for staging oesophageal and gastro-oesophageal junction cancers than gastric cancer. This finding is consistent with their response to the type of GOC patients they usually referred for EUS staging.

With respect to clinicians' current use of EUS, more than half of these respondents (n = 47 out of 81) responded that they would not necessarily routinely request EUS for staging all newly diagnosed GOC patients although the majority supported the use of EUS for GOC staging. About three-quarters of these respondents (n = 59 out of 81) reported that EUS is available within their hospital. The remaining respondents who reported EUS is not available within their hospital stated that they usually referred patients to another hospital situated on average about 21 miles away (SD 21.7 miles) from their hospital for EUS screening. Benefits may be gained if EUS could be offered in more clinical settings to avoid delays in the diagnosis and treatment planning of GOCs. However, it is acknowledged that there are logistical, cost and training challenges in providing EUS at every hospital in the UK.

Based on the responses to the pre-validated clinical vignettes, clinicians' recommendations in treatment decisions (resection versus less invasive or palliative therapy) varied depending upon detailed differences in tumour stage resulting from EUS scan/staging.

Various factors could influence the referral of GOC patients for EUS in clinical practice. These include factors such as variations in clinical management styles, perception of clinical utility of EUS, availability of EUS, whether clinicians attend MDT meeting, clinicians' age and years in practice (Kim and Koch, 1999; personal communications with Dr Hasan Haboubi, Consultant Gastroenterologist and Clinical Advisor to this thesis). When the association of these potential factors with referral for EUS were explored, two factors – 'attend Multidisciplinary Team (MDT) meeting' and 'clinicians' age' – demonstrated significant association ($p<0.05$) with referral of GOC patients for EUS, with 'attend MDT meeting' was the most important factor. The odd ratios of these two factors indicated that clinicians who attend MDT meeting appear to be more likely to refer GOC patients for EUS, and clinicians who are younger in age appear to be less likely to refer GOC patients for EUS.

Given that Teaching Hospitals potentially having more funding (and staff) and therefore more access to specialist tests, and may be more radical in their approach to patient care compared to General Hospitals which may be more conservative, the factor 'hospital type' (Teaching Hospital versus District General Hospital (DGH)) was also assessed and it was found to have no significant association with referral for EUS for GOC staging though Teaching Hospitals may have better outcomes than General Hospitals in cancer care (Burke et al., 2017). Hence, this means that 'hospital type' is not an important factor for EUS referral.

Thesis Chapter 5

Research Question 4: Has disease-specific measure got potential use in QALY estimations for cost-utility analysis?

Main Findings: This exploratory study is the expansion of the concept of the "hybrid QALY technique" first tried in the MORTISE trial (on which I was a Research Officer working with the Trial Statistician Dr Daphne Russell) (Edwards et al., 2015) to other disease areas (e.g. cancer (COGNATE trial) and ophthalmology (CLARITY trial)). The exploratory study in this Chapter was to further generalise the methodology in the MORTISE trial to explore the transferability of the hybrid approach to other disease areas. Data from two large clinical trials, each with a 12-month follow-up [COGNATE, a HTA-funded cancer trial (Trial registration: ISRCTN1444215), and CLARITY, a NIHR-EME-funded non-cancer trial (Trial registration: ISRCTN32207582)] were utilised for the purpose of this exploratory study. In these two clinical datasets, participants' responses to the disease-specific measures of the two trials were converted into utilities on the original EQ-5D scale, using a novel innovative approach in which disease-specific measures guided QALYs were generated for cost-utility analysis. The incremental cost-effectiveness planes and cost-effectiveness acceptability curves constructed from cost-utility analysis using disease-specific measures guided QALY were compared with cost-utility analysis using conventional QALY.

In general, the effect of the disease-specific measures guided QALY on the findings of cost-utility/cost-effectiveness is unique for each trial. For example, in this exploratory study, it was found that, for the COGNATE trial, there was a decrease in the mean difference in the disease-specific measure guided QALYs ($p < 0.05$) compared to the mean difference in the conventional QALY ($p < 0.05$). However, the cost-effectiveness acceptability curves for all the disease-specific measures guided QALYs showed that the probability of the new intervention

being cost-effective at the NICE thresholds of £20,000 and £30,000 per QALY remained very closely the same as the probabilities for the conventional QALY.

In both trials, cost-utility analysis results showed not only more certainty around the estimates of incremental cost-effectiveness ratios when conventional QALY was replaced with disease-specific measures guided QALY but also a shifting towards the direction of the respective disease-specific measures.

6.3 Comparisons with other studies

Thesis Chapter 2

***Title:* Cost-effectiveness of EUS in the management of patients with GOC: findings from the COGNATE trial**

No direct comparisons could be made as, to date, this is the first study of its kind in this area in the evaluation of the economic values of EUS in managing patients with GOC. Furthermore, this is also the first economic evaluation study in this area that used primary economic data collected prospectively alongside a pragmatic randomised controlled trial (namely 'COGNATE' trial); hence this economic study offers novel contributions to existing literatures, particularly to the economic evidence of EUS in GOC staging as well as to the health technology assessments in the UK. This economic evaluation study was published in the HTA in 2013 as part of the COGNATE trial (Russell et al., 2013).

Thesis Chapter 3

***Title:* Systematic review of economic evidence of EUS in GOC staging**

No comparisons to other systematic review studies could be made as, to date, this is the one and only robust systematic review of economic evidence of EUS in GOC staging. Though a non-economic specific systematic review in this field was published 20 years ago by Harris et al. (1998), there was no health economics-related study identified in their systematic review. Hence, after all the searches had been updated to the year 2021, this systematic review is considered to be the most up-to-date collection of economic literature in this area.

Thesis Chapter 4

***Title:* The utilisation of EUS for GOC staging and the current practice in the UK: a healthcare professional survey**

The reported findings of this survey study in relation to clinicians' knowledge and feelings about the use of EUS in GOC staging are in line with the findings reported in the recent survey undertaken by Jones and colleagues (Jones et al., 2021) investigating the use of EUS in the diagnosis and treatment of oesophageal cancer in the UK. Their survey study found that (1) EUS is mostly used as a complementary staging technique to PET-CT, either following PET-CT or at the same time as PET-CT, in the diagnostic pathway of oesophageal cancer; (2) for surgical planning, EUS is most commonly used to assess unresectable T4 disease, which is in line with the reported greater accuracy of EUS over PET-CT for T-staging (Thakkar and Kaul, 2020; DaVee et al., 2017; Valero and Robles-Medranda, 2017; Allum et al., 2011; Puli et al., 2008; Pfau et al., 2007); and (3) for treatment planning, EUS is most commonly used for selecting between surgery and chemoradiotherapy for more advanced but potentially resectable tumour which it helps to avoid unnecessary surgery for oesophageal patients (Jones et al., 2021).

In terms of clinicians' perceptions about the clinical utility of EUS in GOC staging, this survey's findings are in line with the review conducted by Lennon and Penman (2007) indicating that EUS is more accurate in staging oesophagus cancer than gastric cancer. With regards to the findings reported by most of the respondents that they would not necessarily routinely request EUS for staging all newly diagnosed GOC patients, this is supported by the findings of the similar UK survey, conducted recently by Jones et al. (2021). Jones and colleagues (2021) found that there is a proportion of their survey respondents would request EUS on a case-specific basis dependent on PET-CT results, and some would request EUS at the same time as PET-CT or routinely following PET-CT.

With regards to the findings relating to factors associated with referral of GOC patients for EUS, the findings of the two factors – 'availability of EUS' and 'perceived utility of EUS' – are in contrast to the results of the US survey study conducted by Kim and Koch (1999) where their study showed that these two factors were both significantly associated with EUS referral. This could be due to the following facts which may account for this difference –

- The availability of EUS has been increased since two decades ago although it varies across the hospitals in the UK.
- Clinicians' perceptions about the clinical utility of EUS is believed to have also increased since two decades ago alongside the evolution of EUS technology in clinical contexts.

- There are also other cultural differences in medical practice between the US health care system which is privately funded versus the UK National Health Service which is publicly funded, as well as the potential differences in access to these services and capacity to provide them in a publicly funded system.
- It is acknowledged that, with time, EUS has become accepted much more widely for various indications including diagnosis/staging of gastro-oesophageal cancers (Thakkar and Kaul, 2020; DaVee at al., 2017; Valero and Robles-Medranda, 2017; Sahai, 2012).

Thesis Chapter 5

<u>Title:</u> An exploratory study of novel methodology in QALY estimations using data from two clinical trials – COGNATE and CLARITY trials

No direct comparisons could be made as this is the first study of its kind in the field of cancer and ophthalmology in exploring the transferability/generalisability of the concept of the "hybrid QALY technique" first tried in the MORTISE trial (Morton's neuroma study) in other disease areas.

Additionally, the cost-utility analyses' results demonstrated in this exploratory study are in line with the findings of the MORTISE trial where the similar methodology was first tried to modify its conventional QALY (Edwards et al., 2015). The MORTISE trial was a study investigating the effectiveness and cost-effectiveness of steroid (Methylprednisolone) injections versus anaesthetic alone for the treatment of Morton's neuroma (Thomson et al., 2013; Edwards et al., 2015) on which I was the Research Officer of the trial. In the MORTISE trial, the Trial Statistician, Dr Daphne Russell and I first tried the concept of the "hybrid QALY technique" and found that the cloud of the scatter plots of the bootstrapped incremental costs and incremental QALYs that had taken account of the Foot Health Thermometer (FHT) scores, the disease-specific measure of the trial, had demonstrated a shift towards the direction of the FHT (Edwards et al., 2015).

6.4 Implications for service commissioning

A novel innovative approach was undertaken in Chapter 5 (the exploratory study chapter) given concerns have been raised over the lack of sensitivity and responsiveness of the preferred generic health-related quality of life tool – the EQ-5D (NICE, 2013), in detecting clinical changes in certain disease areas such as alcohol dependency, schizophrenia, limb

reconstruction, hearing impairment, visual disorders and dementia (Pennington et al., 2020; Payakachat et al. 2015; Yang et al., 2015; Mulhern et al., 2013). Furthermore, it is widely acknowledged that generic health-related quality of life instruments, such as the EQ-5D, are not always sensitive enough to capture particular aspects of certain conditions, such as functional and symptomatic gains from a new health care intervention or technology (Whitehead and Ali, 2010; Tosh et al., 2012; Pettitt et al., 2016; Wichmann et al., 2017; Pennington et al., 2020).

The novel innovative approach explored in Chapter 5 involves applying a holistic approach to modifying conventional QALYs by converting participants' responses to the disease-specific measures of trials into utilities on the original EQ-5D scale. This concept of the "hybrid QALY technique" uses the greater discrimination achieved by the respective disease-/condition-specific measures used in trials to fill gaps in the EQ-5D scales – a method of strengthening the EQ-5D, which we called the disease-/condition-specific-enhanced EQ-5D (Edwards et al., 2015). By expanding the novel innovative approach to other disease areas, e.g. cancer and ophthalmology areas as demonstrated in this thesis, it not only allows us to explore the transferability/generalisability of the "hybrid QALY technique" first tried in the MORTISE trial in other disease areas but also the effect of incorporating disease-/condition-specific measures into QALY estimations on cost-utility analysis results. Hence, with further studies to cover a broader range of disease areas, this may offer a more comprehensive evidence-based source of information in facilitating resource allocation decisions in health care setting.

The collective findings from the four independently related studies in this thesis are believed to not only have important implications for further improvements in the clinical area of EUS in GOC staging in the UK but also provide useful economic evidence-based information to policy makers and commissioners for informing policy and improving clinical practice in the UK.

6.5 Strengths and limitations of this thesis

To fill gaps in the literature about the use of EUS in the management of GOC patients, this thesis was developed to explore several important research questions related to the use of EUS in GOC staging from an economic perspective. These include questions about the economic value of EUS in GOC staging, the economic evidence and the utilisation of EUS in

GOC staging in the current practice in the UK, as well as the potential use of disease-/condition-specific measures in QALY estimations for cost-utility analysis. And to do so, a range of research methods were employed. These included a quantitative study applying cost-utility analysis to economic data collected alongside the COGNATE trial (Chapter 2), a systematic literature review (Chapter 3) and a survey-based quantitative study (Chapter 4). In addition, given the concerns over the insensitivity and unresponsiveness of the EQ-5D instrument (Pennington et al., 2020; Pettitt et al., 2016; Payakachat et al., 2015; Yang et al., 2015; Mulhern et al., 2013; Tosh et al., 2012; Whitehead and Ali, 2010), a quantitative-based exploratory study was undertaken to explore the transferability/generalisability of the "hybrid QALY technique" first tried in the MORTISE trial in other disease areas e.g. cancer and ophthalmology (Chapter 5), providing a further strength to this thesis. Within the chapters of this thesis, various novel contributions to clinical and economic literature have been made, adding to clinical and economic knowledge in this field.

In addition, this thesis had several other notable strengths which included the economic evaluation study was piggybacked onto the COGNATE randomised controlled trial, study sample used in the economic analysis of the COGNATE trial were the same as in its clinical effectiveness analysis, conduct of systematic literature review retrospectively, inclusion of different types of quantitative studies and use of strong and transparent methodology in each component of study in this thesis. Further details are described as below –

- ***Piggyback economic evaluation onto clinical randomised controlled trial***

  The economic evaluation study in this thesis was piggybacked onto the COGNATE randomised controlled trial; by doing so, it had several advantages which included (1) primary economic data, that collected prospectively alongside the COGNATE trial, were able to use for its analysis compared to other economic assessments studies such as modelling studies that use secondary data; (2) collection of patient-level data on costs and effects was made possible especially for a student with a limited research resources; (3) in terms of costs, it is more viable to do so especially for a student with a limited research budget.

- ***Same study sample used in the clinical and economic analysis of the COGNATE trial***

The economic analysis in the COGNATE trial used the same study sample as for the clinical effectiveness analysis (n=213 participants, see Chapter 2), leading to valid findings.

- ***Retrospective systematic literature review***

  Due to resource constraints, a literature review instead of a systematic review was undertaken during the COGNATE trial. In view of that, a retrospective systematic review was planned and conducted as part of this thesis (Chapter 3), with agreement from supervisors, after the completion of the COGNATE trial. This delay gives additional strength to the scope and knowledge in this field in a way that it enables more existing economic literature relevant to EUS in GOC staging (including our published COGNATE HTA report) be captured, identified, screened, synthesised and appraised together as a whole, leading to a more comprehensive and up-to-date systematic literature review of economic evidence of EUS in GOC staging.

- ***Inclusion of different types of quantitative studies***

  Different types of quantitative research were used in this thesis which included an economic-based (Chapter 2), a survey-based (Chapter 4) and a methodological-based (Chapter 5) quantitative studies. The combination of different types of quantitative studies is useful for gaining an overall picture of not only the economic value, but also the use of the new intervention/technology in the current clinical practice in the UK complimenting the aim of the study in evaluating the economic value of the new intervention/technology. With respect to the methodological-based quantitative study in Chapter 5, it allows for the concept of the "hybrid QALY technique" first tried in the MORTISE trial be explored and expanded to other disease areas for cost-utility analysis in response to the growing concerns over the insensitivity and unresponsiveness of the generic health-related quality of life instruments, such as the EQ-5D (Pennington et al., 2020; Pettitt et al., 2016; Payakachat et al., 2015; Yang et al., 2015; Mulhern et al., 2013; Tosh et al., 2012; Whitehead and Ali, 2010). It is acknowledged that the five-level version of EQ-5D (EQ-5D-5L) was developed to address the concerns to some extent (Emrani et al., 2020; EuroQoL Group, 2021). Although no studies (other than MORTISE) have used the same method, several studies have used alternative methods to try to achieve

the goal of improving the disease specific sensitivity of preference-based measures in cancer or vision research. For example, Herdman et al. (2020) compared the validity and responsiveness of FACT-8D and ED-5D-5L and found that the EQ-5D performed as well as or better than the FACT-8D in patients with lymphoma. Luo et al. (2015) and Gandhi et al. (2020) concluded that the vision bolt on EQ-5D was more discriminative than the standard EQ-5D in vision problems. Brazier et al (2017a) found that the VFQ-UI was more sensitive than the EQ-5D in diabetic macular edema.

The hybrid approach explored in Chapter 5 provides several benefits – This innovative approach can be performed as a sensitivity analysis like the MORTISE trial without extra resource if disease-specific quality of life measure is collected alongside EQ-5D. In terms of being able to assess disease-specific cost-utility quickly as a sensitivity analysis, this is very resource conservative. As disease-specific sensitivity analysis (i.e. the hybrid approach) is not going to be resource intensive compared with all the other methods where they require development, validation and verification, this is another way of looking at the data collected in trial in a more disease-specific way. It will not necessary be performed as a primary analysis but maybe as a sensitivity analysis. This would give indication directly for patients in those disease populations within the context of the trial that is being analysed in.

Compared with other methods, in terms of the impact on patient treatment, the hybrid approach would allow analysts to get two estimates of cost-utility analysis – One would be generic that would be able to compare across all disease areas which is advocated by the NICE guidance in the UK (i.e. QALY) and the other one would be more disease specific (i.e. hybrid QALY or disease-specific measure guided QALY) and would optimise information generated in the trial.

Another benefit of the hybrid approach is that analysts would be able to build regression algorithm from the actual data itself within the respective trial for use to generate disease-specific measure guided QALY (hybrid QALY) directly for use in cost-utility analysis for a more disease-specific sensitive result.

- ***Use of strong and transparent methodology***

  Each component of study in this thesis exhibits a strong and transparent methodology, of which they were designed in accordance with the aims, research questions and available resources. Each study in this thesis was conducted and reported in accordance with the relevant methodological guidance. The economic analysis in COGNATE study (Chapter 2) was conducted and reported in accordance with the published standard operating procedure for economic evaluation alongside randomised controlled trials (Edwards et al., 2008), published guidance (Ramsey et al., 2005; Glick et al., 2014; Drummond et al., 2015) and published CHEERS checklist (Husereau et al., 2013). The systematic review (Chapter 3) was conducted and reported in accordance with the updated PRISMA guidelines (Moher et al., 2009; Liberati et al., 2009). The online survey (Chapter 4) was designed and conducted in accordance with the published guidance for online questionnaire surveys (Regmi et al., 2016). The exploratory study of expanding the concept of the "hybrid QALY technique" first tried in the MORTISE trial to other disease areas (e.g. cancer and ophthalmology) for cost-utility analysis (Chapter 5) was conducted and reported in accordance with published guidance (Glick et al., 2014; Drummond et al., 2015) and published CHEERS checklist (Husereau et al, 2013), where appropriate.

- ***Use of additional clinical trial's datasets leads to more robust results***

  Datasets from another clinical trial (CLARITY trial) were used to conduct an additional exploratory study of expanding the concept of the "hybrid QALY technique" to other disease areas (e.g. cancer and ophthalmology) for cost-utility analysis (Chapter 5). This added further value to the exploratory study. Whilst using additional datasets from another clinical trial supports the results of the primary exploratory analysis study, this expansion of the exploratory study allows for comparison of results within this thesis to be made, leading to more robust results.

On the other hand, there are also several limitations to this PhD study. These included-

- ***Survey study sample marginally below the minimum target number of sample size***

  The total number of responses to the online survey in Chapter 4 is marginally below the minimum target number of sample size calculated for the survey although it provides an adequate level of precision. This may limit the power of the study to

identify factor less strongly associated with referral for EUS and may affect the generalisability of the study findings to GOC clinicians in the UK.

- ***Burden of survey questionnaire***

  The service evaluation questionnaire designed for the online survey was a 17-page long questionnaire that consisted of five sections. Although the online survey could be completed in approximately less than 15 minutes, the number of sections may have been overly burdening for participants (Appendix 4.2).

- ***No direct comparisons to other studies***

  For the exploratory methodological study in relation to advancing cost-utility analysis (Chapter 5), no direct comparisons to other studies could be made in the field of cancer and ophthalmology. This is because this is not only the first study of its kind in these disease areas but also the hybrid methodology explored for a more disease-specific sensitive result in cost-utility analysis (Chapter 5) is novel and innovative in the field of cancer and ophthalmology besides several studies have explored alternative methods to try to achieve the goal of improving the disease specific sensitivity of preference-based measures in cancer and vision research.

- ***No comparison to other randomised controlled studies of EUS staging in GOC***

  Besides the COGNATE trial, there are no other randomised controlled studies of EUS staging in GOC found in the UK as well as in other countries, based on the studies identified in the systematic review (Chapter 3). Thus, the economic evaluation study alongside the COGNATE trial in this thesis is the first study of its kind in this area. Hence, comparisons to other randomised studies in the similar clinical area are not possible.

6.6 Future research and recommendations

Given that there was a lack of health economics research collecting primary data on costs and effects (such as utility values to construct QALYs) of EUS staging in GOC, more primary health economic research in this area is recommended. In particular, integrated clinical and economic trials of EUS in GOC staging that can offer robust evidence of costs and effects.

In terms of consolidating the knowledge and experiences clinicians have acquired to date, additional cancer research studies including other survey studies in the field of EUS for GOC staging are recommended. This can help to strengthen the link between clinicians' knowledge and the clinical utility of EUS in GOC staging, which this will positively help influence the clinical practice in this field in the UK. For example, increased uptake of EUS for GOC staging will be a result of increased clinicians' knowledge and confidence in the clinical utility of EUS in GOC staging. Consequently, this will help improve clinical efficiency and quality of care in the delivery of diagnosis and treatment planning of GOC.

In addition, increased education efforts are required to encourage utilisation of EUS in GOC staging given EUS has been shown to be cost saving and offers greater QALYs (Yeo et al., 2019; Russell et al., 2013) and is recommended as a complementary staging technique to other staging techniques in the evaluation of GOCs (NICE, 2018a; Allum et al., 2011; Pfau et al., 2007). Concomitant efforts must be made to offer more EUS-related trainings and workshops to increase clinicians' awareness, perception, and knowledge of clinical utility of EUS.

Given that performing EUS requires technical proficiency, increased intensive yet comprehensive trainings in clinical utility of EUS for GOC staging is recommended to help develop technically trained and skilled experts in using EUS for GOCs diagnosis and treatment planning. This includes trainings in using EUS for staging GOC and performing EUS in the diagnostic and therapeutic procedures for GOCs as well as evaluating and interpreting EUS images (Dietrich et al., 2019; Penman et al., 2011; Carroll and Penman, 2004; Catalano et al., 1995). This allows clinicians to rely upon the detailed information resulting from EUS imaging of tumour staging to help guide their ongoing management. On the whole, increased availability of EUS more widely across the hospitals in the UK is vital to increase the accessibility and utilisation of EUS for GOC staging. Consequently, this will help avoid delays in the diagnosis and treatment planning of GOCs.

With regards to the concerns on the insensitivity and unresponsiveness of the EQ-5D tool in measuring individuals' health-related quality of life, more research on methodological development in modifying the conventional QALY by incorporating disease-/condition-specific measures in QALY estimations for use in cost-utility analysis is recommended. This will allow for more comparisons of results of cost-utility analysis using conventional QALY

versus hybrid QALY to be made, and hence, improve the research strength in this area. This could be achieved by either undertaking a stand-alone methodological study in this area or a sensitivity analysis embedded in a clinical trial. Undertaking a sensitivity analysis as part of trial is useful to assess the effect of incorporating disease-/condition-specific measures into QALY estimations on the outcomes of cost-utility analysis. Further research on applying this innovative hybrid approach to other disease areas is vital to provide comprehensive assessment of the effect of this approach and hence allow a robust conclusion to be made. As a result, a conclusive decision can be made on the usefulness of incorporating disease-/condition-specific measures in QALY estimations for use in cost-utility analysis in different disease areas.

Besides the hybrid QALY method explored in this thesis, it is acknowledged that several studies have used alternative methods (e.g. disease-specific preference-based measures and EQ-5D disease-specific bolt-on) to try to achieve the goal of improving the disease specific sensitivity of preference-based measures in cancer or vision research. Hence, question as to which approach(es) might be most appropriate in future studies where there is concern that the conventional EQ-5D is not sensitive, studies could choose to use approach(es) such as using 'condition-specific preference-based quality of life measure' (e.g. VFQ-UI, DemQoL-U), or 'EQ5D-disease-specific bolt on' (e.g. EQ5D-vision bolt on), or a mapping approach but choosing which approaches to use depends on the available research resources and budget in a study/trial. This is because the intensity of these approaches is variable. Different approaches require different amount of resources and budget.

6.7 Novel contribution of this thesis

The studies reported in this thesis offer multiple novel contributions to both health economics and policy.

6.7.1 Contribution to methodology

The studies reported in this thesis used different types of quantitative research of which the combination offers a holistic insight into the economic value of EUS in managing patients with GOCs and the use of EUS in GOC staging in the current practice in the UK complimenting the overarching aim of this thesis in evaluating the economic value of EUS in GOC staging. Together, these build a holistic economic study of EUS in GOC staging which has not previously

been conducted. The various types of quantitative studies used in this thesis included an economic-based quantitative study (Chapter 2), a survey-based quantitative study (Chapter 4) and a methodological-based quantitative study (Chapter 5).

The methodological-based quantitative study presented in Chapter 5 is the first study of its kind in the area of cancer and ophthalmology in health economic research. The study used the concept of the "hybrid QALY technique" first tried in the MORTISE trial to explore the transferability/generalisability of the hybrid approach in other disease areas in cost-utility analysis, using data from the cancer trial (COGNATE trial) and ophthalmology trial (CLARITY trial). This allows for the development of cost-utility analysis being explored in response to the growing concerns over the insensitivity and unresponsiveness of the generic health-related quality of life instruments, such as the EQ-5D (Pennington et al., 2020; Pettitt et al., 2016; Payakachat et al., 2015; Yang et al., 2015; Mulhern et al., 2013; Tosh et al., 2012; Whitehead and Ali, 2010), despite the EQ-5D-5L, a new version of the EQ-5D, was developed to address this concern to some extent (EuroQoL group, 2021; Herdman et al., 2011). The hybrid QALY approach explored in this thesis may serve as a complementary/supporting approach in resolving the uncertainty on the sensitivity and responsiveness of the EQ-5D in QALY estimations. In some disease areas, dementia and visual disorders for example, have adapted condition-specific tools to resolve the insensitivity and unresponsiveness of EQ-5D in QALY estimations. In their quest to resolve this issue in dementia, Mulhern and colleagues (2013) have developed condition-specific tool for dementia (i.e. DEMQOL-U and DEMQOL-PROXY-U) for QALY estimations in economic evaluation. Yang and colleagues (2015) meanwhile developed the bolt-on vision dimension to the standard EQ-5D, namely the ED-5D-Vis, to resolve the issue in the area of visual disorders.

6.7.2 Contribution to policy

The study reported in Chapter 2 is the first prospective economic evaluation study conducting alongside a randomised trial (the COGNATE trial, a UK-led randomised controlled trial) to evaluate whether EUS was cost-effective in the management of patients with GOC. The economic findings from the COGNATE trial contribute significantly to literature as there are no other prospective economic evaluation study conducted alongside randomised trial found in this area, apart from the COGNATE trial, based on the studies identified in the systematic review in Chapter 3. Thus, the prospective economic evaluation study reported in Chapter 2

offers a novel opportunity to inform evidence-based health care policy decision-making in the area of EUS in GOC staging within scarce resources.

The conduct of the economic-specific systematic literature review of EUS staging in patients with GOCs (Chapter 3) is novel in the areas of GOC and health economic research. This systematic review study fills gap in the literature, showing a significant contribution to the research in this area. The extensive search terms designed for the systematic review and together with the detailed and comprehensive search strategy developed for each search database in the systematic review aid the design of future systematic reviews of EUS staging within the context of health economics. The detailed, comprehensive, and transparent reporting of the systematic review in Chapter 3 not only aids future work on conducting a systematic review in similar if not the same field but also provides policy makers and commissioners with robust evidence-based economic information on the use of EUS in GOC staging in which to help inform health care policy decision making (Liberati et al., 2009; Aromataris and Pearson, 2014; Munn et al., 2018).

The study reported in Chapter 4 is the first survey study, up to January 2021, of MDT GOC-interested clinicians on the use of EUS in GOC staging in the UK. This survey offers an insight into the utilisation of EUS in GOC staging and the current clinical practice in the UK, including the potential factors associated with referral for EUS. The findings from this survey provide useful evidence-based information to policy makers and commissioners in which to inform health care policy decision-making on the use of EUS for GOC staging in clinical practices in the UK.

The exploratory methodological study of developing cost-utility analysis reported in Chapter 5 is novel. This is the first exploratory study of its kind in the area of cancer and ophthalmology. This exploratory study used data from a cancer trial (COGNATE trial) and an ophthalmology trial (CLARITY trial), leading to more robust and conclusive findings. This exploratory methodological study offers an opportunity to further explore the transferability/generalisability of the “hybrid QALY technique” first tried in the MORTISE trial (on which Dr Daphne Russell, the Trial Statistician and I, the Research Officer in Health Economics were on the trial) in other disease areas for cost-utility analysis in clinical trials. To date, there are no other studies exploring similar methodology, except in the MORTISE trial

(Edwards et al., 2015) where the same hybrid approach was undertaken to modify its conventional QALY and assessed the effect of the resulting hybrid QALY, that had taken account of the disease-specific measure of the MORTISE trial (i.e. the Foot Health Thermometer), on cost-utility analysis. Our research to date shows interesting findings that disease-specific measures are potentially useful in developing cost-utility analysis. Nevertheless, more studies covering a wider range of conditions/disease areas are required to allow for comprehensive assessment on the effect of this hybrid approach on cost-utility analysis in clinical trials. As a result, robust conclusion can be drawn on the usefulness of this approach in resolving the insensitivity and unresponsiveness issue of the EQ-5D to specific condition/disease areas.

6.8 Conclusion

This thesis evaluated and identified the economic value and evidence of EUS in GOC staging. Evidence from this thesis suggests that EUS saves costs, provides greater QALYs and is cost-effective. These encouraging findings provide strong evidence in favour of EUS scans for all GOC patients who have the potential to benefit. For a more complete picture of the use of EUS in GOC staging in the UK, a survey of MDT GOC members was undertaken. This ensured that GOC clinicians' views on the use of EUS in GOC staging in current clinical practice in the UK was explored, providing greater benefit to policy makers and health care service commissioners in their decision making. Attend MDT meeting appears to be the most important factor associated with referral for EUS. Clinicians who attend MDT meetings appear to be more likely to refer GOC patients for EUS. The exploratory study using a novel concept of the hybrid approach to modifying conventional QALY for cost-utility analysis of new intervention/technology suggests that condition-/disease-specific measures are found to be potentially useful in developing cost-utility analysis in clinical trials, allowing policy makers and commissioners to make decisions more efficiently and rationally on health care resource allocation, at least within specific conditions or disease areas, which use the same disease-specific guided QALY (i.e. hybrid QALY). In addition, the insights gained from the exploratory study serve to expand and guide future research in improving economic evaluation methodology.

On the whole, each element of this thesis together with the findings from the thesis make multiple novel contributions to both health economic methodology and policy. This thesis

provides evidence-based information to policy makers and health care service commissioners within the need for an effective and efficient management of gastro-oesophageal cancer, to improve efficiency of care in the delivery of diagnosis and treatment planning of gastro-oesophageal cancer within limited resources.

References

Adams, M. E., McCall, N. T., Gray, D. T., Orza, M. J. and Chalmers, T. C. (1992). Economic analysis in randomized control trials. *Medical Care*, 30(3), 231-243.

Agborsangaya, C. B., Lahtinen, M., Cooke, T., Johnson, J. A. (2014). Comparing the EQ-5D 3L and 5L: measurement properties and association with chronic conditions and multimorbidity in the general population. *Health Qual. Life Outcomes,* 12(1), 74.

Allum, W. H., Blazeby, J. M., Griffin, S. M., Cunningham, D., Jankowski, J. A., Wong, R. on behalf of the Association of Upper Gastrointestinal Surgeons of Great Britain and Ireland, the British Society of Gastroenterology and the British Association of Surgical Oncology (2011). Guidelines for the management of oesophageal and gastric cancer. *Gut,* 60(11), 1449-1472. doi:10.1136/gut.2010.228254

Allum, W. H., Griffin, S. M., Watson, A., Colin-Jones, D. on behalf of the Association of Upper Gastrointestinal Surgeons of Great Britain and Ireland, the British Society of Gastroenterology, and the British Association of Surgical Oncology. (2002). Guidelines for the management of oesophageal and gastric cancer. *Gut*, 50(Suppl V), v1–v23.

Althubaiti, A. (2016). Information bias in health research: definition, pitfalls, and adjustment methods. *Journal of Multidisciplinary Healthcare,* 9, 211–217. doi: 10.2147/JMDH.S104807.

Aromataris, E. and Pearson, A. (2014). The systematic review: An overview. *The American Journal of Nursing*, 114(3), 55-58.

Arroyos-Calvera, D., Covey, J., Loomes, G. and McDonald, R. (2019). The efficiency-equity trade-off, self-interest, and moral principles in health and safety valuation. *Social Science and Medicine*, 238, 112477. doi: 10.1016/j.socscimed.2019.112477.

Association of Upper Gastrointestinal Surgery of Great Britain and Ireland (AUGIS). (2020). Outcome data 2020. Retrieved from: https://www.augis.org/Guidelines/NOGCA-Audit/Outcomes-Data-2020

Birch, S. and Donaldson, C. (2003). Valuing the benefits and costs of health care programmes: where's the 'extra' in extra-welfarism? *Social Science and Medicine*, 56(5), 1121–1133.

Boadway, R. and Bruce, N. (1984). *Welfare economics*. Oxford: Basil Blackwell.

Bosetti, C., Levi, F., Ferlay, J., Garavello, W., Lucchini, F., Bertuccio, P., Negri, E. and La Vecchia, C. (2008). Trends in oesophageal cancer incidence and mortality in Europe. *International Journal of Cancer*, 122(5), 1118-1129. doi: 10.1002/ijc.23232.

Botet, J. F., Lightdale, C. J., Zauber, A. G., Gerdes, H., Urmacher, C. and Brennan, M. F. (1991a). Preoperative staging of esophageal cancer: Comparison of endoscopic US and dynamic CT. *Radiology*, 181(2), 419–425. doi: 10.1148/radiology.181.2.1924783.

Botet, J. F., Lightdale, C. J., Zauber, A. G., Gerdes, H., Winawer, S. J., Urmacher, C. and Brennan, M. F. (1991b). Preoperative staging of gastric cancer: comparison of endoscopic US and dynamic CT. *Radiology*, 181(2), 426-432. doi: 10.1148/radiology.181.2.1924784.

Bray, N., Noyes, J., Edwards, R.T. and Harris, N. (2014). Wheelchair interventions, services and provision for disabled children: a mixed-method systematic review and conceptual framework. *BMC Health Services Research*, 14, 309. DOI: https://doi.org/10.1186/1472-6963-14-309

Brazier, J., Roberts, J., Deverill, M. (2002). The estimation of a preference-based single index measure for health from the SF-36. *J Health Econ*, 21:271–92.

Brazier, J., Ratcliffe, J., Salomon, J. A. and Tsuchiya, A. (2007). *Measuring and valuing health benefits for economic evaluation*. Oxford: Oxford University Press.

Brazier, J. E., Rowen, D., Mavranezouli, I., Tsuchiya, A., Young, T., Yang, Y., Barkham, M. and Ibbotson, R. (2012). Developing and testing methods for deriving preference-based measures of health from condition-specific measures (and other patient-based measures of outcome). Health Technology Assessment, 16(32).

Brazier, J., Connell, J., O'Cathain, A. (2014). Do EQ-5D and SF-6D ask the right questions in mental health? A content validation using interviews with patients. *Value Health*, 17(3), A194.

Brazier, J., Ara, R., Rowen, D., Chevrou-Severac, H. (2017). A review of generic preference-based measures for use in cost-effectiveness models. *PharmacoEconomics*, 35 (Suppl 1): S21–S31. DOI 10.1007/s40273-017-0545-x

Brazier, J., Muston, D., Konwea, H., Sarah Power, G., Barzey, V., Lloyd, A., Sowade, O., Vitti, B., Gerlinger, C. and Roberts, J. (2017a). Evaluating the relationship between visual acuity and utilities in patients with diabetic macular edema enrolled in intravitreal aflibercept studies. *Investigative Ophthalmology & Visual Science*, 58(11), 4818-4825.

Brazier, J., Ratcliffe, J., Salomon, J. A. and Tsuchiya, A. (2017b). *Measuring and valuing health benefits for economic evaluation (*2nd ed.). Oxford: Oxford University Press.

Brazier et al. (2017d). Measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. *Health Technol Assess*. 18(9), 1–224.

Briggs, A. H. and Gray, A. (1999). Handling uncertainty when performing economic evaluation of healthcare interventions. *Health Technology Assessment*, 3(2), 1–134.

Briggs, A., Clark, T., Wolstenholme, J. and Clarke, P. (2003). Missing.... presumed at random: cost-analysis of incomplete data. *Health Economics*, 12, 377-92. doi: http://dx.doi.org/10.1002/hec.766

Brooks, R. (1996). EuroQol: the current state of play. *Health Policy*, 37, 53-72.

Brouwer, W. B. F., Culyer, A. J., van Exel, N. J. A. and Rutten, F. F. H. (2008). Welfarism vs. extra-welfarism. *Journal of Health Economics*, 27, 325-338.

Burke, L. G., Frakt, A. B., Khullar, D., Orav, E. J., Jha, A. K. (2017). Association Between Teaching Status and Mortality in US Hospitals. *JAMA*, 317, 20, 2105-2113. doi:10.1001/jama.2017.5702

Cancer Research UK (CRUK). (2019). *TNM staging for oesophageal cancer*. Retrieved from https://www.cancerresearchuk.org/about-cancer/oesophageal-cancer/stages-types-and-grades/tnm-staging

Cancer Research UK (CRUK). (2021a). *Oesophageal Cancer Statistics.* Retrieved from https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/oesophageal-cancer?

Cancer Research UK (CRUK). (2021b). *Stomach Cancer Statistics.* Retrieved from https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/stomach-cancer?

Carpenter, J. R, Kenward, M. G. (2007). *Missing data in randomised controlled trials: a practical guide.* Southampton: National Evaluation Trials and Studies Coordinating Centre.

Catalano, M. F., Sivak, M. V. Jr, Bedford, R. A., Falk, G. W., van Stolk, R., Presa, F. and Van Dam, J. (1995). Observer variation and reproducibility of endoscopic ultrasonography. *Gastrointestinal Endoscopy,* 41(2), 115-120. doi: 10.1016/s0016-5107(05)80592-9.

Carroll, N. and Penman, I. (2004). UK EUS Users Group: Recommendations for training in endosonography (EUS). Retrieved from https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.bsg.org.uk%2Fwp-content%2Fuploads%2F2019%2F12%2FRecommendations-for-Training-in-Endosonography-EUS-1.doc&wdOrigin=BROWSELINK

Chang, K. J., Soetikno, R. M., Bastas, D., Tu, C. and Nguyen, P. T. (2003). Impact of Endoscopic Ultrasound Combined with Fine-Needle Aspiration Biopsy in the Management of Esophageal Cancer. *Endoscopy*, 35(11), 962-965.

Chevreul, K., Brigham, K. B., Brunn, M., des Portes, V. (2015). Fragile X syndrome: economic burden and health-related quality of life of patients and caregivers in France. *Journal of Intellectual Disability Research*, 59, 1108–20.

Cleary, P. D. (1997). Editorial- Subjective and objective measures of health: which is better when? *Journal of Health Service Research Policy*, 2(1), 3-4.

Coast, J., Smith, R. D. and Lorgelly, P. (2008a). Welfarism, extra-welfarism and capability: The spread of ideas in health economics. *Social Science and Medicine*, 67(7), 1190-1198.

Conner-Spady, B., Suarez-Almazor, M. E. (2001). *A comparison of preference-based health status tools in patients with musculoskeletal disease*. In: 18th Plenary Meeting of the EuroQol Group. p. 235–245.

Cook, M. B., Chow, W. H. and Devesa, S. S. (2009). Oesophageal cancer incidence in the US by race, sex and histologic type 1977-2005. *British Journal of Cancer,* 101, 855-859.

Critical Appraisal Skills Programme (CASP). (2017). *CASP Economic Evaluation Checklist*. Retrieved from http://www.casp-uk.net/casp-tools-checklists. (2018, August 9).

Culyer, A. J. (1989). The normative economics of health care finance and provision. *Oxford Review of Economic Policy*, 5, 34-58.

Culyer, A. J. (1990). *Commodities, characteristics of commodities, characteristics of people, utilities and the quality of life.* In Baldwin, S., Godfrey, C. & Propper, C. (Eds.). The quality of life: perspectives and policies. London: Routledge.

Cunningham, S. J. (2001). An introduction to economic evaluation of health care. *Journal of Orthodontics*, 28(3), 246-250.

Curtis, L. and Burns, A. (2015). *Unit Costs of Health and Social Care 2015.* Kent: Personal Social Services Research Unit (PSSRU), University of Kent.

Curtis, L. and Burns, A. (2016). *Unit Costs of Health and Social Care 2016.* Kent: Personal Social Services Research Unit (PSSRU), University of Kent.

Curtis, L. and Burns, A. (2017). *Unit Costs of Health and Social Care 2017*. Kent: Personal Social Services Research Unit (PSSRU), University of Kent.

Curtis, L. and Netten, A. (2005). *Unit Costs of Health and Social Care 2005.* Kent: Personal Social Services Research Unit (PSSRU), University of Kent.

Curtis, L. (2008). *Unit costs of health and social care 2007-2008*. Kent: Personal Social Services Research Unit (PSSRU), University of Kent.

Curtis, L. and Burns, A. (2019). *Unit Costs of Health and Social Care 2019*. Kent: Personal Social Services Research Unit (PSSRU), University of Kent.

DaVee, T., Ajani, J. A. and Lee, J. H. (2017). Is endoscopic ultrasound examination necessary in the management of esophageal cancer? *World Journal of Gastroenterology*, 23(5), 751–762. doi:10.3748/wjg.v23.i5.751.

De Ieso, P. B., Coward, J. I., Letsa, I., Schick, U., Nandhabalan, M., Frentzas, S. and Gore, M. E. (2013). A study of the decision outcomes and financial costs of multidisciplinary team meetings (MDMs) in oncology. *British Journal of Cancer*, 109, 2295–2300, DOI: 10.1038/bjc.2013.586

Department of Health (DoH). (2009). *NHS Reference Costs 2007-08*. London: DoH.

Devlin, N. J., Brooks, R. (2017). EQ-5D and the EuroQol Group: Past, present and future. *Appl Health Econ Health Policy*, 15:127–137.

Devlin, N., Shah, K. K., Feng, Y., Mulhern, B., van Hout, B. (2016). *Valuing Health-Related Quality of Life: An EQ-5D-5L Value Set for England.* Discussion Paper. HEDS Discussion Paper Series (16.02). Health Economics and Decision Science, School of Health and Related Research (ScHARR), University of Sheffield, Sheffield. http://eprints.whiterose.ac.uk/97964/

Dietrich, C. F., Arcidiacono, P. G., Braden, B., Burmeister, S., Carrara, S., Cui, X., Di Leo, M., Dong, Y., Fusaroli, P., Gilja, O. H., Healey, A. J., Hocke, M., Hollerbach, S., Garcia, J. I., Ignee,

A., Jürgensen, C., Kahaleh, M., Kitano, M., Kunda, R., Larghi, A., Möller, K., Napoleon, B., Oppong, K. W., Petrone, M. C., Saftoiu, A., Puri, R., Sahai, A. V., Santo, E., Sharma, M., Soweid, A., Sun, S., Teoh, A. Y. B., Vilmann, P., Jenssen, C. (2019). What should be known prior to performing EUS? *Endoscopic Ultrasound*, 1, 3-16. doi: 10.4103/eus.eus_54_18.

Dittler, H. J. (1995). Assessment of resectability of gastrointestinal cancers by endoscopic ultrasonography. *Gastrointestinal Endoscopy Clinics of North America*, 5, 569.

Dolan P. (1997). Modeling valuations for EuroQol health states. *Med Care*, 35(11), 1095–108.

Drummond, M. F. and Davies, L. M. (1991). Economic analysis alongside clinical trials: revisiting the methodological issues. *International Journal of Technology Assessment in Health Care*, 7(4), 561-573.

Drummond, M. F., Cooke, J. and Walley, T. (1996). *Economic evaluation in healthcare decision-making: evidence from the UK*. York: University of York Centre for Health Economics.

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L. and Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes.* (4th ed.) Oxford: Oxford University Press.

Dubecz, A., Solymosi, N., Stadlhuber, R. J., Schweigert, M., Stein, H. J. and Peters, J. H. (2013). Does the incidence of adenocarcinoma of the esophagus and gastric cardia continue to rise in the twenty-first century? ASEER database analysis. *Journal of Gastrointestinal Surgery,* 18, 2345-2348.

Dudley-Brown, S., Nag, A., Cullinan, C., Ayers, M., Hass, S. and Panjabi, S. (2009). Health-related quality-of-life evaluation of Crohn disease patients after receiving natalizumab therapy. *Gastroenterology Nursing*, 32, 327–339.

Eckardta, A. J. and Jenssen, C. (2015). Current endoscopic ultrasound-guided approach to incidental subepithelial lesions: optimal or optional? *Annals of Gastroenterology,* 28, 160-172.

Edwards, R. T., Hounsome, B., Linck, P. and Russell, I. T. (2008). Economic evaluation alongside pragmatic randomised trials: developing standard operating procedure for clinical trials unit. *Trials*, 9, 64. doi: http://dx.doi.org/10.1186/1745-6215-9-64

Edwards, R. T., Yeo, S. T., Russell, D., Thomson, C. E., Beggs, I., Alastair Gibson, J. N., McMillan, D., Martin, D. J. and Russell, I. T. (2015). Cost-effectiveness of steroid (methylprednisolone) injections versus anaesthetic alone for the treatment of Morton's neuroma: economic evaluation alongside a randomised controlled trial (MortISE trial). *Journal of Foot and Ankle Research,* 8, 6-16.

Emrani, Z., Akbari Sari, A., Zeraati, H., Olyaeemanesh, A. and Daroudi, R. (2020). Health-related quality of life measured using the EQ-5D–5L: population norms for the capital of Iran. Health Qual Life Outcomes, 18, 108. doi: https://doi.org/10.1186/s12955-020-01365-5

EuroQoL Group (2021). *EQ-5D*. Retrieved from https://euroqol.org/publications/user-guides/

EuroQol Group (2021). EQ-5D-5L User Guide: Basic information on how to use the EQ-5D-5L instrument. EuroQoL Research Foundation: The Netherlands. Retrieved from https://euroqol.org/publications/user-guides/

FACIT.org. (2021a). *FACT-G: Functional Assessment of Cancer Therapy – General. Constitutes the core of all subscales; can be used with patients of any tumor type.* Retrieved from https://www.facit.org/measures/FACT-G

FACIT.org. (2021b). *FACT-E: Functional Assessment of Cancer Therapy – Esophageal. For patients with Esophageal cancer*. Retrieved from https://www.facit.org/measures/FACT-E

FACIT.org. (2021c). *FACT-Ga: Functional Assessment of Cancer Therapy – Gastric. For patients with Gastric cancer*. Retrieved from https://www.facit.org/measures/FACT-Ga

Feeny, D. H., Furlong, W. J., Torrance, G. W., Goldsmith, C. H., Zenglong, Z., Depauw, S., Denton, M., Boyle, M. (2002). Multiattribute and single-attribute utility function: the Health Utility Index Mark 3 system. *Med Care*, 40:113–28.

Fenwick, E., Marshall, D. A., Levy, A. R. and Nichol, G. (2006). Using and interpreting cost-effectiveness acceptability curves: an example using data from a trial of management strategies for atrial fibrillation. *BMC Health Service Research*, 6, 52.

Fenwick, E., O'Brien, B. J. and Briggs, A. (2004). Cost-effectiveness acceptability curves – facts, fallacies and frequently asked questions. *Health Economics*, 13, 405–415.

Finch, A.P., Brazier, J., Mukuria, C. (2019). Selecting Bolt-On Dimensions for the EQ-5D: Examining Their Contribution to Health-Related Quality of Life. *Value In Health*, 22, 50-61.

Finch, A.P., Brazier, J., Mukuria, C. (2021). Selecting bolt-on dimensions for the EQ-5D: testing the impact of hearing, sleep, cognition, energy, and relationships on preferences using pairwise choices. *Medical Decision Making*, 41 (1). pp. 89-99. DOI: https://doi.org/10.1177/0272989X20969686

Findlay, J. M., Bradley, K. M., Maile, E. J., Braden, B., Maw, J., Phillips-Hughes, J., Gillies, R. S., Maynard, N. D. and Middleton, M. R. (2015). Pragmatic staging of oesophageal cancer using decision theory involving selective endoscopic ultrasonography, PET and laparoscopy. *British Journal of Surgery,* 102(12), 1488-1499. doi: 10.1002/bjs.9905. PMID: 26458070.

Forrest, L. M, McMillan, D. C, McArdle, C. S., Dunlop, D. J. (2005). An evaluation of the impact of a multidisciplinary team, in a single centre, on treatment and survival in patients with inoperable non-small-cell lung cancer. *Br J Cancer*, 93, 9, 977–978.

Franssen, T., Stijnen, M., Hamers, F., and Schneider, F. (2020). Age differences in demographic, social and health-related factors associated with loneliness across the adult life

span (19–65 years): a cross-sectional study in the Netherlands. *BMC Public Health,* 20, 1118. DOI: https://doi.org/10.1186/s12889-020-09208-0

Friedman, M. (1953). *Essays in positive economics.* Chicago: The University of Chicago Press.

Gandhi, M., Ang, M., Teo, K., Wong, C. W., Wei, Y. C., Tan, R. L., Janssen, M. F., Luo, N. (2020). A vision 'bolt-on' increases the responsiveness of EQ-5D: preliminary evidence from a study of cataract surgery. The *European Journal of Health Economics,* 21(4), 501-511. doi:10.1007/s10198-019-01156-w.

Gao, Z., Wang, C., Xue, Q., Wang, J., Shen, Z., Jiang, K., Shen, K., Liang, B., Yang, X., Xie, Q., Wang, S., and Ye, Y. (2017). The cut-off value of tumor size and appropriate timing of follow-up for management of minimal EUS-suspected gastric gastrointestinal stromal tumors. *BMC Gastroenterology,* 17, 8. DOI 10.1186/s12876-016-0567-4.

Glick, H. A., Doshi, J. A., Sonnad, S. S. and Polsky, D. (2007). *Economic evaluation in clinical trials*. Oxford: Oxford University Press.

Glick, H. A., Doshi, J. A., Sonnad, S. S. and Polsky, D. (2014). *Economic evaluation in clinical trials*. Oxford: Oxford University Press.

Gold, M. R., Siegel, J. E., Russell, L. B. and Weinstein, M. C. (1996). *Cost-effectiveness in health and medicine*. New York: Oxford University Press.

Goodacre, S. and McCabe, C. (2002). Review: An introduction to economic evaluation. *Emergency Medicines Journal*, 19, 198–201.

Gravelle, H., Rees, R. (1992). *Microeconomics.* London: Longman.

Gray, A. M., Clarke, P. M., Wolstenholme, J. L. and Wordsworth, S. (2011). *Applied methods of cost-effectiveness analysis in health care.* Oxford: Oxford University Press.

Gress, F., Ikenberry, S., Conces, D., Wonn, J., Khruso, Q. and Hawes, R. (1995). Endoscopic ultrasound (EUS) staging in patients with esophageal cancer is more accurate than CT and correlates with survival. *Gastrointestinal Endoscopy*, 41, 349.

Grimm, H., Binmoeller, K. F., Hamper, K., Koch, J., Henne-Bruns, D. and Soehendra, N. (1993). Endosonography for preoperative locoregional staging of esophageal and gastric cancer. *Endoscopy*, 25(3), 224-230. doi: 10.1055/s-2007-1010297.

Hadzijahic, N., Wallace, M. B., Hawes, R. H., VanVelse, A., LeVeen, M., Marsi, V., Hoffman, B. J. and Sahai, A. V. (2000). CT or EUS for the initial staging of esophageal cancer? A cost minimization analysis. *Gastrointestinal Endoscopy*, 52(6), 715-720. doi:10.1067/mge.2000.108481.

Haldar, P., Brightling, C. E., Hargadon, B., Gupta, S., Monteiro, W., Sousa, A., Marshall, R. P., Bradding, P., Green, R. H., Wardlaw, A. J. and Pavord, I. D. (2009). Mepolizumab and exacerbations of refractory eosinophilic asthma. *New England Journal of Medicine*, 360, 973–984.

Harewood, G. C., and Wiersema, M. J. (2002). A cost analysis of endoscopic ultrasound in the evaluation of esophageal cancer. *American Journal of Gastroenterology*, 97, 452-458.

Harper, R., Brazier, J. E., Waterhouse, J. C., Walters, S. J., Jones, N. M., Howard, P. (1997). Comparison of outcome measures for patients with chronic obstructive pulmonary disease (COPD) in an outpatient setting. *Thorax*, 52(10), 879–87.

Harris, K. M., Kelly, S., Berry, E., Hutton, J., Roderick, P., Cullingworth, J., Gathercole, L., O'Connor, P. J., Boyce, J. C. and Smith, M.A. (1998). Systematic review of endoscopic ultrasound in gastro-oesophageal cancer. *Health Technology Assessment*, 2(18).

Hausman, D., and McPherson, M. (2006). *Rationality in Positive and Normative Economics*. In Economic Analysis, Moral Philosophy and Public Policy (pp. 60-77). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511754289.007

Healthcare Quality Improvement Partnership (HQIP) (2016). *National Oesophago-Gastric Cancer Audit 2016: An audit of the care received by people with Oesophago-Gastric Cancer in England and Wales 2016 Annual Report.* NHS Digital, UK. Retrieved from: https://digital.nhs.uk/data-and-information/publications/statistical/national-oesophago-gastric-cancer-audit/national-oesophago-gastric-cancer-audit-2016-annual-report.

Heijink, R., Xe, K., Saksena, P. and Evans, D. (2011). Validity and Comparability of Out-of pocket Health Expenditure from Household Surveys: A Review of the Literature and Current Survey Instruments**.** Geneva: *World Health Organization.*

Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., Bonsel, G. and Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research*, 20(10), 1727–1736. doi:10.1007/s11136-011-9903-x

Herdman, M., Kerr, C., Pavesi, M., Garside, J., Lloyd, A., Cubi-Molla, P., Devlin, N. (2020). Testing the validity and responsiveness of a new cancer-specific health utility measure (FACT-8D) in relapsed/refractory mantle cell lymphoma, and comparison to EQ-5D-5L. *J Patient Rep Outcomes,* 27, 4(1), 22. doi:10.1186/s41687-020-0185-3.

Hicks, J. R. (1939). The foundations of welfare economics. *The Economic Journal,* 49(196), 696-712.

Hounsome, N., Orrell, M., Edwards, R.T. (2011). EQ-5D as a quality of life measure in people with dementia and their carers: evidence and key issues. *Value Health*, 14(2), 390–9.

Hubbard, G., and O'Brien A. P. (2006). *Microeconomics*. Upper Saddle River, New Jersey: Pearson Prentice Hall.

Hurley, J. (1998). *Welfarism, extra-welfarism and evaluative economic analysis in the health care sector*. In: Barer, M. L., Getzen, T. E., Stoddard, G. L. (eds.) Health, Health Care and Health Economics: Perspectives on Distribution. Chichester: John Wiley & Sons Ltd.

Hurley, J. (2000). *An overview of the normative economics of the health care sector*. In: Culyer, A. J., Newhouse, J. (eds.) Handbook of Health Economics. Amsterdam: Elsevier Science.

Hurley, J. and Culyer, A. J. (2014). *Welfarism and Extra-Welfarism*. Encyclopedia of Health Economics (eds). Amsterdam: Elsevier. https://doi.org/10.1016/B978-0-12-375678-7.00212-1.

Husereau, D., Drummond, M., Petrou, S., Carswell, C., Moher, D., Greenberg, D., Augustovski, F., Briggs, A. H., Lindsay, W. R., Mauskopf, J., Loder, E., on behalf of the CHEERS Task Force (2013). Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *BMJ,* 346, f1049. doi:10.1136/bmj.f1049

International Monetary Fund (IMF). (2017). *Exchange Rate Archives by Month*. Retrieved from: http://www.imf.org/external/np/fin/data/param_rms_mth.aspx.

Janssen, M., Pickard, A. S., Golicki, D., Gudex, C., Niewada, M., Scalone, L., Swinburn, P., Busschbach, J. (2013). Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Qual. Life Res*. 22(7), 1717–1727.

Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E. and Forman, D. (2011). Global cancer statistics. *CA: A Cancer Journal for Clinicians*, 61, 69-90.

Jones, C. M., Lyles, A. and Foley, K. G. (2021). A national cross-sectional survey investigating the use of endoscopic ultrasound in the diagnosis and treatment of oesophageal cancer in the UK. *Clinical Radiology*, 76(6), 458-464. https://doi: 10.1016/j.crad.2021.02.021.

Kang, H. (2013). The prevention and handling of the missing data. *Korean J Anesthesiol,* 64(5), 402–406.

Kaldor, N. (1939). Welfare propositions and interpersonal comparisons of utility. *The Economic Journal*, 49(195), 549-552.

Kim, J., Kim, S. G., Chung, H., Lim, J. H, Choi, J. M., Park, J. Y., Yang, H-J, Han, S. J., Oh, H., Kim, M. S., Kim, H. J., Hong, H., Lee, H. J., Kim, J. L., Lee, E., and Jung, H. C. (2018). Clinical efficacy of endoscopic ultrasonography for decision of treatment strategy of gastric cancer. *Surgical Endoscopy,* 32, 3789–3797. https://doi.org/10.1007/s00464-018-6104-5

Kim, L. S. and Koch, J. (1999). Do we practice what we preach? Clinical decision making and utilisation of endoscopic ultrasound for staging esophageal cancer. *American Journal of Gastroenterology*, 94(7), 1847-1852.

Kind, P., Hardman, G., Macran, S. (1999). *UK Population Norms for EQ-5D.* The University of York: Centre for Health Economics.

Kind, P., Elston Lafata, J., Matuszewski, K. and Raisch, D. (2009). The use of QALYs in clinical and patient decision-making: issues and prospects. *Value Health*, 12 (Suppl 1), S27–30. doi: 10.1111/j.1524-4733.2009.00519.x.

King, M. T., Norman, R., Mercieca-Bebber, R., Costa, D. S. J., McTaggart-Cowan, H., Peacock, S., Janda, M., Müller, F., Viney, R., Pickard, A. S., Cella, D. (2021). Multi-Attribute Utility in Cancer Consortium. The Functional Assessment of Cancer Therapy Eight Dimension (FACT-8D), a Multi-Attribute Utility Instrument Derived from the Cancer-Specific FACT-General (FACT-G) Quality of Life Questionnaire: Development and Australian Value Set. *Value Health*, 24(6), 862-873. doi: 10.1016/j.jval.2021.01.007.

Knapp, M. and Beecham, J. (1990). Costing mental health services. *Psychological Medicine*, 20, 893-908.

Lang, E. (2004). The well-built clinical question: a key to evidence-based decisions. *McGill Journal of Medicine*, 8(1), 90-93.

Laramée, P., Wonderling, D., Cahen, D. L., Dijkgraaf, M. G., Gouma, D. J., Bruno, M. J. and Pereira S. P. (2013). Trial-based cost-effectiveness analysis comparing surgical and endoscopic drainage in patients with obstructive chronic pancreatitis. *BMJ Open*, 3, e003676.

Lennon, A. M. and Penman, I. D. (2007). Endoscopic ultrasound in cancer staging. *British Medical Bulletin,* 84, 81–98. doi:10.1093/bmb/ldm033

Leurent, B., Gomes, M. and Carpenter, J. R. (2018). Missing data in trial-based cost-effectiveness analysis: An incomplete journey. *Health Economics*, 27, 1024–1040.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gotzche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Klwijnen, J. and Moher, D. (2009). The PRISMA Statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *Annals of Internal Medicine*, 6(7), 1-28.

Lin, D. Y., Feuer, E. J., Etzioni, R. and Wax, Y. (1997). Estimating medical costs from incomplete follow-up data. *Biometrics*, 53, 419-34. doi: http://dx.doi.org/10.2307/2533947

Longworth L, Bryan S. (2003). An empirical comparison of EQ-5D and SF-6D in liver transplant patients. *Health Econ.*, 12(12), 1061–7.

Longworth, L., Yang, Y., Young, T., Mulhern, B., Hernández Alava, M., Mukuria, C., Rowen, D., Tosh, J., Tsuchiya, A., Evans, P., Devianee Keetharuth, A., Brazier, J. (2014). Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. *Health Technol Assess*., 18(9), 1–224. doi: https://doi. org/10.3310/hta18090

Lowe, V. J., Booya, F., Fletcher, J. G., Nathan, M., Jensen, E., Mullan, B., Rohren, E., Wiersema, M. J., Vazquez-Sequeiros, E., Murray, J. A., Allen, M. S., Levy, M. J. and Clain, J. E. (2005). Comparison of positron emission tomography, computed tomography, and endoscopic ultrasound in the initial staging of patients with esophageal cancer. *Molecular Imaging and Biology*, 7(6), 422-430. doi: 10.1007/s11307-005-0017-0.

Luo, N., Wang, X., Ang, M., Finkelstein, E.A., Aung, T., Wong, T.Y., Lamoureux, E. (2015). A Vision "Bolt-On" Item Could Increase the Discriminatory Power of the EQ-5D Index Score. *Value Health,* 18(8),1037-1042. doi:10.1016/j.jval.2015.08.002.

Manca, A., Hawkins, N. and Sculpher, M. J. (2005). Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Economics*, 14, 487–496.

Maynard, A. and Kanavos, P. (2000). *Health economics: an evolving paradigm. Health Economics*, 9(3), 183-190.

Mazzaferro V, Majno P (2011). Principles for the best multidisciplinary meetings. *Lancet Oncol*, 12, 4, 323–325.

McIntosh, E., Clarke, P. M., Frew, E. J. and Louviere, J. J. (2011). *Applied methods of cost-benefit analysis in health care*. Oxford: Oxford University Press.

Mishan, E. J. (1981). *Do economic evaluations of allocative changes have any validity in the West today?* In: Economic efficiency and social welfare: selected essays on fundamental aspects of the economic theory of social welfare. London: Allen & Unwin.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine*, 6(7), e1000097. doi:10.1371/journal.pmed1000097

Morris, S., Devlin, N., Parkin, D. and Spencer, A. (2012). *Economic analysis in health care*. Chichester: John Wiley & Sons Ltd.

Mulhern, B., Rowen, D., Brazier, J., Smith, S., Romeo, R., Tait, R., Watchurst, C., Chua, K. C., Loftus, V., Young, T., Lamping, D., Knapp, M., Howard, R. and Banerjee, S. (2013). Development of DEMQOL-U and DEMQOL-PROXY-U: generation of preference-based indices from DEMQOL and DEMQOL-PROXY for use in economic evaluation. *Health Technology Assessment,* 17(5), v-xv, 1-140. doi: 10.3310/hta17050.

Mulhern, B.J., Sampson, C., Haywood, P., Addo, R., Page, K., Mott, D., Shah, K., Janssen, M. F. and Herdman, M. (2022). Criteria for developing, assessing and selecting candidate EQ-5D bolt-ons. *Quality of Life Research*. doi:10.1007/s11136-022-03138-7.

Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C, McArthur, A., and Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, 18(1), 143.

National Eye Institute (NEI). (2021). *Visual Function Questionnaire 25 (2000 version)*. Retrieved from https://www.nei.nih.gov/learn-about-eye-health/resources-for-health-educators/outreach-materials/visual-function-questionnaire-25

NICE (National Institute for Health and Care Excellence) (2008). *Guide to the methods of technology appraisal.* London: National Health Service. (Replaced by the 2013 version). Available from: www.nice.org.uk

National Institute for Health and Care Excellence (NICE). (2013). *Guide to the methods of technology appraisal 2013: Process and method.* London: National Health Service. Retrieved from https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781

NHS Information Centre. (2009). *Prescription cost analysis 2008 England*. Retrieved from https://digital.nhs.uk/data-and-information/publications/statistical/prescription-cost-analysis/prescription-cost-analysis-2008

NICE (2018a). *NICE Guideline NG83: Oesophago-gastric cancer: assessment and management in adults.* Retrieved from https://www.nice.org.uk/guidance/ng83/resources/oesophagogastric-cancer-assessment-and-management-in-adults-pdf-1837693014469

NICE (2018b). *NICE Guideline NG83: Oesophago-gastric cancer: assessment and management in adults: Appendix L: Cost-effectiveness analyses.* Retrieved from https://www.nice.org.uk/guidance/ng83/evidence/appendix-l-pdf-170036297751

Nickl, N., Bhutani, M., Catalano, M., Hoffman, B., Hawes, R., Chak, A., Roubein, L. D., Kimmey, M., Johnson, M., Affronti, J., Canto, M., Sivak, M., Boyce, H. W., Lightdale, C. J., Stevens, P. and Schmitt, C. (1996). Clinical implications of endoscopic ultrasound: The American Endosonography Club Study. *Gastrointestinal Endoscopy*, 44, 371. doi: 10.1016/s0016-5107(96)70083-4.

O'Sullivan, A. K., Thompson, D. and Drummond, M. F. (2005). Collection of health-economic data alongside clinical trials: Is there a future for piggyback evaluations? *Value in Health,* 8(1), 67-79.

Office for National Statistics (ONS). (2018). *Living longer: how our population is changing and why it matters.* Retrieved from https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing/articles/livinglongerhowourpopulationischangingandwhyitmatters/2018-08-13

Palmer, S. and Raftery, J. (1999). Opportunity cost. *BMJ,* 318, 1551–1552.

Palmer, S. and Torgerson, D. J. (1999). Definitions of efficiency. *BMJ*, 318, 1136.

Parkin, M. (2000). *Economics*. Harlow: Addison Wesley.

Payakachat, N., Ali, M. M. and Tilford, J. M. (2015). Can the EQ-5D detect meaningful change? A systematic review. *Pharmacoeconomics,* 33(11), 1137-1154. doi:10.1007/s40273-015-0295-6

Peixoto, R. D. S., Krstic, L., Hill, S. C. L. and Foss, A. J. E. (2021). Predicting quality of life in AMD patients—insights on the new NICE classification and on a bolt-on vision dimension for the EQ-5D. *Eye*, 35, 3333–3341. https://doi.org/10.1038/s41433-021-01414-3

Penman, I., Norton, S. and Harris, K. (2011). *Staging of Oesophago-Gastric Carcinoma by Endoscopic Ultrasonography: Guidance and Minimum Standards.* 1e19. EUS Users Group and the British Society of Gastroenterology (BSG).

Pennington, B., Hernández-Alava, M., Hykin, P., Sivaprasad, S., Flight, L., Alshreef, A. and Brazier, J. (2020). Mapping From Visual Acuity to EQ-5D, EQ-5D With Vision Bolt-On, and VFQ-UI in Patients With Macular Edema in the LEAVO Trial. *Value in Health,* 23(7), 928–935. doi: 10.1016/j.jval.2020.03.008.

Pettitt, D. A., Raza, S., Naughton, B., Roscoe, A., Ramakrishnan, A., Ali, A., Davies, B., Dopson, S., Hollander, G., Smith, J. A. and Brindley, D. A. (2016). The limitations of QALY: A Literature Review. *Journal of Stem Cell Research & Therapy,* 6, 4. doi: 0.4172/2157-7633.1000334.

Pfau, P. R., Perlman, S. B., Stanko, P., Frick, T. J., Gopal, D. V., Said, A., Zhang, Z. and Weigel, T. (2007). The role and clinical value of EUS in a multimodality esophageal carcinoma staging program with CT and positron emission tomography. *Gastrointestinal Endoscopy*, 65, 377-84. doi: 10.1016/j.gie.2006.12.015.

Philips, Z., Ginnelly, L., Sculpher, M., Claxton, K., Golder, S., Riemsma, R., Woolacott, N. and Glanville, J. (2004). Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technology Assessment*, 8(36).

Phillips, C. and Thompson, G. (2001). *What is cost-effectiveness?* Kent: Hayward Medical Communication.

Polkowski, M., Palucki, J., Wronska, E., Szawlowski, A., Nasierowska-Guttmejer, A. and Butruk, E. (2004). Endosonography versus helical computed tomography for staging of gastric cancer. *Endoscopy*, 36, 617–623.

Puli, S. R., Reddy, J. B., Bechtold, M. L., Antillon, D., Ibdah, J. A. and Antillon, M. R. (2008). Staging accuracy of esophageal cancer by endoscopic ultrasound: a meta-analysis and systematic review. *World Journal of Gastroenterology*, 14, 1479-1490. doi: 10.3748/wjg.14.1479.

Pyne, J. M., Sieber, W. J., David, K., Kaplan, R. M., Rapaport, M. H., Williams, D. K. (2003). Use of the quality of well-being self-administered version (QWB-SA) in assessing health-related quality of life in depressed patients. *Journal of Affective Disorder,* 76(1-3):237-47. doi: 10.1016/s0165-0327(03)00106-x.

Ramsey, S., Willke, R., Briggs, A., Brown, R., Buxton, M., Chawla, A., Cook, J., Glick, H., Liljas, B., Petitti, D. and Reed, S. (2005). Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force Report, *Value in Health,* 8(5), 521-533.

Regmi, P. R., Waithaka, E., Paudyal, A., Simkhada, P. and van Teijlingen, E. (2016). Guide to the design and application of online questionnaire surveys. *Nepal Journal of Epidemiology*, 6(4), 640-644.

Rentz, A. M., Kowalski, J. W., Walt, J. G., Hays, R. D., Brazier, J. E., Yu, R., Lee, P., Bressler, N., and Revicki, D. A. (2014). Development of a Preference-Based Index from the National Eye Institute Visual Function Questionnaire-25. *JAMA Ophthalmol.*, 32(3), 310–318. doi: 10.1001/jamaophthalmol.2013.7639

Richardson J, McKie J, Bariola E. (2014). *Multiattribute utility instruments and their use*. In: Culyer AJ, editor. Encyclopaedia of health economics, vol. 2. San Diego: Elsevier.

Richardson, J., Sinha, K., Iezzi, A., Khan, M. A. (2014). Modelling utility weights for the Assessment of Quality of Life (AQoL) 8D. *Qual Life Res*, 23, 2395–2404. doi: 10.1007/s11136-014-0686-8.

Ridyard, C. H. and Hughes, D. A. (2010). Methods for the collection of resource use data within clinical trials: a systematic review of studies funded by the UK Health Technology Assessment Programme. *Value in Health,* 13, 867-872.

Riphaus, A., Steckstor, M. and Adam, B. (2013a). Endoscopic image and endoscopic ultrasound-guided staging of adenocarcinoma of the gastroesophageal junction. *Video*

Journal and Encyclopedia of GI Endoscopy, 1(1), 25-26. https://doi.org/10.1016/S2212-0971(13)70013-X.

Riphaus, A., Pech, O., Steckstor, M., and Adam, B. (2013b). Gastrointestinal Stromal Tumors (GIST) – Endoscopic Image, Endoscopic Ultrasound, and Value of Endoscopic-Guided Fine-Needle Aspiration. *Video Journal and Encyclopedia of GI Endoscopy,* 1(1), 183-184. https://doi.org/10.1016/S2212-0971(13)70076-1

Robinson, R. (1993a). Economic analysis and health care. What does it mean? *BMJ*, 307, 670-673.

Robinson, R. (1993b). Costs and cost-minimisation analysis. *BMJ*, 307, 726-728.

Robinson, R. (1993c). Cost-effectiveness analysis. *BMJ*, 307, 793-795.

Robinson, R. (1993d). Cost-utility analysis. *BMJ*, 307, 859-862.

Robinson, R. (1993e). Cost-benefit analysis. *BMJ*, 307, 924-926.

Ronellenfitsch, U., Schwarzbach, M., Hofheinz, R., Kienle, P., Kieser, M., Slanger, T. E., Jensen, K., GE Adenocarcinoma Meta-analysis Group. (2013). Perioperative chemo(radio)therapy versus primary surgery for resectable adenocarcinoma of the stomach, gastroesophageal junction, and lower esophagus. *Cochrane Database of Systematic Reviews*, (5):CD008107. https://doi.org/10.1002/14651858.CD008107.pub2

Rösch, T., Lorenz, R., Zenker, K., von Wichert, A., Dancygier, H., Höfler, H., Siewert, J. R. and Classen, M. (1992). Local staging and assessment of resectability in carcinoma of the esophagus, stomach, and duodenum by endoscopic ultrasonography. *Gastrointestinal Endoscopy*, 38, 460.

Rudmik, L. and Drummond, M. (2013). Health economic evaluation: important principles and methodology. *Laryngoscope*, 123(6), 1341-1347. doi: 10.1002/lary.23943.

Russell, I. T., Edwards, R. T., Gliddon, A. E., Ingledew, D. K., Russell, D., Whitaker, R., Yeo, S. T., Attwood, S. E., Barr, H., Nanthakumaran, S. and Park, K. G. M. (2013). Cancer of Oesophagus or Gastricus – New Assessment of Technology of Endosonography (COGNATE): report of pragmatic randomised trial. *Health Technology Assessment,* 17(39). https://doi.org/10.3310/hta17390

Russell, L. B. (1992). Opportunity costs in modern medicine. *Health Affair,* 11, 162-169.

Ryan, M., Gerard, K. and Amaya-Amaya, M. (eds.) (2008). *Using discrete choice experiments to value health and health care* (The economics of non-market goods and resources, 11), vol. 11, Netherlands Antilles, AN: Springer Academic Publishers, 256pp.

Sahai, A. V. (2012). Is EUS here to stay? Accuracy is not an indication. *Endoscopic ultrasound*, 1(3), 117–118.

Sampson, C., Addo, R., Haywood, P., Herdman, M., Janssen, B., Mulhern, B., Page, K., Reardon, O., Sanchez, M.R., Schneider, J., and others (2019). PMU141 Development of EQ-5D-5L Bolt-ons for Cognition and Vision. *Value in Health*, 22, p.S733. doi: 10.1016/j.jval.2019.09.1759.

Schardt, C., Adams, M. B., Owens, T., Keitz, S. and Fontelo, P. (2007). Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Medical Informatics and Decision Making*, 7, 16. Retrieve from https://doi.org/10.1186/1472-6947-7-16

Scottish Intercollegiate Guidelines Network (SIGN). (2006). *SIGN 87: Management of oesophageal and gastric cancer.* Retrieved from https://www.sign.ac.uk/our-guidelines/management-of-oesophageal-and-gastric-cancer/

Seixas, B. V. (2017). Welfarism and extra-welfarism: a critical overview. *Cad Saude Publica*, 33(8), e00014317. doi: 10.1590/0102-311X00014317.

Shenfine, J., McNamee, P., Steen, N., Bond, J. and Griffin, S. M. (2005). A pragmatic randomised controlled trial of the cost-effectiveness of palliative therapies for patients with inoperable oesophageal cancer. *Health Technology Assessment*, 9(5).

Shiell, A., Donaldson, C., Mitton, C. and Currie, G. (2002). Health economic evaluation. *Journal of Epidemiology Community Health,* 56, 85–88.

Shimpi, R. A., George, J., Jowell, P. and Gress, F. G. (2007). Staging of Esophageal cancer by EUS: staging accuracy revisited. *Gastrointestinal Endoscopy*, 66 (3), 475-483.

Shumaker, D. A., de Garmo, P. and Faigel, D. O. (2002). Potential impact of preoperative EUS on esophageal cancer management and cost. *Gastrointestinal Endoscopy*, 56(3), 391-396. doi: 10.1067/mge.2002.126831

Sintonen H. (2001). The 15D instrument of health-related quality of life: properties and applications. *Ann Med.*, 33(5):328–36

Sivaprasad, S., Hykin, P., Prevost, A. T., Vasconcelos, J., Riddell, A., Ramu, J., Murphy, C., Kelly, J., Edwards, R. T., Yeo, S. T., Bainbridge, J., Hopkins, D. and White-Alao, B. (2018). Intravitreal aflibercept compared with panretinal photocoagulation for proliferative diabetic retinopathy: the CLARITY non-inferiority RCT. *Efficacy Mechanism Evaluation*, 5(5).

Smee, C. H. (1997). Bridging the gap between public expectations and public willingness to pay. *Health Economics*, 6, 1–10.

Smith, S. C., Lamping, D. L., Banerjee, S., Harwood, R., Foley, B., Smith, P., Cook, J. C., Murray, J., Prince, M., Levin, E., Mann, A. and Knapp, M. (2005). Measurement of health-related quality of life for people with dementia: development of a new instrument (DEMQOL) and an evaluation of current methodology. *Health Technology Assessment,* 9(10), 1–93. doi:10.3310/hta9100.

Smyth, E. C., Lagergren, J., Fitzgerald, R. C., Lordick, F., Shah, M. A., Lagergren, P., and Cunningham, D. (2017). Oesophageal cancer. *Nature Reviews Disease Primers*, *3*, 17048. doi: 10.1038/nrdp.2017.48.

SPSS Inc. (2007). *SPSS 16.0 user's guide.* Chicago, IL: SPSS Inc.

Stata Inc. (2008). *Stata 10.1 update*. College Station, TX: StatCorp LP.

Stephens, M. R., Lewis, W. G., Brewster, A. E., Lord, I., Blackshaw, G. R., Hodzovic, I., Thomas, G. V., Roberts, S. A., Crosby, T. D., Gent, C., Allison, M. C., Shute, K. (2006). Multidisciplinary team management is associated with improved outcomes after surgery for esophageal cancer. *Dis Esophagus*, 19, 3, 164–171.

Streiner, D. L., Norman, G. R. (2008). *Health measurement scales: a practical guide to their development and use*. 4th edn. Oxford: Oxford University Press.

Takizawa, K., Matsuda, T., Kozu, T., Eguchi, T., Kato, H., Nakanishi, Y., Hijikata, A. and Saito, D. (2009). Lymph node staging in esophageal squamous cell carcinoma: a comparative study of endoscopic ultrasonography versus computed tomography. *Journal of Gastroenterology Hepatology*, 24, 1687-1691. doi: 10.1111/j.1440-1746.2009.05927.x

Thakkar, S. and Kaul, V. (2020). Endoscopic ultrasound staging of esophageal cancer. *Gastroenterology and Hepatology*, 16(1), 14-20.

The Royal College of Surgeons of England (RCSE). (2013). *National Oesophago-Gastric Cancer Audit 2013: An audit of the care received by people with Oesophago-Gastric Cancer in England and Wales.* Retrieved from https://files.digital.nhs.uk/publicationimport/pub11xxx/pub11093/clin-audi-supp-prog-oeso-gast-2013-rep.pdf

Thomson, C. E., Beggs, I., Martin, D. J., McMillan, D., Edwards, R. T., Russell, D., Yeo, S. T., Russell, I. T. and Alastair Gibson, J. N. (2013). Methylprednisolone Injections for the Treatment of Morton Neuroma: A Patient Blinded Randomised Trial. *The Journal of Bone and Joint Surgery,* 95, 790-798.

Tosh, J., Brazier, J., Evans, P. and Longworth, L. (2012). A review of generic preference-based measures of health-related quality of life in visual disorders. *Value Health*, 15(1), 118-127.

Tsuchiya, A. and Williams, A. (2001). *Welfare economics and economic evaluation*. In McGuire, A. and Drummond, M.F. (eds) Economic Evaluation in Health Care: Merging Theory with Practice. Oxford: Oxford University Press.

University of York Centre for Reviews and Dissemination. (2009). *Systematic Reviews: CRD's guidance for undertaking reviews in health care*. York: Centre for Reviews and Dissemination, University of York.

Valero, M. and Robles-Medranda, C. (2017). Endoscopic ultrasound in oncology: An update of clinical applications in the gastrointestinal tract. *World Journal of Gastrointestinal Endoscopy,* 9(6), 243-254. doi: 10.4253/wjge.v9.i6.243

van Teijlingen, E. and Hundley, V. (2002). The importance of pilot studies. *Nursing Standard,* 16(40), 33-36.

van Vliet, E. P., Heijenbrok-Kal, M. H., Hunink, M. G., Kuipers, E. J. and Siersema, P. D. (2008). Staging investigations for oesophageal cancer: a meta-analysis. *British Journal of Cancer*, 98(3), 547-557.

Vilgrain, V., Mompoint, D., Palazzo, L., Menu, Y., Gavet, B., Ollier, P., Nahum, H. and Fekete, F. (1990). Staging of esophageal carcinoma: Comparison of results with endoscopic sonography and CT. *American Journal of Roentgenology*, 155, 277–281.

Wailoo, A. J., Hernandez-Alava, M., Manca, A., Mejia, A., Ray, J., Crawford, B., Botteman, M., Busschbach, J. (2017). Mapping to estimate health-state utility from non-preference-based outcome measures: An ISPOR good practices for outcomes research task force report. *Value in Health, 20, 18-27.*

Wallace, M. B., Nietert, P. J., Earle, C., Krasna, M. J., Hawes, R. H., Hoffman, B. J. and Reed, C. E. (2002). An analysis of multiple staging management strategies for carcinoma of the esophagus: Computed tomography, endoscopic ultrasound, positron emission tomography, and thoracoscopy/laparoscopy. *Annals of Thoracic Surgery*, 74, 1026–1032.

Weinstein, M. C. and Stason, W. B. (1977). Foundations of cost-effectiveness analysis for health and medical practices. *New England Journal of Medicine,* 296, 716-721.

White, I. R., Royston, P. and Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistic in Medicine,* 30, 377–399.

Whitehead, S. J. and Ali, S. (2010). Health outcomes in economic evaluation: the QALY and utilities. *British Medical Bulletin,* 96, 5-21. doi: 10.1093/bmb/ldq033.

Whynes, D. K., McCahon, R. A., Ravenscroft, A., Hodgkinson, V., Evley, R. and Hardman, J. G. (2013). Responsiveness of the EQ-5D Health-Related Quality-of-Life Instrument in Assessing Low Back Pain. *Value in Health*, 16, 124-132.

Wichmann, A. B., Adang, E. M. M., Stalmeier, P. F. M., Kristanti, S., Van den Block, L., Vernooij-Dassen, M. J. F. J., Engels, Y. On behalf of PACE. (2017). The use of Quality-Adjusted Life Years in cost-effectiveness analyses in palliative care: Mapping the debate through an integrative review. *Palliative Medicine,* 31(4), 306-322. doi: 10.1177/0269216316689652.

Willems, D. C., Joore, M. A., Nieman, F. H., Severens, J. L., Wouters, E. F. and Hendriks, J. J. (2009). Using EQ-5D in children with asthma, rheumatic disorders, diabetes, and speech/language and/or hearing disorders. *International Journal Technology Assess Health Care*, 25, 391–399.

Williams, I., Mclver, S., Moore, D. and Bryan, S. (2008). The use of economic evaluations in NHS decision-making: a review and empirical investigation. *Health Technology Assessment*, 12(7).

Yang, Y., Rowen, D., Brazier, J., Tsuchiya, A., Young, T. and Longworth, L. (2015). An exploratory study to test the impact on three “bolt-on” items to the EQ-5D. *Value in Health*, 18, 52–60.

Yeo, S. T., Bray, N., Haboubi, H., Hoare, Z. and Edwards, R. T. (2016). Economic evidence for EUS staging in patients with gastro-oesophageal cancer (GOC): protocol for a systematic

review. PROSPERO 2016: CRD42016043700. Retrieved from http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42016043700.

Yeo, S. T., Bray, N., Haboubi, H., Hoare, Z. and Edwards, R. T. (2019). Endoscopic ultrasound staging in patients with gastro-oesophageal cancers: a systematic review of economic evidence. *BMC Cancer,* 19, 900. doi: https://doi.org/10.1186/s12885-019-6116-0

Yeo, S. T., Williams, H. L. and Edwards, R. T. (2019). *The supply and demand of preventive goods and services and the need for economic analysis.* In Edwards, R. T. and McIntosh, E. (eds) Applied Health Economics for Public Health Practice and Research. Oxford: Oxford University Press.

Appendices

Appendices relating to Chapter 2 (Economic evaluation)

Appendix 2.1: Estimated mean frequency of contacts with secondary health care and hospital-prescribed drugs by 213 participants over 54 months

	EUS group							Non-EUS group							Total sample						
	Estimated mean frequency in complete interval						Estimated mean frequency for complete case	Estimated mean frequency in complete interval						Estimated mean frequency for complete case	Estimated mean frequency in complete interval						Estimated mean frequency for complete case
Time period (months)[a]	**0-12**	**12-18**	**18-24**	**24-36**	**36-48**	**48-60**		**0-12**	**12-18**	**18-24**	**24-36**	**36-48**	**48-60**		**0-12**	**12-18**	**18-24**	**24-36**	**36-48**	**48-60**	
Effective sample size, n[b]	**106.5**	**97.5**	**83**	**63**	**33.5**	**10**		**105**	**96**	**80.5**	**60**	**31**	**10**		**211.5**	**193.5**	**163.5**	**123**	**64.5**	**20**	
Outpatient visits																					
Number of visits	11.31	2.47	2.90	2.97	1.28	0.80	**21.73**	11.29	2.06	1.45	1.75	1.16	0.00	**17.71**	11.30	2.27	2.19	2.37	1.22	0.40	**19.75**
Inpatient stay and day case																					
Inpatient stay, cancer-related causes (no. of bed-days)	12.79	2.68	1.80	3.02	0.57	0.00	**20.84**	17.82	1.99	3.19	3.38	3.16	0.00	**29.55**	15.29	2.34	2.48	3.20	1.81	0.00	**25.11**
Inpatient stay, non-cancer-related causes (no. of bed-days)	0.49	0.24	0.00	0.03	0.00	0.00	**0.76**	0.18	0.03	0.04	0.08	0.00	0.00	**0.33**	0.34	0.13	0.02	0.06	0.00	0.00	**0.55**
Inpatient hospital stay, all causes (no. of bed-days)	13.28	2.91	1.80	3.05	0.57	0.00	**21.60**	18.00	2.02	3.23	3.47	3.16	0.00	**29.88**	15.62	2.47	2.50	3.25	1.81	0.00	**25.66**
Day-case EUS (no.)	0.90	0.00	0.00	0.00	0.00	0.00	**0.90**	0.03	0.00	0.00	0.00	0.00	0.00	**0.03**	0.47	0.00	0.00	0.00	0.00	0.00	**0.47**
Treatment																					
Surgery (count, no. of bed-days)	0.61/ 13.68	0.00	0.00	0.00	0.00	0.00	**0.61/ 13.68**	0.62/ 15.74	0.00	0.00	0.00	0.00	0.00	**0.62/ 15.74**	0.61/ 14.70	0.00	0.00	0.00	0.00	0.00	**0.61/ 14.70**
EMR day case	0.08	0.02	0.01	0.00	0.00	0.00	**0.11**	0.08	0.01	0.00	0.00	0.00	0.00	**0.09**	0.08	0.02	0.01	0.00	0.00	0.00	**0.10**
EMR inpatient (count, no. of bed-days)	0.03/ 0.06	0.00	0.00	0.00	0.00	0.00	**0.03/ 0.06**	0.01/ 0.03	0.00	0.00	0.00	0.00	0.00	**0.01/ 0.03**	0.02/ 0.04	0.00	0.00	0.00	0.00	0.00	**0.02/ 0.04**
Chemotherapy (no. of cycles)	2.24	0.25	0.35	0.33	0.00	0.00	**3.17**	2.82	0.14	0.26	0.13	0.13	0.00	**3.48**	2.53	0.19	0.31	0.24	0.06	0.00	**3.32**
Radiotherapy (no. of fractions)	5.44	0.35	0.48	0.17	0.00	0.00	**6.44**	5.35	0.54	0.17	0.43	0.00	0.00	**6.50**	5.39	0.44	0.33	0.30	0.00	0.00	**6.47**
Other procedures																					
UGIE or dilatation day case (no.)	0.65	0.37	0.11	0.11	0.00	0.00	**1.24**	0.47	0.03	0.05	0.22	0.00	0.00	**0.76**	0.56	0.20	0.08	0.16	0.00	0.00	**1.00**

UGIE or dilatation inpatient (count, no. of bed-days)	0.09/ 0.35	0.04/ 0.11	0.01/ 0.02	0.03/ 0.08	0.00/ 0.00	0.00/ 0.00	**0.18/ 0.56**	0.08/ 0.33	0.01 / 0.02	0.00/ 0.00	0.00/ 0.00	0.06/ 0.16	0.00/ 0.00	**0.15/ 0.52**	0.09/ 0.34	0.03 / 0.07	0.01/ 0.01	0.02/ 0.04	0.03/ 0.08	0.00/ 0.00	**0.16/ 0.54**
STENT insertion day case	0.05	0.00	0.01	0.00	0.00	0.00	**0.06**	0.08	0.01	0.00	0.00	0.03	0.00	**0.12**	0.06	0.01	0.01	0.00	0.02	0.00	**0.09**
STENT insertion inpatient (count, no. of bed-days)	0.11/ 0.47	0.02/ 0.07	0.02/ 0.11	0.05/ 0.13	0.00/ 0.00	0.00/ 0.00	**0.20/ 0.78**	0.06/ 0.20	0.02 / 0.07	0.02/ 0.10	0.05/ 0.45	0.03/ 0.10	0.00/ 0.00	**0.19/ 0.92**	0.09/ 0.34	0.02 / 0.07	0.02/ 0.10	0.05/ 0.28	0.02/ 0.05	0.00/ 0.00	**0.19/ 0.84**
Other procedures day case	0.27	0.02	0.01	0.03	0.00	0.00	**0.34**	0.28	0.01	0.06	0.05	0.06	0.00	**0.46**	0.27	0.02	0.04	0.04	0.03	0.00	**0.40**
Other procedures inpatient (count, no. of bed-days)	0.10/ 0.38	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	**0.10/ 0.38**	0.10/ 0.64	0.00 / 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	**0.10/ 0.64**	0.10/ 0.51	0.00 / 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	**0.10/ 0.51**
Tests and investigations																					
Count	66.03	8.69	7.41	7.63	2.90	4.60	**97.26**	78.00	7.34	7.48	11.32	9.32	0.00	**113.46**	71.97	8.02	7.44	9.43	5.98	2.30	**105.15**
PAMs																					
Dietician	4.86	0.41	0.01	0.08	0.15	0.00	**5.51**	6.25	0.02	0.40	0.10	0.00	0.00	**6.77**	5.55	0.22	0.20	0.09	0.08	0.00	**6.14**
Occupational therapist	0.08	0.02	0.04	0.00	0.00	0.00	**0.14**	0.05	0.02	0.00	0.00	0.00	0.00	**0.07**	0.07	0.02	0.02	0.00	0.00	0.00	**0.11**
Physiotherapist	5.65	0.09	0.01	0.08	0.00	0.00	**5.84**	9.01	0.06	0.16	0.10	0.00	0.00	**9.33**	7.32	0.08	0.09	0.09	0.00	0.00	**7.57**
Social worker	0.02	0.00	0.00	0.00	0.00	0.00	**0.02**	0.01	0.00	0.00	0.00	0.00	0.00	**0.01**	0.01	0.00	0.00	0.00	0.00	0.00	**0.01**
Other PAMs (e.g. district nurse, palliative care team)	1.03	0.04	0.07	0.40	0.00	0.00	**1.54**	2.13	0.22	0.29	0.08	0.00	0.00	**2.72**	1.58	0.13	0.18	0.24	0.00	0.00	**2.13**
All the above PAMs	11.65	0.56	0.13	0.56	0.15	0.00	**13.05**	17.45	0.32	0.84	0.28	0.00	0.00	**18.90**	14.53	0.44	0.48	0.42	0.08	0.00	**15.96**
Drugs																					
Hospital prescribed drugs (no. of items)[c]	28.93	5.26	4.67	5.40	3.01	5.90	**53.18**	31.79	4.31	4.09	3.70	4.29	2.40	**50.58**	30.35	4.79	4.39	4.57	3.63	4.15	**51.87**

PAMs, Professionals Allied to Medicine; UGIE, upper gastrointestinal endoscopy; EMR, endometrial mucosal resection.

a 0–12 months is the first year after randomisation; 12–18 months is the period between 12 and 18 months after randomisation, etc.

b The sample size decreases to include only those who were still in the trial: survivors observed throughout the interval, and those who died before the end of the interval but were randomised early enough for the end of the interval to occur before 31 July 2009. Participants (dead or alive) for whom the end of the trial occurred during the interval were included but given half of the weight of those in the trial for the full interval.

c Not including those prescribed for surgery or chemotherapy.

Appendix 2.2: Estimated mean cost of secondary health care and hospital-prescribed drugs by 213 patients over 54 months

	EUS group							Non-EUS group							Total sample						
	Estimated mean cost in complete interval						Estimated mean cost for complete case	Estimated mean cost in complete interval						Estimated mean cost for complete case	Estimated mean cost in complete interval						Estimated mean cost for complete case
Time period (months)[a]	**0-12**	**12-18**	**18-24**	**24-36**	**36-48**	**48-60**		**0-12**	**12-18**	**18-24**	**24-36**	**36-48**	**48-60**		**0-12**	**12-18**	**18-24**	**24-36**	**36-48**	**48-60**	
Effective sample size, n[b]	**106.5**	**97.5**	**83**	**63**	**33.5**	**10**		**105**	**96**	**80.5**	**60**	**31**	**10**		**211.5**	**193.5**	**163.5**	**123**	**64.5**	**20**	
Total outpatient visits																					
Number of visits	1290	262	324	318	133	70	2397	1317	226	165	187	113	0	2007	1303	244	246	254	123	35	2205
Inpatient stay and day case																					
Inpatient hospital stay (all cancer-related causes)	6003	973	741	1103	239	0	9059	8161	779	1366	1230	1556	0	13092	7074	877	1049	1165	872	0	11037
Inpatient hospital stay (all non-cancer-related causes)	176	87	0	14	0	0	278	67	12	31	66	0	0	176	122	50	15	40	0	0	227
Inpatient hospital stay (all causes, cancer- and non-cancer-related)	6179	1060	741	117	239	0	9336	8228	791	1397	1296	1556	0	13268	7196	927	1064	1205	872	0	11263
Day case EUS	497	0	0	0	0	0	497	16	0	0	0	0	0	16	258	0	0	0	0	0	258
Inpatient stay and day case costs																					
Total cost	6676	1060	741	1117	239	0	9833	8244	791	1397	1296	1556	0	13284	7454	927	1064	1205	872	0	11521
Treatment																					
Surgery (all types including palliative surgery)	8226	0	0	0	0	0	8226	8976	0	0	0	0	0	8976	8598	0	0	0	0	0	8598
EMR day case	39	10	6	0	0	0	56	40	5	0	0	0	0	45	40	8	3	0	0	0	51
EMR inpatient	25	0	0	0	0	0	25	13	0	0	0	0	0	13	19	0	0	0	0	0	19
Chemotherapy	1278	172	250	230	0	0	1929	1613	60	170	82	76	0	2001	1445	116	211	158	36	0	1965
Radiotherapy	712	40	62	22	0	0	836	660	69	24	64	0	0	816	686	55	43	42	0	0	826
Other procedure																					

UGIE or dilatation day case	339	187	55	54	0	0	636	245	16	25	106	0	0	392	292	102	40	80	0	0	514
UGIE or dilatation inpatient	154	48	10	33	0	0	245	148	9	0	0	64	0	221	151	29	5	17	31	0	233
Stent insertion day case	71	0	20	0	0	0	91	133	18	0	0	51	0	201	102	9	10	0	24	0	145
Stent insertion inpatient	346	55	75	107	0	0	582	158	56	72	243	74	0	603	253	55	73	173	36	0	590
Other procedures day case	111	16	6	14	0	0	147	91	4	24	19	26	0	165	101	10	15	17	13	0	155
Other procedures inpatient	168	0	0	0	0	0	168	330	0	0	0	0	0	330	248	0	0	0	0	0	248
Total treatment cost	11470	528	484	460	0	0	12942	12405	237	315	514	291	0	13763	11934	384	401	486	140	0	13345
Total tests and investigations cost	1003	103	124	97	28	44	1399	1041	109	126	132	77	0	1485	1022	106	125	114	52	22	1440
PAMs																					
Dietician	81	6	0.19	1	2	0	91	101	0.32	6	1	0	0	109	91	3	3	1	1	0	100
Occupational therapist	4	1	2	0	0	0	6	2	1	0	0	0	0	3	3	1	1	0	0	0	4
Physiotherapist	124	2	0.25	2	0	0	128	197	2	3	2	0	0	205	160	2	2	2	0	0	166
Social worker	0.69	0	0	0	0	0	0.69	0.35	0	0	0	0	0	0.35	0.52	0	0	0	0	0	0.52
Other PAMs (e.g. district nurse, palliative care team)	65	4	8	43	0	0	120	133	25	18	9	0	0	184	99	14	13	26	0	0	152
Total PAMs cost	274	13	10	46	2	0	345	434	28	27	13	0	0	502	354	20	19	30	1	0	423
Total NHS secondary sector costs	**20714**	**1966**	**1683**	**2037**	**402**	**114**	**26916**	**23441**	**1392**	**2030**	**2141**	**2037**	**0**	**31041**	**22068**	**1681**	**1854**	**2088**	**1188**	**57**	**28935**
Hospital drugs[c]																					
Cost	997	178	162	379	269	227	2212	1816	201	165	236	310	28	2756	1403	189	163	310	289	128	2482
Grand total cost (total secondary care cost + hospital prescribed drugs cost)	***21710***	***2144***	***1844***	***2417***	***671***	***341***	***29128***	***25257***	***1592***	***2195***	***2378***	***2346***	***28***	***33797***	***23471***	***1870***	***2017***	***2398***	***1476***	***185***	***31417***

PAMs, Professionals Allied to Medicine; UGIE, upper gastrointestinal endoscopy; EMR, endometrial mucosal resection.

a 0–12 months is the first year after randomisation; 12–18 months is the period between 12 and 18 months after randomisation, etc.

b The sample size decreases to include only those who were still in the trial: survivors observed throughout the interval, and those who died before the end of the interval, but were randomised early enough for the end of the interval to occur before 31 July 2009. Participants (dead or alive) for whom the end of the trial occurred during the interval were included but given half of the weight of those in the trial for the full interval.

c Not including those prescribed for surgery or chemotherapy.

Appendix 2.3: Treatment plan at randomisation by plan amended after EUS. This table is quoted directly from the published COGNATE report.

TABLE 28 Treatment plan at randomisation by plan amended after EUS

Group	Treatment plan at randomisation	Treatment plan as amended after EUS or post-randomisation MDT				
		EMR	Surgery	Neo-adjuvant chemotherapy before surgery	Multimodal including palliative	Total
EUS	EMR	0	1	0	0	**1**
	Surgery	2	10	1	5	**18**
	Neo-adjuvant chemotherapy before surgery	4	7	49	7	**67**
	Multimodal	0	2	0	21	**23**
	Total	**6**	**20**	**50**	**33**	**109**
Non-EUS	EMR	3	1	0	1	**5**
	Surgery	0	9	0	1	**10**
	Neo-adjuvant chemotherapy before surgery	2	6	55	13	**76**
	Multimodal	0	1	2	16	**19**
	Total	**5**	**17**	**57**	**31**	**110**

Shaded cells represent changes of plan.

Appendix 2.4: Treatment plan by actual treatment. This table is quoted directly from the published COGNATE report.

TABLE 32 Treatment plan by actual treatment

Group	Treatment plan at randomisation	Treatment delivered				
		EMR	Surgery	Neo-adjuvant chemotherapy before surgery	Multimodal including palliative	Total
EUS	EMR	0	0	0	1	**1**
	Surgery	2	10	1	5	**18**
	Neo-adjuvant chemotherapy before surgery	4	8	36	19	**67**
	Multimodal	1	3	1	18	**23**
	Total	**7**	**22**	**38**	**43**	**109**
Non-EUS	EMR	2	1	0	2	**5**
	Surgery	0	9	0	1	**10**
	Neo-adjuvant chemotherapy before surgery	2	6	43	25	**76**
	Multimodal	1	1	2	15	**19**
	Total	**5**	**17**	**45**	**43**	**110**

Shaded cells represent changes between plan and treatment.

Appendices relating to Chapter 3 (Systematic review)

Appendix 3.1: PROSPERO systematic review protocol

Citation

Seow Tien Yeo, Nathan Bray, Hasan Haboubi, Zoe Hoare, Rhiannon Tudor Edwards. Economic evidence for EUS staging in patients with gastro-oesophageal cancer (GOC): protocol for a systematic review. PROSPERO 2016 CRD42016043700 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42016043700

Review question

What are the costs, benefits (in terms of health-related quality of life (HRQoL) and cancer-specific quality of life), and economic implications of the use of EUS staging in the management of gastro-oesophageal cancer (GOC) patients?

Searches

The electronic databases that will be used by the primary author (STY) for conducting the searches are: NHS EED, DARE, HTA, MEDLINE, EMBASE, The Cochrane Collaboration Register and Library (including CCRCT, Cochrane Central Register of Controlled Trials (CENTRAL)), CINAHL, Web of Science (Core Collection). In order to ensure a comprehensive search is achieved and any relevant research has not been missed, firstly, online searches will be conducted through the following internet search engines and appropriate websites to identify grey literature, reports, ongoing and unpublished studies from conference papers and abstracts: Google, The Google Scholar, The Department of Health (DoH), The National Institute for Health and Clinical Excellence (NICE), The National Institute for Health Research (NIHR) Journals Library, The NIHR UK Clinical Trials Gateway, The National Cancer Research Institute (NCRI), The Cancer Research Wales (CRW), The Wales Cancer Research Centre (WCRC), The Welsh Government (WG), The Health and Care Research Wales (HCRW), The Cancer Research UK (CRUK) and other relevant charitable organisation websites; lastly, gathering possible additional relevant papers through searching reference lists of all included papers will be conducted. Contacts with study authors will also be made, if necessary, to locate further relevant literature and publications. All health economic literatures relating to EUS staging of GOC from the last 20 years (from 1996 to 2016) will be searched. Due to resources constraints, only studies written in English will be included. International studies will be included if they are translated or written in English.

Types of study to be included

All full economic evaluation studies i.e. cost-effectiveness, cost-utility and cost-benefit will be considered for inclusion in this review. Other economic studies that contain partial economic evaluation or no evaluation context e.g. outcome description in terms of health-related quality of life and cancer-specific quality of life measures e.g. QALYs and FACT-G score, cost description, cost analysis, cost-outcome description and budgetary studies will also be considered for inclusion in this review. In terms of study design, economic evaluation studies conducted alongside randomised controlled trial (RCT) will be of greatest relevance as it will be directly linked with the economic evaluation evidence stream of this review and the results of costs, outcomes and the joint cost-outcome evaluation yielded from the primary data source (i.e. RCT) will have high degree of reliability. If a randomised controlled trial does not have a full economic evaluation alongside but does have estimates of resource use and/or costs alone or outcomes alone, it may be considered for inclusion as evidence of resource use only and/or costs only or outcome only as they will give valuable insight. In addition to randomised controlled trial, economic evaluation conducted alongside non-randomised controlled trials, quasi-experimental trials, quality of life research (in the context of health-related or cancer specific quality of life measures), observation studies, longitudinal studies and cohort studies will also be considered for inclusion.

Condition or domain being studied

Use of endoscopic ultrasound (EUS) in staging of gastro-oesophageal cancer.

Participants/population

Inclusion: Adults (aged 19 or over) who had cancer (i.e. localised tumour) of the oesophagus, stomach or gastro-oesophageal junction; free of metastatic disease.
Exclusion: Aged below 19, metastatic oesophagus, stomach or gastro-oesophageal junction cancer.

Intervention(s), exposure(s)
Use of endoscopic ultrasound (EUS) staging in addition to standard staging algorithm (e.g. CT scan) in patient with gastro-oesophageal cancer.

Comparator(s)/control
Standard staging algorithm e.g. trans-abdominal ultrasound scan, computed tomography (CT) scan.

Context
Studies of endoscopic ultrasound (EUS) staging in patients with oesophagus, stomach or gastro-oesophageal cancer will all be considered relevant. International study will be included if translated into English.

Main outcome(s)
All relevant outcomes of full economic evaluation studies including (but not be restricted to) cost per QALY and cost per life-year gained; and all other relevant economic outcomes (including partial economic evaluation studies outcomes in which costs or consequences alone of a single intervention (e.g. EUS staging of GOC) were described) including (but not be restricted to) resource use, direct and indirect costs, incremental benefits e.g. quality-adjusted survival or quality-adjusted life years (QALYs), health-related quality of life, cancer-specific quality of life and utility gained. Effect on costs of healthcare resource use, frequency of healthcare resource use, quality-adjusted survival/quality-adjusted life years gained, life-years gained, utility scores, cancer-specific quality of life scores, quality of life measures, incremental cost-effectiveness ratio and incremental cost-utility ratio will inform the economic outcomes.
Health-related quality of life measures will include validated health-related quality of life measure tools such as EQ-5D, Health Utilities Index (HUI) and SF-36. Cancer-specific quality of life measures will include validated cancer-specific quality of life measure tools such as FACT-G scale, FACT-AC scale and EORTC QLQ-C30.

Additional outcome(s)
None

Data extraction (selection and coding)
Electronic search results from all databases will be downloaded into RefWork bibliographic software; then, the combined search results will be screened for duplicates. Duplicates will be removed if there are any. Prior to data extraction, there are two stages of initial screening (i.e. selection process) for study eligibility. Firstly, titles and abstracts from the search results will be screened for potential eligibility for inclusion in the review – they will be examined against the inclusion criteria by two independent reviewers (STY and NB) separately; reasons for exclusion will be noted for studies that do not meet inclusion criteria. Secondly, full papers for the abstracts that meet the inclusion criteria will be obtained and examined against the inclusion criteria by STY and NB independently; reasons for exclusion will be noted for studies that do not meet inclusion criteria. Any disagreements will be resolved through discussion in accordance with protocol. If decisions cannot be reached, further discussions will be carried out with the review team until consensus is reached and if necessary, authors will be contacted for clarity.
Once all the eligible full papers have been included, reviewing, assessing for methodological quality, extracting and synthesising data will be the next important challenging stages in conducting a systematic review of research evidence. All the retrieved eligible full papers will be categorised into two main streams: Economic evaluation evidence stream, and other economic evidence (including quality of life evidence) stream (see Figure 1). Then, the included full papers from each stream will be reviewed and assessed for its quality, and evidence will be extracted and synthesised.
In terms of data extraction, data from each of the included papers will be extracted onto the relevant data extraction form by the primary reviewer (STY) and the second reviewer (NB) separately. In order to ensure the validity and reliability of this review is achieved, results from the data extraction form from the two independent reviewers (STY and NB) will be assessed for any discrepancies. Any discrepancies will be resolved through discussion. If any previous similar systematic reviews are found and if all their included economic literatures fell in the period of 1996 to 2016, they will be used as a unit of paper for comparison; if not all fell in that period, only individual papers that fell in that period will be retrieved and considered for inclusion in this review. Data to be extracted will include (but are not limited to): Study design, country, perspective, methods, populations, interventions, outcomes and results. Tables depicting the extracted data

will be produced to summarise key information.

Risk of bias (quality) assessment
Two independent reviewers (STY and NB) will critically appraise the included papers for methodological quality using validated checklist for economic studies i.e. Drummond's checklist or Critical Appraisal Skills Programme (CASP) economic evaluation checklist, for their reporting standards and risk of bias in order to limit bias. A table presenting a list of included papers with quality appraisal ratings will be produced.

Strategy for data synthesis
Two independent reviewers (STY and NB) will review and synthesise the evidence of all the included papers separately. The data will be reviewed and summarised across the different evidence streams in a narrative form to answer the review objectives. The narrative synthesis will be supported by information presented in the tables which describe the characteristics, methods, outcomes and quality of the included papers. Meta-analysis is unlikely to be appropriate as it is anticipated that the search will identify heterogeneous study type and outcomes across the studies.

Analysis of subgroups or subsets
None

Contact details for further information
Ms Yeo
s.t.yeo@bangor.ac.uk

Organisational affiliation of the review
None
None

Review team members and their organisational affiliations
Ms Seow Tien Yeo. Centre for Health Economics and Medicines Evaluation, Bangor University, UK.
Dr Nathan Bray. Centre for Health Economics and Medicines Evaluation, Bangor University, UK.
Dr Hasan Haboubi. Cancer Biomarkers Group, Swansea University, UK.
Dr Zoe Hoare. North Wales Organisation for Randomised Trials in Health and Social Care (NWORTH), Bangor University, UK.
Professor Rhiannon Tudor Edwards. Centre for Health Economics and Medicines Evaluation, Bangor University, UK.

Type and method of review
Cost effectiveness, Diagnostic, Intervention, Systematic review

Anticipated or actual start date
02 May 2016

Anticipated completion date
31 January 2017

Funding sources/sponsors
This systematic review will be undertaken as part of a PhD thesis. The PhD is funded by Tenovus Cancer Care, registered charity number 1054015. The funder has not had any role in developing the protocol. Bangor University is the sponsor for this systematic review.

Conflicts of interest
None known

Language
English

Country
Wales

Stage of review
Review Ongoing

Subject index terms status
Subject indexing assigned by CRD

Subject index terms
Esophageal Neoplasms; Esophagogastric Junction; Humans

Date of registration in PROSPERO
27 July 2016

Date of first submission

Stage of review at time of this submission

Stage	Started	Completed
Preliminary searches	Yes	Yes
Piloting of the study selection process	Yes	No
Formal screening of search results against eligibility criteria	No	No
Data extraction	No	No
Risk of bias (quality) assessment	No	No
Data analysis	No	No

The record owner confirms that the information they have supplied for this submission is accurate and complete and they understand that deliberate provision of inaccurate information or omission of data may be construed as scientific misconduct.
The record owner confirms that they will update the status of the review when it is completed and will add publication details in due course.

Versions
27 July 2016

Appendix 3.2: PRISMA Guidelines (Updated version)

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	

Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	
Synthesis of results	21	Present the main results of the review. If meta-analyses are done, include for each, confidence intervals and measures of consistency.	
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

Appendix 3.3: Search strategy for MEDLINE via Ovid and the results

#	Searches	Results
1	exp Endosonography/	10030
2	endosono$.tw.	2227
3	endoscopic ultraso$.tw.	8020
4	endoscopic-ultraso$.tw.	8020
5	EUS.tw.	6441
6	(echoendoscop$ or echo-endoscop$).tw.	610
7	((endosono$ or endoscopic ultraso$ or endoscopic-ultraso$ or EUS) adj6 aspiration).tw.	2176
8	1 or 2 or 3 or 4 or 5 or 6 or 7	16409
9	staging.tw.	59177
10	((Preoperative or pre-operative) adj6 staging).tw.	3687
11	9 or 10	59177
12	8 and 11	2530
13	exp Adenocarcinoma/	317200
14	adenocarcinoma$.tw.	112263
15	13 or 14	356486
16	exp Esophagus/	46067
17	exp Esophagogastric Junction/	7423
18	(gastroesophag$ adj3 junction$).tw.	2068
19	(gastro-esophag$ adj3 junction$).tw.	201
20	(gastrooesophag$ adj3 junction$).tw.	32
21	(gastro-oesophag$ adj3 junction$).tw.	412
22	esophagogastric junction$.tw.	1450
23	esophago-gastric junction$.tw.	165

24	oesophagogastric junction$.tw.	217
25	oesophago-gastric junction$.tw.	85
26	exp Stomach/	116444
27	16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26	151578
28	15 and 27	8985
29	exp Esophageal Neoplasms/	42331
30	exp Stomach Neoplasms/	82204
31	(esophag$ adj5 neoplas$).tw.	1487
32	(oesophag$ adj5 neoplas$).tw.	255
33	(esophag$ adj5 cancer$).tw.	19631
34	(oesophag$ adj5 cancer$).tw.	3915
35	(esophag$ adj5 carcin$).tw.	14966
36	(oesophag$ adj5 carcin$).tw.	2763
37	(esophag$ adj5 tumo$).tw.	5026
38	(oesophag$ adj5 tumo$).tw.	902
39	(esophag$ adj5 metasta$).tw.	1972
40	(oesophag$ adj5 metasta$).tw.	279
41	(esophag$ adj5 malig$).tw.	2500
42	(oesophag$ adj5 malig$).tw.	645
43	(esophag$ adj5 adenocarcinoma$).tw.	5071
44	(oesophag$ adj5 adenocarcinoma$).tw.	1308
45	(stomach adj5 neoplas$).tw.	783
46	(stomach adj5 cancer$).tw.	11123
47	(stomach adj5 carcin$).tw.	4498
48	(stomach adj5 tumo$).tw.	3823
49	(stomach adj5 metasta$).tw.	1056

50	(stomach adj5 malig$).tw.	1181
51	(stomach adj5 adenocarcinoma$).tw.	1786
52	(gastric adj5 neoplas$).tw.	2040
53	(gastric adj5 cancer$).tw.	48385
54	(gastric adj5 carcin$).tw.	18391
55	(gastric adj5 tumo$).tw.	9450
56	(gastric adj5 metasta$).tw.	6111
57	(gastric adj5 malig$).tw.	3296
58	(gastric adj5 adenocarcinoma$).tw.	7474
59	29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58	148829
60	28 or 59	149629
61	(gut$ adj5 (neoplas$ or cancer$ or carcin$ or tumo$ or adenocarcinoma$ or metasta$ or malig$)).tw.	1011
62	(gullet$ adj5 (neoplas$ or cancer$ or carcin$ or tumo$ or adenocarcinoma$ or metasta$ or malig$)).tw.	13
63	(food pipe adj5 (neoplas$ or cancer$ or carcin$ or tumo$ or adenocarcinoma$ or metasta$ or malig$)).tw.	0
64	(("upper GI" or "upper-GI") adj5 (neoplas$ or cancer$ or carcin$ or tumo$ or adenocarcinoma$ or metasta$ or malig$)).tw.	315
65	(("upper gastrointestin$" or "upper-gastrointestin$") adj5 (neoplas$ or cancer$ or carcin$ or tumo$ or adenocarcinoma$ or metasta$ or malig$)).tw.	1733
66	((upper digestive tract$ or upper-digestive tract$) adj5 (neoplas$ or cancer$ or carcin$ or tumo$ or adenocarcinoma$ or metasta$ or malig$)).tw.	250
67	61 or 62 or 63 or 64 or 65 or 66	3212

68	60 or 67	151393
69	12 and 68	991
70	exp Economics/	526348
71	health economics.mp.	2853
72	Economic evaluation.mp.	6595
73	exp Cost-Benefit Analysis/	65882
74	(cost$ adj2 (effective$ or utilit$ or consequence$ or benefit$ or minimi$)).tw.	112506
75	Cost effectiveness analysis.mp.	6951
76	cost utility analysis.mp.	1657
77	cost consequences analysis.mp.	47
78	cost minimisation analysis.mp.	128
79	cost minimization analysis.mp.	405
80	exp "Costs and Cost Analysis"/	197564
81	(unit cost or unit-cost or unit-costs or unit costs or drug cost or drug costs or hospital costs or health-care costs or health care cost or medical cost or medical costs).tw.	25430
82	(cost$ adj2 (efficac$ or analys$ or allocation$ or control$ or illness$ or affordable$ or fee$ or charge$)).tw.	32778
83	exp Models, Economic/	11684
84	(decision adj1 (tree$ or analys$ or model$)).tw.	10447
85	Markov$.tw.	16615
86	exp Economics, Pharmaceutical/ or exp Economics, Medical/ or exp Economics, Hospital/	37277
87	(econom$ or cost$ or price$ or pricing or pharmacoeconomic$ or pharmaeconomic$ or pharmaco-economic$).tw.	595818
88	exp "Fees and Charges"/	28183

89	exp Budgets/	12815
90	(financ$ or fee$).tw.	490886
91	((value or values or valuation) adj2 (money or monetary or life or lives or costs or cost)).tw.	4987
92	exp Health Expenditures/	17219
93	(low adj cost).mp.	32881
94	(high adj cost).mp.	9687
95	(health?care adj cost$).mp.	6233
96	(cost adj estimate$).mp.	1683
97	exp Hospital Costs/	8804
98	exp "Cost Savings"/	9754
99	exp "Quality of Life"/	136989
100	*"Quality of Life"/	61982
101	70 or 71 or 72 or 73 or 74 or 75 or 76 or 77 or 78 or 79 or 80 or 81 or 82 or 83 or 84 or 85 or 86 or 87 or 88 or 89 or 90 or 91 or 92 or 93 or 94 or 95 or 96 or 97 or 98 or 99 or 100	1517091
102	69 and 101	58
103	limit 102 to (english language and humans and yr="1996 - 2016" and "all adult (19 plus years)")	19

Appendix 3.4: Search strategy for CINAHL with Full Text via EBSCOhost and the results

#	Searches	Results
S1	(MH "Endosonography")	1,657
S2	AB endosono*	98
S3	AB endoscopic ultraso*	404
S4	AB endoscopic-ultraso*	365
S5	AB EUS	573
S6	AB echoendoscop* or echo-endoscop*	57
S7	AB (endosono* or endoscopic ultraso* or endoscopic-ultraso* or EUS) N6 aspiration	118
S8	S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7	2,013
S9	AB staging	3,162
S10	AB (Preoperative or pre-operative) N6 staging	127
S11	S9 OR S10	3,162
S12	S8 AND S11	161
S13	(MH "Adenocarcinoma+")	9,429
S14	AB adenocarcinoma*	2,496
S15	S13 OR S14	10,809
S16	(MH "Esophagus")	2,150
S17	"Esophagogastric Junction"	70
S18	AB gastroesophag* N3 junction*	120
S19	AB gastro-esophag* N3 junction*	1

S20	TX gastrooesophag* N3 junction*	7
S21	AB gastro-oesophag* N3 junction*	16
S22	AB esophagogastric junction*	55
S23	AB esophago-gastric junction*	2
S24	AB oesophagogastric junction*	13
S25	AB oesophago-gastric junction*	2
S26	(MH "Stomach+")	1,977
S27	S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26	4,126
S28	S15 AND S27	290
S29	(MH "Esophageal Neoplasms+")	2,248
S30	(MH "Stomach Neoplasms+")	2,506
S31	AB esophag* N5 neoplas*	74
S32	AB oesophag* N5 neoplas*	5
S33	AB esophag* N5 cancer*	611
S34	AB oesophag* N5 cancer*	281
S35	AB esophag* N5 carcin*	232
S36	AB oesophag* N5 carcin*	79
S37	AB esophag* N5 tumo*	93
S38	AB oesophag* N5 tumo*	37
S39	AB esophag* N5 metasta*	31

S40	AB oesophag* N5 metasta*	17
S41	AB esophag* N5 malig*	68
S42	AB oesophag* N5 malig*	20
S43	AB esophag* N5 adenocarcinoma*	252
S44	AB oesophag* N5 adenocarcinoma*	94
S45	AB stomach N5 neoplas*	13
S46	AB stomach N5 cancer*	372
S47	AB stomach N5 carcin*	36
S48	AB stomach N5 tumo*	58
S49	AB stomach N5 metasta*	19
S50	AB stomach N5 malig*	20
S51	AB stomach N5 adenocarcinoma*	40
S52	AB gastric N5 neoplas*	75
S53	AB gastric N5 cancer*	1,136
S54	AB gastric N5 carcin*	252
S55	AB gastric N5 tumo*	174
S56	AB gastric N5 metasta*	97
S57	AB gastric N5 malig*	82
S58	AB gastric N5 adenocarcinoma*	194
S59	S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38 OR S39 OR S40 OR S41 OR S42 OR S43 OR S44 OR S45 OR S46 OR S47 OR S48 OR S49 OR S50 OR S51 OR S52 OR S53 OR S54 OR S55 OR S56 OR S57 OR S58	5,566

S60	S28 OR S59	5,589
S61	AB gut* N5 (neoplas* or cancer* or carcin* or tumo* or adenocarcinoma* or metasta* or malig*)	32
S62	TX gullet* N5 (neoplas* or cancer* or carcin* or tumo* or adenocarcinoma* or metasta* or malig*)	4
S63	TX food pipe N5 (neoplas* or cancer* or carcin* or tumo* or adenocarcinoma* or metasta* or malig*)	1
S64	AB ("upper GI" or "upper-GI") N5 (neoplas* or cancer* or carcin* or tumo* or adenocarcinoma* or metasta* or malig*)	30
S65	AB ("upper gastrointestin*" or "upper-gastrointestin*") N5 (neoplas* or cancer* or carcin* or tumo* or adenocarcinoma* or metasta* or malig*)	83
S66	AB (upper digestive tract* or upper-digestive tract*) N5 (neoplas* or cancer* or carcin* or tumo* or adenocarcinoma* or metasta* or malig*)	17
S67	S61 OR S62 OR S63 OR S64 OR S65 OR S66	155
S68	S60 OR S67	5,677
S69	S12 AND S68	50
S70	(MH "Economics+")	501,302
S71	"health economics"	3,531
S72	"Economic evaluation"	1,348
S73	(MH "Cost Benefit Analysis")	15,358
S74	AB cost* N2 (effective* or utilit* or consequence* or benefit* or minimi*)	16,294
S75	"Cost effectiveness analysis"	1,264
S76	"cost utility analysis"	273
S77	"cost consequences analysis"	7

S78	"cost minimisation analysis"	26
S79	"cost minimization analysis"	59
S80	(MH "Costs and Cost Analysis+")	60,866
S81	AB unit cost or unit-cost or unit-costs or unit costs or drug cost or drug costs or hospital costs or health-care costs or health care cost or medical cost or medical costs	11,510
S82	AB cost* N2 (efficac* or analys* or allocation* or control* or illness* or affordable* or fee* or charge*)	5,244
S83	"economic model*"	473
S84	AB decision N1 (tree* or analys* or model*)	1,985
S85	AB Markov*	875
S86	(MH "Economics, Pharmaceutical")	1,355
S87	AB econom* or cost* or price* or pricing or pharmacoeconomic* or pharmaeconomic* or pharmaco-economic*	77,571
S88	(MH "Fees and Charges+")	9,271
S89	(MH "Budgets")	7,004
S90	AB financ* or fee*	65,793
S91	AB (value or values or valuation) N2 (money or monetary or life or lives or costs or cost)	1,174
S92	"Health Expenditures"	297
S93	low N1 cost	2,722
S94	high N1 cost	2,521
S95	healthcare N5 cost*	3,004

S96	health care N5 cost*	28,350
S97	health-care N5 cost*	28,013
S98	cost N5 estimate*	3,522
S99	(MH "Health Care Costs+") OR (MH "Health Facility Costs")	27,259
S100	(MH "Cost Savings")	9,343
S101	(MH "Health Resource Allocation") OR (MH "Health Resource Utilization")	16,105
S102	(MH "Health Services Needs and Demand+")	13,803
S103	(MH "Health Care Delivery+")	183,276
S104	(MH "Quality of Life+") OR (MH "Quality-Adjusted Life Years")	52,692
S105	*"Quality of Life"	67,519
S106	*"Quality-adjusted life year*"	2,436
S107	S70 OR S71 OR S72 OR S73 OR S74 OR S75 OR S76 OR S77 OR S78 OR S79 OR S80 OR S81 OR S82 OR S83 OR S84 OR S85 OR S86 OR S87 OR S88 OR S89 OR S90 OR S91 OR S92 OR S93 OR S94 OR S95 OR S96 OR S97 OR S98 OR S99 OR S100 OR S101 OR S102 OR S103 OR S104 OR S105 OR S106	752,484
S108	S69 AND S107	8
S109	S69 AND S107 (Limiters - Published Date: 1996-2016; English Language; Age Groups: All Adult; Language: English)	5

Appendix 3.5: Search strategy for EMBASE Ovid and the results

#	Searches	**Results**
1	exp endoscopic echography/	22483
2	endosono$.tw.	3296
3	endoscopic ultraso$.tw.	13042
4	endoscopic-ultraso$.tw.	13042
5	EUS.tw.	12766
6	(echoendoscop$ or echo-endoscop$).tw.	1241
7	((endosono$ or endoscopic ultraso$ or endoscopic-ultraso$ or EUS) adj6 aspiration).tw.	3690
8	1 or 2 or 3 or 4 or 5 or 6 or 7	30501
9	staging.tw.	89632
10	((Preoperative or pre-operative) adj6 staging).tw.	5470
11	9 or 10	89632
12	8 and 11	3843
13	exp adenocarcinoma/	85749
14	adenocarcinoma$.tw.	156037
15	13 or 14	189413
16	exp esophagus/	67964
17	Esophagogastric Junction.mp. or exp lower esophagus sphincter/	12901
18	(gastroesophag$ adj3 junction$).tw.	3374
19	(gastro-esophag$ adj3 junction$).tw.	504
20	(gastrooesophag$ adj3 junction$).tw.	67
21	(gastro-oesophag$ adj3 junction$).tw.	591
22	esophagogastric junction$.tw.	2190
23	esophago-gastric junction$.tw.	347

24	oesophagogastric junction$.tw.	288
25	oesophago-gastric junction$.tw.	135
26	exp stomach/	144323
27	16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26	201289
28	15 and 27	9911
29	Esophageal Neoplasms.mp. or exp esophagus tumor/	65893
30	Stomach Neoplasms.mp. or exp stomach tumor/	124738
31	(esophag$ adj5 neoplas$).tw.	2022
32	(oesophag$ adj5 neoplas$).tw.	379
33	(esophag$ adj5 cancer$).tw.	27294
34	(oesophag$ adj5 cancer$).tw.	5437
35	(esophag$ adj5 carcin$).tw.	19981
36	(oesophag$ adj5 carcin$).tw.	3316
37	(esophag$ adj5 tumo$).tw.	6873
38	(oesophag$ adj5 tumo$).tw.	1238
39	(esophag$ adj5 metasta$).tw.	2925
40	(oesophag$ adj5 metasta$).tw.	377
41	(esophag$ adj5 malig$).tw.	3481
42	(oesophag$ adj5 malig$).tw.	858
43	(esophag$ adj5 adenocarcinoma$).tw.	7512
44	(oesophag$ adj5 adenocarcinoma$).tw.	1894
45	(stomach adj5 neoplas$).tw.	875
46	(stomach adj5 cancer$).tw.	12696
47	(stomach adj5 carcin$).tw.	5230
48	(stomach adj5 tumo$).tw.	4830
49	(stomach adj5 metasta$).tw.	1404

50	(stomach adj5 malig$).tw.	1529
51	(stomach adj5 adenocarcinoma$).tw.	2227
52	(gastric adj5 neoplas$).tw.	2811
53	(gastric adj5 cancer$).tw.	65949
54	(gastric adj5 carcin$).tw.	23647
55	(gastric adj5 tumo$).tw.	12791
56	(gastric adj5 metasta$).tw.	8435
57	(gastric adj5 malig$).tw.	4484
58	(gastric adj5 adenocarcinoma$).tw.	10114
59	29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58	205633
60	28 or 59	206828
61	(gut$ adj5 (neoplas$ or cancer$ or carcin$ or tumo$ or adenocarcinoma$ or metasta$ or malig$)).tw.	1362
62	(gullet$ adj5 (neoplas$ or cancer$ or carcin$ or tumo$ or adenocarcinoma$ or metasta$ or malig$)).tw.	10
63	(food pipe adj5 (neoplas$ or cancer$ or carcin$ or tumo$ or adenocarcinoma$ or metasta$ or malig$)).tw.	0
64	(("upper GI" or "upper-GI") adj5 (neoplas$ or cancer$ or carcin$ or tumo$ or adenocarcinoma$ or metasta$ or malig$)).tw.	670
65	(("upper gastrointestin$" or "upper-gastrointestin$") adj5 (neoplas$ or cancer$ or carcin$ or tumo$ or adenocarcinoma$ or metasta$ or malig$)).tw.	2448
66	((upper digestive tract$ or upper-digestive tract$) adj5 (neoplas$ or cancer$ or carcin$ or tumo$ or adenocarcinoma$ or metasta$ or malig$)).tw.	311
67	61 or 62 or 63 or 64 or 65 or 66	4614
68	60 or 67	209313

69	12 and 68	1634
70	exp economics/	233559
71	exp health economics/	692434
72	exp economic evaluation/	242818
73	exp "cost benefit analysis"/	71793
74	(cost$ adj2 (effective$ or utilit$ or consequence$ or benefit$ or minimi$)).tw.	151624
75	exp "cost effectiveness analysis"/	114521
76	exp "cost utility analysis"/	6741
77	cost consequences analysis.mp.	68
78	exp "cost minimization analysis"/	2807
79	cost minimisation analysis.mp.	235
80	"Costs and Cost Analysis".mp. or exp "cost"/	288134
81	(unit cost or unit-cost or unit-costs or unit costs or drug cost or drug costs or hospital costs or health-care costs or health care cost or medical cost or medical costs).tw.	38072
82	(cost$ adj2 (efficac$ or analys$ or allocation$ or control$ or illness$ or affordable$ or fee$ or charge$)).tw.	46497
83	Models, Economic.mp.	31
84	(decision adj1 (tree$ or analys$ or model$)).tw.	14321
85	Markov$.tw.	20127
86	exp pharmacoeconomics/	179007
87	Economics, Medical.mp.	318
88	Economics, Hospital.mp.	20
89	(econom$ or cost$ or price$ or pricing or pharmacoeconomic$ or pharmaeconomic$ or pharmaco-economic$).tw.	775104
90	"Fees and Charges".mp.	61
91	exp budget/	22363

92	(financ$ or fee$).tw.	617072
93	((value or values or valuation) adj2 (money or monetary or life or lives or costs or cost)).tw.	6789
94	Health Expenditures.mp. or exp "health care cost"/	234299
95	(low adj cost).mp.	37828
96	(high adj cost).mp.	11791
97	(health?care adj cost$).mp.	10395
98	(cost adj estimate$).mp.	2416
99	exp "hospital cost"/	29609
100	Cost Savings.mp.	14604
101	exp "quality of life"/	338827
102	*"Quality of Life"/	70024
103	70 or 71 or 72 or 73 or 74 or 75 or 76 or 77 or 78 or 79 or 80 or 81 or 82 or 83 or 84 or 85 or 86 or 87 or 88 or 89 or 90 or 91 or 92 or 93 or 94 or 95 or 96 or 97 or 98 or 99 or 100 or 101 or 102	2143805
104	69 and 103	124
105	limit 104 to (english language and yr="1996 - 2016")	113

Appendix 3.6: Search strategy for the Cochrane Library (databases include Cochrane Reviews, DARE, HTA, NHS EED, CMR, CENTRAL) and the results

#	Searches	Results
1	MeSH descriptor: [Endosonography] explode all trees	389
2	endosono*:ti,ab,kw (Word variations have been searched)	469
3	"endoscopic ultraso*":ti,ab,kw (Word variations have been searched)	376
4	"endoscopic-ultraso*":ti,ab,kw (Word variations have been searched)	376
5	"EUS":ti,ab,kw (Word variations have been searched)	355
6	("echoendoscop*") or ("echo-endoscop*"):ti,ab,kw (Word variations have been searched)	41
7	((endosono*) or ("endoscopic ultraso*") or ("endoscopic-ultraso*") or ("EUS")) near/6 ("aspiration"):ti,ab,kw (Word variations have been searched)	154
8	#1 or #2 or #3 or #4 or #5 or #6 or #7	768
9	"staging":ti,ab,kw (Word variations have been searched)	41222
10	(("Preoperative") or ("pre-operative")) near/6 ("staging"):ti,ab,kw (Word variations have been searched)	409
11	#9 or #10	41222
12	#8 and #11	177
13	MeSH descriptor: [Adenocarcinoma] explode all trees	5182
14	adenocarcinoma*:ti,ab,kw (Word variations have been searched)	4544
15	#13 or #14	6949
16	MeSH descriptor: [Esophagus] explode all trees	1193
17	MeSH descriptor: [Esophagogastric Junction] explode all trees	364
18	gastroesophag* near/3 junction*:ti,ab,kw (Word variations have been searched)	187
19	gastro-esophag* near/3 junction*:ti,ab,kw (Word variations have been searched)	18

20	gastrooesophag* near/3 junction*:ti,ab,kw (Word variations have been searched)	3
21	gastro-oesophag* near/3 junction*:ti,ab,kw (Word variations have been searched)	49
22	"esophagogastric junction*":ti,ab,kw (Word variations have been searched)	358
23	"esophago-gastric junction*":ti,ab,kw (Word variations have been searched)	7
24	"oesophagogastric junction*":ti,ab,kw (Word variations have been searched)	15
25	"oesophago-gastric junction*":ti,ab,kw (Word variations have been searched)	10
26	MeSH descriptor: [Stomach] explode all trees	2990
27	#16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26	3921
28	#15 and #27	258
29	MeSH descriptor: [Esophageal Neoplasms] explode all trees	1134
30	MeSH descriptor: [Stomach Neoplasms] explode all trees	1944
31	esophag* near/5 neoplas*:ti,ab,kw (Word variations have been searched)	1302
32	oesophag* near/5 neoplas*:ti,ab,kw (Word variations have been searched)	16
33	esophag* near/5 cancer*:ti,ab,kw (Word variations have been searched)	1219
34	oesophag* near/5 cancer*:ti,ab,kw (Word variations have been searched)	337
35	esophag* near/5 carcin*:ti,ab,kw (Word variations have been searched)	847
36	oesophag* near/5 carcin*:ti,ab,kw (Word variations have been searched)	155
37	esophag* near/5 tumo*:ti,ab,kw (Word variations have been searched)	122
38	oesophag* near/5 tumo*:ti,ab,kw (Word variations have been searched)	29
39	esophag* near/5 metasta*:ti,ab,kw (Word variations have been searched)	92

40	oesophag* near/5 metasta*:ti,ab,kw (Word variations have been searched)	27
41	esophag* near/5 malig*:ti,ab,kw (Word variations have been searched)	62
42	oesophag* near/5 malig*:ti,ab,kw (Word variations have been searched)	27
43	esophag* near/5 adenocarcinoma*:ti,ab,kw (Word variations have been searched)	247
44	oesophag* near/5 adenocarcinoma*:ti,ab,kw (Word variations have been searched)	77
45	stomach near/5 neoplas*:ti,ab,kw (Word variations have been searched)	2077
46	stomach near/5 cancer*:ti,ab,kw (Word variations have been searched)	1147
47	stomach near/5 carcin*:ti,ab,kw (Word variations have been searched)	189
48	stomach near/5 tumo*:ti,ab,kw (Word variations have been searched)	189
49	stomach near/5 metasta*:ti,ab,kw (Word variations have been searched)	106
50	stomach near/5 malig*:ti,ab,kw (Word variations have been searched)	14
51	stomach near/5 adenocarcinoma*:ti,ab,kw (Word variations have been searched)	186
52	gastric near/5 neoplas*:ti,ab,kw (Word variations have been searched)	148
53	gastric near/5 cancer*:ti,ab,kw (Word variations have been searched)	2761
54	gastric near/5 carcin*:ti,ab,kw (Word variations have been searched)	480
55	gastric near/5 tumo*:ti,ab,kw (Word variations have been searched)	132
56	gastric near/5 metasta*:ti,ab,kw (Word variations have been searched)	270
57	gastric near/5 malig*:ti,ab,kw (Word variations have been searched)	79
58	gastric near/5 adenocarcinoma*:ti,ab,kw (Word variations have been searched)	285
59	#29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39 or #40 or #41 or #42 or #43 or #44 or #45 or #46 or #47 or #48 or #49 or #50 or #51 or #52 or #53 or #54 or #55 or #56 or #57 or #58	6272
60	#28 or #59	6284

61	gut* near/5 (neoplas* or cancer* or carcin* or tumo* or adenocarcinoma* or metasta* or malig*):ti,ab,kw (Word variations have been searched)	24
62	gullet* near/5 (neoplas* or cancer* or carcin* or tumo* or adenocarcinoma* or metasta* or malig*):ti,ab,kw (Word variations have been searched)	3
63	"food pipe" near/5 (neoplas* or cancer* or carcin* or tumo* or adenocarcinoma* or metasta* or malig*):ti,ab,kw (Word variations have been searched)	2
64	("upper GI" or "upper-GI") near/5 (neoplas* or cancer* or carcin* or tumo* or adenocarcinoma* or metasta* or malig*):ti,ab,kw (Word variations have been searched)	25
65	("upper gastrointestin*" or "upper-gastrointestin*") near/5 (neoplas* or cancer* or carcin* or tumo* or adenocarcinoma* or metasta* or malig*):ti,ab,kw (Word variations have been searched)	117
66	("upper digestive tract*" or "upper-digestive tract*") near/5 (neoplas* or cancer* or carcin* or tumo* or adenocarcinoma* or metasta* or malig*):ti,ab,kw (Word variations have been searched)	6
67	#61 or #62 or #63 or #64 or #65 or #66	156
68	#60 or #67	6379
69	#12 and #68	56
70	MeSH descriptor: [Economics] explode all trees	26554
71	MeSH descriptor: [Health Care Economics and Organizations] explode all trees	34088
72	"Economic evaluation*":ti,ab,kw (Word variations have been searched)	3552
73	"health economic*":ti,ab,kw (Word variations have been searched)	922
74	MeSH descriptor: [Cost-Benefit Analysis] explode all trees	17715
75	cost* near/2 (effective* or utilit* or consequence* or benefit* or minimi*):ti,ab,kw (Word variations have been searched)	29373

76	"cost effectiveness analysis":ti,ab,kw (Word variations have been searched)	6024
77	"cost utility analysis":ti,ab,kw (Word variations have been searched)	998
78	"cost consequences analysis":ti,ab,kw (Word variations have been searched)	91
79	"cost minimisation analysis":ti,ab,kw (Word variations have been searched)	257
80	"cost minimization analysis":ti,ab,kw (Word variations have been searched)	257
81	MeSH descriptor: [Costs and Cost Analysis] explode all trees	24576
82	"unit cost" or "unit-cost" or "unit-costs" or "unit costs" or "drug cost" or "drug costs" or "hospital costs" or "health-care costs" or "health care cost" or "medical cost" or "medical costs":ti,ab,kw (Word variations have been searched)	11389
83	cost* near/2 (efficac* or analys* or allocation* or control* or illness* or affordable* or fee* or charge*):ti,ab,kw (Word variations have been searched)	28467
84	MeSH descriptor: [Models, Economic] explode all trees	1987
85	decision near/1 (tree* or analys* or model*):ti,ab,kw (Word variations have been searched)	2429
86	Markov*:ti,ab,kw (Word variations have been searched)	2553
87	MeSH descriptor: [Economics, Pharmaceutical] explode all trees	243
88	MeSH descriptor: [Economics, Medical] explode all trees	105
89	MeSH descriptor: [Economics, Hospital] explode all trees	1731
90	econom* or cost* or price* or pricing or pharmacoeconomic* or pharmaeconomic* or "pharmaco-economic*":ti,ab,kw (Word variations have been searched)	54801
91	MeSH descriptor: [Fees and Charges] explode all trees	502
92	MeSH descriptor: [Budgets] explode all trees	71

93	financ* or fee*:ti,ab,kw (Word variations have been searched)	31277
94	(value or values or valuation) near/2 (money or monetary or life or lives or costs or cost):ti,ab,kw (Word variations have been searched)	761
95	MeSH descriptor: [Health Expenditures] explode all trees	315
96	low near cost:ti,ab,kw (Word variations have been searched)	4821
97	high near cost:ti,ab,kw (Word variations have been searched)	3148
98	("healthcare" or "health-care" or "health care") near cost*:ti,ab,kw (Word variations have been searched)	8280
99	cost near estimate*:ti,ab,kw (Word variations have been searched)	1972
100	MeSH descriptor: [Hospital Costs] explode all trees	1470
101	MeSH descriptor: [Cost Savings] explode all trees	984
102	MeSH descriptor: [Health Care Costs] explode all trees	7128
103	MeSH descriptor: [Health Resources] explode all trees	561
104	MeSH descriptor: [Health Services Needs and Demand] explode all trees	503
105	MeSH descriptor: [Quality of Life] explode all trees	17758
106	"*Quality of Life":ti,ab,kw (Word variations have been searched)	41881
107	MeSH descriptor: [Quality-Adjusted Life Years] explode all trees	4078
108	"quality-adjusted life year*" or "quality adjusted life year*" or "quality-adjusted-life-year*":ti,ab,kw (Word variations have been searched)	5150
109	#70 or #71 or #72 or #73 or #74 or #75 or #76 or #77 or #78 or #79 or #80 or #81 or #82 or #83 or #84 or #85 or #86 or #87 or #88 or #89 or #90 or #91 or #92 or #93 or #94 or #95 or #96 or #97 or #98 or #99 or #100 or #101 or #102 or #103 or #104 or #105 or #106 or #107 or #108	123635
110	#69 and #109	6
111	#69 and #109 Publication Year from 1996 to 2016	6

Appendix 3.7: Search strategy for Web of Science Core Collection and the results

#	Searches	Results
1	TS=((endosono* OR endoscopic ultraso* OR endoscopic-ultraso* OR EUS OR echoendoscop* OR "echo-endoscop*") AND (staging))	3,683
2	TS=((endosono* OR endoscopic ultraso* OR endoscopic-ultraso* OR EUS OR echoendoscop* OR "echo-endoscop*") AND (Preoperative NEAR/6 staging))	636
3	TS=((endosono* OR endoscopic ultraso* OR endoscopic-ultraso* OR EUS OR echoendoscop* OR "echo-endoscop*") AND ("Pre-operative" NEAR/6 staging))	46
4	TS=(((endosono* OR "endoscopic ultraso*" OR "endoscopic-ultraso*" OR EUS) NEAR/6 aspiration) AND staging)	602
5	TS=(((endosono* OR "endoscopic ultraso*" OR "endoscopic-ultraso*" OR EUS) NEAR/6 aspiration) AND (Preoperative NEAR/6 staging))	71
6	TS=(((endosono* OR "endoscopic ultraso*" OR "endoscopic-ultraso*" OR EUS) NEAR/6 aspiration) AND ("Pre-operative" NEAR/6 staging))	4
7	OR/1-6	3,683
8	TS=(((gastroesophag* NEAR/3 junction*) OR ("gastro-esophag*" NEAR/3 junction*) OR (gastrooesophag* NEAR/3 junction*) OR ("gastro-oesophag*" NEAR/3 junction*) OR "esophagogastric junction*" OR "esophago-gastric junction*" OR "oesophagogastric junction*" OR "oesophago-gastric junction*" OR "Esophagogastric Junction" OR "Esophagus" OR "Stomach") AND adenocarcinoma*)	14,254

9	TS=("Esophageal Neoplasms" OR "Stomach Neoplasms" OR ("esophag*" Near/5 "neoplas*") OR ("oesophag*" NEAR/5 "neoplas*") OR ("esophag*" NEAR/5 "cancer*") OR ("oesophag*" NEAR/5 "cancer*") OR ("esophag*" NEAR/5 "carcin*") OR ("oesophag*" NEAR/5 "carcin*") OR ("esophag*" NEAR/5 "tumo*") OR ("oesophag*" NEAR/5 "tumo*") OR ("esophag*" NEAR/5 "metasta*") OR ("oesophag*" NEAR/5 "metasta*") OR ("esophag*" NEAR/5 "malig*") OR ("oesophag*" NEAR/5 "malig*") OR ("esophag*" NEAR/5 "adenocarcinoma*") OR ("oesophag*" NEAR/5 "adenocarcinoma*") OR (stomach NEAR/5 neoplas*) OR (stomach NEAR/5 cancer*) OR (stomach NEAR/5 carcin*) OR (stomach NEAR/5 tumo*) OR (stomach NEAR/5 metasta*) OR (stomach NEAR/5 malig*) OR (stomach NEAR/5 adenocarcinoma*) OR (gastric NEAR/5 neoplas*) OR (gastric NEAR/5 cancer*) OR (gastric NEAR/5 carcin*) OR (gastric NEAR/5 tumo*) OR (gastric NEAR/5 metasta*) OR (gastric NEAR/5 malig*) OR (gastric NEAR/5 adenocarcinoma*))	124,890
10	TS=((gut* OR gullet* OR "food pipe" OR "upper GI" OR "upper-GI" OR "upper gastrointestin*" OR "upper-gastrointestin*" OR "upper digestive tract*" OR "upper-digestive tract*") NEAR/5 (neoplas* OR cancer* OR carcin* OR tumo* OR adenocarcinoma* OR metasta* OR malig*))	4,089
11	OR/8-10	128,996
12	WC=Economics	648,798
13	TS=("health economics" OR "Economic evaluation" OR "Cost-Benefit Analysis" OR "Cost-effectiveness analysis" OR "cost-utility analysis" OR "cost-consequences analysis" OR "cost-minimi?ation analysis" OR (cost* NEAR/2 (effective* OR utilit* OR consequence* OR benefit* OR minimi*)))	244,724

14	TS=(Cost$ OR "Cost Analys*" OR "unit cost" OR "unit-cost" OR "unit-costs" OR "unit costs" OR "drug cost" OR "drug costs" OR "hospital costs" OR "health-care costs" OR "health care cost" OR "medical cost" OR "medical costs" OR (cost$ NEAR/2 (efficac* OR analys* OR allocation* OR control* OR illness* OR affordable* OR fee* OR charge*)))	1,057,473
15	TS=("Economic Model*" OR (decision NEAR/1 (tree* OR analys* OR model*)) OR Markov* OR econom* OR cost* OR price* OR pricing OR pharmacoeconomic* OR pharmaeconomic* OR pharmaco-economic* OR financ* OR fee* OR ((value or values or valuation) NEAR/2 (money or monetary or life or lives or costs or cost)) OR (low NEAR cost) OR (high NEAR cost) OR ("health-care" NEAR cost*) OR (cost NEAR estimate*) OR "Hospital Cost*" OR "Cost Saving*")	3,110,411
16	TS=("Quality of Life" OR "Quality-Adjusted Life Year$" OR "QALY$")	246,323
17	OR/12-16	3,672,335
18	AND/7,11,17	83
19	Limit #18 to articles, meeting abstracts, proceedings papers and correction; English language and year="1996-2016"	54

Appendix 3.8: Data extraction form for economic studies (1a) – Economic evaluations

DATA EXTRACTION FORM (1a)- Economic Evaluations				
Study ID				
Reviewer				
Checked by				
Study Title				
Author(s)				
Publication year				
Aims/objectives/hypotheses				
Study perspective				
Price year/currency (unit)				
METHOD				
Type of study				
Randomised?	Yes	No	If yes, allocation type:	
No. of groups				
No. in each group				
No. completed in each group/response rate				
Data collection time points				
Measure of benefit (QALY, life years gained etc)				
PARTICIPANTS				
Types of participants				
Age range				

Inclusion criteria	
Exclusion criteria	
Ethnicity/country	
Baseline characteristics	
INTERVENTIONS	
Type of intervention(s)	
Content of intervention(s)	
Duration of intervention(s)	
Control intervention(s)	
Follow-up period	
Outcome and measure(s) 1	
Outcome and measure(s) 2	
Outcome and measure(s) 3	
Outcome and measure(s) 4	
Outcome and measure(s) 5	
Outcome and measure(s) 6	
Outcome and measure(s) 7	
Outcome and measure(s) 8	
Outcome and measure(s) 9	
Outcome and measure(s) 10	
Summary of findings (including statistical significance, significance	

level, confidence intervals and effect size)	
Cost per QALY/Incremental Cost-Effectiveness Ratio (ICER) conclusions	
Inflated cost per QALY/ICER to 2016 price year	
CONCLUSIONS	
Study conclusions	

Appendix 3.9: Data extraction form for economic studies (1b) – Other economic studies

DATA EXTRACTION FORM (1b) – Other Economic Data				
Study ID				
Reviewer				
Checked by				
Study Title				
Author(s)				
Publication year				
Aims/objectives/hypotheses				
Study perspective				
Price year/currency (unit) (if applicable)				
METHOD				
Type of study/methodology				
Randomised?	Yes	No	If yes, allocation type:	
No. of groups				
No. in each group				
No. completed in each group/response rate				
Data collection time points				
PARTICIPANTS				
Types of participants (e.g. adult diagnosed with GOC stage 1 or 2 or 3)				
Age range				

Inclusion criteria	
Exclusion criteria	
Ethnicity/country	
Type of diagnostic tool	
Follow-up period	
Outcomes and measures	
Outcomes and measures	
Outcomes and measures	
Outcomes and measures	
INTERVENTIONS	
Type of intervention(s)	
Content of intervention(s)	
Duration of intervention(s)	
Control intervention(s)	
Follow-up period	
SUMMARY OF RESULTS	
Summary of findings	
Identified economic costs/consequences/outcomes/benefits and implications	
Converted cost into pound sterling (£) and inflated it to 2016 price year (current price year)	
CONCLUSIONS	
Study conclusions	

Appendix 3.10: Data extraction form for economic modelling studies

DATA EXTRACTION FORM (2) – Economic Modelling	
Study ID (i.e. paper no.)	
Reviewer	
Checked by	
Study Title	
Author(s)	
Publication year	
Aims/objectives/hypotheses	
METHOD	
Structure	
Statement of decision problem/objective of the evaluation and of the model	
Statement of scope/perspective of the model	
Rationale for structure of the model	
Structural assumptions	
Strategies/comparators	
Model type	
Time horizon	
Disease states/pathways	
Cycle length	
Summary of structure section	

Data	
Data identification	
Data modelling	
Baseline data	
Treatment effects	
Costs, source for all costs and discount rates	
Quality of life weights (utilities)	
Data incorporation	
Assessment of uncertainty	
Methodological uncertainty	
Structural uncertainties	
Heterogeneity	
Parameter uncertainty	
Summary of data section	
Consistency	
Internal consistency	
External consistency	
Summary of consistency section	
PARTICIPANTS	
Types of participants	
Age range	

Inclusion criteria	
Exclusion criteria	
Ethnicity/country	
OUTCOMES	
Outcome and measure(s) 1	
Outcome and measure(s) 2	
Outcome and measure(s) 3	
Outcome and measure(s) 4	
Outcome and measure(s) 5	
Outcome and measure(s) 6	
Outcome and measure(s) 7	
Outcome and measure(s) 8	
Outcome and measure(s) 9	
Outcome and measure(s) 10	
SUMMARY OF RESULTS	
Summary of findings (including statistical significance, significance level, confidence intervals and effect size)	
Cost per QALY/Incremental Cost-Effectiveness Ratio (ICER) conclusions	
Inflated cost per QALY/ICER to 2015 price year (current price year)	
CONCLUSIONS	

Study conclusions	

Appendix 3.11: Data extraction results of Shumaker et al (2002)

DATA EXTRACTION FORM 3 – Other Economic Data	
Study ID	14
Reviewer	STY (Primary reviewer)
Checked by	NB (Second reviewer)
Study Title	Potential impact of preoperative EUS on oesophageal cancer management and cost
Author(s)	Shumaker DA, de Garmo P and Faigel DO.
Publication year	2002
Aims/objectives/hypotheses	Aim – To determine the relative proportions of each oesophageal cancer stage in a group of patients referred for preoperative staging with EUS. Proportion of patients with EUS stage 1 and 4 tumours that would not be treated with combined modality therapy was determined and potential cost impact estimated. Hypothesis – Preoperative staging of oesophageal cancer with EUS identifies a significant proportion of patients with Stage 1 and 4 disease that would not be treated with combined modality therapy (chemoradiotherapy plus surgery), resulting in a potential cost savings.
Study perspective	Not stated specifically, the authors described US Medicare data.
Price year/currency (unit) (if applicable)	2000 price year; US dollars ($)
METHOD	
Type of study/methodology	Cost analysis using a retrospective review of a large multicentre national computerised endoscopic database – The potential cost savings of performing preoperative EUS in oesophageal cancer patients.

Randomised?	Yes	No	If yes, allocation type:	
No. of groups	NA. Data from three study sites (the Phoenix VAMC, Portland Oregon and Richmond Virginia) were included with a total of 180 data pertaining to EUS examinations for oesophageal cancer were reviewed and analysed.			
No. in each group	NA. No. of EUS examinations in each included site – Of the 180 examinations reviewed, 10 cases (6%) were from Phoenix VAMC, 135 cases (75%) from Portland, Oregon and 35 (19%) from Richmond, Virginia.			
No. completed in each group/response rate	NA			
Data collection time points	Not stated. Data between Feb 4, 1998 and October 31, 2000 were extracted and entered into the database for review and analysis.			
PARTICIPANTS				
Types of participants (e.g. adult diagnosed with GOC stage 1 or 2 or 3)	Patients with oesophageal cancer received preoperative staging with EUS. Sample size = 180			
Age range	Of the 180 patients referred for preoperative staging of oesophageal cancer by EUS, 82% were men and mean age was 66.5 years.			
Inclusion criteria	Complete staging data of which preoperative staging of oesophageal tumours by EUS performed to patients with oesophageal cancer were included for review and analysis (n=180).			
Exclusion criteria	Data were excluded for incomplete staging information for more than two thirds of examinations or no staging information was recorded in the database.			

Ethnicity/country	USA, ethnicity not stated.
Type of diagnostic tool	Preoperative EUS staging for oesophageal cancer
Follow-up period	Not stated
Outcomes and measures	Proportion of stage 1 and 4 tumours
Outcomes and measures	Cost saving associated with avoidable treatment following EUS staging
Outcomes and measures	
Outcomes and measures	
INTERVENTIONS	
Type of intervention(s)	NA, because it is a retrospective review of a large national endoscopic database.
Content of intervention(s)	NA
Duration of intervention(s)	NA
Control intervention(s)	NA
Follow-up period	NA
SUMMARY OF RESULTS	
Summary of findings	For every 100 patients staged before surgery with EUS (cost $63,420), 14 (14%) patients with Stage I disease would be spared neoadjuvant chemoradiotherapy (saving $122,192) and 12 (12%) patients with Stage IV cancer would be spared surgery (saving $285,600) for an average cost savings of $3443 per patient ([$122,192 + $285,600 − $63,420]/100). And, sensitivity analysis shows that preoperative staging of oesophageal cancer with EUS results in cost savings across a wide range of clinical scenarios [1) EUS saves the cost of surgery and chemoradiotherapy in 1 out of 20 patients, 2) EUS saves these costs in 1 out of 10 patients (which is similar to the results in the present study) and 3) EUS saves these costs in 1 out of 5 patients] and cost assumptions

	(cost of EUS and cost of combined chemoradiotherapy & esophagectomy).
Identified economic costs/consequences/outcomes/benefits and implications	See above in 'Summary of findings' section.
Converted cost into pound sterling (£) and inflated it to 2016 price year (current price year)	For every 100 patients staged before surgery with EUS (cost $63,420) – This equals to £43,140 (2000 price year) which this is converted using the currency exchange rates (i.e. the currency units per SDR for September 2000) from International Monetary Fund (IMF, 2017). This £43,140 (2000 price year) is then inflated using Hospital and Community Health Services (HCHS) Pay and Prices Index (Curtis and Netten, 2005; Curtis and Burns, 2015; Curtis and Burns, 2016; Curtis and Burns, 2017) to £69,750 (2017 price year). Saving $122,192 – This equals to £83,118 (2000 price year) which this is converted using the currency exchange rates (i.e. the currency units per SDR for September 2000) from International Monetary Fund (IMF, 2017). This £83,118 (2000 price year) is then inflated using Hospital and Community Health Services (HCHS) Pay and Prices Index (Curtis and Netten, 2005; Curtis and Burns, 2015; Curtis and Burns, 2016; Curtis and Burns, 2017) to £134,388 (2017 price year). Saving $285,600 – This equals to £194,272 (2000 price year) which this is converted using the currency exchange rates (i.e. the currency units per SDR for September 2000) from International Monetary Fund (IMF, 2017). This £194,272 (2000 price year) is then inflated using Hospital and Community Health Services (HCHS) Pay and Prices

	Index (Curtis and Netten, 2005; Curtis and Burns, 2015; Curtis and Burns, 2016; Curtis and Burns, 2017) to £314,105 (2017 price year). $3,443 per patient – This equals to £2,342 (2000 price year) which this is converted using the currency exchange rates (i.e. the currency units per SDR for September 2000) from International Monetary Fund (IMF, 2017). This £2,342 (2000 price year) is then inflated using Hospital and Community Health Services (HCHS) Pay and Prices Index (Curtis and Netten, 2005; Curtis and Burns, 2015; Curtis and Burns, 2016; Curtis and Burns, 2017) to £3,786 (2017 price year). **Inflation indices:** 1999/2000 – 188.5; 2004/05 – 234.2 (Curtis and Netten, 2005) 2004/05 – 232.3; 2005/06 – 240.9 (Curtis and Burns, 2015) 2005/06 – 240.9; 2015/16 – 297.0 (Curtis and Burns, 2016) 2015/16 – 297.0; 2016/17 – 302.3 (Curtis and Burns, 2017) **References:** International Monetary Fund (IMF). Currency unit per SDR for September 2000. Available from: https://www.imf.org/external/np/fin/data/rms_mth.aspx?SelectDate=2000-09-30&reportType=CVSDR (Accessed October 2017).

	Curtis L and Netten A. Unit costs of health and social care 2005. Personal Social Services Research Unit: University of Kent, 2005. Available from: http://www.pssru.ac.uk/pdf/uc/uc2005/uc2005.pdf (Accessed October 2017). Curtis L and Burns A. Unit costs of health and social care 2015. Personal Social Services Research Unit: University of Kent, 2015. Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2015/ (Accessed October 2017). Curtis L and Burns A. Unit costs of health and social care 2016. Personal Social Services Research Unit: University of Kent, 2016. Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2016/ (Accessed October 2017). Curtis, Lesley A. and Burns, Amanda. Unit Costs of Health and Social Care 2017. Personal Social Services Research Unit: University of Kent, 2017. Available from: https://doi.org/10.22024/UniKent/01.02/65559 (Accessed February 2018)
CONCLUSIONS	
Study conclusions	Preoperative staging of oesophageal cancer with EUS can facilitate cost savings by reducing the need for additional treatments in stage 1 and 4 oesophageal cancer (a significant proportion of patients – 26% in this series).

Appendix 3.12: Data extraction results of Chang et al (2003)

DATA EXTRACTION FORM 3 – Other Economic Data	
Study ID	1
Reviewer	STY (Primary reviewer)
Checked by	NB (Second reviewer)
Study Title	Impact of Endoscopic Ultrasound Combined with Fine-Needle Aspiration Biopsy in the Management of Oesophageal Cancer
Author(s)	Chang KJ, Soetikno RM, Bastas D, Tu C and Nguyen PT
Publication year	2003
Aims/objectives/hypotheses	Aim – To determine the impact of EUS in combination with FNA on patients' choice of therapy and on the cost of care. Objective – (1). To investigate the role of EUS in combination with FNA in guiding the choice of therapy made by patients who had oesophageal cancer who otherwise were surgical candidates; (2). To explore whether or not the use of EUS decreased the cost of care for these patients. Hypothesis – It was hypothesized that the EUS staging results influence patients' choice of therapy, and that this influence on decision-making decreases the cost of care.
Study perspective	Not stated specifically – It's a US study but did not state specifically as to which perspective the cost analysis took from.
Price year/currency (unit) (if applicable)	Price year – The authors did not state specifically. The authors described their cost analyses were based on the published direct costs of endosonography-guided aspiration biopsy and thoracotomy published in year 1997 (Gress et al, 1997); Currency – US dollars (USD$)
METHOD	

Type of study/methodology	Cost analysis alongside prospective case series. Patient Selection – All patients with a diagnosis of oesophageal cancer (squamous-cell carcinoma or adenocarcinoma) who were referred to the University of California's Irvine Medical Center for local staging using EUS between August 1993 and August 1997. These patients were all being considered for surgical resection and had undergone standard evaluation, including computed tomography (CT) which showed no evidence of distant metastases. Patient Follow-up – All patients were informed of their EUS staging results. The overall staging was determined according to the TNM classification. The referring physicians saw patients within 2 weeks after completion of EUS. Patients decided in consultation with their referring physicians whether or not to undergo surgery. All surgical pathology reports were obtained for the patients who did undergo surgery. In all of the patients, the clinical status was periodically followed via phone calls to the patients and to the referring physicians.			
Randomised?	Yes	No ✓	If yes, allocation type:	
No. of groups	NA			
No. in each group	NA. Data of 60 consecutive patients with oesophageal cancer who were referred to the University of California's Irvine Medical Center for preoperative EUS staging were used in the cost analyses.			
No. completed in each group/response rate	NA			

Data collection time points	Average follow-up period = 17 months. Time points not explicitly defined.
PARTICIPANTS	
Types of participants (e.g. adult diagnosed with GOC stage 1 or 2 or 3)	Patients diagnosed with oesophageal cancer (squamous-cell or adenocarcinoma) who were referred to the University of California's Irvine Medical Center for preoperative EUS staging between August 1993 and August 1997. These patients were all being considered for surgical resection and had undergone standard evaluation including CT which showed no evidence of distant metastases.
Age range	Of the 60 consecutive patients referred for preoperative EUS staging of oesophageal cancer, 39 were men, 21 were women and mean age was 68±10 years.
Inclusion criteria	All patients referred for preoperative EUS staging of oesophageal cancer. Diagnosed with oesophageal cancer (squamous-cell or adeno) referred for EUS staging, having already undergone CT and considered for surgical resection. No evidence of distant metastases.
Exclusion criteria	Not specified.
Ethnicity/country	USA; Ethnicity not stated.
Type of diagnostic tool	EUS: Radial scanning or linear scanning EUS-guided FNA: linear-array echo endoscope
Follow-up period	Average of 17 months (range 1-51 months).
Outcomes and measures	No. of patients choosing not to undergo thoracotomy following EUS findings. Potential cost savings of performing preoperative EUS staging of oesophageal cancer.
Outcomes and measures	
Outcomes and measures	
Outcomes and measures	

INTERVENTIONS	
Type of intervention(s)	NA, because it is a cost analysis study alongside prospective case series.
Content of intervention(s)	NA
Duration of intervention(s)	NA
Control intervention(s)	NA
Follow-up period	NA, based on the data used in the cost analyses, the length of follow-up was, on average, 17 months (range 1-51 months).
SUMMARY OF RESULTS	
Summary of findings	Patients' medical decisions on whether to undergo medical therapy or surgical treatment correlated significantly (p=0.0051) with the results of their EUS staging, but not with age, sex, or referring physicians (surgeons vs. non-surgeons). EUS-guided therapy potentially decreased the cost of care by USD$740,424 (USD$12,340/patient) due to reduced number of thoracotomies undertaken (patient choice).
Identified economic costs/consequences/outcomes/benefits and implications	See above in 'Summary of findings' section.
Converted cost into pound sterling (£) and inflated it to 2016 price year (current price year)	EUS-guided therapy potentially decreased the cost of care by USD$740,424 (USD$12,340/patient) – This equals to £442,415 (£7,373/patient) (2003 price year) which this is converted using the currency exchange rates (i.e. the currency units per SDR for September 2003) from International Monetary Fund (IMF, 2017). This £442,415 (£7,373/patient) (2003 price year) is then inflated using Hospital and Community Health Services (HCHS) Pay and Prices Index (Curtis and Netten, 2005; Curtis and Burns,

	2015; Curtis and Burns, 2016; Curtis and Burns, 2017) to £630,664 (£10,510/patient) (2017 price year). **Inflation indices:** 1996/97 – 170.6; 2004/05 – 234.2 (Curtis and Netten, 2005) 2002/03 – 213.8; 2004/05 – 234.2 (Curtis and Netten, 2005) 2004/05 – 232.3; 2005/06 – 240.9 (Curtis and Burns, 2015) 2005/06 – 240.9; 2015/16 – 297.0 (Curtis and Burns, 2016) 2015/16 – 297.0; 2016/17 – 302.3 (Curtis and Burns, 2017) **References:** International Monetary Fund (IMF). Currency unit per SDR for September 2003. Available from: https://www.imf.org/external/np/fin/data/rms_mth.aspx?SelectDate=2003-09-30&reportType=CVSDR (Accessed October 2017). Curtis L and Netten A. Unit costs of health and social care 2005. Personal Social Services Research Unit: University of Kent, 2005. Available from: http://www.pssru.ac.uk/pdf/uc/uc2005/uc2005.pdf (Accessed October 2017). Curtis L and Burns A. Unit costs of health and social care 2015. Personal Social Services Research Unit: University of Kent, 2015. Available from:

	http://www.pssru.ac.uk/project-pages/unit-costs/2015/ (Accessed October 2017). Curtis L and Burns A. Unit costs of health and social care 2016. Personal Social Services Research Unit: University of Kent, 2016. Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2016/ (Accessed October 2017). Curtis, Lesley A. and Burns, Amanda. Unit Costs of Health and Social Care 2017. Personal Social Services Research Unit: University of Kent, 2017. Available from: https://doi.org/10.22024/UniKent/01.02/65559 (Accessed February 2018).
CONCLUSIONS	
Study conclusions	Patients' decisions on whether to undergo medical or surgical treatment correlated significantly with their overall tumour staging, suggesting that the information provided by EUS played a significant role in patients' decision-making. EUS-guided therapy potentially reduces the cost of managing patients with oesophageal cancer by USD$12,340 per patient due to reduced number of thoracotomies undertaken (patient choice).

Appendix 3.13: Data extraction results of Russell et al (2013)

DATA EXTRACTION FORM 1- Economic Evaluations				
Study ID	13			
Reviewer	STY (Primary reviewer)			
Checked by	NB (Second reviewer)			
Study Title	Cancer of Oesophagus or Gastricus – New Assessment of Technology of Endosonography (COGNATE): report of pragmatic randomised trial			
Author(s)	Russell IT, Edwards RT, Gliddon AE, Ingledew DK, Russell D, Whitaker R, Yeo ST, Attwood SE, Barr H, Nanthakumaran S and Park KGM			
Publication year	2013			
Aims/objectives/hypotheses	Examine whether the addition of EUS to usual staging uses resources cost-effectively			
Study perspective	NHS perspective, focusing on health-care resources used by participants including investigation, treatment and palliation, and other elements of secondary and pharmaceutical care			
Price year/currency (unit)	Price year 2008; Currency – Pounds sterling (£)			
METHOD				
Type of study	Cost-effectiveness			
Randomised?	√Yes	No	If yes, allocation type:	Stratified by centre and tumour location, equal proportions EUS and not [EUS group (EUS + standard staging algorithm) versus non-EUS group (standard staging algorithm)].
No. of groups	Two – EUS and no EUS			
No. in each group	223 patients in total were randomised → EUS (n = 111); No EUS (n = 112).			

No. completed in each group/response rate	Of 223 randomised patients, 213 yielded enough data for primary analysis → EUS (n = 107); No EUS (n = 106).
Data collection time points	EQ-5D data was collected at discharge from hospital after initial treatment (Baseline) and at follow-up clinics after 1, 3, 6, 12, 18, 24 and 36 months post baseline. Use of NHS resources by participants was recorded throughout the period of the trial (54 months) by the local co-ordinator at each trial site using an electronic database.
Measure of benefit (QALY, life years gained etc)	QALY, survival, quality of life (generic and cancer specific), health care resource use
PARTICIPANTS	
Types of participants	Patients with proven cancer of the oesophagus, stomach or gastro-oesophageal junction; had not started treatment.
Age range	All adults, age not specifically stated but analysis showed sample has an average age of 64.4
Inclusion criteria	To be eligible for the trial, patients had to be medically fit for both surgery (even if not planned) and chemoradiotherapy, free of metastatic disease and had not started treatment. Both their ASA (American Society of Anesthesiologists) grading and their WHO performance status had to be 1 or 2 (as shown in *Figure 1*). Following initial staging, clinicians could exclude patients from the trial for clinical reasons.
Exclusion criteria	a. Patients of WHO performance status 3 or 4 or medically unsuitable for either surgery or chemotherapy were excluded. b. Patients found to have metastatic liver disease were excluded. c. Patients who had evidence of metastases or then had plans for palliative treatment or were then known to be medically unfit for surgery were excluded.

Ethnicity/country	UK, ethnicity not stated
Baseline characteristics	Well documented, see demographics tables on page 34 – **Gender:** Male (n=165, 77%) in the overall 213 patients, with n=83 (78%) in intervention group and n=82 (77%) in control group. **Age:** Mean age is 64.4 year-old for the overall 213 patients (mean age of 64.4 Intervention group, 64.3 control group). For further baseline characteristics details, see Table 9 entitled "Demographic and baseline data by allocated group for 213 participants in main analyses" on page 34 of the COGNATE HTA report.
INTERVENTIONS	
Type of intervention(s)	Patients randomised to intervention group received EUS scan in addition to conventional (standard) staging investigations within the staging process – All patients with gastro-oesophageal cancer received standard staging algorithms, after which the relevant multidisciplinary team (MDT) chose a provisional management plan from: endoscopic mucosal resection (EMR); immediate surgery; surgery after neo-adjuvant chemotherapy; and chemotherapy and radiotherapy. In principle patients randomised to the intervention group then received EUS, while those randomised to the control group continued with their agreed management plan.
Content of intervention(s)	1. All patients should receive biochemistry, haematology, pulmonary function tests and cardiac assessment, not least to exclude patients whose World Health Organization (WHO) status is 3 or 4, or who are medically unsuitable for either surgery or chemotherapy.

	2. Patients who are medically fit for surgery without evidence of metastases should undergo CT following an agreed protocol using spiral scanner and intravenous contrast. 3. Patients with any suspicion of peritoneal disease should undergo laparoscopy as the best means of detecting peritoneal tumour deposits. 4. Fit patients with localised tumours and no contraindications were eligible for randomisation to EUS (intervention group) or not (control group). Those randomised to the control group continued with their agreed management plan (see above). Those randomised to the intervention group (EUS group), the final choice of treatment followed the EUS scan. The authors explained what EUS is – "Endoscopic ultrasound (EUS; or endosonography) is a medical procedure performed by gastroenterologists, radiologists or surgeons with specialised training. Endosonography combines endoscopy – the insertion of a probe into the upper gastrointestinal tract – with ultrasonography. It places a high-frequency ultrasound probe mounted on the end of the endoscope in direct contact with oesophageal or gastric tumours. This provides good images of the structures of the bowel wall and local lymph nodes, but is less good at identifying distant metastases. To patients it feels very similar to normal endoscopy, unless it includes ultrasound-guided biopsy of deeper structures. Although biopsy may increase risk, the basic procedure is no more risky than an endoscopy."

Duration of intervention(s)	The duration of intervention is not specified specifically but it explained that eligible and medically fit patients randomised to intervention group received EUS scan following their initial (standard/conventional) staging investigations.
Control intervention(s)	Patient in the control group received conventional staging investigations where appropriate i.e. biochemistry, haematology, pulmonary function tests and cardiac assessment, CT scan using spiral scanner and intravenous contrast, and laparoscopy. Those randomised to the control group continued with their agreed management plan (i.e. the choice of treatment depended on the results of the completed initial staging investigations, revisited if necessary). At the end of staging, with or without EUS, MDTs assigned patients to one of three treatment options. Patients with: 1. tumours that were adjudged to be mucosal underwent **EMR** with or without argon-beam ablation of the surrounding mucosa. 2. tumours that were adjudged to be resectable underwent **surgery with or without neo-adjuvant chemotherapy**, typically with cisplatin and 5FU. 3. advanced localised disease, for which complete resection was adjudged to be impossible, received **multimodal treatment, possibly including palliative surgery for gastric cancers.**

Follow-up period	Follow-up period was 54 months or until death, whichever comes first; Main analyses of the study including health economic analysis used 48 months.
OUTCOMES	
Outcome and measure(s) 1	Outcome – QALYs (Measurement of effectiveness) Measures – Health related quality of life measure using EQ-5D-3L tool.
Outcome and measure(s) 2	Outcome – Healthcare resource use and cost (Measurement of costs) Measures – The local co-ordinators at each of the study sites uses an electronic database to record the main uses of NHS healthcare resources by trial patients throughout the study period.
Outcome and measure(s) 3	
Outcome and measure(s) 4	
Outcome and measure(s) 5	
Outcome and measure(s) 6	
Outcome and measure(s) 7	
Outcome and measure(s) 8	
Outcome and measure(s) 9	
Outcome and measure(s) 10	
SUMMARY OF RESULTS	
Summary of findings (including statistical significance, significance level, confidence intervals and effect size)	EUS reduced net use of health-care resources by £2,860 (95% 'bootstrapped' CI from –£2200 to £8000). Combining the estimated benefits and savings shows that EUS is likely to be cost-effective, with 96.6% probability of achieving the National Institute for Health and Care Excellence (NICE) criterion of costing less than £20,000 to gain a QALY.

	EUS yielded an increase of 72 days in estimated mean quality-adjusted survival [i.e. 0.1969 QALYs (bootstrapped 95%CI -0.0640 to 0.4575)] – from 1.1647 QALYs (SD 0.9756) in the control group to 1.3616 QALYs (SD 0.9989) in the intervention group. The benefits of EUS were significantly greater for those with poor initial quality of life, but did not differ between centres. Those with poor initial quality of life had a much higher mean QALY gain (0.3067) than those with better initial quality of life (0.0183) but a smaller cost saving (£1918 vs £4259). Combining these effectiveness and cost findings by bootstrapping, EUS in participants with lower self-reported HRQoL at baseline has much higher probability of being cost-effective at NICE thresholds of £20,000 and £30,000 per QALY (97.1% and 97.9% respectively) than healthier participants (76.6% and 72.5% respectively).
Cost per QALY/Incremental Cost-Effectiveness Ratio (ICER) conclusions	More effect and less cost: cost per QALY not calculated. EUS reduced net use of health-care resources by £2,860 and had an increase of 0.1969 in estimated mean QALYs – Combining these estimated benefits and savings yields probability of 96.6% that EUS is cost-effective in the sense of achieving the NICE criterion of costing less than £20,000 to gain a QALY.
Inflated cost per QALY/ICER to 2016 price year	EUS saves £2,860 (2008 price year) – This equal to £3,364 (2017 price year) which £2,860 is inflated using Hospital and Community Health Services (HCHS) Pay and Prices Index (Curtis and Burns, 2017). **Inflation indices:**

	2007/08 – 257.0; 2015/16 – 297.0 (Curtis and Burns, 2016) 2015/16 – 297.0; 2016/17 – 302.3 (Curtis and Burns, 2017) **Reference:** Curtis L and Burns A. Unit costs of health and social care 2016. Personal Social Services Research Unit: University of Kent, 2016. Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2016/ (Accessed October 2017). Curtis, Lesley A. and Burns, Amanda. Unit Costs of Health and Social Care 2017. Personal Social Services Research Unit: University of Kent, 2017. Available from: https://doi.org/10.22024/UniKent/01.02/65559 (Accessed February 2018).
CONCLUSIONS	
Study conclusions	EUS achieved significant improvements in survival (121 days) and quality-adjusted survival (i.e. QALY) (66 days); a substantial, although non-significant, net saving of £2,860 per patient; and combining these quality-adjusted survival and economic findings, EUS is probably cost-effective (with 96.6% probability of being cost effective by NICE criteria (£20,000 to £30,000 threshold)).

Appendix 3.14: Data extraction results of Hadzijahic et al (2000)

DATA EXTRACTION FORM 2 – Economic Modelling	
Study ID (i.e. paper no.)	6
Reviewer	STY (Primary reviewer)
Checked by	NB (Second reviewer)
Study Title	CT or EUS for the initial staging of oesophageal cancer? A cost minimization analysis
Author(s)	Hadzijahic N, Wallace MB, Hawes RH, VanVelse A, LeVeen M, Marsi V, Hoffman BJ, Sahai AV
Publication year	2000
Aims/objectives/hypotheses	To determine whether it is less costly to request CT or EUS first to identify advanced oesophageal cancer; to determine which variables most affect the overall cost of identifying advanced disease.
METHOD	
Structure	
Statement of decision problem/objective of the evaluation and of the model	**Statement of decision problem** – Currently not known which of the initial staging strategy (EUS first or CT first) costs less to diagnose advance oesophageal cancer. **Objective of the evaluation and of the model** – Determine whether it is less costly to request CT or EUS first to identify advanced oesophageal cancer and to determine which variables most affect the overall cost of identifying advanced disease.
Statement of scope/perspective of the model	**Statement of scope** – Type of patients and 2 staging strategies were stated within the objectives. **Perspective of the model** – Not stated specifically, the study took local referral centre perspective.
Rationale for structure of the model	The model (as shown in Figure 1) was built to compare the direct facility costs of 2 staging strategies aimed at identifying advanced disease, CT first or EUS first. Rationale for structure of the decision model –

	Each strategy started after oesophageal cancer was diagnosed endoscopically. In the EUS first strategy, dilation might or might not be required and the probability of finding advanced disease could differ depending on whether dilation is required. If EUS showed advanced disease, no further testing was required; if not, CT was performed. Liver metastases were not sought during EUS. In the CT first strategy, EUS (with or without dilation) was required only if CT failed to demonstrate unresectable disease. The aim of this study was not to compare the accuracy of CT with that of EUS. Instead, it was to study the short-term outcomes that would result from decisions based on CT and EUS results when staging is performed to look for advanced disease. It was therefore assumed that CT and EUS results were correct.
Structural assumptions	The structural assumptions are transparent and justified – 1. It was assumed that CT and EUS results were correct. This is because the aim of the study was to study the short-term outcomes (yes or no advanced disease found) that would result from decisions based on CT and EUS results when staging is performed to look for advanced disease but not to compare the accuracy of CT with that of EUS. 2. Secondly, the model assumed that the latest small diameter echoendoscopes are used; therefore, the risk of perforation was not included in the decision tree.
Strategies/comparators	Two initial staging strategies to look for advanced oesophageal cancer were compared: CT first or EUS first.
Model type	Decision tree model
Time horizon	Not stated specifically.
Disease states/pathways	Decision tree pathways reflect the impact of finding out which of the two initial staging strategies (EUS first or CT first) would cost less to detect advanced disease in patients diagnosed endoscopically with oesophageal cancer.

Cycle length	Not applicable
Summary of structure section	Structural assumptions of the decision tree model are transparent and justified. The assumptions made seem to be appropriate and reflect the initial staging strategies for finding advanced disease in patients diagnosed endoscopically with oesophageal cancer. Without prior clinical knowledge of the area it is difficult to determine whether the assumptions made were adequate. However, the structure of the model seems to be an appropriate presentation of the initial staging strategies for finding advanced disease in patients with oesophageal cancer.
Data	
Data identification	The details of probabilities and costs identification methods are described as below – **Probabilities –** Initial probabilities of finding advanced disease and of requiring dilation were obtained from 124 consecutive patients with oesophageal cancer from their institution who underwent both CT and EUS (routine for both tests to be used at institution). Staging results, presenting symptoms, laboratory results, and endoscopic tumor characteristics were obtained retrospectively. **Costs –** Procedural costs for EUS and dilation included 1999 direct hospital costs for sedation, monitoring, nursing, nonreusable supplies and equipment, and physician fees were all presented in Table 1. EUS-FNA cost included the cost of diagnostic EUS plus the costs of a disposable FNA needle and a fee for cytologic interpretation. EUS (with or without dilation and/or FNA) and CT are both outpatient procedures. Similar to EUS and dilation costs, CT cost was also used 1999 direct hospital costs.
Data modelling	Statistical techniques used were specified – 1. Variables analysed for univariate associations were listed;

	2. Stepwise multiple logistic regression analysis was used to identify significant predictor variables of advanced disease. The decision model suggests that the variables that are most likely to affect the overall cost of identifying advanced disease are, in descending order of importance, (1) the relative costs of EUS and CT, (2) the probability of finding advanced disease by EUS (with or without dilation), and (3) the probability of finding advanced disease by CT (as shown in Figure 2). All other variables had little or no effect on overall costs (Figure 2).
Baseline data	The paper did not explicitly describe and justify the choice of baseline data for the decision tree. The paper mentioned about initial input data and all initial input key variables data were presented in Table 2. The initial model inputs were based on data from a consecutive series of patients with oesophageal cancer seen at the authors' institution. Because of the referral center population is likely not representative of all practice settings, the authors undertook sensitivity analysis and threshold analysis to provide results that can be applied in other settings. Therefore, I think the initial input data means baseline data in this paper.
Treatment effects	Not applicable. Oesophageal dilation was required to perform EUS in 44 (35%) cases, with no complications. Advanced disease was found more often by EUS (Table 3). In 4 (3%) cases, CT showed advanced disease (all T3 M1) that was not diagnosed by EUS. In 46 (37%) cases, EUS showed advanced disease that was not diagnosed by CT (27 T4 M0; 19 T1 M1, T2 M1, or T3 M1).
Costs, source for all costs and discount rates	The costs incorporated into the model are justified: EUS costs (with or without dilation and/or FNA) and CT costs; these costs were sourced from 1999 direct hospital costs i.e. local hospital/institutional costs. See table 2 for breakdown of costs.

	Discount rates are not specified.
Quality of life weights (utilities)	Not applicable – Utilities were not evaluated for the model of the study as it's a cost-minimisation analysis and not a cost-utility analysis, thus, incorporating utilities into the model is not appropriate in this study.
Data incorporation	Probabilities and costs data incorporated into the model have been described accordingly in detail and all are presented in Table 2 in the paper.
Assessment of uncertainty	Sensitivity analyses performed
Methodological uncertainty	Not performed
Structural uncertainties	Not performed
Heterogeneity	Heterogeneity has not been dealt with by running the model separately for different subgroups; Model ran for all patients with oesophageal cancer.
Parameter uncertainty	All key variables were addressed in sensitivity analyses under all possible ranges of value. One-way sensitivity analysis with threshold analysis was performed for all variables (Fig. 2). The ranges used for sensitivity analysis are stated clearly and presented in Table 2 in the paper. Initial estimates for costs and probabilities were varied between 50% and 200% of their original estimates (Table 2). If no thresholds were found with these ranges, they were increased to a maximum of 0% and 500%. Two- and three-way sensitivity analyses were performed as needed by using the most important determinants of overall cost. However, no justification as to why and how the ranges are chosen.
Summary of data section	All key data (Probabilities and costs data) incorporated into the decision tree model were described accordingly. And all data sources were described. Structural and methodological

	uncertainty was not addressed; however, parameter uncertainty were addressed by sensitivity analyses.
Consistency	
Internal consistency	Not specified
External consistency	Not addressed
Summary of consistency section	Both internal and external consistency was not addressed. No previous models were specified, and therefore no comparison of results with previous models was able to be made to validate the model; however the model results seem not counter-intuitive.
PARTICIPANTS	
Types of participants	Oesophageal cancer patients who underwent both CT and EUS.
Age range	Mean age was 62.7 years.
Inclusion criteria	A consecutive series of patients with oesophageal cancer seen at the authors' institution (a referral center).
Exclusion criteria	Not explicitly stated.
Ethnicity/country	USA, 58% white.
OUTCOMES	
Outcome and measure(s) 1	Cost, dependent on initial test
Outcome and measure(s) 2	Detection of advanced disease
Outcome and measure(s) 3	
Outcome and measure(s) 4	

Outcome and measure(s) 5	
Outcome and measure(s) 6	
Outcome and measure(s) 7	
Outcome and measure(s) 8	
Outcome and measure(s) 9	
Outcome and measure(s) 10	
SUMMARY OF RESULTS	
Summary of findings (including statistical significance, significance level, confidence intervals and effect size)	Initial CT is the least costly strategy if the probability of finding advanced disease by initial CT is greater than 20%, if the probability of finding advanced disease by initial endoscopic ultrasound (EUS) is less than 30%, or if the cost of EUS is greater than 3.5 times the cost of CT. However, in our referral center population, endosonography found advanced disease more frequently than CT (44% vs. 13%; $p < 0.0001$) and the least costly strategy was initial endosonography (expected cost $804 vs. $844).
Cost per QALY/Incremental Cost-Effectiveness Ratio (ICER) conclusions	Not applicable for cost-minimisation analysis study.

Inflated cost per QALY/ICER to 2015 price year (current price year)	Initial EUS strategy expected cost was USD804 (1999 price year) – This equals to £488 (1999 price year) which this is converted using the currency exchange rates (i.e. the currency units per SDR for September 1999) from International Monetary Fund (IMF, 2017). This £488 (1999 price year) is then inflated using Hospital and Community Health Services (HCHS) Pay and Prices Index (Curtis and Netten, 2005; Curtis and Burns, 2015; Curtis and Burns, 2016; Curtis and Burns, 2017) to £824 (2017 price year). Initial CT strategy expected cost was USD844 (1999 price year) – This equals to £513 (1999 price year) which this is converted using the currency exchange rates (i.e. the currency units per SDR for September 1999) from International Monetary Fund (IMF, 2017). This £513 (1999 price year) is then inflated using Hospital and Community Health Services (HCHS) Pay and Prices Index (Curtis and Netten, 2005; Curtis and Burns, 2015; Curtis and Burns, 2016; Curtis and Burns, 2017) to £867 (2017 price year). **Inflation indices:** 1998/99 – 180.4; 2004/05 – 234.2 (Curtis and Netten, 2005) 2004/05 – 232.3; 2005/06 – 240.9 (Curtis and Burns, 2015) 2005/06 – 240.9; 2015/16 – 297.0 (Curtis and Burns, 2016) 2015/16 – 297.0; 2016/17 – 302.3 (Curtis and Burns, 2017) **References:** International Monetary Fund (IMF). Currency unit per SDR for September 1999. Available from:

	http://www.imf.org/external/np/fin/data/rms_mth.aspx?SelectDate=1999-09-30&reportType=CVSDR (Accessed October 2017). Curtis L and Netten A. Unit costs of health and social care 2005. Personal Social Services Research Unit: University of Kent, 2005. Available from: http://www.pssru.ac.uk/pdf/uc/uc2005/uc2005.pdf (Accessed October 2017). Curtis L and Burns A. Unit costs of health and social care 2015. Personal Social Services Research Unit: University of Kent, 2015. Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2015/ (Accessed October 2017). Curtis L and Burns A. Unit costs of health and social care 2016. Personal Social Services Research Unit: University of Kent, 2016. Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2016/ (Accessed October 2017). Curtis, Lesley A. and Burns, Amanda. Unit Costs of Health and Social Care 2017. Personal Social Services Research Unit: University of Kent, 2017. Available from: https://doi.org/10.22024/UniKent/01.02/65559 (Accessed February 2018).
CONCLUSIONS	
Study conclusions	CT remains as the initial staging test of choice in most clinical settings. However, in referral centers, initial EUS may be reasonable, but individualized model inputs must be obtained before reliable conclusions can be drawn.

Appendix 3.15: Data extraction results of Harewood and Wiersema (2002)

DATA EXTRACTION FORM 2 – Economic Modelling	
Study ID (i.e. paper no.)	8
Reviewer	STY (Primary reviewer)
Checked by	NB (Second reviewer)
Study Title	A cost analysis of endoscopic ultrasound in the evaluation of oesophageal cancer
Author(s)	Harewood GC and Wiersema MJ
Publication year	2002
Aims/objectives/hypotheses	Decision analysis to examine which staging/management technique was the least costly: EUS FNA, CT-guided FNA or surgical management of oesophageal tumours. A cost-minimisation approach was employed and undertaken from the third party payer perspective. Sensitivity analysis performed to examine critical factors defining relative costs.
METHOD	
Structure	
Statement of decision problem/objective of the evaluation and of the model	**Statement of decision problem** – The economic impact of endoscopic ultrasound (EUS) with guided fine needle aspiration (FNA) staging strategy of suspicious lymph nodes has not yet been described. **Objective of the evaluation and of the model** – To quantitate the relative financial value of the addition of EUS FNA of CLNs to the preoperative evaluation of patients with apparently "resectable" oesophageal cancer on CT. A cost-minimisation approach was employed using a decision tree model.
Statement of scope/perspective of the model	**Statement of scope** – Type of patients, 2 addition staging options (CT-guided FNA and EUS-FNA) and one 'proceed straight to surgery' options were stated within the objectives of the study – Management of oesophageal tumours. **Perspective of the model** – Third party payer perspective.

Rationale for structure of the model	Rationale for structure of the decision model is to compare the direct medical costs of 3 management options: a) Obtain CT-guided biopsy of any suspicious-appearing CLNs detected on CT, b) Obtain EUS with FNA of any suspicious-appearing CLNs, or c) Proceed directly to surgery. By attaching costs to each management strategy of the model, the total costs and outcomes associated with particular health care strategies could be modelled. Success rates and potential risks of each procedure are included in the model.
Structural assumptions	The structural assumptions are transparent and justified – At the initial decision node patient is assumed to have undergone CT of abdomen and thorax, deemed to have resectable tumour. Three main branches of decision tree represent management options: CT guided FNA, EUS-FNA, direct to surgery. The detection of metastatic CLNs on FNA implied unresectability and prompted palliative management rather than surgery. 3. It was assumed that patients referred for either EUS-FNA or CT-guided FNA are assumed to have the same comorbidities, and all other diagnostic and additional clinical decision making is assumed to be similar. 4. Pathology interpretation costs are not included because these are also assumed similar in all arms. 5. The cost of combined chemotherapy is not taken into account as the proportion of early and advanced tumours in each management arm is similar. 6. It is assumed that enlarged CLNs detected on CT will be visualised by EUS because its sensitivity is significantly better than that of CT for detection of nodal metastases.

	7. The positive predictive value of a cytological finding of malignancy is assumed to be 100% for both EUS-FNA and CT-guided FNA. 8. The morbidity and long term side effects of radiation or chemotherapy were not included in the primary analysis as these remain poorly quantified. 9. The complication rates of EUS-FNA, CT-guided FNA, and exploratory laparotomy were considered to be small, having negligible effects on model results and were excluded to simply the model.
Strategies/comparators	Three management options of oesophageal tumours - a) CT-guided FNA – Obtain CT-guided biopsy of any suspicious-appearing CLNs detected on CT, b) EUS-FNA – Obtain EUS with FNA of any suspicious-appearing CLNs, or c) Surgery – Proceed directly to surgery.
Model type	Decision tree model
Time horizon	Not stated specifically.
Disease states/pathways	Decision tree pathways (see Figure 1) reflect the impact of finding out which of the three management options is least costly strategy in the management of patients with apparently "resectable" oesophageal cancer on CT.
Cycle length	Not applicable
Summary of structure section	Structural assumptions of the decision tree model are transparent and justified. The assumptions made seem to be appropriate and reflect the management options in patients with apparently "resectable" oesophageal cancer on CT. The structure of the model seems to be an appropriate presentation of the management options for patients with apparently "resectable" oesophageal cancer on CT.

	A decision analytic model was created, using DATA 3.5 software, to simulate the clinical problem and the costs assigned to each screening strategy. The three management strategies modelled were CT FNA, EUS FNA and immediate surgery. Palliative management was considered to be combined chemotherapy and radiotherapy with endoscopic insertion of a metallic oesophageal stent. False-negative patients with FNA, and also patients with previously undetermined lymph node metastases, would undergo exploratory laparotomy and would then proceed to palliative oesophageal stenting with chemo-radiation.
Data	
Data identification	The probabilities and costs data identification methods are transparent and appropriate given the objectives of the model. The details of probabilities and costs identification methods are described as below – **Probabilities –** The probabilities were obtained from an extensive literature search. The probabilities for EUS FNA and CT-guided FNA sensitivity were obtained from studies that had similar patient characteristics (i.e. patients with nonmetastatic oesophageal cancer). **Costs –** Direct medical costs were estimated from the Medicare ambulatory patient classification (APC) plus professional fees for hospital-based outpatient procedures and are outlined in Table 2.
Data modelling	• The authors did not justify the data modelling methodology whether it was based on justifiable statistical and epidemiological techniques. The authors only mentioned that they used the DATA 3.5 (TreeAge Software) software package to model their decision analysis. • Alternative assumptions have been explored through sensitivity analysis – The authors described that baseline

	probabilities were varied through plausible ranges using sensitivity analysis. • Probabilities and costs data incorporated into the decision tree model have been described accordingly in detail and all are presented in Table 1 and Table 2 in the paper. • Assumptions were described and justified.
Baseline data	The authors justified and described the choice of baseline data – Baseline probability of all the variables used in the decision tree model was described and presented in Table 1. Amongst the variables presented in Table 1, only the variable 'EUS-FNA sensitivity' data is justified.
Treatment effects	Not applicable
Costs, source for all costs and discount rates	• The costs incorporated into the model are justified and presented in Table 2. • The source for all costs has been described – Outpatient costs – Direct medical costs were estimated from the Medicare ambulatory patient classification (APC) plus professional fees for hospital-based outpatient procedures and are presented in Table 2. Inpatient costs – Diagnosis Related Groups (DRGs) costs were used for hospital inpatient services cost plus professional fees and are presented in Table 2. The cost areas included were those relating to the staging procedures, oesophageal stent and associated hospitalisation, laparotomy, oesophagectomy and postsurgical care. The costs of pathology interpretation and palliative care were not included, as they were considered similar in each management arm. All the costs related to the procedures were obtained from Medicare reimbursement rates. The Medicare ambulatory patient classification plus

	professional fees for hospital-based outpatient procedures were used. It appears that the costs have been estimated using actual data, and the average costs were reported. The unit costs and the quantities of resources used were not presented separately. • Discounting was not performed as the evaluation takes only several days.
Quality of life weights (utilities)	• Not applicable – Utilities were not evaluated for the model of the study as it's a cost-minimisation analysis and not a cost-utility analysis, thus, incorporating utilities into the model is not appropriate in this study.
Data incorporation	Probabilities and costs data incorporated into the model are all presented in Table 1 and 2 in the paper.
Assessment of uncertainty	One, two and three-way sensitivity analysis. Altering of parameters in sensitivity analysis was undertaken, with single and multiple adjustments made. Only done for probabilities, not costs.
Methodological uncertainty	The authors did not run alternative versions of the model with different methodological assumptions.
Structural uncertainties	Not specified
Heterogeneity	Heterogeneity has not been dealt with by running the model separately for different subgroups; Model ran for all patients with apparently "resectable" oesophageal cancer on CT.
Parameter uncertainty	The methods of assessment of parameter uncertainty are appropriate, the following sensitivity analyses were performed – 1. One-way sensitivity analysis varying prevalence of malignant CLNs, 2. One-way sensitivity analysis varying sensitivity of EUS-FNA, and

	3. Two-way sensitivity analysis varying prevalence of malignant CLNs and sensitivity of EUS-FNA simultaneously. All key variables were addressed in sensitivity analyses under all possible ranges of value. The ranges used for sensitivity analysis are stated in Table 1 in the paper. No justification as to why and how the ranges are chosen with the exception of variable "EUS-FNA sensitivity".
Summary of data section	Key data (Probabilities and costs data of variable "EUS-FNA sensitivity") incorporated into the decision tree model were described accordingly. And the appropriate data sources were described. Structural uncertainty was not addressed; however, methodological and parameter uncertainty were addressed by sensitivity analyses.
Consistency	
Internal consistency	Not specified
External consistency	Not addressed
Summary of consistency section	Both internal and external consistency was not addressed. No previous models were specified, and therefore no comparison of results with previous models was able to be made to validate the model; however the model results seem not counter-intuitive.
PARTICIPANTS	
Types of participants	Patients with apparently "resectable" oesophageal cancer on CT (i.e. patients with non-metastatic oesophageal cancer).
Age range	No age range specified
Inclusion criteria	Patients with verified cancer of the esophagus that is apparently "resectable" according to the T and M stages as determined by CT -- that is, no evidence of invasion of adjacent organs or distant metastases (i.e. non-metastatic oesophageal cancer).

Exclusion criteria	With regard to costs, indirect health and institutional costs such as the cost to society for lost work, quality of life, and institutional administration or maintenance of buildings were not included.
Ethnicity/country	Ethnicity – Not specified; Country – USA
OUTCOMES	
Outcome and measure(s) 1	Cost – The Relative cost of the addition of EUS FNA of CLNs to the preoperative evaluation of patients with apparently "resectable" oesophageal cancer on CT.
Outcome and measure(s) 2	Least costly management option
Outcome and measure(s) 3	Number of unnecessary surgeries avoided.
Outcome and measure(s) 4	
Outcome and measure(s) 5	
Outcome and measure(s) 6	
Outcome and measure(s) 7	
Outcome and measure(s) 8	
Outcome and measure(s) 9	
Outcome and measure(s) 10	
SUMMARY OF RESULTS	
Summary of findings (including statistical significance,	The total costs were $13,811 for EUS FNA, $14,350 for CT-guided FNA, and $13,992 for surgery only.

significance level, confidence intervals and effect size)	EUS FNA remained the least costly option, provided that the prevalence of CLN involvement was greater than 16%. Below this value, surgery became the least costly strategy. The final outcome of the model was also sensitive to variation in the sensitivity of EUS FNA. Provided that the sensitivity of EUS-FNA was greater than 66%, EUS-FNA remained the least costly staging option in the management of oesophageal tumours. Despite changing the values of two or three variables simultaneously in the two- and three-way sensitivity analyses, the result still showed that EUS-FNA remained the least costly strategy.
Cost per QALY/Incremental Cost-Effectiveness Ratio (ICER) conclusions	Not applicable for cost-minimisation analysis.
Inflated cost per QALY/ICER to 2015 price year (current price year)	EUS-FNA strategy costs $13,811 (2001 price year) – This equals to £9,394 (2001 price year) which this is converted using the currency exchange rates (i.e. the currency units per SDR for September 2001) from International Monetary Fund (IMF, 2017). This £9,394 (2001 price year) is then inflated using Hospital and Community Health Services (HCHS) Pay and Prices Index (Curtis and Netten 2005; Curtis and Burns, 2015; Curtis and Burns, 2016; Curtis and Burns, 2017) to £14,578 (2017 price year). CT-FNA strategy costs $14,350 (2001 price year) – This equals to £9,761 (2001 price year) which this is converted using the currency exchange rates (i.e. the currency units per SDR for September 2001) from International Monetary Fund (IMF, 2017). This £9,761 (2001 price year) is then inflated using Hospital and Community Health Services (HCHS) Pay and Prices Index (Curtis

	and Netten 2005; Curtis and Burns, 2015; Curtis and Burns, 2016) to £15,147 (2017 price year). Surgery costs $13,992 (2001 price year) – This equals to £9,517 (2001 price year) which this is converted using the currency exchange rates (i.e. the currency units per SDR for September 2001) from International Monetary Fund (IMF, 2017). This £9,517 (2001 price year) is then inflated using Hospital and Community Health Services (HCHS) Pay and Prices Index (Curtis and Netten, 2005; Curtis and Burns, 2015; Curtis and Burns, 2016) to £14,768 (2017 price year). **Inflation indices:** 2000/01 – 196.4; 2004/05 – 234.2 (Curtis and Netten, 2005) 2004/05 – 232.3; 2005/06 – 240.9 (Curtis and Burns, 2015) 2005/06 – 240.9; 2015/16 – 297.0 (Curtis and Burns, 2016) 2015/16 – 297.0; 2016/17 – 302.3 (Curtis and Burns, 2017) **References:** International Monetary Fund (IMF). Currency unit per SDR for September 2001. Available from: https://www.imf.org/external/np/fin/data/rms_mth.aspx?SelectDate=2001-09-30&reportType=CVSDR (Accessed October 2017). Curtis L and Netten A. Unit costs of health and social care 2005. Personal Social Services Research Unit: University of Kent, 2005. Available from: http://www.pssru.ac.uk/pdf/uc/uc2005/uc2005.pdf (Accessed October 2017).

	Curtis L and Burns A. Unit costs of health and social care 2015. Personal Social Services Research Unit: University of Kent, 2015. Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2015/ (Accessed October 2017). Curtis L and Burns A. Unit costs of health and social care 2016. Personal Social Services Research Unit: University of Kent, 2016. Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2016/ (Accessed October 2017). Curtis, Lesley A. and Burns, Amanda. Unit Costs of Health and Social Care 2017. Personal Social Services Research Unit: University of Kent, 2017. Available from: https://doi.org/10.22024/UniKent/01.02/65559 (Accessed February 2018).
CONCLUSIONS	
Study conclusions	By minimizing unnecessary surgery, primarily by detecting CLN involvement, EUS FNA was the least costly staging strategy in the workup of patients with non-metastatic oesophageal cancer. Under certain circumstances, surgery was the least costly strategy.

Appendix 3.16: Data extraction results of Wallace et al (2002)

DATA EXTRACTION FORM 2 – Economic Modelling	
Study ID (i.e. paper no.)	15
Reviewer	STY (Primary reviewer)
Checked by	NB (Second reviewer)
Study Title	An analysis of multiple staging management strategies for carcinoma of the esophagus: Computed tomography, endoscopic ultrasound, positron emission tomography, and thoracoscopy/laparoscopy
Author(s)	Michael B. Wallace, Paul J. Nietert, Craig Earle, Mark J. Krasna, Robert H. Hawes, Brenda J. Hoffman and Carolyn E. Reed
Publication year	2002
Aims/objectives/hypotheses	To compare the health care costs and effectiveness of multiple staging options for patients with oesophageal cancer.
METHOD	
Structure	
Statement of decision problem/objective of the evaluation and of the model	**Statement of decision problem** – Currently not known which of the available procedures (CT, EUS-FNA, TL and PET) or which combinations of these procedures offers the most cost-effective approach for the detection of locally advanced or metastatic disease for patients with oesophageal cancer. **Objective of the evaluation and of the model** – To compare the costs and effectiveness of six strategies for staging patients with oesophageal cancer: CT alone, CT+EUS-fine-needle aspiration biopsy (FNA), CT+thoracoscopy and laparoscopy (TL), CT+EUS-FNA+TL, CT+PET+EUS-FNA, and PET+EUS-FNA.
Statement of scope/perspective of the model	**Statement of scope** – Type of patients and staging strategies options were stated within the objectives. **Perspective of the model** – Third-party payer perspective
Rationale for structure of the model	The structure of the decision model reflects the clinical algorithm for the management of oesophageal cancer patients,

	which the staging algorithm was specified according to the protocol.
Structural assumptions	Assumption was made that staging tests were performed sequentially in each strategy case, along with justification according to the protocol for no further staging was performed if distant metastases were found and confirmed. Their model assumed that patients with known metastatic disease (M1b) or tumours that had invaded regional organs (T4) did not undergo surgical esophagectomy but only palliative therapy including stents or radiotherapy and chemotherapy.
Strategies/comparators	Six strategies for staging patients with oesophageal cancer were compared: CT alone, CT+EUS-fine-needle aspiration biopsy (FNA), CT+thoracoscopy and laparoscopy (TL), CT+EUS-FNA+TL, CT+PET+EUS-FNA, and PET+EUS-FNA.
Model type	Decision tree model
Time horizon	Not stated specifically
Disease states/pathways	Decision tree represents the staging algorithm that patients may follow when they are diagnosed with oesophageal cancer (see Figure 1).
Cycle length	Not applicable
Summary of structure section	Assumptions made on the structure of the decision tree model were clearly specified. The assumptions made seem to be appropriate and reflect the clinical algorithm for management of patients with oesophageal cancer. Without prior knowledge of the area it is difficult to determine whether the assumptions made were adequate. However, the structure of the model seems to be an appropriate presentation of the staging strategies for esophagus cancer.
Data	
Data identification	Test characteristics for the different staging procedures were derived from published literature (but didn't specify whether

	systematic) and from a prospective study. Life expectancies were derived from 1973-1996 SEER data and separate life expectancies were derived from median survival times for patients in various stages at diagnosis as well as for patients who had and who had not been treated surgically for the disease. Costs associated with the different staging techniques and inpatient procedure for the diagnostically related group for TL with lymph node biopsy linked to the ICD-9 code for malignant oesophageal neoplasm were based on Medicare reimbursement. Outpatient procedure for EUS-FNA costs were based on the CPT (Current Procedural Terminology) code for EUS with FNA linked to the ICD-9 code for malignant oesophageal neoplasm. See Table 1 for list of data included.
Data modelling	Standard methods of cost-effectiveness analysis were used for the decision analysis.
Baseline data	All baseline values for the model variables came from several sources, as shown in Table 1 in the paper.
Treatment effects	Not applicable
Costs, source for all costs and discount rates	Costs associated with management of patients with esophagus cancer were incorporated into the decision model (i.e. costs associated with the different staging techniques, and costs associated with the treatment and care of patients identified as having local, regional, and distant disease). All costs associated with management of patients with esophagus cancer were derived from SEER-Medicare linked databases and from other Medicare reimbursement rates. Cost and effectiveness measures were discounted at 0% and 3% per year.

Quality of life weights (utilities)	Quality of life utilities values come directly from expert opinion, estimated for local/regional disease unresected, local/regional disease resected, distant disease unresected, distant disease resected. No details about the source and methods of derivation for the utilities weights
Data incorporation	All data incorporated into the model have been described and referenced accordingly and all are presented in Table 1 in the paper. However, expert opinion not referenced as to who and number of experts involved.
Assessment of uncertainty	Ranges used in Sensitivity Analysis. Each model variable was subjected to sensitivity analysis to determine which variables had the greatest effects on the cost-effectiveness measures for each strategy.
Methodological uncertainty	Not assessed
Structural uncertainties	Not assessed
Heterogeneity	Not assessed – Heterogeneity has not been dealt with by running the model separately for different subgroups; Model ran for patients of all ages.
Parameter uncertainty	The methods of assessment of parameter uncertainty are appropriate – All key variables were addressed in sensitivity analyses under all possible ranges of value.
Summary of data section	Wide range of data used to populate model, although QALY data is not robust. All data used in the decision tree model were detailed and referenced appropriately. And all data sources were stated and presented clearly in Table. Methodological and structural uncertainties were not addressed; however, parameter uncertainty was addressed in the sensitivity analyses.
Consistency	
Internal consistency	Not specified
External consistency	Not addressed

Summary of consistency section	Both internal and external consistency was not addressed. No previous models were specified, and therefore no comparison of results with previous models was able to be made to validate the model; however the model results seem not counter-intuitive.
PARTICIPANTS	
Types of participants	Adults (younger patients and patients older than 65 years)
Age range	No age range specified
Inclusion criteria	Using the SEER–Medicare databases, a retrospective cohort was created consisting of all Medicare-eligible patients whose invasive oesophageal cancer was diagnosed between January 1, 1991, and December 31, 1996.
Exclusion criteria	Patients were excluded if they had had any prior cancer or if they were enrolled in a health maintenance organization at some time during the study period and therefore did not have complete treatment information. Patients were also excluded if the dates of diagnosis (or death) differed by more than 2 months in the SEER and Medicare databases, or if the cancer was first identified at the time of death or post-mortem examination.
Ethnicity/country	Ethnicity – Not specified; Country studied – USA.
OUTCOMES	
Outcome and measure(s) 1	Costs
Outcome and measure(s) 2	QALY
Outcome and measure(s) 3	
Outcome and measure(s) 4	

Outcome and measure(s) 5	
Outcome and measure(s) 6	
Outcome and measure(s) 7	
Outcome and measure(s) 8	
Outcome and measure(s) 9	
Outcome and measure(s) 10	
SUMMARY OF RESULTS	
Summary of findings (including statistical significance, significance level, confidence intervals and effect size)	Under baseline assumptions, CT+EUS-FNA was the least costly strategy and offered more quality-adjusted life years, on average, than all other strategies with the exception of PET+EUS-FNA. The latter was slightly more effective but also more costly.
Cost per QALY/Incremental Cost-Effectiveness Ratio (ICER) conclusions	The marginal cost-effectiveness ratio for PET+EUS-FNA was $60,544 per quality-adjusted life-year (QALY). These findings were robust and changed very little in all of the sensitivity analyses.
Inflated cost per QALY/ICER to 2015 price year (current price year)	USD$60,544 per QALY (2000 price year) – This equals to £41,184 per QALY (2000 price year) which this is converted using the currency exchange rates (i.e. the currency units per SDR for September 2000) from International Monetary Fund (IMF, 2017). This £41,184 per QALY (2000 price year) is then inflated using Hospital and Community Health Services (HCHS) Pay and Prices Index (Curtis and Netten, 2005; Curtis and Burns, 2015; Curtis and Burns, 2016) to £66,588 per QALY (2017 price year).

	Inflation indices: 1999/00 – 188.5; 2004/05 – 234.2 (Curtis and Netten, 2005) 2004/05 – 232.3; 2005/06 – 240.9 (Curtis and Burns, 2015) 2005/06 – 240.9; 2015/16 – 297.0 (Curtis and Burns, 2016) 2015/16 – 297.0; 2016/17 – 302.3 (Curtis and Burns, 2017) **References:** International Monetary Fund (IMF). Currency unit per SDR for September 2000. Available from: http://www.imf.org/external/np/fin/data/rms_mth.aspx?SelectDate=2000-09-30&reportType=CVSDR (Accessed October 2017). Curtis L and Netten A. Unit costs of health and social care 2005. Personal Social Services Research Unit: University of Kent, 2005. Available from: http://www.pssru.ac.uk/pdf/uc/uc2005/uc2005.pdf (Accessed October 2017). Curtis L and Burns A. Unit costs of health and social care 2015. Personal Social Services Research Unit: University of Kent, 2015. Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2015/ (Accessed October 2017). Curtis L and Burns A. Unit costs of health and social care 2016. Personal Social Services Research Unit: University of Kent, 2016. Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2016/ (Accessed October 2017). Curtis, Lesley A. and Burns, Amanda. Unit Costs of Health and Social Care 2017. Personal Social Services Research Unit:

	University of Kent, 2017. Available from: https://doi.org/10.22024/UniKent/01.02/65559 (Accessed February 2018).
CONCLUSIONS	
Study conclusions	The combination of PET-EUS-FNA should be the recommended staging procedure for patients with oesophageal cancer, unless resources are scarce or PET is unavailable. In these instances, CT-EUS-FNA can be considered the preferred strategy.

Appendix 3.17: Quality assessment results of Shumaker et al (2002)

Question no.	CASP economic evaluation checklist questions*¥	Response (√, x or NC/NA)	Comments (i.e. a description that explains how the judgement was reached)
1	Was a well-defined question posed?	√	The aim of the study was to determine the relative proportions of each oesophageal cancer stage in a group of patients referred for preoperative staging with EUS.
2	Was a comprehensive description of the competing alternatives given?	NA	Not applicable as this was a cost analysis study to assess the potential cost savings of performing preoperative EUS.
3	Does the paper provide evidence that the programme would be effective (i.e. would the programme do more good than harm)?	√	The authors provided evidences from the literatures that (1) EUS has been shown to be a minimally invasive, safe, and accurate method of determining the preoperative stage of oesophageal cancer, and (2) studies have also demonstrated a potential impact of EUS staging on patient management.
4	Were the effects of the intervention identified, measured and valued appropriately?	NA	Not applicable as this was a cost analysis study.

5a	Were all important and relevant resources required and health outcome costs for each alternative identified?	NC	EUS cost (£634 per patient) and the costs of combined modality therapy [neoadjuvant chemoradiotherapy (£8728 per patient) plus surgery (£23,800 per patient)] were identified. The cost analysis took a Medicare perspective. Price year 2000 was stated. Most costs derived from Medicare reimbursement rates. The authors did not include costs for consulting physicians or treatment complication. Estimated esophagectomy from previous literature rather than Medicare.
5b	Were all important and relevant resources required and health outcome costs for each alternative measured in appropriate units?	√	All costs were measured in US dollar ($). Potential cost savings of performing preoperative EUS was defined as the total cost of the treatment saved (cost of chemoradiotherapy multiplied by the number of patients not receiving neoadjuvant chemoradiotherapy plus esophagectomy cost multiplied by the number of patients with Stage IV cancer not treated with esophagectomy) minus the cost of the EUS examinations.

5c	Were all important and relevant resources required and health outcome costs for each alternative valued credibly?	√	The estimated cost of EUS, chemoradiotherapy and esophagectomy are realistic. The cost of EUS ($634 per patient, price year 2000) was based on Medicare reimbursement rates at Oregon Health and Science University for the year 2000. The cost of neoadjuvant chemoradiotherapy ($8728 per patient, price year 2000) was based on the protocol of a study by Walsh et al (1996). The estimated cost of an esophagectomy for oesophageal cancer ($23,800 per patient, price year 1999) was derived from a previously published estimate by Provenzale et al (1999).
6	Were costs and consequences adjusted for different times at which they occurred (discounting)?	x	The authors did not inflate the cost of esophagectomy from 1999 price year to 2000 price year in order to be in line with the price year for the cost of EUS and chemoradiotherapy (which both these costs are 2000 price year). Discounting was not stated.

7	What were the results of the evaluation?	√	For every 100 patients staged before surgery with EUS (cost $63,420), 14 patients with Stage I disease would be spared neoadjuvant chemoradiotherapy (saving $122,192) and 12 patients with Stage IV cancer would be spared surgery (saving $285,600) for an average cost savings of $3443 per patient.
8	Was an incremental analysis of the consequences and cost of alternatives performed?	NA	Cost saving was performed.
9	Was an adequate sensitivity analysis performed?	√	Sensitivity analysis for potential cost savings with preoperative EUS in oesophageal cancer was performed to account for variability in the cost of EUS and combined modality therapy and any effect of varying the proportion of patients whose management would be altered by preoperative staging with EUS.
10	Is the programme likely to be equally effective in your context or setting?	√	

11	Are the costs translatable to your setting?	x	No as it is a US study that took a Medicare perspective. Thus, the costs are not translatable to UK setting.
12	Is it worth doing in your setting?	√	It is worth doing in UK local hospital setting to assess the similar context (potential cost savings with preoperative EUS in oesophageal cancer) with the similar population.

*Available from: http://www.casp-uk.net/casp-tools-checklists

¥Adapted from: Drummond MF, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press, 1987

Appendix 3.18: Quality assessment results of Chang et al (2003)

Question no.	CASP economic evaluation checklist questions*¥	Response (√, x or NC/NA)	Comments (i.e. a description that explains how the judgement was reached)
1	Was a well-defined question posed?	√	The aim of the study was to determine the impact of EUS in combination with FNA on patients' choice of therapy and on the cost of care.
2	Was a comprehensive description of the competing alternatives given?	NA	Not applicable as this was a cost analysis study to assess whether or not the use of EUS decreased the cost of care for patients who had oesophageal cancer who otherwise were surgical candidates.
3	Does the paper provide evidence that the programme would be effective (i.e. would the programme do more good than harm)?	√	The authors provided evidences from the literatures that (1) Based on a review of 739 reported cases, EUS is accurate for evaluating tumour depth in 85% of cases and nodal stage in 79%, (2) recent development of techniques for performing fine-needle aspiration (FNA) biopsy through the echo endoscope has further improved the specificity of EUS in diagnosing lymph-node metastasis, (3) EUS-guided FNA also provides pathological evidence of metastasis to peritoneal or pleural fluid, or liver.

4	Were the effects of the intervention identified, measured and valued appropriately?	NA	Not applicable as this was a cost analysis study.
5a	Were all important and relevant resources required and health outcome costs for each alternative identified?	NC	It appears that all relevant clinical costs were included, the authors stated that cost analyses were based on the published direct costs of endosonography-guided aspiration biopsy and thoracotomy, but the description of the cost analyses are brief, potentially limited. Price year was not stated.
5b	Were all important and relevant resources required and health outcome costs for each alternative measured in appropriate units?	√	All costs were measured in US dollar ($). Cost analyses were carried out in this cohort of patients to calculate the savings that occurred when patients chose not to undergo thoracotomy based on the EUS findings.
5c	Were all important and relevant resources required and health outcome costs for each alternative valued credibly?	NC	Depends on quality of reference used to cost. Cost analyses not described in sufficient detail.

6	Were costs and consequences adjusted for different times at which they occurred (discounting)?	x	The authors did not mention whether they inflated the costs of endosonography-guided aspiration biopsy and thoracotomy. And they did not mention about discounting.
7	What were the results of the evaluation?	√	The use of EUS in these patients potentially saved $740,424 (as shown in Table 1). The average cost savings per patient referred for pre-operative EUS staging were approximately $12,340.
8	Was an incremental analysis of the consequences and cost of alternatives performed?	NA	Cost saving was performed.
9	Was an adequate sensitivity analysis performed?	x	No sensitivity analysis was performed.
10	Is the programme likely to be equally effective in your context or setting?	√	

11	Are the costs translatable to your setting?	x	It is a US study so the costs are not translatable to UK setting.
12	Is it worth doing in your setting?	√	It is worth doing in UK local hospital setting to assess the similar context (potential cost savings when oesophageal cancer patients chose not to undergo thoracotomy based on the EUS-guided therapy findings) with the similar population.

*Available from: http://www.casp-uk.net/casp-tools-checklists

¥Adapted from: Drummond MF, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press, 1987

Appendix 3.19: Quality assessment results of Russell et al (2013)

Question no.	CASP economic evaluation checklist questions*¥	Response (√, x or NC)	Comments (i.e. a description that explains how the judgement was reached)
1	Was a well-defined question posed?	√	The objective of the health economic analysis was to assess the cost-effectiveness of EUS staging in the management of patients with gastro-oesophageal cancer compared with its absence.
2	Was a comprehensive description of the competing alternatives given?	√	Eligible and medically fit patients with localised tumours randomised to intervention group (EUS group) received EUS in addition to standard staging algorithms, while those randomised to control group (non-EUS group) received standard staging algorithms without EUS. **The staging algorithm from usual practice, as identified by SAGOC, were developed by the COGNATE team before the trial began-** a. Chest x-ray, pulmonary function tests, haematology and biochemistry, together with assessment of cardiac status. Patients of WHO performance status 3 or 4 or medically unsuitable for either surgery or chemotherapy will be excluded. b. Patients who are medically fit will undergo a trans-abdominal USS. Those found to have metastatic liver disease will be excluded. c. Patients without evidence of metastases will undergo a CT scan following an agreed protocol using a spiral scanner, oral water contrast and

			intravenous contrast. Laparoscopy will be undertaken in patients with any suspicion of peritoneal disease, as this remains the best means of detecting peritoneal deposits of tumour. d. Only patients with localised tumours will be randomised between EUS or not.
3	Does the paper provide evidence that the programme would be effective (i.e. would the programme do more good than harm)?	√	In the overall report, the findings of their literature review shows that EUS has potential to provide accurate staging of gastric and oesophageal tumours. They acknowledged that this can therefore provide important prognostic information to guide management; however, as the link between better staging and better management is not proven, the benefit of EUS is not clear. So the COGNATE trial was initiated and designed to evaluate, not the accuracy of EUS, but the effect it has on patient management and thus outcome. Also, the economic evidence shows that combination of EUS-FNA with CT staging strategy was the least costly strategy compared to all other strategies with the exception of combination of EUS-FNA with PET which was slightly more effective but also more expensive.

4	Were the effects of the intervention identified, measured and valued appropriately?	√	Yes the effects of the intervention were identified, measured and valued appropriately.
5a	Were all important and relevant resources required and health outcome costs for each alternative identified?	√	The health economic analysis was undertaken from an NHS perspective. Price year 2008 was stated. Resources – All secondary healthcare resource use and prescribing were identified. Health outcome – Patient's health-related quality of life was measured using EQ-5D tool.
5b	Were all important and relevant resources required and health outcome costs for each alternative measured in appropriate units?	√	Resources – All healthcare resource use and prescriptions were measured in appropriate unit e.g. number of bed days, number of cycles. Health outcome – Patient's health-related quality of life was measured in the common unit – QALY.
5c	Were all important and relevant resources required and health outcome costs for each alternative valued credibly?	√	All the values reported for healthcare resource use and costs, and health-related quality of life outcome are realistic and derived appropriately.

6	Were costs and consequences adjusted for different times at which they occurred (discounting)?	√	In the primary analysis at 48 months, costs and QALYs were discounted at 3.5% per year. All costs were inflated to 2008 prices, as appropriate, using the Hospital and Community Health Services Pay and Price Index obtained from Curtis (2008). <u>Reference:</u> Curtis L. Unit costs of health and social care 2008. Canterbury: PSSRU, University of Kent; 2008.
7	What were the results of the evaluation?	√	The results suggest that, 48 months after randomisation, EUS is still more effective and less costly than usual management.
8	Was an incremental analysis of the consequences and cost of alternatives performed?	√	Incremental analysis of the cost and outcome of alternatives was performed – The intervention group (EUS group) on average has gained 0.1969 more QALYs and cost £2,860 less than the control group (non-EUS group).
9	Was an adequate sensitivity analysis performed?	√	Sensitivity analyses were performed for the variability of EUS costs with discounted and undiscounted QALYs in the primary analysis at 48 months.
10	Is the programme likely to be equally effective in your context or setting?	√	

11	Are the costs translatable to your setting?	√	
12	Is it worth doing in your setting?	√	

*Available from: http://www.casp-uk.net/casp-tools-checklists

¥Adapted from: Drummond MF, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press, 1987

Appendix 3.20: Quality assessment results of Hadzijahic et al (2000)

Quality Criterion	Economic modelling checklist questions*	Response (√, x, NC or NA)	Comments (i.e. a description that explains how the judgement was reached)
S1	Is there a clear statement of the decision problem?	√	Studies found that CT is sensitive for distant metastases but less sensitive than EUS for T4 disease and celiac node involvement. The decision problem is not known whether initial CT or EUS costs less to diagnose advanced oesophageal cancer.
	Is the objective of the evaluation and model specified and consistent with the stated decision problem?	√	The objective is to perform a cost-minimization analysis by using a decision model to determine whether it is less costly to request CT or EUS first to identify advanced disease and to determine which variables most affect the overall cost of identifying advanced disease.
	Is the primary decision-maker specified?	NC	The primary decision-maker is assumed to be the local referral center but this was not explicitly specified.
S2	Is the perspective of the model stated clearly?	x	Not stated explicitly but it seems local referral center perspective.
	Are the model inputs consistent with the stated perspective?	NC	

	Has the scope of the model been stated and justified?	√	The scope of the decision tree model has been stated and justified – A decision model compared the costs of the two strategies. Each strategy (EUS first or CT first strategy) in the decision tree started once an oesophageal cancer was diagnosed endoscopically. EUS may or may not require dilation/ Initial CT or initial EUS may show advanced (T4 and/or M1) or not advanced disease. If either CT or EUS does not show advanced disease, the alternative test was performed.
	Are the outcomes of the model consistent with the perspective, scope and overall objective of the model?	√	The outcomes (1) which strategy could find advanced disease more frequently, and (2) which strategy is the least costly strategy reflects the scope and overall objective of the model. It is not clear whether the outcomes of the model consistent with the perspective, this is because it didn't specified the perspective. However, it seems local referral center perspective. If that's the case, the outcomes would seem reflecting the perspective of local referral center.
S3	Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	√	The structure of the model is consistent with the staging algorithm for finding advanced disease in patients with oesophageal cancer as classified by AJCC (American Joint Commission on Cancer).

	Are the sources of data used to develop the structure of the model specified?	√	The authors specified the sources of data – 1. Probabilities of finding advanced disease and of requiring dilation were obtained from 124 consecutive patients with oesophageal cancer from their institution who underwent both CT and EUS.] 2. Costs data (EUS (with or without dilation and/or FNA) and CT) used 1999 direct hospital costs i.e. local hospital/institutional costs.
	Are the causal relationships described by the model structure justified appropriately?	NA	-
S4	Are the structural assumptions transparent and justified?	√	The structural assumptions are transparent and justified – 1. It was assumed that CT and EUS results were correct. This is because the aim of the study was to study the short-term outcomes (yes or no advanced disease found) that would result from decisions based on CT and EUS results when staging is performed to look for advanced disease but not to compare the accuracy of CT with that of EUS. 2. Secondly, the model assumed that the latest small diameter echoendoscopes are used; therefore, the risk of perforation was not included in the decision tree.

	Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	√	Yes reasonable.
S5	Is there a clear definition of the options under evaluation?	√	The definition of both the EUS or CT first strategies was given. Each strategy started after oesophageal cancer was diagnosed endoscopically. 1. EUS first strategy – In the EUS first strategy, dilation might or might not be required and the probability of finding advanced disease could differ depending on whether dilation is required. If EUS showed advanced disease, no further testing was required; if not, CT was performed. Liver metastases were not sought during EUS. 2. CT first strategy – In the CT first strategy, EUS (with or without dilation) was required only if CT failed to demonstrate unresectable disease i.e. advanced disease.
	Have all feasible and practical options been evaluated?	√	It seems not all possible staging strategies been evaluated e.g. PET, laparoscopy but it's hard to say for this case. This is because this study aimed to assess which initial staging strategy is less costly for detecting

			advanced disease in oesophageal cancer patients. And therefore maybe due to this reason and for achieving the aim of the study, only CT and EUS was evaluated for initial staging strategy as to which one is less costly. With regard to costs, indirect costs for CT and EUS (e.g. cost of time off from work, hospital parking, etc.) were excluded from the model because they were considered comparable seeing that in clinical practice EUS is rarely performed immediately after the initial diagnostic endoscopy which if this happens this could reduce indirect costs associated with EUS.
	Is there justification for the exclusion of feasible options?	NA	-
S6	Is the chosen model type appropriate given the decision problem and specified causal relationships within the model?	√	Decision tree model demonstrates the staging sequences for each initial staging strategy (EUS first or CT first). The chosen model type (decision tree model) seems appropriate given the aim is to find out which of these two initial staging is least costly in detecting advanced disease in patients with oesophageal cancer. Causal relationships are not relevant.

S7	Is the time horizon of the model sufficient to reflect all important differences between options?	x	Though the authors explained the sequence of each staging option in finding advanced disease after oesophageal cancer was diagnosed endoscopically, it did not specify as to what is the time horizon of the model which it may span days or weeks or months.
	Are the time horizon of the model, the duration of treatment and the duration of treatment effect described and justified?	x	The time horizon of the decision tree model is not specified. Also, none specified for the duration of treatment and the duration of treatment effect as I think these are not applicable for the study's model. This is because the decision tree model was designed to reflect the initial staging strategy for detecting advanced disease in patients diagnosed endoscopically with oesophageal cancer.
S8	Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions?	√	Decision tree pathways reflect the impact of finding out which of the two initial staging strategies (EUS first or CT first) would cost less to detect advanced disease in patients diagnosed endoscopically with oesophageal cancer.

S9	Is the cycle length defined and justified in terms of the natural history of disease?	NA	-
D1	Are the data identification methods transparent and appropriate given the objectives of the model?	√	The probabilities and costs data identification methods are transparent and appropriate given the objectives of the model. The details of probabilities and costs identification methods are described as below – **Probabilities –** Initial probabilities of finding advanced disease and of requiring dilation were obtained from 124 consecutive patients with oesophageal cancer from their institution who underwent both CT and EUS. Staging results, presenting symptoms, laboratory results, and endoscopic tumor characteristics were obtained retrospectively. **Costs –** Procedural costs for EUS and dilation included 1999 direct hospital costs for sedation, monitoring, nursing, nonreusable supplies and equipment, and physician fees were all presented in Table 1. EUS-FNA cost included the cost of diagnostic EUS plus the costs of a disposable FNA needle and a fee for cytologic interpretation. EUS (with or without dilation and/or FNA) and CT are both outpatient procedures. Similar to EUS and dilation costs, CT cost was also used 1999 direct hospital costs.

Where choices have been made between data sources, are these justified appropriately?	NA	-
Has particular attention been paid to identifying data for the important parameters in the model?	√	Particular attention has been paid to identifying advanced disease data. Definitions for advanced disease were given according to the AJCC criteria; the detections of advanced disease by CT or EUS were clearly defined – **For CT,** this was defined as radiologic evidence of distant metastases to solid organs or celiac lymph nodes (M1) and/or invasion of adjacent solid organs (T4). Cytologic confirmation was not required for M1 disease detected by CT. **For EUS**, advanced disease was defined as cytologically positive celiac lymph nodes (M1) and/or invasion of adjacent solid organs (T4).
Has the quality of the data been assessed appropriately?	x	Not specified
Where expert opinion has been used, are the methods described and justified?	NA	-

D2	Is the data modelling methodology based on justifiable statistical and epidemiological techniques?	√	Yes, the data modelling methodology is based on justifiable statistical techniques. Statistical techniques used were specified – 1. Variables analysed for univariate associations were listed; 2. Stepwise multiple logistic regression analysis was used to identify significant predictor variables of advanced disease. The decision model suggests that the variables that are most likely to affect the overall cost of identifying advanced disease are, in descending order of importance, (1) the relative costs of EUS and CT, (2) the probability of finding advanced disease by EUS (with or without dilation), and (3) the probability of finding advanced disease by CT (as shown in Figure 2). All other variables had little or no effect on overall costs (Figure 2).
D2a	Is the choice of baseline data described and justified?	√	The paper did not explicitly describe and justify the choice of baseline data for the decision tree. The paper mentioned about initial input data and all initial input key variables data were presented in Table 2. The initial model inputs were based on data from a consecutive series of patients with oesophageal cancer seen at the authors' institution. Because of the referral center population is likely not representative of all practice settings, the authors undertook sensitivity analysis and threshold analysis to provide

			results that can be applied in other settings. Therefore, I think the initial input data means baseline data in this paper.
	Are transition probabilities calculated appropriately?	NA	
	Has a half-cycle correction been applied to both cost and outcome?	NA	
	If not, has this omission been justified?	NA	
D2b	If relative treatment effects have been derived from trial data, have they been synthesised using appropriate techniques?	NA	
	Have the methods and assumptions used to extrapolate short-term results to final outcomes been documented and justified?	NA	
	Have alternative assumptions been explored through sensitivity analysis?	√	Because the initial model inputs were based on data from patients seen at the author's institution and this is likely not representative of all practice

			settings, sensitivity analysis and threshold analysis were used to provide results that can be applied in other settings.
	Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?	NA	
	Have alternative assumptions regarding the continuing effect of treatment been explored through sensitivity analysis?	NA	
D2c	Are the costs incorporated into the model justified?	√	Yes, the costs incorporated into the model are justified: EUS costs (with or without dilation and/or FNA) and CT costs.
	Has the source for all costs been described?	√	These costs were sourced from 1999 direct hospital costs i.e. local hospital/institutional costs (as shown in Table 1).
	Have discount rates been described and justified given the target decision-maker?	NC	

D2d	Are the utilities incorporated into the model appropriate?	NA	Utilities were not evaluated for the model of the study as it's a cost-minimisation analysis and not a cost-utility analysis, thus, incorporating utilities into the model is not appropriate in this study.
	Is the source for the utility weights referenced?	NA	-
	Are the methods of derivation for the utility weights justified?	NA	-
D3	Have all data incorporated into the model been described and referenced in sufficient detail?	NC	Probabilities and costs data incorporated into the decision tree model have been described and presented in Table 2 in the paper but not explicitly cleared.
	Has the use of mutually inconsistent data been justified (i.e. are assumptions and choices appropriate)?	NC	
	Is the process of data incorporation transparent?	√	They explained that for the stepwise multiple logistic regression analysis procedure, the independent variables were added to the model one at a time. In the final model, variables were removed if the retention criterion of $p \leq 0.05$ was not met. No details as to whether data incorporated as distributions and/or as point estimates.

	If data have been incorporated as distributions, has the choice of distribution for each parameter been described and justified?	NA	
	If data have been incorporated as distributions, is it clear that second order uncertainty is reflected?	NA	
D4	Have the four principal types of uncertainty been addressed?	x	Not specified
	If not, has the omission of particular forms of uncertainty been justified?	x	Not specified
D4a	Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?	x	Not specified

D4b	Is there evidence that structural uncertainties have been addressed via sensitivity analysis?	x	Not specified.
D4c	Has heterogeneity been dealt with by running the model separately for different subgroups?	x	Model ran for all patients with oesophageal cancer.
D4d	Are the methods of assessment of parameter uncertainty appropriate?	√	Sensitivity analysis and threshold analysis were used to provide results that can be applied in other settings. All key variables were addressed in sensitivity analyses under all possible ranges of value.
	If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	NC	The ranges used for sensitivity analysis are stated clearly and presented in Table 2 in the paper. But, no justification as to why and how the ranges are chosen.
C1	Is there evidence that the mathematical logic of the model has been tested thoroughly before use?	x	Not specified.

C2	Are any counterintuitive results from the model explained and justified?	NA	The decision model results do not appear to be counterintuitive.
	If the model has been calibrated against independent data, have any differences been explained and justified?	NA	
	Have the results of the model been compared with those of previous models and any differences in results explained?	x	No previous models are specified; no comparison of results with previous models is made.

*****Available from: Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, Woolacott N and Glanville J. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004;**8**(36).

Appendix 3.21: Quality assessment results of Harewood and Wiersema (2002)

Quality Criterion	Economic modelling checklist questions*	Response (√, x or NA)	Comments (i.e. a description that explains how the judgement was reached)
S1	Is there a clear statement of the decision problem?	√	The economic impact of endoscopic ultrasound (EUS) with guided fine needle aspiration (FNA) staging strategy of suspicious lymph nodes has not yet been described.
	Is the objective of the evaluation and model specified and consistent with the stated decision problem?	√	The objective of the study is to compare the costs of EUS-FNA, CT-guided FNA, and surgery in the management of oesophageal tumours using a decision analysis model. A cost-minimisation approach was employed and undertaken from the third party payer perspective.
	Is the primary decision-maker specified?	√	Third party payer
S2	Is the perspective of the model stated clearly?	√	The model took the third party payer perspective
	Are the model inputs consistent with the stated perspective?	√	Yes, the model inputs are consistent with the stated perspective. The decision tree has three main branches represent the management options: a) obtain CT-guided biopsy of any suspicious-appearing CLNs detected on CT, b) obtain EUS with FNA of any suspicious-appearing CLNs, or c) proceed directly to surgery.

	Has the scope of the model been stated and justified?	√	The scope of the decision tree model has been stated and justified to quantitate the relative financial value of the addition of EUS-FNA of CLNs to the preoperative evaluation of a patient with apparently “resectable” oesophageal cancer on CT.
	Are the outcomes of the model consistent with the perspective, scope and overall objective of the model?	√	Yes, the outcomes of the model are consistent with the perspective, scope and overall objective of the model. The outcomes are to find out which of the three management options (CT-FNA, EUS-FNA or surgery) is least costly strategy.
S3	Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	√	The structure of the model is consistent with the staging algorithm for finding advanced disease (detection of metastatic CLNs) in patients with oesophageal cancer.
	Are the sources of data used to develop the structure of the model specified?	√	The authors specified the sources of data – 3. The probabilities were obtained from an extensive literature search. The probabilities for EUS-FNA and CT-guided FNA sensitivity were obtained from studies that had similar patient characteristics (i.e. patients with non-metastatic oesophageal cancer).

			4. Costs data – The cost of a medical procedure varies with the perspective taken. Direct medical costs were estimated from the Medicare ambulatory patient classification (APC) plus professional fees for hospital-based outpatient procedures and are shown in Table 2. Direct costs were used in preference to charges or total costs because they reflect true resource utilisation better and tend to be more generalizable.
	Are the causal relationships described by the model structure justified appropriately?	NA	-
S4	Are the structural assumptions transparent and justified?	√	The structural assumptions are transparent and justified – 3. It was assumed that patients referred for either EUS-FNA or CT-guided FNA are assumed to have the same comorbidities, and all other diagnostic and additional clinical decision making is assumed to be similar. 4. Pathology interpretation costs are not included because these are also assumed similar in all arms.

		5. The cost of combined chemotherapy is not taken into account as the proportion of early and advanced tumours in each management arm is similar. 6. It is assumed that enlarged CLNs detected on CT will be visualised by EUS because its sensitivity is significantly better than that of CT for detection of nodal metastases. 7. The positive predictive value of a cytological finding of malignancy is assumed to be 100% for both EUS-FNA and CT-guided FNA. 8. The morbidity and long term side effects of radiation or chemotherapy were not included in the primary analysis as these remain poorly quantified. 9. The complication rates of EUS-FNA, CT-guided FNA, and exploratory laparotomy were considered to be small, having negligible effects on model results and were excluded to simply the model.
Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	√	Yes reasonable.

S5	Is there a clear definition of the options under evaluation?	√	The definition of three management options was given. Each strategy started after patient has undergone CT of the abdomen and thorax and has been deemed to have a resectable oesophageal tumour, based on T and M stages (except in the case of celiac nodal involvement, which, by strict definition, qualifies as stage M1). The authors mentioned that they have disregarded the N stage as determined by CT because CT sensitivity for nodal metastases has been demonstrated to be poor. 3. First management option – Obtain CT-guided biopsy of any suspicious-appearing CLNs detected on CT. 4. Second management option – Obtain EUS with FNA of any suspicious-appearing CLNs, or 5. Third management option – Proceed directly to surgery.
	Have all feasible and practical options been evaluated?	√	It seems not all possible staging strategies been evaluated e.g. PET, laparoscopy but it's hard to say for this case. This is because this study aimed to assess which initial preoperative staging strategy is least costly for detecting advanced disease (malignant celiac lymph node) in oesophageal cancer patients. And therefore maybe due to this reason and for achieving the aim of the study, CT-guided FNA, EUS-FNA were evaluated for initial

			preoperative staging strategy and compared them to 'proceed directly to surgery' to evaluate which management option is least costly.
	Is there justification for the exclusion of feasible options?	NA	-
S6	Is the chosen model type appropriate given the decision problem and specified causal relationships within the model?	√	Decision tree model demonstrates the preoperative staging management options and the clinical consequences of making the choice of each management option (CT-FNA or EUS-FNA or proceed directly to surgery). The chosen model type (decision tree model) seems appropriate given the aim is to find out which of the three management options is least costly in detecting malignant CLNs in patients with apparently 'resectable' oesophageal cancer on CT. Causal relationships are not relevant.
S7	Is the time horizon of the model sufficient to reflect all important differences between options?	x	Though the paper aimed to look at the relative financial value of the addition of EUS FNA of CLNs to the preoperative evaluation of a patient with apparently "resectable" oesophageal cancer on CT, it did not specify as

			to what is the time horizon of the model used, which this may span days, weeks or months.
	Are the time horizon of the model, the duration of treatment and the duration of treatment effect described and justified?	x	The time horizon of the decision tree model is not specified. Also, none was specified for the duration of treatment and the duration of treatment effect as these are not applicable for the study's model. This is because the decision tree model was designed to model the relative financial value of diagnostic management options for patients with apparently "resectable" oesophageal cancer on CT.
S8	Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions?	√	The pathways of the decision tree model presented in the paper do reflect the impact of finding out which of the three management options (CT-guided FNA, EUS-FNA and 'proceed directly to surgery') is least costly to the preoperative evaluation of a patient with apparently "resectable" oesophageal cancer on CT.
S9	Is the cycle length defined and justified in terms of the natural history of disease?	NA	-

D1	Are the data identification methods transparent and appropriate given the objectives of the model?	NC	The probabilities and costs data identification methods are transparent and appropriate given the objectives of the model. The details of probabilities and costs identification methods are described as below – **Probabilities –** The probabilities were obtained from an extensive literature search. The probabilities for EUS FNA and CT-guided FNA sensitivity were obtained from studies that had similar patient characteristics (i.e. patients with nonmetastatic oesophageal cancer). **Costs –** Direct medical costs were estimated from the Medicare ambulatory patient classification (APC) plus professional fees for hospital-based outpatient procedures and are outlined in Table 2. Although all papers used to gather data are presented, it is not clear if they were found through a systematic review or chosen at random/purposefully.
	Where choices have been made between data sources, are these justified appropriately?	√	The authors explained that direct costs were used in preference to charges or total costs because they reflect true resource utilisation better and tend to be more generalizable. The authors further explained that indirect health and institutional costs such as the cost to society for lost work, quality of life, and institutional administration or maintenance of buildings were not included. Also, the authors noted that the costs represent the average

			payments allowed for each coded procedure by the United States Center for Medicare and Medicaid Services during the fiscal year 2001. Reimbursement by DRG codes is based on the average length of stay for patients with a given diagnosis. Costs involved in the original diagnosis were not included. Discounting was not performed as the evaluation of a patient with oesophageal cancer takes only several days.
	Has particular attention been paid to identifying data for the important parameters in the model?	√	Yes, the authors seem to pay particular attention to identifying data for EUS-FNA sensitivity.
	Has the quality of the data been assessed appropriately?	x	Not specified
	Where expert opinion has been used, are the methods described and justified?	NA	-
D2	Is the data modelling methodology based on justifiable statistical and epidemiological techniques?	NC	Not clear. The authors only mentioned that they used the DATA 3.5 (TreeAge Software) software package to model their decision analysis.

D2a	Is the choice of baseline data described and justified?	√	Baseline probability of all the variables used in the decision tree model is described and presented in Table 1. Amongst the variables presented in Table 1, only the variable 'EUS-FNA sensitivity' data is justified.
	Are transition probabilities calculated appropriately?	NA	
	Has a half-cycle correction been applied to both cost and outcome?	NA	
	If not, has this omission been justified?	NA	
D2b	If relative treatment effects have been derived from trial data, have they been synthesised using appropriate techniques?	NA	
	Have the methods and assumptions used to extrapolate short-term results to final outcomes been documented and justified?	NA	

	Have alternative assumptions been explored through sensitivity analysis?	√	Yes, baseline probabilities were varied through plausible ranges using sensitivity analysis.
	Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?	NA	
	Have alternative assumptions regarding the continuing effect of treatment been explored through sensitivity analysis?	NA	
D2c	Are the costs incorporated into the model justified?	√	Yes, the costs incorporated into the model are justified and presented in Table 2.
	Has the source for all costs been described?	√	Outpatient costs – Direct medical costs were estimated from the Medicare ambulatory patient classification (APC) plus professional fees for hospital-based outpatient procedures and are presented in Table 2.

			Inpatient costs – Diagnosis Related Groups (DRGs) costs were used for hospital inpatient services cost plus professional fees and are presented in Table 2.
	Have discount rates been described and justified given the target decision-maker?	NA	Discounting was not performed as the evaluation takes only several days.
D2d	Are the utilities incorporated into the model appropriate?	NA	Utilities were not evaluated for the model of the study as it's a cost-minimisation analysis and not a cost-utility analysis, thus, incorporating utilities into the model is not appropriate in this study.
	Is the source for the utility weights referenced?	NA	-
	Are the methods of derivation for the utility weights justified?	NA	-
D3	Have all data incorporated into the model been described and referenced in sufficient detail?	√	Probabilities and costs data incorporated into the decision tree model have been described accordingly in detail and all are presented in Table 1 and Table 2 in the paper.

	Has the use of mutually inconsistent data been justified (i.e. are assumptions and choices appropriate)?	NC	
	Is the process of data incorporation transparent?	x	Not specified
	If data have been incorporated as distributions, has the choice of distribution for each parameter been described and justified?	NA	
	If data have been incorporated as distributions, is it clear that second order uncertainty is reflected?	NA	
D4	Have the four principal types of uncertainty been addressed?	x	Not specified
	If not, has the omission of particular forms of uncertainty been justified?	x	Not specified

D4a	Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?	x	
D4b	Is there evidence that structural uncertainties have been addressed via sensitivity analysis?	x	Not specified.
D4c	Has heterogeneity been dealt with by running the model separately for different subgroups?	x	Model ran for all patients with apparently “resectable” oesophageal cancer on CT.
D4d	Are the methods of assessment of parameter uncertainty appropriate?	√	All key variables were addressed in sensitivity analyses under all possible ranges of value. 1. One-way sensitivity analysis varying prevalence of malignant CLNs, 2. One-way sensitivity analysis varying sensitivity of EUS-FNA, and 3. Two-way sensitivity analysis varying prevalence of malignant CLNs and sensitivity of EUS-FNA simultaneously.

	If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	√	The ranges used for sensitivity analysis are stated clearly and presented in Table 1 in the paper. Justification for the range of EUS-FNA sensitivity was described.
C1	Is there evidence that the mathematical logic of the model has been tested thoroughly before use?	x	Not specified.
C2	Are any counterintuitive results from the model explained and justified?	NA	
	If the model has been calibrated against independent data, have any differences been explained and justified?	NA	
	Have the results of the model been compared with those of previous	X	No previous models are specified; no comparison of results with previous models is made.

models and any differences in results explained?		

*****Available from: Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, Woolacott N and Glanville J. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004;**8**(36).

Appendix 3.22: Quality assessment results of Wallace et al (2002)

Quality Criterion	Economic modelling checklist questions*	Response (√, x or NA)	Comments (i.e. a description that explains how the judgement was reached)
S1	Is there a clear statement of the decision problem?	√	The decision problem is currently not known which of the available procedures (CT, EUS-FNA, TL and PET) or which combinations of these procedures offers the most cost-effective approach for the detection of locally advanced or metastatic disease for patients with oesophageal cancer.
	Is the objective of the evaluation and model specified and consistent with the stated decision problem?	√	The objective of the model is to compare the costs and effectiveness of six strategies for staging patients with oesophageal cancer: CT alone, CT+EUS-fine-needle aspiration biopsy (FNA), CT+thoracoscopy and laparoscopy (TL), CT+EUS-FNA+TL, CT+PET+EUS-FNA, and PET+EUS-FNA.
	Is the primary decision-maker specified?	√	Third party payer, Medicare
S2	Is the perspective of the model stated clearly?	√	The model was based on a third-party payer perspective.
	Are the model inputs consistent with the stated perspective?	√	Inputs relate to the test characteristics for each of the staging techniques; the prevalence of local, regional, and distant disease among patients with oesophageal cancer; life expectancies for oesophageal cancer patients with

			local, regional, and distant disease; costs associated with the different staging techniques; costs associated with the treatment and care of patients identified as having local, regional, and distant disease; and probability of death for patients undergoing TL and those undergoing oesophageal resection.
	Has the scope of the model been stated and justified?	√	Type of patients and staging strategies options were stated within the objectives. Assumption was made that staging tests were performed sequentially in each strategy case, along with justification for no further staging was performed if distant metastases were found and confirmed.
	Are the outcomes of the model consistent with the perspective, scope and overall objective of the model?	√	Outcomes: costs and life expectancies
S3	Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	√	The structure of the model reflects the clinical algorithm for the management of oesophageal cancer patients, which the staging algorithm was specified according to the protocol.

	Are the sources of data used to develop the structure of the model specified?	√	Data used to develop the structure of the model came from several sources, which are specified in Table 1.
	Are the causal relationships described by the model structure justified appropriately?	NA	-
S4	Are the structural assumptions transparent and justified?	√	It was assumed that staging tests were performed sequentially in each of the six staging strategies and the staging algorithm was specified according to the protocol. And if distant metastases were found and confirmed, no further staging was performed.
	Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	√	Yes reasonable.
S5	Is there a clear definition of the options under evaluation?	√	All six strategies details are given.

	Have all feasible and practical options been evaluated?	√	Seems to be the six strategies are all the possible staging strategies for management of oesophageal cancer patients, although without the clinical knowledge of the area it's hard to say.
	Is there justification for the exclusion of feasible options?	NA	-
S6	Is the chosen model type appropriate given the decision problem and specified causal relationships within the model?	√	Decision tree model represents the sequences of combinations of each of the six staging strategies incorporating: the prevalence of disease; life expectancies, costs associated with the different staging techniques, costs associated with the treatment and care of patients, probability of death for patients undergoing TL and oesophageal resection. Seems that the chosen model type is appropriate given the aim is to identify which of these or the combinations of these staging strategies offers the most cost-effective approach in detecting oesophageal cancer (locally advanced or metastatic disease).
S7	Is the time horizon of the model sufficient to reflect all important differences between options?	x	Cannot say as to what is the time horizon of the decision tree model from the paper though the authors mentioned that using the SEER-Medicare databases, a retrospective cohort was created consisting of all Medicare-

			eligible patients whose invasive oesophageal cancer was diagnosed between January 1, 1991 and December 31, 1996.
	Are the time horizon of the model, the duration of treatment and the duration of treatment effect described and justified?	x	The time horizon of the decision tree model are not specified. Also, none specified for the duration of treatment and the duration of treatment effect as I think these are not applicable for the study's model as the decision tree model was designed to reflect the staging algorithm for the management of patients diagnosed with carcinoma of esophagus.
S8	Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions?	√	Decision tree represents the staging algorithm that patients may follow when they are diagnosed with oesophageal cancer.
S9	Is the cycle length defined and justified in terms of the natural history of disease?	NA	

D1	Are the data identification methods transparent and appropriate given the objectives of the model?	√	Test characteristics for the different staging procedures were derived from published literature (but didn't specify whether systematic) and from a prospective study. Life expectancies were derived from 1973-1996 SEER data and separate life expectancies were derived from median survival times for patients in various stages at diagnosis as well as for patients who had and who had not been treated surgically for the disease. Costs associated with the different staging techniques and inpatient procedure for the diagnostically related group for TL with lymph node biopsy linked to the ICD-9 code for malignant oesophageal neoplasm were based on Medicare reimbursement. Outpatient procedure for EUS-FNA costs were based on the CPT (Current Procedural Terminology) code for EUS with FNA linked to the ICD-9 code for malignant oesophageal neoplasm.
	Where choices have been made between data sources, are these justified appropriately?	√	The reason as to why they estimated the mix of patients with and without comorbid conditions from surveys of patients undergoing staging for esophagus cancer at the Medical University of South Carolina over 1 year period was given – Surveys were carried out due to the fact that the Medicare inpatient reimbursement is higher for patients with comorbid conditions; the choice made between data sources seems to have been justified appropriately.

	Has particular attention been paid to identifying data for the important parameters in the model?	x	Life expectancies data were derived from SEER and separate life expectancies data were derived from median survival times for patients in various stages at diagnosis and for who had and had not been treated surgically for the disease. A key uncertainty was found to be associated with the QALY values associated with patients with resected/unresected local/regional/distant disease. QALY values are based on expert opinion; no details are given regarding the expert panel or methods.
	Has the quality of the data been assessed appropriately?	x	Not specified
	Where expert opinion has been used, are the methods described and justified?	x	Expert opinion was used to identify quality of life utilities scores. No details are given regarding the expert panel or methods.
D2	Is the data modelling methodology based on justifiable statistical and epidemiological techniques?	√	Standard methods of cost-effectiveness analysis were used for the decision analysis
D2a	Is the choice of baseline data described and justified?	√	All baseline values for the model variables came from several sources, as shown in Table 1 in the paper.

	Are transition probabilities calculated appropriately?	NA	
	Has a half-cycle correction been applied to both cost and outcome?	NA	
	If not, has this omission been justified?	NA	
D2b	If relative treatment effects have been derived from trial data, have they been synthesised using appropriate techniques?	NA	
	Have the methods and assumptions used to extrapolate short-term results to final outcomes been documented and justified?	NA	
	Have alternative assumptions been explored through sensitivity analysis?	√	It's noted that all key variables for each model were subjected to sensitivity analyses to determine which variables had the greatest effects on the cost-effectiveness measures for each strategy.

	Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?	NA	
	Have alternative assumptions regarding the continuing effect of treatment been explored through sensitivity analysis?	NA	
D2c	Are the costs incorporated into the model justified?	√	Costs associated with management of patients with esophagus cancer were incorporated into the decision model (i.e. costs associated with the different staging techniques, and costs associated with the treatment and care of patients identified as having local, regional, and distant disease).
	Has the source for all costs been described?	√	All costs associated with management of patients with esophagus cancer were derived from SEER-Medicare linked databases.
	Have discount rates been described and justified given the target decision-maker?	√	Both a 0% and 3% discount rate were used to estimate costs.

D2d	Are the utilities incorporated into the model appropriate?	√	Come directly from expert opinion.
	Is the source for the utility weights referenced?	x	Not detailed
	Are the methods of derivation for the utility weights justified?	x	Not detailed
D3	Have all data incorporated into the model been described and referenced in sufficient detail?	√	Described and referenced accordingly and all are presented in Table 1.
	Has the use of mutually inconsistent data been justified (i.e. are assumptions and choices appropriate)?	√	The paper justified why surveys data were used to determine what percentage was reimbursed at the rate for those patients with and without morbid conditions instead of taking it directly from Medicare inpatient reimbursement. The reason is because the Medicare reimbursement is higher for patients with comorbid conditions.
	Is the process of data incorporation transparent?	x	Not detailed.
	If data have been incorporated as distributions, has the choice of	NA	

	distribution for each parameter been described and justified?		
	If data have been incorporated as distributions, is it clear that second order uncertainty is reflected?	NA	
D4	Have the four principal types of uncertainty been addressed?	x	Not specified
	If not, has the omission of particular forms of uncertainty been justified?	x	Not specified
D4a	Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?	x	Not specified
D4b	Is there evidence that structural uncertainties have been addressed via sensitivity analysis?	x	Not specified.

D4c	Has heterogeneity been dealt with by running the model separately for different subgroups?	x	Model ran for patients of all ages.
D4d	Are the methods of assessment of parameter uncertainty appropriate?	√	All key variables were addressed in sensitivity analyses under all possible ranges of value. Ranges of value were used in sensitivity analyses to assess how they would affect the baseline results.
	If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	√	The ranges used for sensitivity analysis are stated clearly and presented in Table 1 in the paper.
C1	Is there evidence that the mathematical logic of the model has been tested thoroughly before use?	x	Not specified.

C2	Are any counterintuitive results from the model explained and justified?	NA	
	If the model has been calibrated against independent data, have any differences been explained and justified?	NA	
	Have the results of the model been compared with those of previous models and any differences in results explained?	x	No previous models are specified; no comparison of results with previous models is made.

*****Available from: Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, Woolacott N and Glanville J. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004;**8**(36).

Appendix 3.23: Published Systematic Review in BMC Cancer Journal (Full paper can be found at https://doi.org/10.1186/s12885-019-6116-0)

Yeo et al. BMC Cancer (2019) 19:900
https://doi.org/10.1186/s12885-019-6116-0

BMC Cancer

RESEARCH ARTICLE Open Access

Endoscopic ultrasound staging in patients with gastro-oesophageal cancers: a systematic review of economic evidence

Seow Tien Yeo[1*], Nathan Bray[1], Hasan Haboubi[2], Zoe Hoare[3] and Rhiannon Tudor Edwards[1]

Abstract

Background: The sensitivity of endoscopic ultrasound (EUS) in staging gastro-oesophageal cancers (GOCs) has been widely studied. However, the economic evidence of EUS staging in the management of patients with GOCs is scarce. This review aimed to examine all economic evidence (not limited to randomised controlled trials) of the use of EUS staging in the management of GOCs patients, and to offer a review of economic evidence on the costs, benefits (in terms of GOCs patients' health-related quality of life), and economic implications of the use of EUS in staging GOCs patients.

Methods: The protocol was registered prospectively with PROSPERO (CRD42016043700; http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42016043700). MEDLINE (ovid), EMBASE (ovid), The Cochrane Collaboration Register and Library (including the British National Health Service Economic Evaluation Database), CINAHL (EBSCOhost) and Web of Science (Core Collection) as well as reference lists were systematically searched for studies conducted between 1996 and 2018 (search update 28/04/2018). Two authors independently screened the identified articles, assessed study quality, and extracted data. Study characteristics of the included articles, including incremental cost-effectiveness ratios, when available, were summarised narratively.

Results: Of the 197 articles retrieved, six studies met the inclusion criteria: three economic studies and three economic modelling studies. Of the three economic studies, one was a cost-effectiveness analysis and two were cost analyses. Of the three economic modelling studies, one was a cost-effectiveness analysis and two were cost-minimisation analyses. Both of the cost-effectiveness analyses reported that use of EUS as an additional staging technique provided, on average, more QALYs (0.0019–0.1969 more QALYs) and saved costs (by £1969–£3364 per patient, 2017 price year) compared to staging strategy without EUS. Of the six studies, only one included GOCs participants and the other five included oesophageal cancer participants.

Conclusions: The data available suggest use of EUS as a complementary staging technique to other staging techniques for GOCs appears to be cost saving and offers greater QALYs. Nevertheless, future studies are necessary because the economic evidence around this EUS staging intervention for GOCs is far from robust. More health economic research and good quality data are needed to judge the economic benefits of EUS staging for GOCs.

PROSPERO Registration Number: CRD42016043700.

Keywords: Costs, Effects, QALYs, Economic review, Endoscopic ultrasound, EUS staging, Staging techniques, Gastro-oesophageal cancers

* Correspondence: s.t.yeo@bangor.ac.uk
[1]Centre for Health Economics and Medicines Evaluation (CHEME), Bangor University, Arduduy, Normal Site, Holyhead Road, Bangor, Gwynedd LL57 2PZ, UK
Full list of author information is available at the end of the article

Appendices relating to Chapter 4 (Healthcare Professional Survey)

Appendix 4.1: Invitation email written to invite GOC clinicians to partake in the online survey

---

Dear BSG/WAGE members,

Subject: Healthcare Professional Survey Goes Online: Use of EUS for Staging Gastro-Oesophageal Cancer (GOC) – Approximately 10-minute Survey

This survey has gained support from Professor Stephen Attwood, the Chairman of Oesophageal Section of British Society Gastroenterology (BSG). Please see the support message from Professor Stephen Attwood in quotation mark below.

"I, Professor Stephen Attwood, the Chairman of Oesophageal Section of British Society Gastroenterology (BSG), am happy to support this survey through the offices of the BSG, as it is a topic of specific interest to the members of the BSG."

Despite the research evidence on EUS sensitivity, specificity, and its effectiveness and cost-effectiveness, evidence on the utilisation of EUS in clinical practice is scarce. To date, only one US study exploring utilisation of EUS for staging oesophageal cancer has been performed. No UK study has ever been undertaken.

You are kindly invited to partake in a survey investigating the utility of EUS in the UK. This survey is undertaken as part of my PhD study, funded by the Tenovus Cancer Care Charity. This survey aims to explore your views about……

- the utilisation of EUS for staging GOC (Gastro-Oesophageal Cancer – oesophageal, gastro-oesophageal junction or gastric cancer)
- the current clinical practice of EUS for staging GOC in your practice
- the usefulness of staging GOC with EUS
- the availability of EUS for staging GOC within your practice or nearby your practice
- how patient with oesophageal cancer is managed based upon small changes in tumour stage.

This survey is voluntary and anonymous and I very much value your participation. To participate, please follow the link below.

Link to the survey: https://bangor.onlinesurveys.ac.uk/healthcare-professionals-survey_eus

If you have questions about this survey please contact;
Seow Tien Yeo
Survey coordinator & PhD student,
Bangor University
Email: s.t.yeo@bangor.ac.uk

In recognition of time given to complete the survey (approx. 10 minutes), you will be offered to enter into a prize draw at the end of your completed survey. If you wish to be entered into a prize draw to win £100 Amazon voucher, please enter your email address at the end of your completed survey.

Thank you very much in anticipation.

Yours sincerely,

Seow Tien Yeo

Survey Coordinator and PhD Student, Bangor University.

Appendix 4.2: Online healthcare professional survey on use of EUS for GOC staging

How is endoscopic ultrasound (EUS) used in gastro-oesophageal cancer (GOC) staging?

About this survey.

This survey has gained support from Professor Stephen Attwood, the Chairman of Oesophageal Section of British Society Gastroenterology (BSG).

Please see the support message from Professor Stephen Attwood in quotation mark below.

"I, Professor Stephen Attwood, the Chairman of Oesophageal Section of British Society Gastroenterology (BSG), am happy to support this survey through the offices of the BSG, as it is a topic of specific interest to the members of the BSG."

Despite the research evidence on EUS sensitivity, specificity, and its effectiveness and cost-effectiveness, evidence on the utilisation of EUS in clinical practice is scarce. To date, only one US study exploring utilisation of EUS for staging oesophageal cancer has been performed. No UK study has ever been undertaken.

This survey is undertaken as part of my PhD study, funded by the Tenovus Cancer Care Charity. This survey aims to explore your views about……

- the utilisation of EUS for staging GOC (Gastro-Oesophageal Cancer – oesophageal, gastro-oesophageal junction or gastric cancer)
- the current clinical practice of EUS for staging GOC in your practice
- the usefulness of staging GOC with EUS
- the availability of EUS for staging GOC within your practice or nearby your practice
- how patient with oesophageal cancer is managed based upon small changes in tumour stage.

The questionnaire will take approximately 10 minutes to complete.

If you have questions about this survey please contact;

Seow Tien Yeo
Survey coordinator & PhD student,
Bangor University
Email: s.t.yeo@bangor.ac.uk

Section A: These questions relate to your demographic details

Please describe yourself:

1. You are a member of (Please tick all that apply);

- ☐ British Society of Gastroenterology (BSG)
- ☐ Welsh Association for Gastroenterology and Endoscopy (WAGE)
- ☐ Other

1.a. If you selected Other, please specify:

2. How did you hear about this survey? Please tick all that apply.

- ☐ BSG website
- ☐ BSG monthly newsletter
- ☐ BSG conference/symposium
- ☐ Email from BSG
- ☐ WAGE website
- ☐ WAGE monthly newsletter
- ☐ WAGE conference/symposium
- ☐ Email from WAGE
- ☐ Other

2.a. If you selected Other, please specify:

3. Which of the following best describes you? Please tick one that best describes you.

- ☐ Gastroenterologist
- ☐ GI surgeon
- ☐ Oncologist
- ☐ Radiologist
- ☐ Other

3.a. If you selected Other, please specify:

4. Age (years)

5. Gender

○ Male ○ Female

6. Years in practice (since primary medical qualification)

7. Please provide the name of the hospital at which you are practicing **and** the name of the Health Board or NHS Trust as appropriate in the box below. This will provide us geographical information as to how well the questionnaire is distributed. **We will not use it for identifying participant and specific hospital or any other purpose.**

8. Is your current hospital a.....(Please tick one that applies)

- Teaching Hospital
- District General Hospital
- Community Hospital

9. Which of the following best describes your exposure and contact with patient with oesophageal, gastro-oesophageal junction (GOJ) or gastric cancer? Please tick one that applies.

- Yes - Regularly (i.e. weekly)
- Yes - Sometimes (i.e. monthly)
- Yes - Rarely (i.e. yearly)
- Yes - But hardly ever (Longer than yearly)
- No - Never at all

10. Does your hospital run an Upper GI cancer Multi-Disciplinary Team (MDT) meeting? Please tick one that applies.

- Yes
- No

11. Do you attend an Upper GI cancer MDT meeting? Please tick one that applies.

- Yes
- No

Section B: These questions relate to your knowledge and feelings about the use of EUS

12. Do you have experience in the staging of GOC (Gastro-Oesophageal Cancer – oesophageal, gastro-oesophageal junction or gastric cancer)?

- ○ Yes
- ○ No

13. Endoscopic ultrasound (EUS) is one of the imaging modalities used for staging GOC. Which of the following can EUS accurately stage? Please tick all that relevant.

- ☐ Oesophageal cancer
- ☐ Gastro-Oesophageal Junction cancer
- ☐ Gastric cancer
- ☐ Other

13.a. If you selected Other, please specify:

13.b. Does EUS help to reduce unnecessary surgery for GOC patients?

- ○ Yes
- ○ No
- ○ I am not sure

13.b.i. If yes, by how much would you estimate EUS reduces unnecessary surgery?

- ○ Very little
- ○ Somewhat little
- ○ No difference
- ○ Somewhat a lot
- ○ A lot

13.c. If EUS is not currently an option used for staging GOC patients and if it was to

become available to you, how likely will you choose to use EUS for staging GOC patients in the future?

- ○ Not applicable
- ○ Never
- ○ Not likely
- ○ Somewhat likely
- ○ Very likely
- ○ Always

14. Do you support the use of EUS for staging GOC?

- ○ Yes
- ○ No

14.a. If yes, please tick all the boxes below that apply

- ☐ EUS offers the best local regional staging of GOC
- ☐ EUS offers opportunity to biopsy during staging
- ☐ Cost-effectiveness of using EUS to stage GOC
- ☐ Other

14.a.i. If you selected Other, please specify:

14.b. If no, please tick all the boxes below that apply

- ☐ Lack of expertise in performing EUS
- ☐ Inter-operator variability in interpreting images
- ☐ Cost-effectiveness of using other staging modalities to stage GOC
- ☐ Other

14.b.i. If you selected Other, please specify:

15. Are you aware of the published UK COGNATE trial HTA report? It is the report regarding a randomised controlled trial evaluating the effectiveness and cost-effectiveness of endoscopic ultrasound in the management of patients with gastro-oesophageal cancer.

- ○ Yes
- ○ No

15.a. If you answer 'Yes' to question 15, have you read the COGNATE trial HTA report?

- ○ Yes
- ○ No

15.a.i. If no, how likely is it that you will read the COGNATE trial HTA report?

- ○ Not likely
- ○ Somewhat likely
- ○ Very likely

15.b. If you answer 'NO' to question 15, how likely is it that you will read the COGNATE trial HTA report?

- ○ Not likely
- ○ Somewhat likely
- ○ Very likely

Experience in the field of EUS

16. Do you have any experience in the field of EUS (i.e. request and/or perform EUS)?

- ○ Yes
- ○ No

Section C: These questions relate to your experience in the field of EUS

17. What best describes your experience in the field of EUS? Please select ***one*** that applies in the Table below ***and*** please ***estimate*** the total number of EUS you request ***and/or*** perform ***per 3-month*** in the table below.

	I request EUS	I perform EUS	I both request and perform EUS	I do not regularly request EUS and do not perform EUS	I do not request EUS and do not regularly perform EUS	**EUS you request per 3-month**	**EUS you perform per 3-month**
Please select one that applies and fill in the appropriate box(es)	○	○	○	○	○		

Section D: These questions relate to your current use of EUS and reasons where this may be limited, the usefulness and availability of EUS.

18. Do you routinely request for EUS for staging ***all newly*** diagnosed GOC patients? Please tick one that applies.

- ○ Not applicable (I am unable to request EUS)
- ○ Yes
- ○ No

18.a. If no, please detail what other staging modalities you would request in the box below.

19. How likely are you to request for EUS for ***suspected relapse*** of GOC cancer?

- ○ Not applicable (I am unable to request EUS)
- ○ Never
- ○ Not likely
- ○ Somewhat likely
- ○ Very likely
- ○ Always

20. Do you have a **standard hospital protocol** regarding the use of EUS for staging GOC to follow?

- ○ Yes
- ○ No

21. What proportion of EUS requests are made through the Upper GI Cancer MDT in your practice or outside? Please select **one** option in the Table below that applies.

	All in MDT	Majority in MDT	Approximately equal request through MDT and outside MDT	Majority outside MDT	All outside MDT
In Your Practice	○	○	○	○	○
Outside Your Practice	○	○	○	○	○

22. Are there specific situations where you don't use EUS?

○ Yes ○ No

22.a. Please tell us what are the specific situations where you don't use EUS. Please tick all that apply.

- ☐ Availability
- ☐ Tight oesophageal stricture
- ☐ Patient morbidity
- ☐ Advanced metastatic disease
- ☐ Other

22.a.i. If you selected Other, please specify:

22.b. Please tell us what are the reasons behind request and utilisation of other imaging modalities compared to EUS in GOC assessment? Please tick all that apply.

- ☐ Availability
- ☐ Cost

- ☐ Expertise
- ☐ Resources e.g. staffing constraints and etc
- ☐ Time taken to perform EUS procedure
- ☐ Patient comfort
- ☐ Other

22.b.i. If you selected Other, please specify:

23. In patients newly diagnosed with following gastro-oesophageal malignancies, how useful is staging with endoscopic ultrasound in assisting your clinical management? Please check **one** box for **each** gastro-oesophageal malignancy in the table below.

	Not useful (0)	Slightly useful (1)	Somewhat useful (2)	Very useful (3)	Essential (4)
a) Oesophageal cancer	○	○	○	○	○
b) Gastro-oesophageal junction cancer	○	○	○	○	○
c) Gastric cancer	○	○	○	○	○

24. Is endoscopic ultrasound available **within your hospital**?

○ Yes ○ No

24.a. **If you answered 'No' to question 24,** which hospital GOC patients are usually referred to for endoscopic ultrasound screening? Please provide your answers to a), b) and c) below. This will provide us the geographical information of the hospital. **We will not use it for identifying participant and specific hospital or any other purpose.**

a) The name of the hospital which GOC patients are usually referred to	
b) The name of the Health Board or NHS Trust as appropriate.	
c) Please estimate how far (in miles) the hospital is situated from your hospital?	

24.a.i. Is this a teaching hospital?

○ Yes ○ No

25. In the past 12 months, approximately how many patients with each of the following gastro-oesophageal malignancies did you **see in total**, and approximately how many of these patients **underwent** endoscopic ultrasound? Please provide your answer in the table below.

	Total Patients Seen (Approx.)	Had EUS (Approx.)
Oesophageal cancer		
Gastro-oesophageal Junction cancer		
Gastric cancer		

Section E: These questions are pre-validated and will help to identify how different clinicians manage different situations. Please try to answer honestly.

A 58 year-old homemaker presents with heartburn and is found to have a 2cm diameter distal oesophageal adenocarcinoma. Staging endoscopic ultrasound is performed. What treatment would you recommend for each of the following EUS results?

Please check one box next to one choice for each of the following questions.

26. Mass penetrates all wall layers with two round 1cm peri-oesophageal node.

- ○ Referral for oesophagectomy (with or without chemo/XRT)
- ○ Referral for chemotherapy and/or radiotherapy without surgery
- ○ No treatment, with endoscopic palliation for future symptoms

27. Mass penetrates all wall layers but no nodes are seen.

- ○ Referral for oesophagectomy (with or without chemo/XRT)
- ○ Referral for chemotherapy and/or radiotherapy without surgery
- ○ No treatment, with endoscopic palliation for future symptoms

28. Mass is confined to muscularis, no nodes are seen.

- ○ Referral for oesophagectomy (with or without chemo/XRT)
- ○ Referral for chemotherapy and/or radiotherapy without surgery
- ○ No treatment, with endoscopic palliation for future symptoms

End of survey (I)

Thank you for your time and efforts in completing this survey.

29. If you would like to be entered into a prize draw where you could win a £100 Amazon voucher for taking part in the survey, please enter your email address in the box below.

Please note:

All information provided in this survey is anonymous and will only be used by me for my PhD research purposes. All responses provided in this survey will remain anonymous in the reporting process.

You have the right to withdraw, but that by submitting your completed survey (i.e. clicking the 'Finish' button) you are agreeing to participate. This survey contains no identifiable information so confidentiality is maintained.

Your completed online responses will be collected and transferred into the University's encrypted computer for analysis. All information provided in this survey will be handled in confidence and in accordance with the Data Protection Act 1998.

End of survey (II)

If you have any further queries about this survey, please contact:

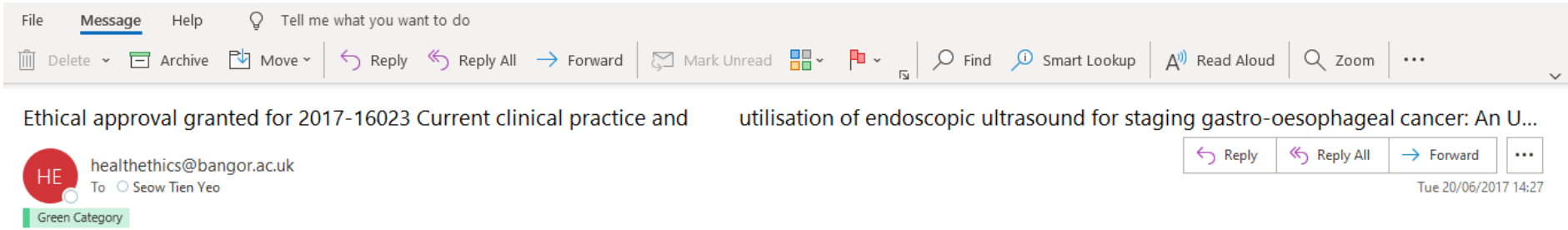
Ms. SeowTien Yeo

Survey coordinator & PhD student

Bangor University, UK

Email: s.t.yeo@bangor.ac.uk

Appendix 4.3: Ethical approval granted for the healthcare professional survey



Ethical approval granted for 2017-16023 Current clinical practice and utilisation of endoscopic ultrasound for staging gastro-oesophageal cancer: An U...

HE healthethics@bangor.ac.uk
To Seow Tien Yeo
Green Category

Reply | Reply All | Forward | ...
Tue 20/06/2017 14:27

Dear Seow Tien,

2017-16023 Current clinical practice and utilisation of endoscopic ultrasound for staging gastro-oesophageal cancer: An UK healthcare professional survey

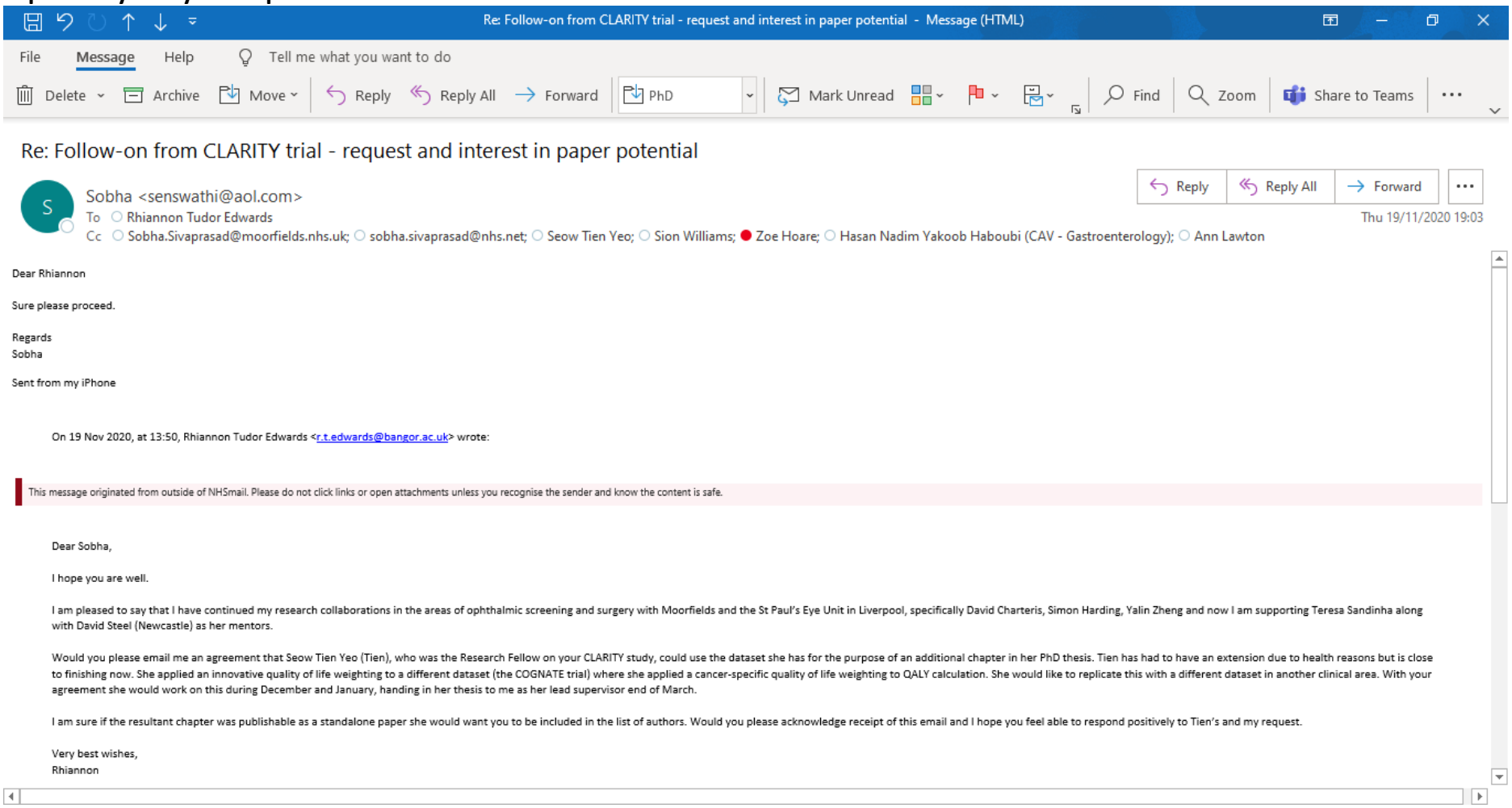
Your research proposal number 2017-16023
has been reviewed by the Healthcare Sciences (Post-reg) Ethics and Research Committee
and the committee are now able to confirm ethical and governance approval for the above research on the basis described in the application form, protocol and supporting documentation. This approval lasts for a maximum of three years from this date.

Ethical approval is granted for the study as it was explicitly described in the application

If you wish to make any non-trivial modifications to the research project, please submit an amendment form to the committee, and copies of any of the original documents reviewed which have been altered as a result of the amendment. Please also inform the committee immediately if participants experience any unanticipated harm as a result of taking part in your research, or if any adverse reactions are reported in subsequent literature using the same technique elsewhere.

Appendices relating to Chapter 5 (Exploratory Study)

Appendix 5.1: Approval granted from the Principal Investigator of the CLARITY trial for use of the CLARITY trial data as part of the exploratory study in Chapter 5 in this thesis



Re: Follow-on from CLARITY trial - request and interest in paper potential - Message (HTML)

File Message Help Tell me what you want to do

Delete Archive Move Reply Reply All Forward PhD Mark Unread Find Zoom Share to Teams

Re: Follow-on from CLARITY trial - request and interest in paper potential

Sobha <senswathi@aol.com>

To Rhiannon Tudor Edwards

Cc Sobha.Sivaprasad@moorfields.nhs.uk; sobha.sivaprasad@nhs.net; Seow Tien Yeo; Sion Williams; Zoe Hoare; Hasan Nadim Yakoob Haboubi (CAV - Gastroenterology); Ann Lawton

Reply Reply All Forward

Thu 19/11/2020 19:03

Dear Rhiannon

Sure please proceed.

Regards
Sobha

Sent from my iPhone

On 19 Nov 2020, at 13:50, Rhiannon Tudor Edwards <r.t.edwards@bangor.ac.uk> wrote:

This message originated from outside of NHSmail. Please do not click links or open attachments unless you recognise the sender and know the content is safe.

Dear Sobha,

I hope you are well.

I am pleased to say that I have continued my research collaborations in the areas of ophthalmic screening and surgery with Moorfields and the St Paul's Eye Unit in Liverpool, specifically David Charteris, Simon Harding, Yalin Zheng and now I am supporting Teresa Sandinha along with David Steel (Newcastle) as her mentors.

Would you please email me an agreement that Seow Tien Yeo (Tien), who was the Research Fellow on your CLARITY study, could use the dataset she has for the purpose of an additional chapter in her PhD thesis. Tien has had to have an extension due to health reasons but is close to finishing now. She applied an innovative quality of life weighting to a different dataset (the COGNATE trial) where she applied a cancer-specific quality of life weighting to QALY calculation. She would like to replicate this with a different dataset in another clinical area. With your agreement she would work on this during December and January, handing in her thesis to me as her lead supervisor end of March.

I am sure if the resultant chapter was publishable as a standalone paper she would want you to be included in the list of authors. Would you please acknowledge receipt of this email and I hope you feel able to respond positively to Tien's and my request.

Very best wishes,
Rhiannon