Quality of Surgery in Oncology Trials

James William Butterworth
MBBS BMedSci MSc (Hons) DIC MRCS (Ed)
Department of Surgery and Cancer, Imperial College London,
St Mary's Hospital, London, UK

2020

A thesis submitted to Imperial College London for the degree of
Doctorate of Philosophy

Supervisors:

Professor G B Hanna[1]
Professor J Reynolds[2]
Dr S Mavroveli[1]
Mr P Boshier[1]

Affiliations:

[1] Department of Surgery and Cancer, Imperial College London, UK

[2] Department of Surgery, Trinity Translational Medicine Institute, St. James's Hospital, Trinity College Dublin, Ireland

ABSTRACT

Randomised controlled trials (RCTs) with surgical interventions often lack a framework to ensure surgical quality. Although recent oncology trials, such as ADDICT (D1+ vs. D2 gastrectomy), have sought to monitor surgery there has been no demonstrably reliable tool to assess surgical quality. We aimed to investigate SQA in oesophagogastric oncology trials and to develop a robust framework of consensus strategies to overcome challenges to design and implementation of SQA. A multi-method approach including both qualitative and quantitative methodologies were applied in order to address the research objectives. On systematic review of previously reported challenges to SQA in trials the most commonly encountered included: constraints of using case volume for credentialing surgeons; inter-centre variation in the definition and execution of interventions, and; insufficient monitoring of surgical quality. A meta-analysis of SQA and protocol utilisation within oesophagogastric RCTs revealed public availability of protocols and Eastern country of origin were associated with improved survival. Semi-structured interviews were subsequently conducted with expert stakeholders examining challenges to SQA in trials. Prominent mitigating strategies included operative monitoring using photographs and/or videos with a structured objective assessment tool. Expert consensus was reached for 59 strategies to overcome challenges to SQA in oncology trials. 14 (74%) of the 19 included expert stakeholder proposed strategies from chapter 4 gained consensus amongst ADDICT trial stakeholders within 2 Delphi rounds, indicating their relevance within oesophagogastric oncology RCTs. A patient focus group and survey, established to gain insight into service user perception of quality of surgery, reinforced the importance of considering operative volume and monitoring surgery using a structured methodology. Robust monitoring methods are required to assess surgical quality and oesophagectomy assessment tools were demonstrated to be reliable using generalisability theory. Condensing the expert Delphi consensus allowed formulation of a 33-item framework of strategies to overcome challenges to implementation of SQA in oncology trials (FOSQAT). Given the relevance of the expert Delphi strategies within ADDICT, in future we recommend trial committees and surgeons will not be required to conduct a Delphi process, but rather will be able to select relevant strategies for implementation from the FOSQAT consensus. Clinical validation of this framework assessing impact of implemented strategies on short and long-term outcomes should be the focus of future research in this area.

Thesis word count: 35,875

DECLARATIONS

I certify that this work is my own and all else is appropriately referenced throughout this thesis. I authorise the library of Imperial College London to lend this thesis to other institutions and individuals.

Signed: J W Butterworth Date: 13.12.20

James W Butterworth

PEER REVIEWED PUBLICATIONS AND PRESENTATIONS

International presentations

Butterworth JW, Boshier PR, Mavrovelli S, Reynolds JV, Hanna GB. 'Surgical Quality Assurance (SQA) in Oncology Trials' presented as an **oral presentation** by Professor GB Hanna at the European Association of Endoscopic Surgery, Seville, Spain, June 2019.

Butterworth JW, Boshier PR, Mavroveli S, Young-woo K, Hanna GB. 'Challenges to recruitment of patients and monitoring of surgery in the ADDICT trial' presented as an **oral presentation** at the KINGCA conference, Korea, 11th April 2019.

Butterworth JW, Boshier PR, Sato M, Vidal-Diez A, Hanna GB. 'Quality Assurance of Surgery in Oesophagogastric Oncology Trials: a Systematic Review and Meta Analysis' presented as **an oral presentation** at the KINGCA conference, 11th April 2019.

Butterworth JW, Boshier PR, Mavroveli M, Vidal-Diez A, Reynolds JV, Hanna GB. 'Quality Assurance of Surgery in Oesophagogastric Oncology trials and trial protocols: a systematic review' presented as an **oral presentation** at the European Society of Surgery in Yerevan, Armenia, on 27th September 2018.

Butterworth JW, Boshier PR, Mavroveli S, Reynolds JV, GHanna GB. Challenges to Surgical Quality Assurance in Oesophagogastric Oncology Trials: a qualitative interview study. Presented at **oral presentation** at European Association of Endoscopic Surgery, London, 31st May to 1st June 2018.

Butterworth JW, Boshier P, Mavroveli S, Hanna GB. Barriers to Quality Assurance of Surgical Intervention in Clinical Trials. **Oral presentation** at the European Association of Endoscopic Surgery, Frankfurt, 16th-18th June 2017.

Publications

Butterworth JW, Boshier PR, Mavroveli, Lanschot JBBV, Sasako M, Reynolds JV, Hanna GB. Challenges to quality assurance of surgical interventions in clinical oncology trials: A systematic review. European Journal of Surgical Oncology. 2020 Oct 02

Harris A, **Butterworth JW**, Boshier PR, MacKenzie H, Tokunaga M, Sunagawa H et al. Development of a Reliable Surgical Quality Assurance System for 2-stage Esophagectomy in Randomized Controlled Trials. Ann Surg. 2020 Mar 27.

ACKNOWLEDGEMENTS

I would like to thank first of my supervisor **Professor George B Hanna (GB)** at Imperial College for their support, inspiration and guidance from October 2016 up until this point. When I felt it was difficult I remembered advice such as 'There are no short-cuts ' and 'Perfect takes a little longer ', and these helped me maintain determination to see the final pages of this thesis. **Professor J V Reynolds** has provided expert guidance and support for this research from its conception. **Dr Stella Mavroveli (SM)** has been a weekly source of care, support, and expert guidance both for my PhD and for my family life, with her in-depth knowledge of child psychology. **Mr Piers R Boshier (PB)** has been an inspiration from the beginning of my work and helped develop the vision and drive I required to lift it off the ground and was particularly influential in facilitating our trip to Japan to interview Japanese surgeons.

I would also like to express my gratitude to the following people who have helped during the course of my research.

Prof. Young-Woo Kim – facilitation of project within ADDICT trial and review of Chapter 6

Mr John Campolini – for facilitation and support of implementation of the project assessing patient perspective of quality of surgery (Chapter 5).

Oesophageal Patient Association – support of project and facilitation of recruitment of patients for project assessing patient perspective.

Jacqueline Cousins – for guidance in the online systematic searches for both systematic reviews (Chapters 2 and 3).

Dr M Sato – for guidance and support in statistical analysis with G-theory and oesophagogastric oncology trial protocol study (Chapters 3 and 7).

Dr A Vidal-Diez – for guidance in statistical analysis of long-term survival and public availability of protocols (Chapter 3).

Dr Alexander Harris – for provision of oesophagectomy assessment tools for assessment of reliability and review of chapter 7.

Dr Naureen Starling – for facilitation of interview study in recruitment of oncologists (Chapter 4)

Mr Christopher Chan – for facilitation of interview study in recruitment of oncologists (Chapter 4)

Miss Alice Tsai – For active participation and facilitation of the 'Patient perspective' focus group in chapter 5.

Ruth Nicholson – for review of projects by ethics committee

Imperial Information Technology department – for facilitating the up-take of essential software required for qualitative analysis and IT support.

I would like to thank my loving wife, **Nadia Butterworth** who has stood by me as my soul-mate, emotional rock and source of unconditional love through one of the most intense 4-year periods of work I have ever endured.

Table of Contents

Table of Figures

Table of Tables

Chapter 1. Introduction

As the primary curative modality for oesophagogastric cancer the quality of surgical intervention within oncology trials can influence outcomes.[1] Unlike pharmacological interventions which can easily be defined and standardised,[2] quality of a surgical interventions depends on multiple technical variables, health care professionals and resources. A surgical intervention can be affected by the skills, experience, decision-making and preferences of the surgeon.[3] Outcomes of surgical interventions are also dependant on institutional team resources such as intensive care unit beds, imaging services, peri-operative management (e.g. ERAS programmes) and rehabilitation programmes. Those factors and associated quality assurance measures are often poorly accounted for in randomised controlled trials (RCTs) involving surgical interventions.[4]

The recommended surgical management, according to the current United Kingdom (UK) National Institute of Clinical Excellence (NICE) guidelines, for oesophagogastric cancer are dependent on the location, staging and histology of the tumour and often involve multi-modality therapy.[5] For those with gastric cancer suitable for radical resection this may involve gastrectomy (with D2 lymph node dissection) with neo-adjuvant chemotherapy and/or adjuvant chemoradiotherapy. [5] For patients with localised oesophageal and junctional gastro-oesophageal adenocarcinoma undergoing surgical resection (with 2-field lymph node dissection to be considered), they may be offered a choice of neo-adjuvant and adjuvant chemotherapy or neo-adjuvant chemoradiotherapy. [5] Alternatively, for squamous cell carcinoma of the oesophagus the NICE guideline recommends offering patients a choice of radical chemoradiotherapy or nea-adjuvant chemoradiotherapy followed by surgical resection. [5] Following gastrectomy for adenocarcinoma, reported UK 5-year survival rates range from 15% for men and 18% for women.[6] Despite the NICE guidelines, there is still considerable dispute amongst UK surgeons regarding the optimal extent of lymphadenectomy with many performing a compromise D1.5-type gastrectomy, i.e. extended lymphadenectomy, omentectomy but with preservation of the spleen and pancreas.[7] 50 years previously, a Canadian surgeon (Appleby LH) whom advocated block dissection of lymph nodes around the coeliac axis for gastric cancer cautioned: ‘If improvement in the results of surgery for cancer of the stomach is to be brought about, a

much wider and more thorough removal of the drainage area of the stomach must be encompassed than has hitherto been thought possible'. Difference in long-term survival from gastric cancer can be seen internationally and In Japan, where the standard operation for gastric cancer is a radical D2 resection, overall 5-year survival reaches 60–70% in most centres. [7]However, these impressive figures cannot be solely attributed to surgical resection technique, as in Japan gastric cancer is diagnosed at an earlier pathological stage than in the West, with over 50% of patients being diagnosed with early gastric cancer through mass screening programmes.[7] For American Joint Committee on Cancer (AJCC) stage I oesophageal cancer 5 year survival rate is approximately 90 per cent, and decreases to 45, 20 and 10 per cent in patients with stage II, III and IV disease respectively.[8] A recent systematic review and meta-analysis of 12 studies (including 19528 patients) supports a link between extent of lymph node dissection during oesophagectomy (which many surgeons consider a surrogate marker for quality of surgical resection) and survival; higher lymph node yields during oesophagectomy after neoadjuvant therapy were associated with improved overall survival (HR = 0·87; 95% CI: 0·79–0·95, p < 0·001).[9]

Previous narrative reviews have broadly reflected on the challenges to quality of surgery within RCTs assessing surgical interventions including the complex, multi-factorial nature of surgical interventions, and surgeon/experience related factors.[3,10] Quality assurance of surgery has been categorised into three broad domains: credentialing of surgeons/centres; standardisation of surgery: and monitoring of surgical interventions.[1,11]

Recognition of the unique challenges to reporting of RCTs with a surgical intervention led to the revised Consolidated Standards of Reporting Trials Nonpharmacologic treatments (CONSORT NPT) guidelines,[12] which specifies that surgical interventions should be described, along with their method of standardisation and monitoring of adherence to protocol. Although is currently no validated method to assess the 'quality' of trial protocols, in 2013 the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) statement was developed through broad consultation with 115 key trial stakeholders, providing guidance for key protocol content with the aim of facilitating the production of high-quality protocols.[13] Although well recognised and established, the CONSORT NPT and SPIRIT guidelines are often inconsistently followed in practice. A previous systematic review

of trial protocols submitted to the European Organization for Research and Treatment of Cancer (EORTC) noted that although commonly cited in protocols, a description of the surgical technique, definition of resectability, surgical margins and methods of assessing adverse events were inconsistently reported in trial manuscripts.[14] The authors recommended that future trials should define and report details of surgical techniques and quality assurance measures at all stages of an RCT, from protocol development to publication of findings.[14]

The Idea, Development, Exploration, Assessment, Long-term Follow-up (IDEAL) framework has previously been developed for the assessment of new surgical interventions and is based on a five-stage description of the surgical development process. Credentialing of surgeons and surgical centres is not however mentioned within the IDEAL framework. Although standardisation of surgery is referred to and monitoring the quality of surgery is mentioned in regard to expertise-based RCTs, no specific guidance is provided for implementation of these measures.[15]

Previous major oesophagogastric RCTs assessing surgical interventions and neoadjuvant therapy have lacked adequate quality assurance measures, resulting in uncertainty in their findings. [16,17-19] In the Dutch D2 versus D1 gastrectomy study insufficient SQA led to partial homogenisation of the study groups, undermining the likelihood of the trial demonstrating superiority of D2 lymphadenectomy in regard to long-term survival.[20] In the US South-West Oncology Group (SWOG) RCT of adjuvant chemo-radiotherapy versus surgery alone 54% of patients were deemed to have undergone inadequate (D0) resection. [16,18] It was therefore not possible to be certain that this RCT compared patients who underwent similar resections, or whether worse survival in the surgery only arms was in fact the result of inadequate initial surgery.[18] Similarly in the UK Medical Research Council (MRC) trial 75% of the D2 group had inadequate lymphadenectomy hence investigators could not conclude the results of this trial were a sustainable argument against D2 gastrectomy.[19]

SQA in oesophagogastric oncology RCTs has been summarised in a previous systematic review. The authors analysed 160 interventions within trials over the years 2010- 2011 and found standardisation of surgical technique was mentioned for only 29% of interventions.

28% of trials reported measurement of adherence to at least one aspect of the intervention, and for 34% some data was provided regarding the expertise of personnel involved. Authors did not attempt a meta-analysis to assess impact of these factor on trial outcomes within this review.[21] In comparison, Markar et al (2015) systematically reviewed oesophagogastric RCTs looking at SQA measures employed, and on meta-analysis found that the use of credentialing and standardisation correlate with reduced in-hospital mortality and reduction in variation in lymph node yield.[1] No study has yet assessed impact of protocol adherence or SQA implementation on long-term survival outcomes in surgical oncology trials.

Whilst, few oesophagogastric RCTs have historically published their protocols or clearly define implemented SQA measures, more recent trials (e.g. the Randomised trial of Open versus Minimally Invasive Oesophagectomy (ROMIO)) have demonstrate a trend towards more clearly defining these processes.[22] The currently active multicentre Clinical Trial of D1+ Versus D2 Distal Gastrectomy for Stage IB & II Advanced Gastric Cancer (ADDICT), is exemplary in respect to the incorporation of photographic and video surgical quality monitoring methods.[23] Within ADDICT trialists have nevertheless faced challenges in monitoring of surgery and recruitment of patients.

Historically RCTs have utilised clinical, radiological and histopathological outcomes as surrogate measures markers of quality of surgical resection. However, a trend towards increased SQA requirements within surgical oncology trials led to a shift in focus to confirm both the extent and completeness of lymphadenectomy, including an objective assessment of the amount of lymphatic tissue remaining around an anatomical landmark that should have been cleared.[24] This assessment can be achieved through use of a structured objective assessment tool of which allows systematic assessment of surgical performance usually through utilising operative videos or photographs. Such real-time monitoring with a structured objective assessment tool has the advantage over retrospective outcome auditing in that it also allows feedback to operators in order to make adjustments and improve practice. [24] One challenge to SQA in oesophagogastric oncology trials is the lack of a published tool to reliably assess surgical quality. Such tools have been previously developed within the National Training Program (NTP) for Laparoscopic Colorectal Surgery (Lapco) and Trans-anal Total Mesorectal Excision (TaTME) in COlon carcinoma Laparoscopic

or Open Resection (COLOR) III.[25,26] Harris et al developed tools to assess quality of open and minimally invasive oesophagectomy achieving content and face validity using videos of the entire operation and/or photographs of key operative stages.[27,28] However, the reliability of this assessment tool had not been previously demonstrated. Within laparoscopic colorectal surgery, competency assessment tools (CAT) have been successfully employed outside of trials to assess technical performance prior to independent practice on a national scale.[29] In further efforts to improve the quality of colorectal surgery, the Laparoscopic 'Training the Trainer' (Lapco TT) programme was utilised by 65 trainers, leading improved training performance in the short- and long-term with an enhanced the learning curve of delegates' trainees.[30] Future incorporation of an oesophagogastric resection assessment tool, once demonstrated to be reliable, within the national training programme may lead to similar improved surgical performance and associated clinical outcomes.

Unlike surgery, radiotherapy has a long and established record of incorporating quality assurance within routine clinical practice and trials, with recognised affects on clinical trial accrual, cost, outcomes, and generalizability.[31] With a national centralised quality assurance group (the National Cancer Institute Work Group on Radiotherapy Quality Assurance (RQA)) recommending evidence based RQA for trials, surgical oncology trialists could potentially learn from the experiences, methods and strategies developed within radiotherapy RCTs. [31]

Despite efforts to increase patient involvement and participation within healthcare planning and decision making, this has often been lacking in oesophagogastric oncology trials. Patient perception of safety of surgery and quality of care have been reported in previous qualitative studies reinforcing the importance of surgeons' provision of holistic care and listening skills [32,33] However there has been no published study assessing patient opinion of quality of surgery or quality assurance of surgery in trials.

There remains therefore an unmet need to investigate challenges to SQA in oncology trials as defined by all relevant stakeholders. Following this it will also be important to elicit appropriate and acceptable strategies for implementation within oncology trials to overcome the identified challenges. Within section 1 of this thesis we seek to assess the existing literature for previously reported challenges to SQA implementation and protocol

utilisation in surgical oncology trials. In section 2 we will utilise qualitative methodologies to examine expert and patient opinion regarding SQA before seeking to investigate SQA and recruitment within an active trial. Finally, in section 3 we move to the assessment of reliability of an SQA tool within an active RCT.

1.1. Aim

To systematically investigate quality assurance of surgery in oesophagogastric oncology trials and develop a robust and feasible framework of consensus strategies to overcome identified challenges in order to aid design and implementation of SQA in trials.

1.2. Objectives

The specific objectives of this thesis were:

i. Assessment of the existing literature for SQA in surgical oncology trials
ii. Examination of expert opinion on challenges to SQA
iii. Exploration of patient opinion on challenges to quality of surgery
iv. Investigation of SQA and recruitment within an active trial
v. Assessment of the reliability of a SQA tool

Section 1.

Chapter 2. Challenges to quality assurance of surgical interventions in oncology trials: a systematic review

2.1. Introduction

Randomised controlled trials remain the gold standard for assessing the efficacy of interventions within the field of surgical oncology. Where surgery forms the primary curative modality in such trials the quality of this intervention has the potential to directly influence outcomes. [34]

There have been several reviews which have looked at SQA measures in oncology trials including the CONSORT Non-pharmacological treatment guidelines and the IDEAL framework, in addition to a previous systematic review of large colorectal trial (Summarised in Table 1).[11,14,15,35] Limitations of these reviews include that the proposed SQA measures have not yet to been validated or received trial stakeholder assessment as to their usability within RCTs. Short term clinical and pathological outcomes have been shown to be improved with utilisation of SQA measures within oesophagogastric oncology trials, [1] however this study was limited by few RCTs reporting adherence to trial protocol or indeed having a trial protocol available for evaluation.[34]

Despite acknowledgement of the importance of quality assurance measures in surgical trials methods of monitoring and standardisation in trials are inconsistently reported.[14] Few have proposed reasons for its inadequate utilisation and to date there has been no comprehensive review of this subject. This systematic review intends to identify published barriers and challenges to implementation of quality assurance of surgical interventions within surgical oncology trials. [34]

2.1.1. Objective

The specific study objective was:

i. Assessment of the existing literature for SQA in oncology trials

Table 1: Published potential strategies to improve quality assurance of surgery in trials [3-5]

Published potential mitigating strategies	
Generic	Culture of cooperation academically
	Secure adequate funding
	Improved epidemiological education of surgeons
Credentialing	Review Operative reports
	Assessing case/procedural volume
	Statistically assess learning curve with hierarchical model
	Live observation first five operations
	Video assessment – objective and validated assessment tools
	Live operating assessment
	Review Operative reports
Standardisation	Pre-trial education through written information, videos or demonstrations and workshops
	Clearly define intervention and comparator interventions and when they can be tailored to individual patients
	Meeting of surgeon investigators – pre-trial and regularly throughout
	Standardisation of surgical approach through trial protocol
	Standardisation of extent of lymphadenectomy
	Clearly define intervention and comparator interventions and when they can be tailored to individual patients
Monitoring	Detail method of assessing adherence
	Video and photographic assessment
	Record indicators for surgical decision making (e.g conversion from laparoscopic to open)
	Regular audit of data including clinical outcome measures
	Record an obvious measure of quality (e.g Lymph Node yield)
	Surgical details captured on standardised data collection form

2.2. Methods

2.2.1. Search Strategy and Selection Criteria

A systematic on-line literature search of the Medline (1946- 4th February 2018) and EMBASE (1974 – 4th February 2018) was conducted using the following key words and relevant MeSH terms: 'challenge'; 'limitation'; 'problem'; 'barrier'; 'quality'; 'quality assurance'; 'randomised controlled trial'; 'clinical trial'; 'surgery'; 'procedure'; 'surgical', and; 'performance' used in combination with the Boolean operators AND and OR. Full details of the search strategy are available (Appendix A). Inclusion criteria defined eligible studies as being oncology RCTs with a surgical intervention and/or associated publications, including published trial protocols and secondary articles presenting data collected from the primary trial, that were relevant to the research question: 'Challenges to quality assurance of surgery in oncology trials.' Exclusion criteria were: studies not published in the English language; not relevant to study question; conference abstracts and proceeding; studies assessing endoscopic interventions; animal studies and non-randomised studies. [34]

Two reviewers (JB and PB) independently screened the titles and abstracts of studies identified through the electronic search using the above criteria. The full texts of potentially relevant articles were obtained and reviewed.

Adherence to this conventional method of performing a systematic literature search identified no directly relevant randomised control trials, necessitating an alternative approach. Eligibility criteria were therefore expanded to include studies reported within review articles, including systematic reviews and meta-analyses identified by the original search.[34] This approach has previously been shown to be effective, particularly in circumstances where the intended research question is not the primary focus of the original RCTs or study.[36,37]

Two independent reviewers (JB and PB) hand searched the reference lists of review articles identified through the primary search to identify potentially relevant primary studies. The

full text articles of identified studies were subsequently obtained to determine eligibility for inclusion. Any disagreement between reviewers was resolved by a third researcher (SM). [34]

Selected articles were screened to identify reported challenges to quality assurance of surgical intervention within their trials. Study protocols identified from the literature search helped in identifying quality assurance measures that were put in place at trials' inception. Quality assurance measures utilised within RCTs were assessed and classified according to 3 previously published domains: credentialing of surgeons/trial centres, standardisation of operative intervention and monitoring of surgery.[11,34]

Methodological quality of included RCTs was evaluated using a validated quality assessment tool, the Cochrane Collaboration Higgins tool (Appendix B), whilst trial protocols were assessed using the SPIRIT guidelines. [38,13] Results were presented in accordance with PRISMA guidelines for the preferred reporting of systematic reviews and meta-analyses.[34,39]

2.3. Results

Results of the online literature search, including assessment of relevant review articles and primary studies, are outlined in Figures 1 and 2. In total 13 review articles (2 systematic reviews, 11 narrative reviews) relevant to the study question were identified (Figure 1). After hand searching the reference lists of these review articles, 34 articles met criteria for inclusion (Figure 2), [16,20,40-71] including 19 RCTs, 11 further analyses of the primary RCTs, and 4 trial protocols. Details of included studies are summarised in table 2. RCTs reflected a range of surgical oncology disciplines including: colorectal (n=11); oesophagogastric (n=5); dermatology (n=1); hepatobiliary (n=1) and gynaecological (n=1). [34]

2.3.1. Assessment of biases

In assessing risk of bias in RCTs (Appendix B), two (10%) were found to be low risk and seven (37%) trials were considered to be high risk. Due to insufficient reporting of trial methodology potential bias within 9 (53%) trials was deemed to be unclear. Within the 19 included RCTs, allocation concealment was rated unclear in 9 (47%) and low risk in ten (53%) trials. Performance bias was rated unclear in 13 (68%), low risk in three (16%) and high risk in three (16%) trials. Finally, blinding of outcome assessment was rated as unclear in 11 (63%), low risk in three (16%) and high risk in four (21%) trials. [34]

Only four published protocols were available following review of published literature and contacting authors of published RCTs. The 33-step SPIRIT guidance tool was utilised to assess the quality of protocols within the EnROL, CLASSIC, LAFA and the Dutch D1 versus D2 gastrectomy trial protocols, achieving scores of 11.6 (39%), 10.8 (33%), 16.3 (49%) and 8.58 (26%) out of a maximum score of 33 (100%) respectively.[44,53,60,68] One retrospective methodological study was not suitable for formal quality assessment.[34,64]

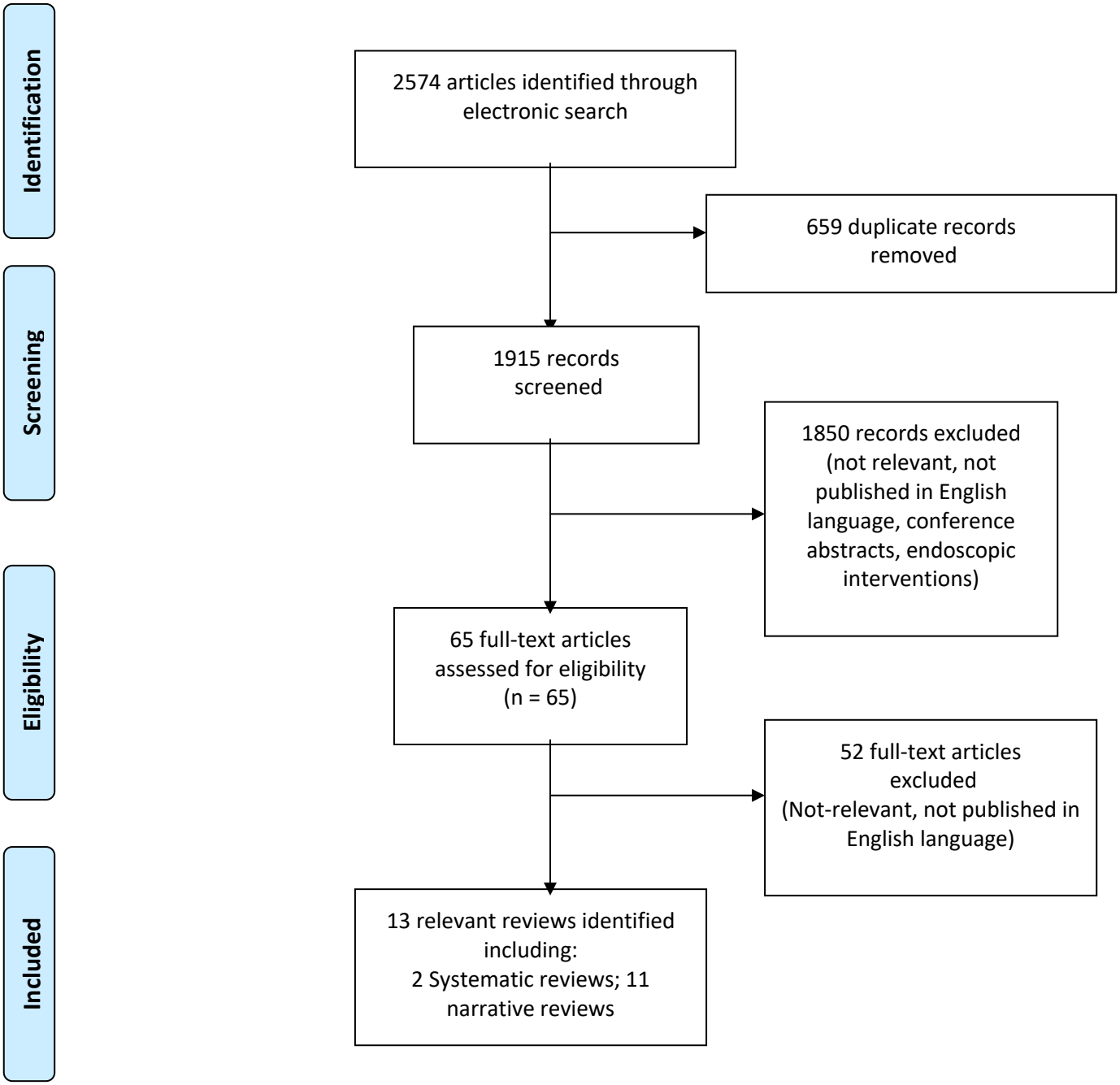

Figure 1: Flow-diagram of search strategy identifying relevant reviews presented in accordance with PRISMA guidelines [34,39]

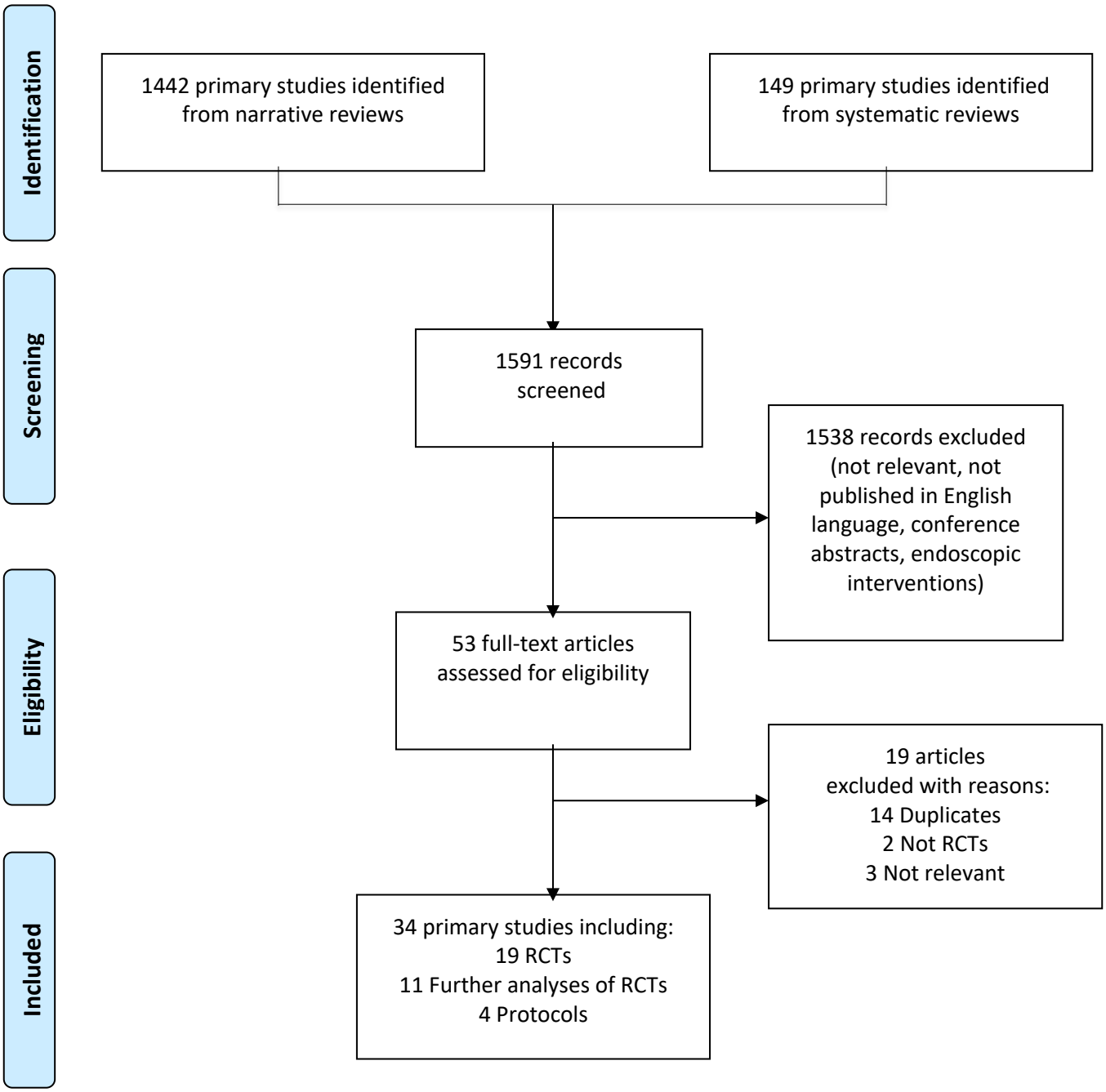

Figure 2: Flow-diagram of search strategy identifying primary studies presented in accordance with PRISMA guidelines[34,39]

2.3.2. Reported SQA measures

Table 3 summarises quality assurance measures utilised in clinical trials and their protocols. Four RCTs (21%) utilised an operation manual whilst in seven (37%) trials surgeons were informed of operative principles required within trial operations. Credentialing by either surgeon case or centre volume was conducted in nine (47%) trials. Alternative credentialing methods utilised by two (11%) studies included the requirement of participation in a previous trial and video assessment. Trial monitoring was accomplished through either: central review of operative and histopathology reports (5 trials, 26%); direct intraoperative monitoring (1 trial, 5%), or; auditing of conversion to open rates (1 trial, 5%).[34]

2.3.3. Reported challenges to SQA

Reported challenges to quality assurance in surgical trials are summarised in Table 4.

Challenges to credentialing methods were reported in six (32%) trials. Two (11%) trials considered selected centres to be 'specialist centres' without information on performance at either an institutional or surgeon specific level. [65,71] Four colorectal trials (21%) credentialed surgeons according to cumulative case volume using 20 laparoscopic colorectal cases per surgeon or centre as the threshold for recruitment, but later reflecting that this limited effective selection of surgeons. [34,53,57,61,66]

Challenges to standardisation of surgical interventions included: variation in definition and execution of surgical interventions between trial centres was reported in five (26%) trials; [16,20,52,55,58] additional unspecified surgical component tasks were found to be commonplace within two (11%) trials investigating D1 versus D2 gastrectomy;[19,49] and key operative steps were left to the discretion of the operating surgeons in two (11%) trials raising concerns over standardisation methods.[42,65] Retrospective case registration and non-specific protocols were also noted as key barriers to standardisation.[16,34,42]

Challenges to monitoring included: one D1 versus D2 gastrectomy trial (5%) reporting insufficient training and monitoring of surgeons;[20,48] and insufficient quality control of pathological assessment of resected specimens that was reported in one trial (5%).[34,46]

Table 2: Included Primary Studies and Risk of Bias assessment/Quality Assessment Summary [34]

Author	Trial	Country	Year	Sample	Design	Single/ multicentre	Risk of bias/QA tool	Assessment Risk Bias/QA	Ref
Balch	Immunotherapy/Melanoma	US	1982	260	RCT	Multicentre	Higgins	High	[40]
Balch§	Immunotherapy/Melanoma	US	1983	136	RCT	Multicentre	Higgins	High	[41]
Gerard	EORTC Rectal Ca	Belgium	1988	466	RCT	Multicentre	Higgins	Unclear	[42]
Krook	Adjuvant therapy Rectal Ca	US	1991	204	RCT	Not specified	Higgins	High	[43]
Sasako*	Dutch D1 vs D2 Gastric Ca	Holland	1992	NA	Protocol	Multicentre	SPIRIT	26%	[44]
Bonenkamp	Dutch D1 vs D2 Gastric Ca	Holland	1995	711	RCT	Multicentre	Higgins	Unclear	[45]
Bunt*	Dutch D1 vs D2 Gastric Ca	Holland	1996	237	RCT	Multicentre	Higgins	Unclear	[46]
Sasako*	Dutch D1 vs D2 Gastric Ca	Holland	1997	711	RCT	Multicentre	Higgins	Unclear	[47]
Bonenkamp*	Dutch D1 vs D2 Gastric Ca	Holland	1998	711	RCT	Multicentre	Higgins	Unclear	[48]
Bonenkamp*	Dutch D1 vs D2 Gastric Ca	Holland	1999	711	RCT	Multicentre	Higgins	Unclear	[49]
Hartgrink*	Dutch D1 vs D2 Gastric Ca	Holland	2004	711	RCT	Multicentre	Higgins	Unclear	[50]
Songun*	Dutch D1 vs D2 Gastric Ca	Holland	2010	711	RCT	Multicentre	Higgins	Unclear	[20]
Cuschieri	MRC D1 vs D2 Gastric Ca	UK	1996	400	RCT	Multicentre	Higgins	Unclear	[51]
Cuschieri†	MRC D1 vs D2 Gastric Ca	UK	1999	400	RCT	Multicentre	Higgins	Unclear	[52]
Gouillou (CLASSIC)	Conventional vs Lap-assisted Colorectal Ca	UK	1996	NA	Protocol	Multicentre	SPIRIT	33%	[53]
Tveit	Adjuvant Surgery vs Surgery Rectal Ca	Norway	1997	144	RCT	Multicentre	Higgins	Unclear	[54]
Holm	Neoadjuvant radiotherapy Rectal Ca	Sweden	1997	1399	RCT	Multicentre	Higgins	Unclear	[55]
MacDonald	Adjuvant chemoradiotherapy GOJ Ca SWOG-9008	US	2001	556	RCT	Multicentre	Higgins	High	[16]
Hundahl‡	Adjuvant chemoradiotherapy GOJ Ca SWOG-9008	US	2002	556	RCT	Multicentre	Higgins	High	[56]
Hazebroek (COLOR)	Colon Cancer Lap vs Open	Netherlands	2002	850	RCT	Multicentre	Higgins	High	[57]
Sano T	D2 vs Para-aortic lymphadenectomy Gastric Ca	Japan	2004	523	RCT	Multicentre	Higgins	Unclear	[58]
Veldkamp||	Colon Cancer Lap vs Open	Netherlands	2005	627	RCT	Multicentre	Higgins	High	[59]
Wind (LAFA)	Open/Lap +/- Fast track Colon Ca	Netherlands	2006	NA	Protocol	Multicentre	SPIRIT	49%	[60]
Fleshman J (COST)	Lap vs Open Colectomy Colon Ca	US	2007	872	RCT	Multicentre	Higgins	Unclear	[61]
Wright	Education LN assessment Colon Ca	Canada	2008	42 centres	Cluster RCT	Multicentre	Higgins	High	[62]
Hewett (ALCCAS)	Lap vs Open Colon Ca	Australasia	2008	601	RCT	Multicentre	Higgins	High	[63]
Allardyce (ALCCAS)	Lap vs Open Colon Ca	Australasia	2008	592	RCT (methodology)	Multicentre	NA	NA	[64]
Kitchener	Endometrial Ca (MRC ASTEC)	UK	2009	191	RCT	Multicentre	Higgins	Unclear	[65]
Neudecker	Lap versus Open Colorectal Ca	Germany	2009	679	RCT	Multicentre	Higgins	Unclear	[66]
Simunovic	Quality initiative Rectal Ca	Canada	2010	105	Cluster RCT	Multicentre	Higgins	Unclear	[67]
Kennedy (EnROL)	Lap vs Open Colorectal Ca	UK	2012	NA	Protocol	Multicentre	SPIRIT	39%	[68]
Van der Pas (COLOR II)	Lap vs Open rectal Ca	Netherlands	2013	1103	RCT	Multicentre	Higgins	High	[69]
Degiuli	D1 vs D2 Gastrectomy	Italy	2014	267	RCT	Multicentre	Higgins	Low	[70]
Primrose	Chemo +/- cetuximab colorectal liver metastasis	UK	2014	257	RCT	Multicentre	Higgins	Low	[71]

*Studies relating to and including Bonenkamp JJ et al (1995).[45] † Studies relating to and including RCT Cuschieri A et al (1996) [19] ‡Studies relating to and including RCT J.S. Macdonald et al (2001).[16] § – studies relating to and including Balch et al (1982).[40] ||studies related to and including Hazebroek EJ et al (2002).[57] Abbreviations: Ca – Cancer; Lap – Laparoscopic; GOJ – Gastro-oesophageal junction; QA – Quality Assurance

Table 3: Reported quality assurance measures [34]

Author	Credentialing	Standardisation	Monitoring	Ref
Balch	none	none	review of operative reports	40
Balch§	none	none	review of operative and pathology reports	41
Gerard	none	specification of operative principles	None	42
Krook	none	none	None	43
Sasako*	none	operative manual	case supervision	44
Bonenkamp	none	operative manual	case supervision	45
Bunt*	none	none	none	46
Sasako*	none	none	none	47
Bonenkamp*	none	none	none	48
Bonenkamp*	none	none	none	49
Hartgrink*	none	none	none	50
Songun*	none	none	external expert review	20
Cuschieri	none	operative manual	none	51
Cuschieri†	none	none	none	52
Gouillou (CLASSIC)	cumulative case volume >20 (surgeon)	surgery according to 'current practice' (none)	central review of operative and pathology reports	53
Tveit	none	none	none	54
Holm	none	none	none	55
MacDonald	none	recommendation of important operative step(s)	review of operative and pathology reports	16
Hundahl‡	none	none	none	56
Hazebroek (COLOR)	cumulative case volume >20 (team)	live and virtual demonstrations for surgeons	none	57
Sano T	cumulative surgeon volume >100, or annual centre volume >80	surgeons agreed to operative steps prior to trial commencement	3 times per year meetings including operative video and LN/pathology review	58
Veldkamp||	cumulative case volume >20 (surgeon), video assessment	operative manual. standardised pathology reporting	none	59
Wind (LAFA)	cumulative case volume >25 (surgeon)	none	none	60
Fleshman J (COST)	submission of 20 operative reports, video assessment	none	none	61
Wright	none	standardised pathology reporting	none	62
Hewett (ALCCAS)	none	none	auditing and analysis of conversion to open	63
Allardyce(ALCCAS)	none	none	auditing and analysis of conversion to open	64
Kitchener	surgeons defined as 'specialist' with 'experience' (none)	specification of important operative step(s)	none	65
Neudecker	cumulative case volume >20 (surgeon)	operative principles agreed by participating surgeons	none	66
Simunovic	none	intraoperative demonstration of optimal resection	none	67
Kennedy (EnROL)	surgeons cumulative case volume > 100 (lap), >50 (open)	surgery carried in 'standard fashion' (none)	none	68
Van der Pas (COLOR II)	submission of operative video(s), direct observation	specification of important operative step(s)	none	69
Degiuli	participation in prior trial	none	appraisal. pathology reports	70
Primrose	centres defined as 'specialist' (none)	operative principles specified	case notes review	71

*Studies relating to and including Bonenkamp JJ et al (1995).[45] † Studies relating to and including RCT Cuschieri A et al (1996) [19] ‡Studies relating to and including RCT J.S. Macdonald et al (2001).[16] § – studies relating to and including Balch et al (1982).[40]|| studies related to and including Hazebroek EJ et al (2002).[57]

Table 4: Reported challenges to quality assurance in surgical trials [34]

Author	Reported challenges to quality assurance in surgical trials	Ref
Balch	Insufficient standardisation of surgical intervention and led to trial participants receiving inappropriate treatment	40
Balch§	Insufficient standardisation of surgical intervention and led to trial participants receiving inappropriate treatment	41
Gerard	Insufficient standardisation: Non-specific Radical dissection at 'discretion of surgeon'	42
Krook	Insufficient standardisation: non-standardised conventional surgical techniques likely contributing to high local recurrence rate	43
Sasako*	Insufficient credentialing: - 'Dutch surgeons unfamiliar with R2 gastrectomy.' Potential for wide variety of methods of lymphadenectomy	44
Bonenkamp	When D2 had worse results it was claimed trial surgeons were still learning the procedure – learning curve causing performance bias.	45
Bunt*	Insufficient trial standardisation of pathological assessment: Significant difference in mean lymph node yield between surgeon and local pathologist	46
Sasako*	Insufficient credentialing: - 'Dutch surgeons unfamiliar with R2 gastrectomy. Insufficient standardisation: Routine use of splenectomy and pancreatectomy in the D2 procedure	47
Bonenkamp*	Insufficient monitoring: insufficient time for supervising Japanese Cancer surgeons to train local Dutch surgeons	48
Bonenkamp*	Insufficient credentialing: 'trial surgeons were still learning the procedure'. Insufficient standardisation of lymph node assessment postoperatively.	49
Hartgrink*	Insufficient standardisation and credentialing: additional procedures in D2 dissection so as not to compromise an adequate dissection of lymph node stations 10 and 11	50
Songun*	Insufficient standardisation: Definition and the execution of a procedure varied between centres. D2 procedures performed in the Dutch trial were actually closer to a D1	20
Cuschieri	Insufficient standardisation led to unnecessary splenectomies and pancreatectomies performed	51
Cuschieri†	Insufficient standardisation led to unnecessary splenectomies and pancreatectomies performed	52
Gouillou (CLASSIC)	Credentialing case threshold underestimated	53
Tveit	Insufficient standardisation and monitoring	54
Holm	Insufficient standardisation: Significant surgeon-related variation in patient outcome likely related to surgical technique	55
MacDonald	Insufficient standardisation and monitoring contributing to 54% of cases in this trial underwent a D-0 lymphadenectomy (i.e. less than D-1 lymphadenectomy)	16
Hundahl‡	Insufficient standardisation: Based on the Maruyama Index analysis - survival is compromised by 'too little' lymphadenectomy for the extent of disease	56
Hazebroek (COLOR)	Credentialing case threshold underestimated	57
Sano	Challenge to standardisation – 8 protocol violations with nodal stations No.13 and/or No.14v not dissected in distal third tumours.	58
Veldkamp||	Insufficient standardisation and credentialing: 'operating time varied substantially between centres' implying differing levels of surgical experience between different centres.	59
Wind (LAFA)	Credentialing case threshold underestimated	60
Fleshman J (COST)	Challenge to credentialing: 'time-consuming for participating surgeons to provide the documents and videos' and for the review team to complete to approve surgeons	61
Wright	Insufficient standardisation: Academic detailing of the local opinion leader or provision of a toolkit did not improve quality of pathological lymph node assessment.	62
Hewett (ALCCAS)	Credentialing case threshold underestimated	63
Allardyce(ALCCAS)	Challenge to credentialing: 'Surgeons have been reluctant to refer cases or become credentialed'.	64
Kitchener	Insufficient credentialing: 'some surgeons less good at pelvic lymphadenectomy'. Insufficient standardisation: 'Periaortic lymph nodes reportedly were left in situ '	65
Neudecker	Credentialing case threshold underestimated	66
Simunovic	Insufficient credentialing: Hospitals eligible to participate if 60 % of their surgeons agreed to participate	67
Kennedy (EnROL)	Challenge to credentialing: Protracted learning curve in laparoscopic colorectal surgery	68
Van der Pas (COLOR II)	Insufficient standardisation: 'Lack of standardisation of perioperative protocols because it was not feasible in a study undertaken at 30 centres and hospitals in eight countries.'	69
Degiuli	Insufficient standardisation: Rate of contamination in the D1 group (18%), and noncompliance in the D2 group (34%)	70
Primrose	Insufficient credentialing: surgeons from all liver centres recruited and only 33% of patients were reported to have R0 (curative) resections.	71

*Studies relating to and including Bonenkamp JJ et al (1995).[45]† Studies relating to and including RCT Cuschieri A et al (1996) [19]‡Studies relating to and including RCT J.S. Macdonald et al (2001).[16]§ – studies relating to and including Balch et al (1982).[40] ||studies related to and including Hazebroek EJ et al (2002).[57]

2.4. Discussion

This is the first systematic review to summarise reported challenges to quality assurance of surgical interventions in oncology trials. Challenges to quality assurance when selecting surgeons included the protracted learning curve, particularly for laparoscopic procedures and the use of low case threshold when credentialing surgeons for trial registration. [53,57,61,66] In respect to standardisation of surgical interventions differences in definition and execution of surgical procedures between centres, [16,20,51,55,58] retrospective case registration and non-specific protocols were noted as key barriers.[16,42] With the exception of efforts to compare operative notes and pathology practices,[16,40,41,53,70] monitoring of surgical interventions was found to be lacking or inadequate with one large, multicentre trial acknowledging that there was insufficient time to adequately monitor complex operations. [20,48] These findings represent only those challenges that are reported within the literature and therefore may represent a narrow window on to the much wider issue of SQA in RCTs. [34]

There are several factors that present challenges in credentialing surgeons to enter oncological RCTs. When the RCTs identified by this review were conducted, there was no agreed methodology for assessing surgical competency. Although case volume reflects experience with its known link to improved clinical outcomes, this relationship is not linear and varies depending on the operation. Case volume therefore remains only a surrogate marker of surgical competency. A number of colorectal trials utilised 20 laparoscopic colorectal resections as the threshold for recruitment. [53,57,61,66] This figure has since been disputed by the finding of more recent trials and cohort studies that have estimated the learning curve for laparoscopic colorectal cancer surgery to be between 150 to 200 cases. [25,68] As illustrated by the ASTEC and EPOCH trials, [65,71] broad credentialing methods do not necessarily confer standards of surgical skills or competence. Within the ASTEC pelvic lymphadenectomy trial any centre that employed one or more ‘gynaecology oncology’ clinician could participate, whilst it was later observed by the author’s, that some surgeons were simply “not good at pelvic lymphadenectomy”. The critical importance of credentialing of surgeons for trials was demonstrated again in the US SWOG 9008/Intergroup 0116 RCT comparing chemoradiotherapy after surgery with surgery alone for adenocarcinomas affecting either the stomach or gastroesophageal junction.[16] To assist recruitment of

surgeons into this trial there was no formal credentialing process adopted. It was thus difficult to standardise the surgical procedures and authors reflected this likely contributed to the high D0 resection rate (54%) and ensuing difficulty in interpreting the results.[72] Provision of an objective methodology and auditing of clinical outcomes are proven methods that can support optimum performance. An inherent sensitivity that surrounds subjecting specialist surgeons to an assessment process of technical performance should be discouraged in surgical culture. Assessment of surgical performance should be predominately about standardising the execution of the intended procedure. The outcome of any assessment should not be considered simply as a pass or fail. Surgeons should be encouraged to participate in RCTs and supported to acquire the necessary standard required for entry into the trial. [34]

Over one third of RCTs utilised no methods of standardisation of surgical intervention. This may have contributed to the high local recurrence rates reported in three colorectal trials that ranged from 20-65%. [43,61,62] In the treatment of gastric cancer the Dutch and MRC trials failed to show a benefit with extended lymphadenectomy, that may have been attributable to high associated perioperative morbidity (43%–46%) and mortality (10%–14%).[20,52] A retrospective review of quality control within the Dutch trial [20] by an expert Japanese surgeon found that some purported D2 gastrectomies performed in this trial were in fact closer to D1 procedures. Insufficient standardisation within the Italian IGCSG-R01 trial was also likely to have contributed to the observed small difference in the number of retrieved lymph nodes between treatment arms with D1 and D2 lymphadenectomies.[70] Protocol violations were reported in the D2 versus D2 plus para-aortic lymphaenectomy trial for gastric cancer RCT in which 8 (1.5%) participants did not have nodal stations No.13 and/or No.14v dissected in distal third tumours.[58] The principle investigator of the MRC gastrectomy trial highlighted the needs for the surgical community to address the challenges in training standards necessary for D2 gastrectomy and the quality issues required for optimum surgical performance in oncological RCTs.[73] Similar challenges in standardising surgery in RCTs have been faced across different specialities. In the ASTEC trial it was later reported that the number of the retrieved lymph nodes was inadequate, and that periaortic lymph nodes were left in situ.[65] Furthermore the inclusion rate of what were considered to

be low-risk patients was recognized to be high. [34]

Approximately two thirds of trials did not report any attempt to monitor of surgical interventions. Within the studies identified by this review monitoring strategies included: review of operative reports and pathology forms for completeness of resection; comparing sites of disease at baseline and within the operative reports, and; central pathology review of trial specimens. Only five (26%) identified trials made efforts to standardise the histopathological assessment of surgical specimens. These methods may be considered surrogate measures of surgical quality but they cannot provide a direct measurement for surgical quality.[34] Further barriers to greater adoption of monitoring within surgical trials are the availability of both trained assessors and validated assessment tools. Assessors should ideally be experienced surgeons or pathologists. Surgeons who are fulfilling the role of an assessors should be blinded to operator identify and should themselves not contribute to study recruitment. Assessment tools should be validated and been shown to demonstrate satisfactory inter-observe agreement. [34] Public attention has recently focussed on surgical complications as one of the 30-40% of major adverse events considered preventable and 'proactive surveillance' is often considered one of the key potential mitigating strategies.[74] In routine clinical practice peer review of operative interventions has not been regularly adopted. However clinical surveillance by peers who understand the clinical process is the only systematic method of assessing adverse events shown to be correlated with subsequent improvement in clinical outcomes.[74,75]

Several publications have recommended the use of existing strategies such as the CONSORT Non-Pharmaceutical Treatment for reporting of clinical trials and the IDEAL framework for evaluation of surgical innovations as platforms that could improve the quality of surgical interventions in RCTs.[12,15] Suggestions for credentialing include assessing case volume; live theatre assessment; submitting videos of a number of operations for assessment and auditing operative records. Recommendations to aid standardisation include pre-trial education using written information, videos, demonstrations or workshops; establishing a trial protocol; defining interventions and determining when they can be tailored to individual patients and standardising the extent of oncological resection. Monitoring proposals have comprised video and photographic assessment; recording indicators for

surgical decision making (e.g conversion from laparoscopic to open); regular audit of practice including clinical outcome measures; recording oncological measures of quality (e.g lymph node yield); capturing surgical details on standardised data collection forms; and centralisation of pathological assessment.[11,12,15] Some trial organisations have already adopted surgical quality control tools such as the SURCARE platform. This has been recommended as a platform with necessary competencies and facilities to organise pragmatic clinical research in surgical oncology. It consists of a collaborative network comprising expert surgical centres/organisations (e.g. EORTC and Japanese Clinical Oncology Group (JCOG)) in Europe and Japan with a focus on quality assurance.[76] Within JCOG, quality control in oncology trials has been developed around central review of pathology and operative images by trial committees.[77] The EORTC surgical quality assurance sub-committee has proposed a risk based approach in quality assurance for surgery in which credentialing, standardisation and intra-operative and central committee review are recommended.[78] However, no published study has yet assessed acceptability or feasibility of those mitigating strategies with key stakeholders and trialists. Without strong stakeholder engagement such strategies will be unlikely to result in the paradigm shift required to significantly improve quality of surgery in clinical trials. Previously suggested methods of stakeholder engagement within oncology trials have included utilising groups of patients, pharmacists and ethicists to help improve trial design, implementation and dissemination of trial findings. [34,79,80] We recommend a qualitive study involving a Delphi consensus process with key oncology trial stakeholders including trial methodologists, oncologists, surgeons and trial managers. This would allow identification of challenges to quality assurance of surgery in oncology trials and development of a framework of mitigating strategies. [34] As the conduct of surgical oncology trials become more complex there is a need to adapt surgical quality assurance measures to make implementation more feasible. To facilitate this goal, a framework of strategies to improve surgical quality assurance (SQA) in oncology trials with broad trial stakeholder consensus is imperative. [34]

This review is limited to the published data reporting challenges to SQA within RCTs. Although published reported challenges to quality assurance are summarised in this review, it is expected that there remain many undocumented constraints experienced by stakeholders conducting those trials. Whilst acknowledging the challenges of real-world

surgical practice we still believe that surgical trials are beholden to demonstrate the efficacy of a surgical technique under optimum circumstances. Non trial surgeons should be encouraged, trained and supported to achieve the same standards. SQA and the challenges encountered during its implementation should be considered a standard element of reporting in RCTs. [34]

2.5. Conclusion

Trial units have strict measures for quality assurance in recruitment, randomisation and data management, however, they do not routinely implement quality measures for the principle intervention (i.e. surgery) that is known to be highly variable and surgeon dependent. The problem is compounded by the fact that craft-based interventions are fundamentally different from pharmacological therapies, commonly employed by trial units. It is important for trial units to employ methodologists who could design and implement SQA measures. The surgical community should also acknowledge the challenges facing quality assurance of surgical interventions and work collaboratively with other trial stakeholders to enable implementation of mitigating strategies to improve surgical quality in oncology trials. [34] Within this chapter we have conducted an in-depth analysis of previously reported challenged to SQA in oncology trials. Due to the likely insufficient reporting of such challenges within published literature, a comprehensive analysis of SQA within oesophagogastric oncology trials requires a different approach. This leads us to the next chapter in which all oesophagogastric RCTs meeting certain criteria within a specific time period will be further examined.

Chapter 3. Quality Assurance of Surgery in Oesophagogastric Oncology trials and trial protocols: a systematic review and meta-analysis

3.1. Introduction

On reviewing published strategies to quality assurance of surgery in oncology trials (Chapter 2), a diverse range of challenges were identified in addition to a paucity of SQA initiatives. Of the 19 RCTs that were identified by the literature search only four (21%) utilised an operation manual, nine (47%) credentialled surgeons and only one trial (5%) utilised direct operative monitoring strategies (Chapter 2).

Consistent with these findings is the recognition that landmark oesophagogastric oncology trials have in retrospect lacked adequate processes for SQA including: credentialing, standardisation and monitoring.[20,81] Despite the lack of adoption within trials, standardisation of surgical techniques and credentialing of surgeons have been found to be associated with reduced adjusted in-hospital mortality in oesophagogastric oncology trials.[1] A comprehensive assessment of the implementation of SQA measures within oesophagogastric oncology trials, and their influence on survival has yet to be reported. Furthermore, no study in oesophagogastric oncology trials has yet assessed the impact of protocol availability on patient survival.

This systematic review and meta-analysis intends to evaluate SQA measures within major oesophagogastric oncology trials and their protocols in order to determine their influence on overall survival.

3.1.1. Objective

The specific study objective was:

i. Assessment of the existing literature for SQA in oncology trials

3.2. Methods

3.2.1. Search strategy and inclusion criteria

A systematic on-line literature search of the Medline (January 2000 to February 2018) and EMBASE (200--February 2018) databases was conducted using the following key words and relevant MeSH terms: 'randomised controlled trial'; 'clinical trial'; 'surgery'; 'oesophagogastric cancer'; 'gastric cancer'; 'oesophageal cancer'; 'gastrectomy'; 'oesophagectomy'; 'carcinoma'; 'oncology'; used in combination with the Boolean operators AND and OR. Details of the search strategy are provided in appendix C. The Scottish Intercollegiate Guideline Network search filter for Randomised Controlled Trials was utilised to maximise search breadth.[82]

Inclusion criteria comprised of the following: Oesophagogastric oncology RCTs (published after 1st January 2000 up until 1st February 2018) with surgery that had curative intent in at least one study arm; a minimum of 100 participants in each study arm, and; reporting of ≥3 years overall survival. Exclusion criteria: review articles; articles not published in the English language, and; conference abstracts. In cases where multiple publications relating a single RCT were identified, their outcomes were considered as a single trial/publication for the purpose of this review.

The original trial protocols for all included RCTs were sought through either: online databases (Medline, EMBASE); online trial registries, and; handsearching reference lists. Where protocols could not be obtained using these search strategies the corresponding author of the principal trial manuscript was contacted directly via email.

Two reviewers (JB and PB) independently screened the titles and abstracts of studies identified through the electronic search using the above criteria. The full texts of potentially relevant articles were obtained and reviewed. Published RCTs and their protocols relevant to the review question were assessed by two independent researchers (JB, PB) for pertinent trial characteristics and use of SQA measures. In RCTs where surgery was performed in only one study arm, only the outcomes of the surgical arm were considered. For RCTs comparing two surgical interventions, only the outcomes of the primary intervention under assessment

were considered. Data extracted from RCTs and their protocols included: date of RCT; region of origin; sample size; average subject age; body mass index; gender; tumour site; tumour histology; use of neoadjuvant therapy; surgical approach; TNM staging; lymph node status; post-operative complications; 30-day mortality; long-term overall survival, and; SQA measures utilised within the trial and trial protocol. Outcomes were reported in accordance with PRISMA guidelines (Figure 2).[39]

3.2.2. Definitions

Oesophagogastric cancer: Any malignant tumour of the upper gastrointestinal tract from the upper third of the oesophagus (starting below the cricopharyngeus) to the pylorus. Randomised controlled trial: Experimental study in which participants are randomly assigned to one of multiple arms receiving different treatments. Long-term overall survival: Overall survival of participants from time of surgery.

3.2.3. Assessment of SQA measures within trial protocols and RCTs

SQA measures were both counted cumulatively within each manuscript and assessed within RCTs and their protocols using a three-point scoring tool developed using existing published literature (Chapter 2, Table 1). This tool allocated points to studies based on their adoption of at least one SQA measure (maximum score: 3/3) within each of the three previously established SQA domains: credentialing, standardisation and monitoring (See Table 5).[11] This allowed both the overall number of SQA measures and the breadth of SQA within recognised categories to be accounted for within each trial.

Table 5: SQA assessment tool

		SQA measure utilised
Credentialing	Review Operative reports	
	Assessing case/procedural volume of surgeon or centre	
	Statistically assess learning curve with hierarchical model	
	Live intraoperative observation	
	Video assessment – objective and validated assessment tools	(…./1)
Standardisation	Pre-trial education through written information, videos or demonstrations and workshops	
	Clearly define intervention and comparator interventions and when they can be tailored to individual patients (CONSORT)	
	Meeting of surgeon investigators – pre-trial and regularly throughout	
	Standardisation of surgical approach (including lymphadenectomy) through trial protocol	(…./1)
Monitoring	Detailed method of assessing adherence (CONSORT)	
	Photographic monitoring	
	Video monitoring	
	Record indicators for surgical decision making (e.g conversion from laparoscopic to open)	
	Regular audit of data including clinical outcome measures	
	Record an obvious measure of quality (e.g LN yield) (IDEAL)	
	Intraoperative monitoring	
	Surgical details captured on standardised data collection form	
	Review of operation notes/operative forms	(…./1)
Total		(…../3)

3.3.4. Assessment of Quality of Trial Protocols and RCTs

The risk of bias within included RCTs was assessed using the Cochrane Collaboration tool described by Higgins *et al*.[38]

There is currently no validated method to assess the 'quality' of trial protocols, although the SPIRIT 2013 Statement, does provide guidance for minimum protocol content.[13] As an indication of adherence to protocol quality guidance the SPIRIT statement was therefore used to assess papers and available trial protocols, giving a total score out of 33 with 1 point given for each SPIRIT guidance component included.

3.2.5. Statistical analysis

Statistical analysis was performed with support of a statistician (AV) using SPSS (Version 24, IBM 1989, 2016) and SAS (Version 9.4, 2013, SAS Institute Inc, USA). An association between trial SQA measures and protocol availability was assessed using the Mann-Whitney U test. RCTs presenting Kaplan-Meier plots and numbers needed to treat (NNT) were included in the survival analysis. The online software Webplotdigitizer (https://automeris.io/WebPlotDigitizer/) was used to extract survival estimates across time from Kaplan-Meier plots presented in the papers by the primary researcher (JB). Aggregated survival data including survival point estimates, standard errors and confidence intervals were obtained based on Guyot's algorithm.[83,84]

Two regression models, following Arends *et al*, [83] were fitted based on publications for gastric cancer and oesophageal cancer separately. The outcome of interest looked at differences between publications with and without available protocols. The basic model was constructed in Statistical Analysis Software (SAS) 9.4 (SAS® UK - Analytics Software & Solutions) as a mixed model with the outcome variable the log(-log(survival)) with covariates protocol, log(time), $\log(\text{time})^2$ and the interaction between log(time) and protocol. The variance-covariance matrix was initially calculated from the variance of each survival curve at multiple time points. It was also structured in two blocks for publications with and without protocols. One of the advantages of this model is that the hazard ratios can be obtained through exponentiating the estimates of the coefficients of the model. The final model included the following covariates: ln(time), the public availability of the

protocol, the interaction between them, a quadratic term for ln(time) and the type cancer, gastric or oesophageal. Arends *et al*, approach was followed to use the available aggregated survival data. All covariates in the model presented a *P*-value < 0.0001. Covariates without missing values and a *P*-value < 0.2 were used in a multivariate model.

3.3. Results

3.3.1. Trial characteristics

Results of the online literature search are presented in Figure 3 according to PRISMA guidelines. From an initial 2589 articles screened, from January 2000 to February 2018, 28 RCTs[81,85-111] with a total of 6937 participants, were identified meeting inclusion criteria. Of the 28 RTC's included in this review 11 had available protocols: one published; eight available via trial registries, and; two were acquired after contacting the corresponding authors. The protocols of 17 RCTs could not be acquired after search on-line and contacting the corresponding author of primary trial publication.

A summary of included oesophagogastric RCTs is presented in Table 6. Of the included RCTs: 18 reported outcomes of patients with gastric cancer; 2 reported outcomes of patients with gastric and oesophageal cancer, and; 8 reported outcomes of patients with oesophageal cancer.

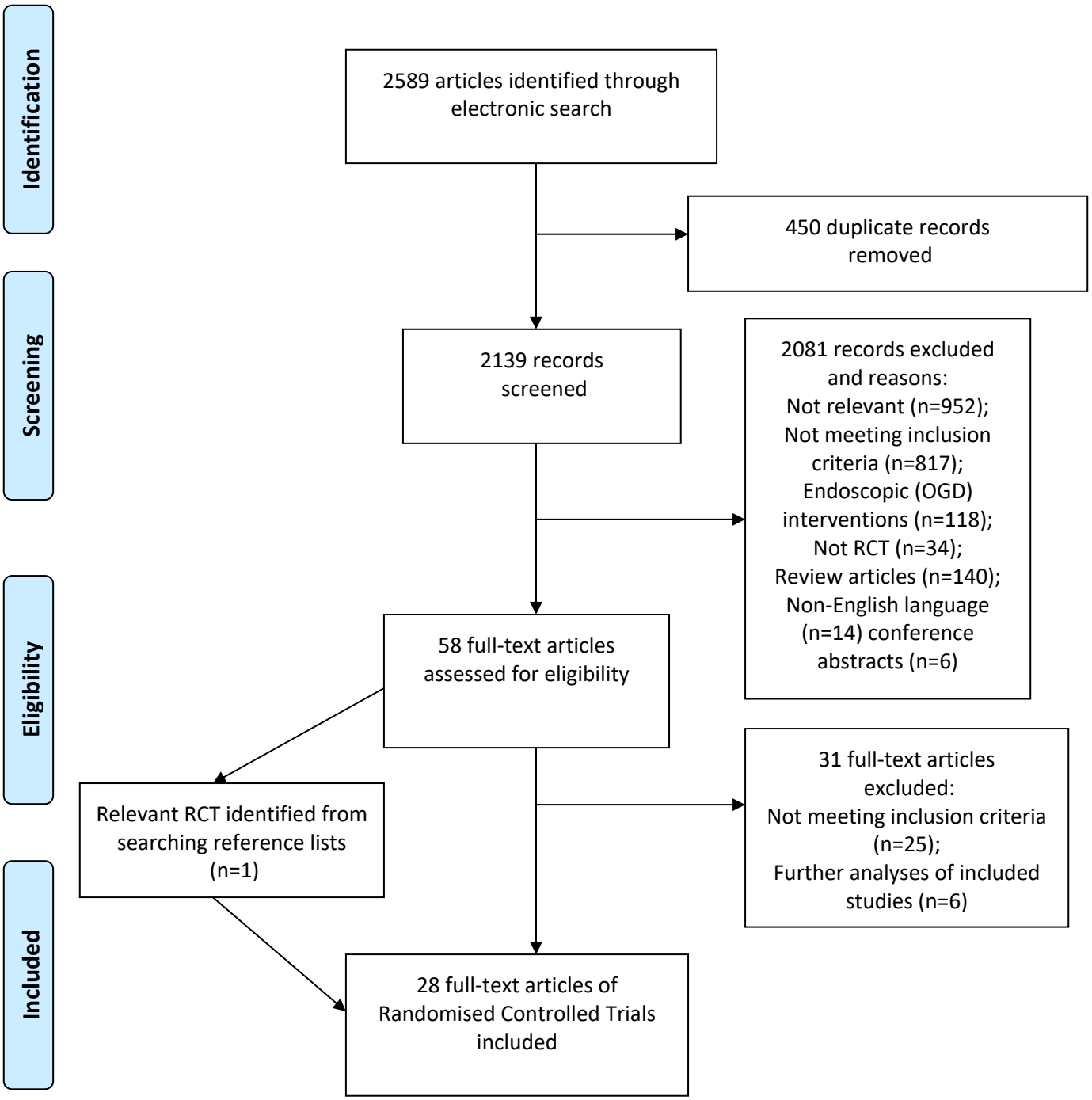

Figure 3: Flow-diagram of search strategy identifying relevant reviews presented in accordance with PRISMA guidelines [39]

Table 6: Summary of included oesophagogastric RCTs

Author	Trial	Cancer	Year	Trial arm utilised	Bias	SQA measures	SQA score	Protocol SQA	Protocol Spirit score	Protocol available	5 yr survival (%)	Ref.
Bajetta	Surgery +/- adjuvant chemotherapy	G	2002		Unclear	1	1	NA	NA	NA	48	[85]
Nashimoto	Surgery +/- adjuvant chemo	G	2003		Unclear	0	0	NA	NA	NA	86.1	[86]
Xiao ZF	Surgery +/- Radiotherapy	O	2003		High	0	0	NA	NA	NA	31.7	[87]
Chipponi	Surgery +/- adjuvant chemotherapy	G	2004		Unclear	0	0	NA	NA	NA	39	[88]
Bouche	Surgery +/- adjuvant chemotherapy	G	2005		Unclear	2	2	NA	NA	NA	41.9	[89]
Burmeister	Surgery +/- neoadjuvant chemoradiotherapy	O	2005		Low	0	0	NA	NA	NA	23.4	[90]
Yu	Gastrectomy +/- splenectomy	G	2006	Splenectomy	Unclear	1	1	NA	NA	NA	54.8	[91]
Cunningham (MAGIC)	Surgery +/- perioperative chemotherapy	O + G	2006		Unclear	1	1	0	6.75	Registry	23	[92]
Wu CW	Gastrectomy: D1 vs. D3 nodal dissection	G	2006	D3 Gastrectomy	Low	3	3	0	6.75	Registry	59.5	[93]
De Vita	Surgery +/- adjuvant chemotherapy	G	2007		Unclear	1	1	NA	NA	NA	43.5	[94]
Kelsen	Surgery +/- neoadjuvant chemotherapy	O	2007		High	0	0	NA	NA	NA	23	[95]
Omloo JM	Transthoracic vs. transhiatal	O	2007	Transthoracic	Unclear	3	2	3	16.24	Procured	36	[96]
Di Costanzo	Surgery +/- adjuvant chemotherapy	G	2008		Unclear	0	0	NA	NA	NA	48.7	[97]
Sasako	D2 Gastrectomy +/- Para-aortic LN dissection	G	2008	D2 + PAND	Unclear	7	3	0	6.75	Registry	70.3	[98]
Allum W (OEO2)	Surgery +/- neoadjuvant chemo	O	2009		Unclear	0	0	NA	NA	NA	17.1	[99]
Kulig	Surgery +/- adjuvant chemotherapy	G	2010		High	1	1	NA	NA	NA	40	[100]
Songun (Dutch D1/D2)	D1 vs. D2 Gastrectomy	G	2010		Unclear	7	2	NA	NA	NA	47	[101]
Miyashiro (JCOG9206-2)	Surgery +/- adjuvant chemotherapy	G	2011		Unclear	0	0	0	6.75	Registry	60.9	[102]
Sasako	Surgery +/- adjuvant chemotherapy	G	2011		High	1	1	0	12.08	Procured	61.1	[103]
Smalley (SWOG 0116)	Surgery +/- adjuvant chemoradiotherapy	G	2012		High	2	1	NA	NA	NA	NS	[81]
Van Hagen P (CROSS)	Surgery +/- neoadjuvant chemoradiotherapy	O+ GOJ	2012		Unclear	1	1	2	12.50	Published	34	[104]
Bass GA	Surgery +/- neoadjuvant chemoradiotherapy	O	2014		Unclear	0	0	NA	NA	NA	11.6	[105]
Noh SH (CLASSIC)	D2 Gastrectomy +/- adjuvant chemotherapy	G	2014		High	1	1	0	7	Registry	69	[106]
Degiuli M	D1 vs. D2 Gastrectomy	G	2014		Low	4	3	0	11.58	Registry	64.2	[107]
Mariette C*	Surgery +/- neoadjuvant chemoradiotherapy	O	2014		Unclear	0	0	0	6.5	ISRTCN	33.8	[108]
Hirao M	Gastrectomy +/- bursectomy	G	2015	Bursectomy	Unclear	3	3	NA	NA	NA	77.5	[109]
Yang Z Q	Open vs. minimally invasive oesophagectomy	O	2016	MIS	High	0	0	NA	NA	NA	67.6	[110]
Sano T	Gastrectomy +/- splenectomy	G	2017	Splenectomy	Unclear	5	2	0	6.92	Registry	75.1	[111]

Oeosophageal (O), Gastric (G), Gastro-oesophageal (GOJ), Not Specified (NS), Minimally Invasive Surgery (MIS). *FFCD 9901 trial, containing 97 patients within the surgical arm, included as seminal trial within oesophageal surgical oncology.[108]

3.3.2. Assessment of bias and SQA measures

Risk of bias was low in 3 (11%) trials, high in 7 (25%) trials and unclear in 18 trials (64.3%) (See Table 6). Full details of the assessment of bias can be found in appendix D. The median SPIRIT score of trial protocols was 6.92 (range 6.5-16.24) indicating poor adherence to the recommended minimum protocol content.

One or more SQA measures were reported in 16 (57%) of the 28 included RCTs. The median number of SQA measures utilised across all trials was 1 (range 0 to 7). SQA measures utilised by RCTs included: credentialing (n=5 RCTs); standardisation of surgical procedures (n=12), and; monitoring of surgical interventions (n=11).

RCT implementation of SQA measures was inadequately reported within trials and protocols, meaning insufficient data was available to robustly analyse this as a co-variable within a meta-analysis. Only two of the 11 acquired RCT protocols mentioned any form of SQA.[112,113] SQA measures in the CROSS trial protocol included: standardisation of operative procedure and pathological assessment of resected specimen including lymph node count.[112] The 'Extented versus transhiatal oesophagectomy' RCT protocol published by van Lanschot et al described plans for operative standardisation, recording of lymph node yield and regular audit of clinical outcomes.[113]

Number of SQA measures utilised in RCTs with protocols (mean rank 17.6) was higher than those trials without publicly available protocols (mean rank 12.5), trending towards significance (1.0 vs. 0.0; *P*=0.091).

Eight trials (29%) that reported SQA measures in their manuscripts had no available protocol (See Table 6). Seven (64%) of the 11 trials with protocols reported SQA measures within manuscripts that were not mentioned in their protocols. One trial protocol planned two SQA measures (standardisation of surgical approach and recording of lymph node yield), whilst only one (lymph node yield) was mentioned within the trial manuscript.[104,112]

3.3.3. Long term survival analysis

Seventeen RCTs presented Kaplan-Meier plots and number needed to treat (NNT) data that was suitable for meta-analysis.[85,87,90,91,93,95-99,101,103,105-108,111] Two studies were excluded from meta-analysis as they included patients with both gastric and oesophageal cancers within the same study cohort.[92,114]

The association between selected covariates and overall survival in both gastric and oesophageal cancer surgical trials are presented in Table 7. For gastric cancer RCTs the availability of a protocol was predictive of long-term overall survival. For oesophageal cancer RCTs the availability of a protocol and trials originating from the Far East (China, Japan, Korea) predicted long-term overall survival (See Figures 4-6).

After adjusting for potential confounding factors, patients within gastric or oesophageal cancer RCTs without available protocols were 3.46 (95% CI: 2.31 to 5.17; p-value:<0.0001) and 2.02 (95% CI 1.56 to 2.61 p<0.001) times more likely to die within the first year of surgery respectively, compared to patients within trials with available protocols. In the oesophageal cancer RCTs, trials from Western countries had a 1.74 (95% CI: 1.26 to 2.44; p-value:0.0011) times greater associated risk of death compared to trials from the Far East (see Tables 8 and 9).

The survival differences between RCTs with and without available protocols reduced as time interval from surgery increased, however it remained statistically significant (Tables 8-9 and Figures 4-6). For oesophageal cancer RCTs year of publication was no longer statistically significant in the multivariate model. The high proportion of oesophageal RCTs that did not report subject age (1 trial), proportion of patients with T3/T4 tumours (14 trials), or number with lymph node involvement (13 trials), meant that reliable statistical inference could not be made for these covariates. No patients in the surgery only arms of trials included in this review received chemoradiotherapy and thus this potential confounding variable was not

considered. Predicted overall survival curves are shown for RCTs with and without available protocols (Figures 4-6).

Table 7: Statistical significance of other covariates on overall survival

Variables	Gastric cancer (p-value)	Observations gastric	Oesophageal cancer (p-value)	Observations oesophageal
Protocol	<0.0001		<0.0001	
Year	0.6701		0.0644	
Age	0.6091		0.0128	1 missing
Eastern country origin	0.9142		0.0012	
Percentage T3/T4	0.6955	3 missing	<0.0001	only 3 papers
Percentage nodal involvement (N1+)	0.5478	2 missing	0.6812	only 4 papers
Percentage male	0.6033		0.1092	

Table 8: Hazard Ratio of death in oesophageal RCTs with no protocol versus publicly available protocol

Time (yr)	HR no protocol vs protocol	HR 95% CI	
1	2.018	1.558	2.614
2	1.958	1.513	2.534
3	1.923	1.486	2.489
4	1.900	1.468	2.458
5	1.881	1.453	2.435

Table 9: Hazard Ratio of death in gastric RCTs with no protocol versus publicly available protocol

Time (yr)	HR no protocol vs protocol	HR 95% CI	
1	3.459	2.313	5.173
2	2.896	1.943	4.315
3	2.610	1.753	3.884
4	2.424	1.629	3.606
5	2.289	1.539	3.405

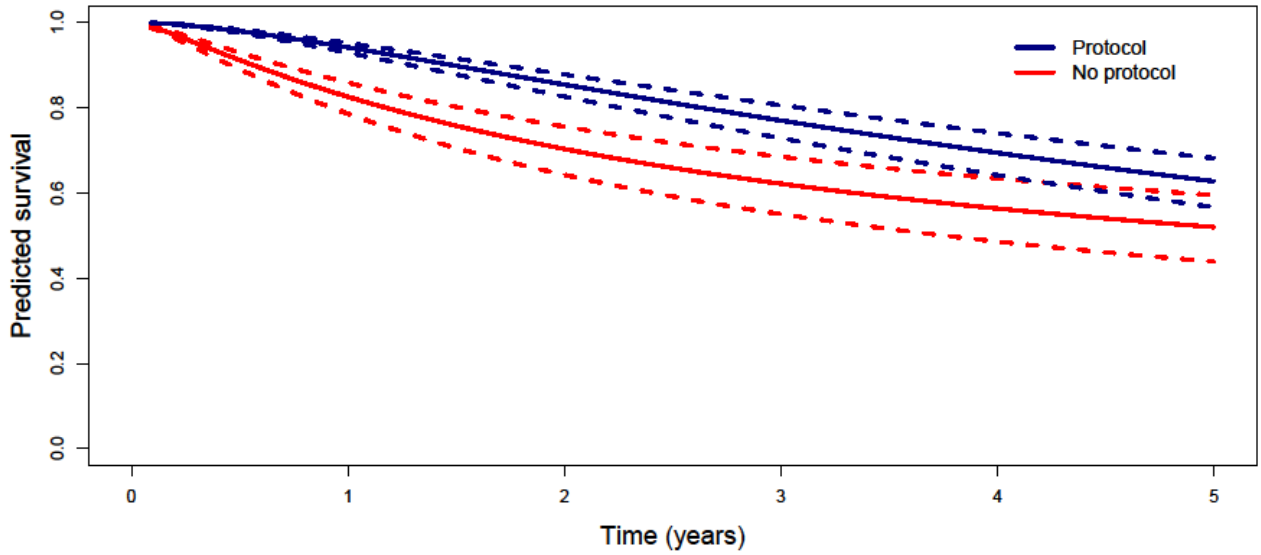

Figure 4: Predicted survival curves for the gastric cancer trials with and without publicly available protocol

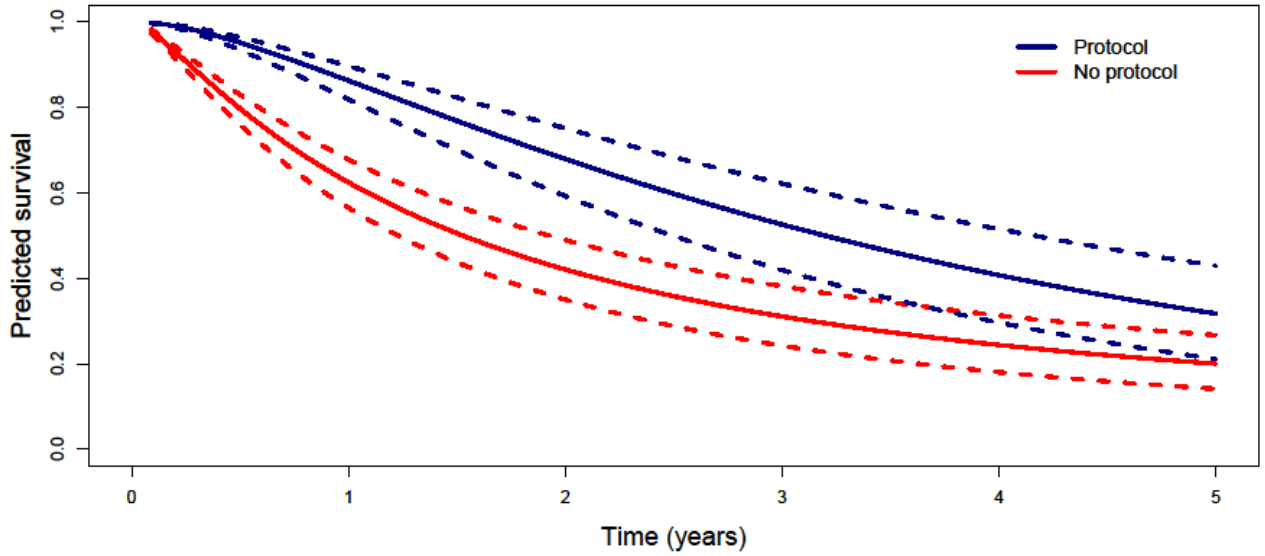

Figure 5: Predicted survival curves for oesophageal cancer with and without publicly available protocol

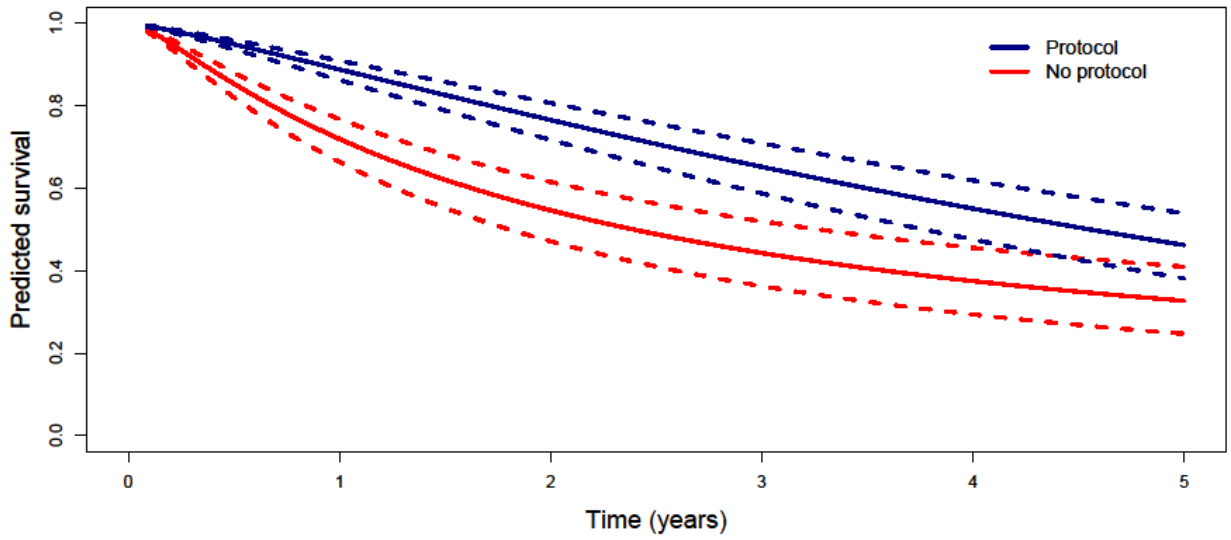

Figure 6: Predicted survival curves for gastric and oesophageal cancer trials with and without publicly available trial protocol

3.4. Discussion

This is the first review to comprehensively assesses SQA measures within oesophagogastric surgical oncology trials and their protocols. Moreover, it is the first review to assess the impact of trial protocol availability on long-term survival. The existence of a study protocol has for the first time been shown to be associated with improved patient survival within an adjusted model. Less than one third of included trials had an available protocol. The majority of trials did not adopt or report what were considered to be adequate SQA measures. Only 7% of all included trials declared an intention to implement SQA measures within an available protocol prior to commencing the study.

Studies assessing outcomes following gastroesophageal cancer resection in Japan have shown the lowest post-operative mortality and longest long-term survival globally.[115,116] A systematic review of studies reporting outcomes of gastrectomy similarly demonstrated improved 5-year survival and reduced rate of cancer recurrence in Eastern compared to Western centres.[117] Differences in tumour biology, patient physiological characteristics and surgical techniques may partly explain this association.[118,119] However equivalent long-term survival following oesophagogastric cancer resection between Eastern and Western centres have also been reported, demonstrating long-term survival can be improved in Western patients to the highest standard achieved in Japan.[120]

Most meta-analyses comparing long-term survival utilise hazard ratios at specific time-points to summarise treatment effects. However, through employing the Arends method within this study the entire survival curve can be exploited. Each survival point has a standard error that can be reflected in the variance-covariance matrix within the model and they may be structured in blocks with respect to variables of interest (e.g. existence of a trial protocol). Prior to this statistical model it was difficult to compare the impact of trial design or other factors on long-term survival between RCTs given the heterogeneous methods for reporting overall survival between trials, often lacking confidence intervals and ‘number of events’.

There has been no previously validated method of assessing SQA implementation within RCTs and thus our three-point scale was developed from published literature. Although this allowed a form of comparison between RCTs within this study, further validation and reliability testing is required before this can be considered a robust tool for SQA assessment.

The association between the availability of a trial protocol and improved overall survival that was observed in this study may be due to the coexistence of SQA measures utilised within those trials. It was not however possible to evaluate the impact of SQA measures on survival within this meta-analysis, as SQA were inconsistently reported within both manuscripts and protocols. Furthermore, there remains considerable ambiguity regarding extent to which intended SQA measure are implemented within trials. One example of this is the Dutch D1 versus D2 gastrectomy trial which reported use of 7 recorded SQA measures within the manuscript,[45] however the authors later reflected that a lack of credentialing and insufficient standardisation and monitoring led to partial homogenisation of the trial arms.[46] Discrepancies between SQA measures that were declared in trial manuscripts but not within their protocols, occurred in 7 studies (64%) of the 11 RCTs with protocols and serves as further evidence for inconsistencies in reporting.

The IDEAL framework has recommended that researchers make trial protocols available in the both development and exploration stages in addition to calling on journals to support the publication of study protocols. [121] The CONSORT checklist also advises that details of how 'adherence of care providers with the protocol was assessed or enhanced' and that authors should 'describe protocol deviations from study as planned together with reasons'.[35] Findings of the current review, further supports this appeal for trial stakeholders to make protocols available publicly either within peer reviewed journals or via online trials registries.

Online trial registries were the most common repository for trial protocols when they were available, of which 'clinicaltrials.gov' was the most commonly utilised website. There are however no specific sections within such registries for the documentation of SQA. To aid further planning of SQA initiatives within trials it is recommend that online trial registries create a separate section specifically for quality assurance of the surgical intervention. This

may prove just as relevant for the other interventions (e.g radiotherapy, endoscopy) in other arms of the trial.

3.5. Conclusion

Chapter 2 of this thesis reported the previously published challenges to quality assurance of surgery in oncology trials, however it was anticipated that many challenges remained undocumented. In this systematic review it was observed that few large oesophagogastric oncology trials have documented or adhered to robust SQA measures. Regarding oesophagogastric oncology trials, the existence of publicly available protocol and trials originating from Eastern centres were found to be associated with improved long-term survival. Stakeholders designing future trials should therefore strive to integrate SQA initiatives within their trial design and make protocols publicly available in endeavour to optimise quality of surgical interventions and long-term outcomes. The reason for poor adherence/implementation of SQA measures within RCTs identified within this chapter has not been previously investigated. Moreover, the best method to overcome challenges to SQA in oncology trials according to expert opinion remains unexplored. The next chapter will permit investigation of expert opinion regarding quality assurance of surgery, its associated challenges and potential mitigating strategies.

Section 2.

Chapter 4. Challenges and proposed mitigating solutions to Quality Assurance of Surgery in Oncology Trials: a qualitative interview study

4.1. Introduction

RCTs remain the gold standard method of assessing the efficacy of interventions in surgical oncology. Where surgery remains the primary curative modality surgical performance and overall quality of surgical intervention can directly influence outcome within oncology trials. However most trials lack a robust framework to assess and ensure surgical quality.[114] On reviewing published literature on challenges to conducting surgical trials the fidelity and quality of surgical interventions became a prominent theme. [10,14,122,123] Within oesophagogastric oncology the importance of quality of surgical performance is demonstrated in several landmark trials.[20 51]

A diverse range of challenges to quality assurance of surgical intervention exist across multiple specialities within clinical trials, including insufficient credentialing of surgeons, and deficient standardisation and monitoring of surgical intervention. [10,14,122,123] SQA measures have been previously proposed within the CONSORT Non-pharmacological treatment guidelines and the IDEAL framework, in addition to a previous systematic review of large colorectal trials (See Chapter 2, Table 1). [11,15,35] However expert trial stakeholder opinion regarding the usability and feasibility of these SQA measures within oncology RCTs remains unexplored within published literature. Similar qualitative research methods have previously been employed to identify barriers to adoption of new medical diagnostics allowing strategy development to overcome identified challenges.[124]

To assess expert opinion on challenges, incentives and disincentives to quality assurance of surgical interventions in oncology trials and possible mitigating strategies, semi-structured interviews were conducted with expert trial stakeholders. Following this, a Delphi process was required in order to gain expert consensus on proposed mitigating solutions and

develop a framework of strategies to overcome challenges to quality assurance of surgery in oncology trials.

4.1.1. Objective

The specific study objective was:

ii. Examination of expert opinion on challenges to SQA

4.2. Method

4.2.1. Participants

Semi-structure interviews: International surgical congresses can result in a coalescence of surgeons with involvement in clinical trials. The respective committee members and presenters at the 2017 congresses of the European Association of Endoscopic Surgery (EAES), Japanese Society of Gastroenterological Surgery (JSGS) and Association of Upper Gastrointestinal Surgeons were reviewed for potential participants. Peer nominated expert trial stakeholders were identified with experience in clinical trials, based upon their contacts, previous publications and affiliations with the aforementioned societies and congresses.

From July 2017 to June 2018 semi-structured interviews were conducted with 71 expert trial stakeholders (35 (49.3%) surgeons, 17 (23.9%) oncologists, 10 (14.1%) trial methodologists, 9 (12.7%) trial managers) with experience in clinical trials. These stakeholders were drawn from 8 countries including: Japan (n=14, 19.7%); United Kingdom (n=29, 40.8%); Netherlands (n=6, 8.5%); Ireland (n=1, 1.4%); France (n=1, 1.4%); United States (n=1, 1.4%); Sweden (n=2, 2.8%); Switzerland (n=2, 2.8%); Italy (n=2, 2.8%). Trial stakeholder median age was 51years (30-71), male; female ratio was 54:17 and 45% (n=35) of participants held a professorial title (See Table 10). 58 (82%) interviews were conducted in-person, 11 (15%) via video conference call, and 2 (3%) by telephone. Median number of oesophagogastric oncology trial publications each stakeholder was associated with was 7 (range 1-64), and 46 (65%) of stakeholders had been principle investigator or chief investigator (PI or CI) within one or more oncology trials (See Table 10).

Delphi process: 21 expert trial stakeholders participated including: 12 surgeons (57%), 6 oncologists (29%), 2 trial methodologists (10%) and 1 trial manager (4%). See Table 11 for further details of expert stakeholder demographics.

Table 10: Demographics of expert trial stakeholders participating in semi-structured interviews

Trial Stakeholder	Age (median)	Median number OG publications	Number with CI/PI experience	Male:Female ratio
Surgeon	58	6.5	27 (77%)	36:0
Trial Manager	36	0	0 (0%)	1:8
Trial Methodologist	46	24	2 (20%)	6:4
Oncologist	43	8	17 (100%)	12:5

Key: CI – Chief Investigator. PI – Principle Investigator.

Table 11: Demographics of expert trial stakeholders involved in Delphi Round 1

Background	**Country**	**Male:female**	**PI/CI experience**	**Trial publications**
Surgeon	Japan	Male	Y	60
Surgeon	Japan	Male	Y	29
Surgeon	France	Male	Y	61
Surgeon	Sweden	Male	Y	7
Surgeon	Switzerland	Male	Y	18
Surgeon	UK	Male	Y	7
Surgeon	UK	Male	Y	36
Surgeon	UK	Male	N	6
Surgeon	UK	Male	N	2
Surgeon	UK	Male	N	0
Surgeon	UK	Male	N	1
Surgeon	UK	Male	N	7
Oncologist	UK	Female	Y	3
Oncologist	UK	Male	Y	68
Oncologist	UK	Male	Y	5
Oncologist	UK	Male	Y	7
Oncologist	UK	Male	Y	31
Oncologist	UK	Male	Y	2
Trial Methodologist	UK	Female	N	61
Trial Methodologist	UK	Male	Y	0
Trial Manager	UK	Male	N	0

Key: CI – Chief Investigator, PI – Principle Investigator. UK – United Kingdom

4.2.2. Study design

Semi-structure interviews: The findings of the previous systematic review and the study objectives were utilised by the primary researcher (JB) to guide development an initial template for the semi-structured interviews. This template was assessed in pilot interviews and underwent multiple adjustments prior to finalisation, in accordance with standard qualitative research methodology. Key areas were first explored with open questions followed by closed questions to clarify participants opinion on specific challenges and strategies. The interviewer (JB) made field notes during and after interviews in addition to recording interviews using an audio device. Stratified purposive sampling was used to identify oncologists, trial methodologists and trial managers with experience of involvement in clinical trials. Participants were contacted by email or in person and informed that responses would be confidential, pseudo-anonymised and that no compensation would be provided. Interviews were conducted in the English language with all participants having satisfactory level of spoken English. Two recording devices were used to for each interview to ensure a back-up copy of recorded data was available if required. Interviews had a median duration of 23.4 minutes (range 10 to 33 minutes). During this phase stakeholders were added where identified throughout the interview and research collaboration process.

Delphi process: Following completion of the expert trial stakeholder interview analysis, an electronic Delphi process was commenced through a series of online surveys utilising Qualtrics XM (NPS software). All strategies proposed by 2 or more of the 71 trial stakeholders previously interviewed were included within the Delphi process. Following the completion of the expert trial interviews, recruitment commenced for the Delphi process to gain expert consensus on proposed mitigating strategies. Seventy-one expert trial stakeholders involved in the previous interview study were contacted electronically by email to request their informed consent for participation in the Delphi process.

4.2.3. Analysis

Semi-structured interviews: Following informed consent, interviews were recorded and then transcribed verbatim though an external transcription company. Any participant

identifiers (e.g. participants names, facility names, contact names, country names) were removed in order to preserve anonymity. For the first set of interview transcripts the primary analyst (JB) and a senior analyst (SM), read them several times to create a coding manual for the latter transcripts using predetermined and other emergent themes. The primary researcher then conducted the coding analysis of the rest of the interviews and modified the coding manual iteratively as required, discussing changes and new emergent themes with the senior analyst (SM) concurrently. The primary analyst was a surgeon with an academic background in research including surgical science and international health. The senior analyst had a background in psychology and research methodology with experience within qualitative research. Transcribed interviews were coded using Nvivo software (12.2.0 QSR International). Interviews were conducted until theme saturation was reached. Concepts expressed were considered emergent themes if they were adopted by one or more participants.

Delphi process: Within the Delphi process of expert consensus on proposed mitigating strategies, an agreement level of 70% was utilised to signify consensus.

4.3. Results of expert trial stakeholder semi-structured interviews

Despite variation in the 389 initial emergent themes, 56 themes were supported by more than one stakeholder on qualitative analysis of stakeholder responses. The most common themes are presented under the sub-categories of credentialing, standardisation, monitoring and generic challenges and proposed mitigating strategies (See Table 12). Please note that the credentialing results section includes both selection of centres and selection of surgeons for participation in oncology trials. For more details regarding the detailed qualitative analysis undertaken and further quotations supporting each emergent theme please see appendix E. Key emergent challenges to SQA and stakeholder proposed mitigating strategies are summarised below. In this survey of international expert trial stakeholder opinion, a wide variety of perceived challenges and potential mitigating solutions for different aspects of SQA within oesophagogastric oncology trials have been identified. The most prominent stakeholder perceived challenges to quality assurance of surgery in trials included: limitations of credentialing centres according to operative volume; differing oncological beliefs and resistance to change adoption leading to inter-surgeons and inter-institutional operative variation, and; surgeons pride and national culture leading to difficulty in providing and receiving feedback.

The variability of surgical quality irrespective of a centres operative volume became a prominent concern amongst study participants. Many stakeholders felt operative quality could be low regardless of the total number of procedures performed within a centre: ‘And there is another problem with threshold, is that you could do 100 bad operations a year and 20 good ones. And so, in isolation, a threshold is not sufficient to make an assessment of quality.’ (Quote, Surgeon 25). The difficulty of gaining national or international agreement within oesophagogastric surgical community due to differing oncological beliefs was highlighted: ‘You know I think particularly with oesophageal surgery it's very, very difficult to even get two separate surgeons to agree what's the best operation. If you open that to different countries and different units, that will be very difficult to get agreement on.’ (Quote, Surgeon 23). The difficulty in persuading experienced surgeons who have performed thousands of operations using a certain method to follow a different protocol is highlighted and explained through the bond each individual surgeon develops with their

own techniques: ‘Well it is difficult to agree on standardisation of techniques, because we all have our own tips and tricks and we believe they are fundamental to the substance of our surgery. I don't think this is true, but I don't think you can force any surgeon to change his technique too much.’ (Quote, Surgeon 32). Stakeholders identified some famous and influential surgeons’ pride as a key challenge in surgeons providing or receiving authentic feedback: ‘So, even if I have a chance to rate some surgeon’s surgical skill it’s difficult to rate, to give bad score to him. Most surgeons participate in the study are very famous surgeons, very powerful surgeons and they have pride so it’s very difficult to rate them……’ (Quote, Surgeon 3). The role of surgeons’ national culture was also felt to contribute to difficulties in providing and receiving feedback within Eastern and Western Societies: ‘You know, if we ask, it maybe could be very embarrassing if you ask somebody to participate in a trial then they send you a video of this operation, you assess it and say well, your quality of surgery is poor. I mean I don’t see… At least in the Western world, I don’t see that it's a very practical way to go.’ (Quote, Surgeon 22)

The most prominent mitigating strategies to overcome challenges to quality assurance of surgery included: considering surgeon’s learning curve in surgeon credentialing process, operative manual utilisation to standardise trial surgical interventions, and monitoring of surgery in trials using unedited video recordings. The particular importance of this within trials comparing new to more established techniques was reinforced by several participants: ‘If this is concerned about a very new technique like laparoscopy or something, then if they are in the learning curve then this video check examination is very useful and important.’ (Quote, Surgeon 2). The importance of an operation manual as an adjunct to aid standardisation of operating in trials became a dominant theme: ‘Yes, I think if you are working to a standard, then a manual and a video would be helpful.’ (Quote, Surgeon 20). Video assessment was regarded by many as the optimal form of monitoring, particularly for laparoscopic operating, due to the clear visualisation of operative events: ‘For something that you can video that would be absolutely the ideal thing. If you had a series of people scoring blindly loads of videos. Yes, that would be ideal.’ (Quote, Surgeon 21). However, in order to develop a robust framework to aid design of SQA measures within future oncology trials, we require an expert consensus using these previously identified mitigating strategies.

We sought to gain expert consensus via a Delphi process on strategies to overcome challenges to quality assurance of surgery in surgical oncology trials.

Table 12: Challenges to Implementation of SQA in Oncology Trials and Proposed Mitigating Strategies: Summary of Emergent Themes

SQA Method		No
Credentialing		
Challenges	Personal factors and bias influencing surgical trial centre selection	11
	Limitations of surgical volume in selecting trial centres	14
	Excessive focus on surgeons/selection bias and difficulty in surgeon recruitment	5
	Challenges to selection of surgeons including retrospective case registration and difficulty in determining surgeon experience or quality of operating	4
Strategies	Case volume should be considered in selecting surgical centres	29
	Case volume should be considered in selecting surgeons	18
	A surgeon's learning curve should be considered in selecting surgeons for participation in oncology trials	34
	Training, education and mentoring of surgeons to gain required trial operative skills to allow accreditation/qualification process prior to trials participation.	5
	Training and education to improve surgeons' understanding of research methodology	5
	Credentialing of centres prior to centre selection for trial participation	5
	Adjust stringency of selection of trial centres according to trial design, phase of trial, disease prevalence or operation complexity	13
	Trial centres required to meet certain criteria for consideration of inclusion in oncology trials including: being specialist centres; part of national audit; having published expertise; not have outlying results	29
	Impartial committee to select surgeons for participation in oncology trials	3
	Adjust stringency of surgeon selection according to the following factors: trial design; phase of trial; disease, and; operation complexity	8
	Graded method of surgeon selection utilising structured objective assessment tools	3
	Selection of surgeons according to postoperative outcomes through audit of: Histopathology; intraoperative bloods loss, postoperative complications and outcomes	15
Standardisation		
Challenges	Challenges to surgeon adherence to protocol and standardisation	7
	Differing oncological beliefs and resistance to change adoption leading to inter-surgeons and inter-institutional operative variation	18
	Overly prescriptive protocols and standardisation causing difficulty in recruitment of surgeons	12
	Concerns regarding consequences of standardisation and operative variation internationally	3
	Regional differences in perception of surgical quality and insufficient quality of data regarding outcomes	4
Strategies	Use of operation manual to aid standardisation	34
	Other methods recommended to aid standardisation including: trial launch days; operative room boards explaining trial intervention; standardisation of histopathology processes and radiology diagnostics, and; use of videos, photographs, virtual reality and webinars	18
	Flexible standardisation depending on trial design and study question, focussing on standardisation of steps affecting primary end-points, safety or survival	22
	Virtual training system, technology, artificial intelligence and robotics to aid credentialing, standardisation and monitoring of surgery in trials	10
	Inter-institutional operative team visit exchanges to observe trial operating	3
Monitoring		
Challenges	Surgeons pride, fame and national culture leading to difficulty in providing feedback and receiving feedback	14
	Limitations of operative videos to monitor surgery	17
	Limitations of photographs to monitor surgery	12
	Feasibility and other concerns with intra-operative monitoring	6
	Stakeholder perceived limitations of use of other methods of monitoring of surgery including post-operative complications and review of operative notes	3
	Dealing with consequences of monitoring of surgery and providing feedback	7
	Insufficient monitoring, surgeon reluctance to be monitored and concerns over adverse outcomes/litigation	8
	Insufficient resources for monitoring of surgery plus potential negative impact on generalizability and surgeon recruitment	5
Strategies	Surgery in oncology trials should be monitored	11
	Surgery in oncology trials should be monitored with video recording	22
	Surgery in oncology trials should be monitored with photographs	11
	Surgery in oncology trials should be monitored using intra-operative monitoring	15
	Surgery in oncology trials should be monitored according to histopathology assessment of trial resection specimens	10
	Other suggested methods of monitoring surgical quality in trials including post-operative complications/outcomes, lymph node yield, operative notes, case report forms	32
	Adjusting stringency and methods of monitoring depending on trial design and research question/surgical procedure	4
	Standardised monitoring process with intraoperative monitoring plus real-time data transfer to Trial Committee	5
	Selective reviewing of recorded trial videos by trial committee assessors	4
	Graded monitoring system, adjusting stringency of monitoring according to performance and monitoring of initial cases	9
	Anonymity and confidentiality of monitoring assessments of surgeons	5
	Structured, regular and appropriately senior feedback mechanism with positive reward system and external peer review	4
Generic		
Challenges	Lack of clinical equipoise for trial surgical interventions amongst surgeons and a lack of trust in trial data/design	3
	Insufficient resources, time and implementation of SQA measures	13
	Regional variability in operative standards, diagnostics, patient factors, concerns over patient access to surgical interventions and insufficient generalizability of trial results	9
Strategies	National, centralised SQA initiative and the importance of strong trial leadership shared between trial centres	5
	Developing shared goals and surgeon consensus on acceptable SQA measures prior protocol development, facilitating the development of research relationships between trial centres	6
	Adequate design and costing/funding of SQA in trials to facilitate surgeon commitment and SQA	7
	Ensuring centre preparedness prior to trial commencement and conducting pre-trial feasibility studies	3
	Teamwork and trials improving quality of clinical practice through skill development and introduction of new technologies	3
	Trial group learning, regular meetings and focus groups facilitating learning of QA from radiotherapy community/other trial groups	7
	Quality assure the SQA process within oncology trials	6

4.4. Results of expert stakeholder Delphi process

70 strategies were included with the online survey in Delphi Round 1, of which 59 strategies (84.3%) had gained consensus (at least 70% agreement) within 3 Delphi Rounds (See Table 13). There was attrition of 1 (4.8%) expert stakeholder following Delphi Round 1. The eleven strategies not reaching consensus within the Delphi process can be found in appendix F.

Table 13: Expert trial stakeholder Delphi consensus on strategies to overcome challenges to SQA (Full List)

CONSENSUS MITIGATING STRATEGIES			
GENERIC MITIGATING STRATEGIES	**Delphi Round 1 (%)**	**Delphi Round 2 (%)**	**Delphi Round 3 (%)**
A centralised SQA initiative is required to aid SQA in oncology trials	95		
Strong trial leadership shared between trial centres is required within oncology trials	95		
Developing shared goals is required to facilitate the development of research relationships between trial centres	91		
Developing surgeon consensus on acceptable SQA measures prior to protocol development	100		
Adequate training and education within national medical surgical training to improve surgeons' understanding of research methodology	91		
Trial group learning with regular meetings and focus groups to facilitate learning regarding SQA from other trial groups	76		
The SQA process within oncology trials should itself be standardised and monitored	95		
Robotics, technology and virtual training systems may be useful for:			
• Standardisation of surgery in oncology trials	57	85	
• Monitoring of surgery in oncology trials	62	75	
CREDENTIALING			
Stringency of selection of trial centres should be adjusted according to:			
• Trial design	71		
• Phase of trial	62	90	
• Disease prevalence	81		
• Trial operation complexity	86		
A national database of operative centres is required to aid selection surgical centres for participation in oncology trials	52	60	85
The following factors should be considered in selection of trial centres for participation in oncology trials:			
• Trial centre operative volume	91		
• Training, education and mentoring of surgeons is required to ensure surgeons have required trial operative skills prior to trial commencement	81		
The following factors should be considered in selecting surgeons for participation in oncology trials:			
• Operative case volume	100		
• Surgeons should undertake an accreditation/qualification processes prior to participation in oncology trials	62	75	
• Position of surgeons on their learning curve of trial specific operative skills	86		
• Audit of histopathology of resected specimens.	81		
• Audit of intraoperative bloods loss, postoperative complications and outcomes	76		
Stringency of surgeon selection should be adjusted according to the following factors:			
• Operation complexity	86		
• Trial design	67	80	
The optimal method of selecting surgeons for participation in oncology trials includes:			
• Structured objective assessment tools	71		
• Review of operative videos	86		
STANDARDISATION			
Trial stakeholder consensus on definition of quality of surgery is required prior to standardisation of surgery within oncology trials	100		
Prospective data collection is required to provide evidence to aid in gaining consensus on definition of quality of surgery	86		
Regular trial stakeholder meetings should be conducted to aid standardisation.	95		
Trial centres should be given time to adequately prepare before trials to aid standardisation of surgical procedures	91		
Stringency of standardisation of surgery should be adjusted according to the following			

factors:			
• **Study question**	**81**		
• **Disease process studied in the trial**	**62**	**85**	
• **Trial design (explanatory versus pragmatic)**	**67**	**90**	
Standardisation of surgery should focus on operative steps affecting primary end-points, safety or survival	**81**		
Regular feedback on SQA within trial centres should be provided by the trial committee to facilitate standardisation of surgery	**95**		
Inter-institutional operative team visit exchanges are useful to facilitate standardisation	**86**		
Trial launch days in which trial surgeons and stakeholders are invited to an event to discuss the trial operative protocol	**95**		
An operation manual should be utilised to aid standardisation of surgery and facilitate training of surgical trainees	**95**		
The following modalities as adjuncts to an operation manual would be useful to reinforce standardisation of surgery:			
• **Operative videos**	**86**		
• **Operative photographs**	**71**		
• **Webinars**	**65**	**65**	**91**
Adherence to trial operation manual should be assessed to aid standardisation	**90**		
To facilitate standardisation of surgery in oncology trials the following methods are recommended:			
• **Histopathology assessment of resected specimens**	**91**		
• **Standardisation of radiology and diagnostic processes**	**86**		
• **Standardisation should be extended to to peri-operative care by the clinical team**	**95**		
• **Display of operation room boards detailing required steps for each specific trial operation**	**48**	**65**	**80**
MONITORING			
The process for monitoring of surgery within oncology trials itself should be standardised	**95**		
Monitoring of surgical quality should consist of a graded monitoring system, adjusting stringency of monitoring according to performance and monitoring of initial cases	**76**		
Monitoring of surgery should be kept anonymous and confidential	**95**		
The monitoring of surgery should include external peer review	**71**		
Stringency and methods of monitoring of surgery in trials should be adjusted according to the following factors:			
• **Trial design and research question**	**71**		
• **Surgical procedure involved within the trial**	**91**		
The optimal method for monitoring of surgery in oncology trials includes:			
• **Use of a structured objective assessment tool involving review of either operative photographs, unedited videos or intraoperative assessment is the optimal method for monitoring of surgery**	**76**		
• **Review of operative photographs**	**65**	**65**	**70**
• **Review of unedited operative videos**	**62**	**80**	
• **Review of random selection of unedited trial videos**	**62**	**75**	
• **Intra-operative monitoring by visiting surgical team**	**62**	**75**	
• **Histopathology assessment of quality of resected specimens**	**86**		
• **Review of post-operative complications/outcomes and lymph node yield**	**100**		
• **There should be a structured, regular and appropriately senior feedback mechanism with positive reward system for monitoring of surgery**	**71**		

4.5. Discussion

Historically oncology trials have not utilised robust SQA within their design and implementation leading to difficulty in results interpretation and affecting outcomes.[14] No published study has yet assessed expert trial stakeholder opinion regarding quality assurance of surgery in oncology trials and acceptability of potential mitigating strategies. Challenges to SQA and proposed mitigating strategies can be categorised into key areas of: [i] generic, [ii] credentialing of surgeons/centres, [iii] standardisation of surgical interventions and [iv] monitoring of surgery as previously described.[11,125] There are few reported challenges to SQA within published oesophagogastric oncology trials alone (Chapters 2 and 3), and thus published challenges to SQA across multiple specialities are considered within this discussion. Following examination of expert opinion regarding challenges to SQA and possible solutions within the interview study, expert trial stakeholders reached consensus on mitigating strategies through a Delphi process.

Limitations of selecting centres according to operative volume was one of the key perceived challenges to credentialing by expert trial stakeholders. Similar challenges to credentialing of surgeons have been reported in four colorectal trials in which surgeons were credentialed according to cumulative case volume using 20 laparoscopic colorectal cases per surgeon or centre as the threshold for recruitment, but later reflecting that this limited effective selection of surgeons.[53,57,61,66] The consequences of insufficient credentialing methods were demonstrated within two trials recruiting patients broadly from all specialist centres performing liver surgery and gynaecology oncology resections resulting in predictable deleterious effects on oncological outcomes.[65,71] However, the interview study within this chapter also revealed credentialing centres and surgeons according to operative volume continues to be considered an important method amongst experts within oncology trials. Some expert stakeholders advocated assessing trial centre volume, whereas others advised reviewing individual trial surgeon operative volume prior to surgeon recruitment. Multiple trial stakeholders felt that trial centre operative volume was associated with surgical expertise, and that through incorporation of such high-volume centres the quality of the trial could be enhanced. This conceptual importance of operative volume is supported in published literature, with one epidemiological study of oesophagogastric cancer resections

within 42 UK hospitals over 2 years showing higher hospital operative volume was associated with lower 30-day mortality and lower anastomotic leakage rates. Higher surgeon volume was found to be associated with lower anastomotic leakage rates.[126] Strategies reaching Delphi consensus to improve surgeon/centre selection included: use of structured objective assessment tools; review of unedited operative videos; review of operative volume, and; consideration of surgeons' learning curve for trial specific operative skills. Although there is a heightened awareness of the need for a structured objective assessment tool to assess quality of surgery and select surgeons for trial participation of surgery, [22] this has not been previously utilised in oesophagogastric oncology trials (Chapter 3). Surgeons learning curves have been extensively studied with regard to safe implementation of surgical procedures, [127] however the credentialing threshold is often decided arbitrarily in previous trials. [53,57,61,66]

Key barriers to standardisation reported by expert trial stakeholders in this chapter included insufficient inter-institutional operation consistency and difficulty in surgeon adherence to protocol. Similar challenges have been encountered when endeavouring to standardise surgical interventions in previously published literature including variation in definition and execution of surgical interventions between trial centres reported in five trials. [16,20,52,55,58] Additional unnecessary procedures were reported to have become commonplace within two of the landmark trials of D1 versus D2 gastrectomy with one trial having a high rate of splenectomy,[45] and the other an excessive rate of both splenectomies and pancreatectomies.[51] Comparable difficulties were reflected upon in the narrative review by Cook et al in which significant non-compliance within some RCTs with surgical interventions is highlighted.[10] Key operative steps were left to the discretion of the operating surgeons in two trials raising concerns over standardisations methods, [42,65] and one trial reported insufficient training and monitoring of trial surgeons.[46,47] The expert interview study in this chapter revealed multiple participants prefer flexible standardisation which focuses on factors effecting safety, survival and outcomes. Overly prescriptive standardisation methods utilised in previous trials may have led to difficulties in compliance manifesting in the above-mentioned difference in surgical procedure execution between centres. Use of an operation manual to aid standardisation was also considered important by stakeholders within the interview study and then reached consensus within the Delphi process. The use of an

operational manual to aid standardisation is becoming more popular and evident in published literature with 12 (43%) of large oesophagogastric RCTs (2000-2018) documenting standardisation of surgical procedures (Chapter 3). The concept of utilising an operation manual with mandated, prohibited or optional steps to aid standardisation has also been referred to in one review of typology of standardisation and monitoring in trials.[128] Other consensus methods to overcome challenges to standardisation included the use of operative videos, operative photographs and webinars as adjuncts to an operation manual. Although becoming more popular methods for monitoring surgical quality in recent trials, [22,129] operation videos and photographs have not been previously used as an adjunct to aid standardisation within large oesophagogastric oncology trials (Chapter 3).

Prominent emergent challenges to monitoring of surgery within the interview study included: surgeons' pride, status and national culture leading to difficulty in both providing and receiving feedback, and; concerns over limitations of utilising videos or photographs to assess surgical quality, citing the rationale of time and resource limitations. Such concepts, although understood by many in the surgical profession, are often considered taboo and are not extensively documented. Previously reported challenges to monitoring of quality of surgery within RCTs involve inconsistencies in pathological assessment of post-operative resected specimens within one report.[46] Narrative reviews have reflected broadly on 'surgeon related factors' such as surgeons' knowledge, operative experience and learning curve as challenges to conducting surgical RCTs, however none have reflected on difficulties experienced in monitoring surgery with videos or photographs.[3,10] The importance of monitoring quality of surgery in clinical trials was a concept shared by the multiple interviewees. However, there were diverging views on the best method available with many advocating the use of video, citing its ability to record the process, tissue handling skills and damage to surrounding structures to a greater extent than photographs. Alternatively, multiple participants preferred photographic monitoring at key stages viewing the dissected operative field claiming this was more feasible, particularly for open surgery and generally less burdensome on the surgeon and the data monitoring team. Within the Delphi process both strategies (photographic and video monitoring) reached consensus. Other prominent consensus strategies to overcome challenges of monitoring surgery in oncology trials included: use of a structured objective assessment tool, and; utilising a graded monitoring

system, adjusting stringency of monitoring according to performance and monitoring of initial cases. Operation videos and photographs were rarely utilised for quality assurance in large oesophagogastric oncology trials (n=1, 3.6%) from 2000-2018 (Chapter 3), however they have begun to be utilised in recent trials. [22,129] Intra-operative monitoring by visiting surgeon teams was adopted more frequently (n=8, 29%) in previous oncology trials (Chapter 3). Adjustment of frequency of monitoring of procedures in trials according to initial performance or perceived risk has long been understood and practiced within the radiotherapy oncology trial community, [130] however this is a novel concept within oesophagogastric oncology trials with surgical interventions and has not been previously implemented.

Generic challenges to SQA in oncology trials included: insufficient resources, time and implementation of SQA measures; regional variability in operative standards and patient factors; concerns over patient access to surgical interventions, and; insufficient generalizability of trial results. Although trialists seldom report insufficient resources to implement SQA within published literature, it is a recognised perception amongst those within the wider UK healthcare service. This was demonstrated in a recent interview study of NHS directors in which under-investment within the NHS has been recognised to reduce the quality of surgical services provided.[131] The expert stakeholders concerns within this study regarding patient access to surgical intervention were related to the development of highly standardised surgery within tertiary centres which may not be available to all. This concern was supported by a systematic review of 19 studies investigating whether centralisation of cancer services had been cost-effective, which revealed evidence from four studies indicating that centralised services increased the costs of accessing care for patients and their carers.[132] Similar concerns regarding generalisability of trial results produced within highly controlled, selected and standardised environments have long been expressed as a limitation to the external validity and applicability of such results to 'real-world' surgical environments. [133] Consensus mitigating strategies to overcome generic challenges to SQA in oncology trials included: adequate training and education to improve surgeons' understanding of research methodology, and; trial group learning with regular meetings and focus groups. Although there are efforts to improve undergraduate medical education regarding medical research,[134] in a survey of 100 post-graduate doctors they still feel that

they lack adequate knowledge of medical research with some attributing this to a lack of resources or poor research training. [135] It is often part of clinical trial units mission statement and objectives to conduct regular trial stakeholder meetings to educate, inform and disperse trial group activities,[82] however implementation of this within trials can be variable.

This study has several limitations, one of which is the inevitable selection bias inherent in recruitment of expert stakeholders. Although we endeavoured to recruit impartially through reviewing expert surgeons for those with publications in oesophagpgastric oncology trials, and those previously holding roles as PIs or CIs in such trials attending large surgical conferences in Japan, Europe and the UK, this process is susceptible to convenience sampling in which clinicians who are known to researchers are more likely to be selected. All interviews were conducted in English and although a good level of spoken English was evident with all interviewees, there may have been some loss of meaning through difficulty in expressing complex ideas in participants' second or third language. Expert trial stakeholders included in the Delphi process were predominately male and had backgrounds as surgeons or oncologists. This meant that female stakeholders, trial managers and methodologists were less well represented. As this study relates primarily to quality assurance of surgical interventions it may be expected that surgeons find this topic of higher importance and would thus be more likely to wish to participate in the Delphi process. In addition, as surgery tends to be a predominately male dominated speciality this trend can also be expected. However, stakeholders from each of the four different background specialities (surgeons, oncologists, methodologists and managers), 8 different countries spanning Eastern and Western hemispheres, and from both male and female genders did participate, making it as representative as possible of the expert oncology trial stakeholder community within the constraints of this study.

4.6. Conclusion

This survey of international expert trial stakeholder opinion identified a wide variety of perceived challenges and potential mitigating solutions for different aspects of SQA within oesophagogastric oncology trials. In the first international expert consensus in this area,

agreement has been reached for 59 strategies to overcome challenges to implementation of SQA within 3 Delphi Rounds. These strategies could both aid in the design of future oncology trials and be utilised to overcome challenges to implementation of SQA within active RCTs. In order to assess the potential feasibility of these strategies, an investigation is required into stakeholder perceived challenges to SQA and possible mitigating strategies within an active oesophagogastric RCT. This would allow the opportunity to gain insight into trial stakeholder opinion regarding the applicability of the expert consensus strategies formulated in this chapter, and possibly their future utilisation within a trial cohort. Within two active RCTs, ADDICT (Chapter 6) and Neo-AEGIS (Appendix R), we proceeded to evaluate this utilising a robust qualitative methodology. However, first we sought to explore patient opinion regarding both quality of surgery generically and the quality of surgery within trials.

Chapter 5. Patient perspective on quality of surgery

5.1. Introduction

The expression "nothing about us without us" was first developed by disability rights activists to convey the idea that no policy should be reached without full participation of all stakeholders, and more recently it has been adopted by patient communities seeking broader patient participation within healthcare system planning.[136] In chapter 4, we have explored expert stakeholder opinion but the most important stakeholder must surely be the end-user of oesophagogastric services – the patient. This is the first study assessing patient perspective quality of surgery.

There is an increased focus on patient involvement in care and decision making, from the strategic government perspective within the NHS England 5-year forward view,[137] to the patient embraced and tech-company led arrival of patient feedback websites, smartphone apps and artificial intelligence ostensibly aiding personalised medical decision making. These advances have certainly contested the traditional healthcare structure and challenged doctors decision making and practice.[138]

Use of SQA in upper gastrointestinal (Upper GI) randomised controlled oncology trials has been shown to reduce lymph node yield and in-hospital mortality rate in recent systematic review.[1] However what do patients perceive as important with regard to quality of surgery or quality of surgery in trials?

There have been multiple reports studying patient satisfaction and quality of life following Upper GI surgery.[139,140] Patient perception of factors effecting safety and quality of care has also been studied. Surveys of patient perception of safety of surgery in one qualitative study indicated physician-patient interactions, relationships and trust were the most positive factors influencing their perception of the safety environment.[32] Similar factors affect patient perception of quality of care with another qualitative study involving 174 surveys

demonstrating 'concern/caring of the attending physician (15.7%), and having a physician who listens (10.1%)' were amongst the most important factors.[33]

Although qualitative methodologies have been used to support patients making decisions on their participation in oncology trials,[141] and in several trials in order to establish standardised reporting of outcomes,[142,143] patient perspective on quality of surgery or quality of surgery in trials has not yet been explored. However, the importance of patient empowerment within the oncological care process is becoming increasingly recognised.[144] Exploring patient perspective is often considered the first step towards empowerment and frequently yields results at odds with traditional expert-derived models of care. This was aptly demonstrated in a focus-group study of patient perception of 'empowerment enablers' within the oncological care process. Unlike traditional conceptions regarding patient empowerment, patient opinion did not focus on the direct patient control over their condition, but rather on maintaining an active role within their relationship with caretakers such as 'the ability to choose their doctor/oncological care team/health organisation'. [144]

Within chapter 4, expert opinion regarding to challenges to various aspects of SQA in trials and their proposed mitigating strategies have been explored. We cannot however extrapolate from these findings to elicit patients' perspective regarding quality of surgery. As shown in the qualitative study of 'empowerment enablers' above, [144] there may be a sharp divide between expert and patient perception regarding aspects of the oncological care process. Therefore, within this study we aim to explore patients' perspective on quality of surgery and quality of surgery in trials through the use of robust qualitative methodologies.

5.1.1. Objective

The specific study objective was:

ii. Examination of patient opinion on challenges to quality of surgery

5.2. Method

5.2.1. Participants

Focus groups: Two in depth focus groups were conducted with 10 participants split into two groups of 6 and 4, male:female ratio 7:3. median age 71.5 (range 60-87), facilitated by the main investigator (JB) and another researcher (AT). Participants' median time since surgery was 4.9 years (range 1.5-23.3) and 5 participants (50%) had received either adjuvant or neo-adjuvant chemotherapy or radiotherapy. No focus group participants had previously partaken in a clinical trial. Location of tumours prior to surgery included 6 oesophageal, 2 gastro-oesophageal junction and 2 participants were unable to specify. Focus groups lasted 1 hour 26 minutes, and 1 hour 12 minutes respectively.

Survey questionnaire: Of 41 participants, the median age was 69 years (51 – 86 years) and male:female ratio 23:18. 25 (61%) patients had neo-adjuvant, 2 (4.9%) had adjuvant, and 14 did not receive any chemotherapy or radiotherapy. 36 (87.8%) patients had oesophageal cancer, 3 (7%) had gastric cancer, and 2 (4.8%) patients did not specify the type of cancer. Median time from surgery was 5.79 years (range 0.23 - 23.6 years), and 6 (14.6%) of participants had previously partaken in a clinical trial.

5.2.2. Study Design

Focus groups: Two researchers (JB and AT) facilitated a focus group discussion regarding their opinion on quality of surgery in trials. Participants were encouraged to share their opinion regarding quality of surgery in trials, participants' concerns/perception of challenges regarding quality of surgery and the quality of surgery in trials. Participants' ideas regarding possible solutions to overcome the concerns/challenges identified were also explored. The focus group was used to aid formulation of a survey questionnaire (Appendix H) with open

and closed questions regarding patient perspective on quality of surgery and surgery in trials. Following ethics committee approval (ICREC Reference: 18IC4857, Appendix G) and gaining informed consent from participants, a convenience sample of participants from the Oesophageal Patient Association (OPA), a UK based charity supporting patients with gastric and/or oesophageal cancer, were contacted to request their participation in a focus group. Prior to commencing the focus groups investigators explained the meaning of important relevant terms/expressions and checked understanding of these amongst participants. This glossary of terms included: 'Upper GI'; 'clinical trials'; 'quality assurance of surgery'; 'standardisation of surgery'; 'credentialing of surgeons and/or centres'; 'operative volume of centre and/or surgeon'; 'surgical outcomes'; 'survival', and; 'monitoring of surgery'.

Survey questionnaire: Following the focus group, a convenience sample of OPA members were then contacted electronically and by post by the head of the OPA, to request their participation in a written on-line semi-structured survey questionnaire. Prior to commencement of the written survey it was piloted with three OPA members following their informed consent to check for comprehensibility and feasibility. They were given the option of completing the survey through the online link or by completing the written survey by post. For those participants who requested to complete the survey by post a participant information sheet, consent form and written survey was sent to them, along with a pre-paid return envelope for their completed surveys and signed informed consent forms. This survey (see Appendix H) contained 14 questions of which 8 were open and 6 closed. Of the 6 closed questions there were a total of 51 items for participants to rank their agreement or disagreement on a Likert scale. Questions 9-11 were based upon emergent themes from the focus group and explored patient opinion regarding participation in clinical trials and their thoughts/concerns on the evening prior to surgery. Although questions 9-11 can be seen in the survey (Appendix H), the results of these individual questions have not been reported in the results as they are not relevant to the study objectives within this thesis. Open questions sought to explore patients' opinion without influencing them through the structure/phrases employed. An example of an open question included in the survey: 'What are your thoughts on quality of surgery in trials?' Closed questions sought to gain clarity on specific aspects of patient perspective identified within the focus group as emergent

themes. An example of a closed question included in the survey in which participants' can rate their disagreement or agreement with different items on a Likert scale:

'Please indicate whether you disagree or agree on whether the following factors are important in your perception of quality of surgery:

- Confidence in the operating surgeon
- Follow-up care by the surgeon following leaving hospital'

5.2.3. Analysis

Focus groups: Group discussions were transcribed verbatim and qualitatively analysed an independent researcher (JB) using Nvivo software (QSR International 12.2.0) to identify emergent themes, according to Grounded theory. Grounded theory, developed in the 1960s by Glaser and Strauss, is a qualitative research methodology that aims to explain social phenomena at a conceptual level.[145] Compared to other qualitative methods it has ability to achieve understanding of, rather than simply describing, a social phenomenon. Analysis focusses on categorising data according to emergent themes (concepts) and describing those categories in terms of their properties or characteristics which give them meaning.[145]

Survey questionnaire: Results of the on-line survey were analysed utilising Nvivo software (QSR International 12.2.0) with the same iterative qualitative methodologies described above, in order to identify and categorise emergent themes.

5.3. Focus Group Results

45 emergent themes (Summarised in Tables 14 and 15) fell within 3 over-arching categories: patients' perspective on quality of surgery (n=16); patient's concerns/perception of challenges regarding quality of surgery, and surgery in Upper GI trials (n=16), and; patient's ideas relating to possible solutions to challenges to quality of surgery in Upper GI cancer trials (n=13).

Table 14: Patients' perspectives on quality of surgery and associated challenges

PATIENT PERSPECTIVE ON QUALITY OF SURGERY	
Confidence in operative surgeon	'If he treats you like a real person then you'll have confidence in him'
Personal and Professional attributes of surgeon	'Relaxed', 'Confident', 'Expert'
Trust in surgeon and surgical team	'I don't think you can separate the surgeon from the team within operations…' Patient awareness of surgical training and inherent trust
Quality of life and survival	Risk of recurrence versus reward of not having such an invasive procedure
Importance of the Multidisciplinary team (MDT)	MDT; nurses; dietician; physio
Aftercare	Postoperative care described as '1st class'
Concept of quality - Return to normal physiological function	'An operation that gets you back to normal'
Outcome statistics – Difficulty in interpretation	Difficulty in interpretation of outcome statistics for patients – 'I said does that mean he gets the most difficult cases?
Communication with patient relatives	'it's tough for the supporters, it really is.' Forewarning relative of expectations – e.g. appearance of relative in ITU
Operative volume of surgeon	'40 oesophagecotmies per year – they must be doing something right'
Operative volume of surgical centre	'You want to be a routine operation'
Surgical team management of complications	'Yes, and the ability because the chances are they would have had more exposure to the complication.''
Standardisation of surgery	Patients felt standardization was important, though had some concerns: '..would it stop surgeons being creative?'
Difficulty in commenting on quality of surgery	Difficulty in defining surgical quality was expressed: 'What is quality of surgery? That is a very, very good question.'
PATIENT PERCEPTION OF CHALLENGES REGARDING QUALITY OF SURGERY	
Concerns regarding cost-effectiveness paradigm	Concerns that cost-effectiveness of operating elderly with oesophagogastric cancer may be considered in resource limited health service
Funding	'If you haven't got the funding nothing flows from it.'
Delayed detection of Upper GI cancer in primary care	Concerns of lack of investigating and awareness of oesophagogastric cancer amongst GPs/other health care professionals
Lack of justification and explanation for treatment options	'I personally would want a tried and trusted and I say that being very selfish because you can only learn by doing the trials.'
Finding surgical team with sufficient experience	'I think the main challenge is the rarity of the disease and the ability to find a team that's had enough exposure to treating it'
Lack of contact/input from certain MDT members	'I was still a bit uneasy that this man was supposed to be my consultant and hadn't bothered to come to see me at all.'
Diagnostic process intimidating	'…I just wondered whether or not that could be made somehow a little bit less intimidating?''
Poor communication with relative	'They mentioned to my partner that I was dying, but I don't think they ever came back to her to say he's coming around again'
Insufficient number of available intensive care beds	Morning of surgery – not knowing if surgery would go ahead due to uncertainty over bed
Fear of loss of control	'I decided it's a waste of time trying to pretend you've got control…'
'Breaking bad news' concerns – other members of MDT	'Sudden jump to oesophagectomies – ….too much'. 'Perhaps a specialist nurse is needed to bridge the gap'
Limitations of operative volume in selecting surgeons	Operative volume of surgery may not be reliable indicator of quality as the surgeon may do the hard/complex cases/different stages ect
Poor mental health support	Long wait -4 months for MacMillan nurses
Lack of clinical equipoise/enthusiasm/confidence with trials	'…one surgery must be better, would I be putting myself at risk?'
Perceived lack of imaging modality in follow-up appointment	No MRI scan in follow-up appointments

Key: GP – General Practitioner

Table 15: Patients' proposed mitigating strategies to challenges to quality of surgery

PATIENTS' PROPOSED SOLUTIONS TO CHALLENGES TO QUALITY OF SURGERY	
Facilitating early diagnosis	Screening program required - possibly non-invasive screening. 'We can cure every cancer if we catch it in time. That's probably true.'
Improving patient pre-operative fitness	Important to improve post-operative recovery
Team and anaesthetic contact pre-operatively	Pre-operative contact helps give patients a sense of control:'You're part of the team as well. You're not just the patient'
Funding	Improved funding (as found abroad in countries like Japan) could help provide more counselling services
Emotional support	Emotional support advocated for relatives and patients. 'Because it's too much for the partner often'
Spiritual support	'Essential to be offered spiritual support'.
Structured follow-up care post-operatively	Patients felt this should be more comprehensible and structured
Surgical team recognition of importance of patient's relative	Partner consultation in private advised to give them realistic pre and post-operative expectations
Centralisation of surgical services	Multiple participants expressed they would be happy to travel for centralised cancer services: 'all the experts all together'
Standardisation of surgery	Standardisation of surgery should be supported within trials with space for anatomical variation
Structured method of assessing surgeons' skills	'Airline pilots get tested and checked out regularly, so I'm sure a surgeon should be checked out as well.'
Monitoring by peer review	Participants supported peer review monitoring, and expect it to be in place with senior surgeons monitoring more junior surgeons

Within their 3 major categories emergent themes are summarised and presented below along with supporting quotations. It became apparent early on within the focus group that patients' had an eclectic perspective regarding 'quality of surgery'. For study participants', all aspects of pre-operative investigation and work-up for surgery including discussions between the patient, their relatives and the surgeon, members of the MDT and/or anaesthetic team, the surgery itself, support for patients' relatives, and all aspects of post-operative care were included within patients' conception of the 'quality of surgery' they had received. Therefore, all of these factors and more were considered relevant and were introduced by participants on asking their opinion regarding 'challenges to quality of surgery' generically, and more specifically regarding 'quality of surgery within clinical trials'.

5.3.1. Patient Perspective On Quality Of Surgery

Prominent emergent themes regarding patient perspective on quality of surgery included: Trust in surgeon and surgical team; communication with patient relatives, and; operative volume of surgeons and centres.

Trust in surgeon and surgical team

The concept of trust in the surgeon and the surgical team was reinforced. For many participants 'The whole thing about feeling confident in the surgeon is absolutely critical.' The concept of trust appeared to be built of belief in the surgeons skills: 'I'm taking this man on trust that he could do it and get me out the other end'; in addition to his/her interpersonal skills and provision of holistic care -'If he treats you as a real person then you'll have confidence in him.'

Communication with patient relatives

Communication with patients' relatives was also considered important to the development of trust in the surgeon, and to the concept of 'quality of the patient

experience'. Universally participants felt the surgical process was most difficult for the patients' relatives: 'it's tough for the supporters, it really is.' The surgeon forewarning the relative of pre-operative expectations including 'the appearance of their relative in ITU' was felt to be a vital component of surgeon-patient communication. Furthermore, the surgeon taking the time to call the relative post-operatively regarding the operation was felt to be important with one participant describing it as the surgeon making an 'extra little difference'.

Operative volume of surgeons and centres

Operative volume was considered important in building patients' conception of trust in their surgeon. One participant made an analogy to his own experience as a 'roofer', explaining the significance of surgical volume and trust: 'He does it all the time. I do it all the time. Trust me. Trust him.' On asking participants which they feel is more important – volume of operating per surgeon or per centre, participants felt that surgeon volume was most essential with one participant responding: 'the surgeon'. Many participants also felt that a specific number corresponded to a sufficient volume of operations annually in order for the surgeon to be a specialist: 'Surgeon…., does like 40 or 50 a year or something so we're going to send you down to him 'cause that's what he specialises in. That's part of the confidence thing.'

5.3.2. Patient Perceived Challenges

Key emergent patient perceived challenges to quality of surgery and quality of surgery in trials included: Delayed detection of Upper GI cancer in primary care; insufficient mental health support; disbelief in the concept of clinical equipoise, and; and a lack of enthusiasm and confidence in clinical trials.

Delayed detection of Upper GI cancer in primary care

Concerns regarding early detection of upper GI cancer within primary care centred on the perception that delayed detection will lead to the condition progressing. This is highlighted with a comparative reference was made in this regard to the prompt investigation of patients with upper GI symptoms in Japan: 'I'm led to believe in Japan that if you go the GP there the moment you tell them you've got a problem in the throat there's no question, you're straight away given the camera down the throat.'

Insufficient mental health support

The emotional and psychological distress caused by the treatment process of upper GI cancer was highlighted with one participant explaining: 'I mean I lost it a few times I really, really used to sit there and just cry my eyes out.' The lack of available counsellors following for patients with upper GI malignancy was highlighted with one participant stating he started going to see his wife's counsellor he was so desperate to talk to a professional who would listen: 'But you feel like you want to talk to somebody, I got to a stage where I was going to see... my wife's a counsellor'

Disbelief in the concept of clinical equipoise

An emergent challenge to quality of surgery in clinical trials was participants' lack of belief in the concept of clinical equipoise. They felt there should not be uncertainty regarding the optimal surgical procedure available for a disease, but rather some participants felt that there must a gold standard for all forms of surgery: 'Surely the concept of a gold standard type of operation exists in all sorts of surgery. It must do, so that's what you want. You don't want anything that's remotely experimental and a bit 'iffy' in outcome.'

A lack of enthusiasm and confidence in clinical trials

Participants felt that a clinical trial may place them at undue risk and that this would affect their confidence in the surgical procedure within the trial: 'But in a trial I don't think I'd be as confident …… Would I be putting myself at risk if I agreed to the trial'

5.3.2. Patients' Proposed Solutions

Key emergent patient proposed strategies to overcome challenges to quality of surgery and quality of surgery in Upper GI trials included: Preoperative fitness; monitoring of surgery; supporting patient's relatives, and; early detection of gastro-oesophageal disease.

Improving Patients' preoperative fitness

Participants felt strongly that improving their pre-operative fitness would aid in their overall physiology and capacity to recover post-operatively: 'I feel that if I hadn't made efforts to get myself stronger and fitter I might not have recovered as well as I did or at all. It's everything from how strong your heart is as to whether your arms are strong enough to push you up in the bed. '

Monitoring of surgery

Participants felt monitoring in the form of peer review was important and that this is in place in tertiary centres: 'I always assumed that there would be a continuing peer group review going on.' However, participants also expressed the idea that it may be difficult to monitor surgery due to the 'inexact nature of surgery'. On being asked regarding methods of monitoring surgery and possible use of an 'assessment tool', patients were in favour of this explaining that similar methods are utilised within the aviation industry: 'Airline pilots get tested and checked out regularly, so I'm sure a surgeon should be checked out as well.'

Supporting patient's relatives

Giving patient's relatives realistic expectations of what to expect peri-operatively was felt to be important. Participants felt given the pivotal role of patient's relatives in this process they should have their own individual appointment with the consultant in order to ask questions they may not be able to ask in the presence of the patient: 'I think in terms of quality surgery and treatment of a partner, a partner or someone who is designated by the patient needs to have a consultation on their own without the patient at some stage or be offered it.'

Early detection of gastro-oesophageal disease

The importance of early detection is felt to be a key strategy to improve quality of surgery and one participant quotes his surgeon to emphasise this: 'He said we can cure every cancer if we catch it in time. That's probably true.' Other participants linked early detection with funding and they correlated this with a higher chance of curing their disease. Reflecting on research one participant was involved in at Imperial College London he reflects on the potential for breath testing as a cheaper surveillance method: 'The blowing into a bag is painless. It can be done in your GP's surgery. Brown paper bags are very cheap.'

5.3.4. Summary

This focus group study revealed a diverse range of 45 emergent themes relating to patient opinion regarding quality of surgery, challenges to quality of surgery and potential mitigating strategies. Key emergent challenges to quality of surgery in upper GI cancer trials included: Delayed detection of Upper GI cancer; insufficient mental health support; disbelief in the concept of clinical equipoise in trials, and; a lack of enthusiasm and confidence in clinical trials. Key proposed strategies to overcome these challenges included: Early detection of gastro-oesophageal disease; improving patients' preoperative fitness; monitoring of surgery, and; adequately supporting patients' relatives. Essentially, this focus group was the first step towards

development of the 14 question semi-structured survey, allowing further exploration of patient perspective on quality of surgery, quality of surgery in trials, associated challenges and potential mitigating strategies. It became clear during the focus group that participants were more comfortable in discussion challenges to ‘quality of surgery’ in general rather than those specific to Upper GI oncology trials, therefore within the survey questionnaire there was an increased focus on challenges to ‘quality of surgery’ in general and potential solutions.

5.4. Survey results and discussion

Patient responses to Likert-response questions are summarised in Table 16 in which 'participant support' corresponds to the proportion of participants who indicated that they either 'agree' or 'strongly agree' with each item listed. Themes from open questions and the comments section of Likert-response questions are outlined within the survey results below. Detailed participants' responses, quotations and all emergent themes can be found in appendix I. A summary of key important themes along with supporting quotations are outlined below and discussed in context of the literature. Seminal findings included a high proportion of participant support within the following sub-categories: Factors of perceived importance in influencing patient perspective on quality of surgery and their confidence in the operating surgeon; patient perceived challenges to quality of surgery, and; strategies recommended to overcome identified challenges to quality of surgery (See Table 16).

Table 16: Summary of patient perception in Likert-response survey questions

Category	Emergent themes	Participant support (%)
Factors important in quality of surgery	Confidence in the operating surgeon	95.2
	Trust in operating surgeon	95.5
	Care received by the multi-disciplinary team (MDT)	95.5
	Care received by the Intensive Care Unit	95.5
	Care received following operation on the ward	90.9
	Follow-up care by surgeon following leaving hospital	95.5
	Follow-up care by other members of MDT (including oncologists, nurses, and physiotherapists)	84.1
	Quality of life patient experiences following the operation including patient experience of discomfort and recovery time following the operation.	86.1
	Surgical centre operative volume (number of cases operated per year)	70.7
Factors contributing towards confidence in the operating surgeon	Perceived self-confidence of the surgeon	92.9
	Surgeon's perceived character	90.5
	A surgeon's explanation of diagnosis and management plan	97.6
	A surgeons' operative volume (number of cases they perform per year)	80.0
	The operative volume of the centre in which the surgeon works (number of cases performed in that centre per year)	77.5
	Survival outcome statistics of that surgeon	85.4
	A surgeon's holistic treatment of the patient	78.6
	A surgeon's treatment of a patient's relative (e.g. allowing relatives to ask questions or setting expectations)	95.2
Challenges to quality of surgery	Lack of funding	75%
	Lack of sufficient beds for operations to be performed (e.g. Intensive Care bed spaces)	90%
	Insufficient communication or clinical contact	56%
	Lack of sufficient screening for Upper GI cancer in the UK	92.5%
	Anxiety due to prolonged waiting for a diagnosis	66%
Strategies to overcome the challenges to quality of surgery	Education and training	94.1
	Raising awareness of Upper GI cancer and postoperative complications	94.1
	UK screening program introduction for Upper GI cancer to aid early detection	97.3
	Non-invasive breath testing as a screening tool for Upper GI cancer	91.2
	Patients receiving physiotherapy and cardiovascular training pre-operatively to enhance fitness	76.47
	Provision of counselling service to patient and relative/next of kin pre and post-operatively	85.3
	Provision of personal appointment for patient relative with operating surgeon before and after surgery	78.79
	Use of a structured method of assessing Upper GI surgeons' skill/competencies on a regular basis	88.23
	Monitoring of surgery in Upper GI centres to check standards of operating are being met	94.12

The majority of participants (95.2%) agreed that their 'trust' and/or 'confidence in their surgeon' influenced their perception of quality of surgery, explaining it is 'vital before surgery'. Multiple participants explained this trust developed through reference to the 'thorough explanations' and reassurance they had received from their surgeons pre-operatively: 'He instilled confidence that he would do a good job, very thorough in explaining procedure, and very sympathetic and professional in his approach'. The prominence of the emergent concepts of patient trust and confidence in the operating surgeon within this study were in keeping with surveys of patient perception of safety of surgery in another qualitative study, in which physician-patient interactions, relationships and trust were the most positive factors influencing their perception of the safety environment.[32] Most participants agreed that the following factors contributed to their confidence in the operating surgeon: a surgeons' operative volume (number of cases they perform per year) (78%), and; the operative volume of the centre in which the surgeon works (number of cases performed in that centre per year) (75.6%). Although participants supported patient access to survival and other outcome statistics, they advised caution in its interpretation without sufficient understanding of the complexity of the cases: 'But probably should not without knowledge of difficulty of cases treated.' Although the surgeons' operative volume may significantly contribute to a patient's confidence in their surgeon, it may not be as closely linked to the confidence that a surgeon has in their own ability. A low operative/case volume was recognised as a contributing factor to some graduating surgeons' low self-confidence within one qualitative study, however other studies found no correlation between case volume and a residents' confidence or perception of competence.[146]

Prominent challenges to quality of surgery included: insufficient resources, and; insufficient investigation and screening for gastro-oesophageal cancer. Patient agreement regarding other barriers to quality of surgery are represented in figure 7 and table 16. The majority of participants felt that there was a lack of funding (73.2%) and insufficient beds for the operations to be performed (e.g. Intensive Care bed spaces) (87.8%). A general lack of resources and support for staff was cited as a

challenge to quality of surgery: 'Lack of support for surgeons and staff. Time and equipment pressures'. Patient opinion is in keeping with a recent interview study of NHS directors in which under-investment within the NHS has been recognised to reduce the quality of surgical and imaging services provided according to a qualitative study in which NHS directors were interviewed. [131] Most participants also expressed a concern regarding insufficient screening (90.3%) and the slow pace of investigation within primary care (66.0%) for Upper GI cancer in the UK. Concerns regarding the risk of disease progression due to late diagnosis secondary to a lack of screening was expressed: 'Until patients have symptoms no preventative screening appears to be done and by then the disease is, sadly, often well advanced.' The importance of screening in detecting oesophagogastric cancer at an earlier stage and the consequent impact on survival has long been recognised and is reinforced within published literature.[147] Indeed the higher 5-year survival rates from gastric cancer reported in Asia (69%) compared to those in the Western World (10-30%), are thought to be attributable to the widespread availability of screening programmes in Asia.[147]

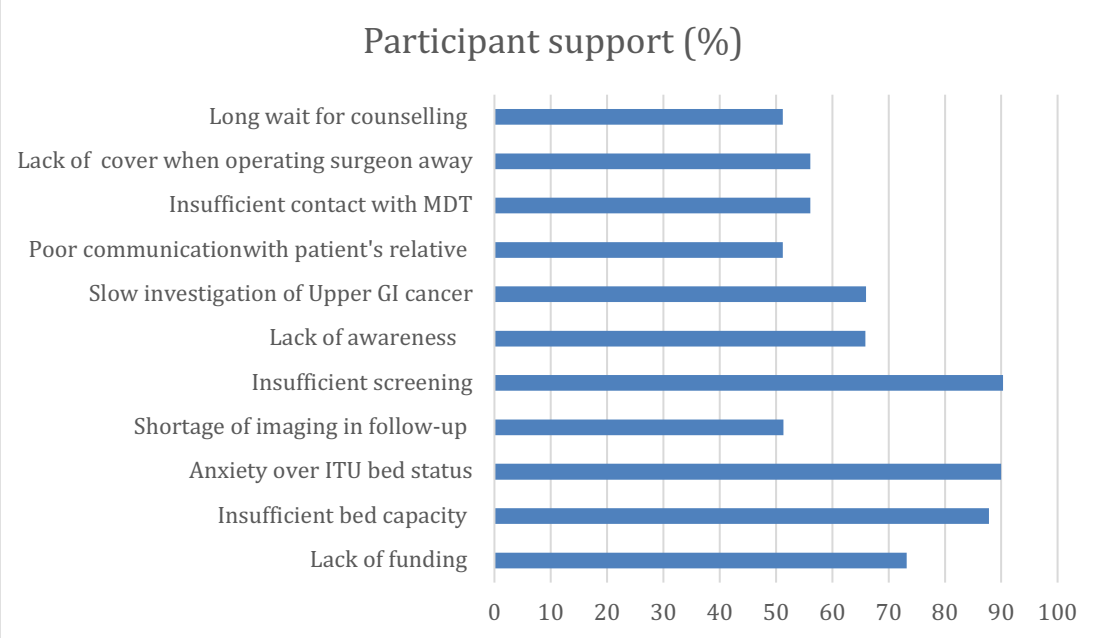

Figure 7: Participant support for specific challenges to quality of surgery

Key emergent challenges to quality of surgery in Upper GI cancer trials included: over stringent standardisation, and; concern over potential for one trial arm to be inferior. The concept that standardising surgery may not allow for operative variation required for each individual or hospital was expressed. Participants cautioned that surgeons may be 'enforced to follow a particular method of surgery as opposed to one which is bespoke to individuals needs'. In keeping with this concern another explained: 'Every patient is different. Every hospital is different'. This thematic challenge is in keeping with the philosophy of certain trialists who advocate for pragmatic trials. In one review of trial typology the author argues that for pragmatic trials, attempts to standardise surgery often create difficulties, and ensuring that each step was delivered as planned would be unrealistic.[128] Participants had concerns one arm of the trial may be inferior with regards to quality of surgery: 'Inclusion in inferior arm of trial'. Similar patient concerns regarding the potential for psychological distress due to randomisation to a trial arm with inferior outcomes has been previously reported within oncology trials.[148]

Seminal strategies to overcome identified challenges to quality of surgery included: further education/training for non-specialist centres in gastro-oesophageal cancer; screening program introduction; credentialing of centres; monitoring of surgery, and; centralisation of services. Participant support for these and other proposed strategies are represented in table 16 and figure 8. The majority of participants (94.1%) supported the strategies of further education or training for community health services and GPs in recognising symptoms of Upper GI cancer to aid in early diagnosis. The same majority (94.1%) agreed with raising awareness of Upper GI cancer and potential postoperative complications amongst other health care professionals and non-specialist centres. Many participants expressed concern regarding GPs prescription of indigestion medication rather than further investigations of symptoms of upper gastro-intestinal cancer as a justification for this strategy: 'Definitely. I had no typical symptoms. Too many GP's prescribe indigestion medications, sometimes on repeat prescriptions, but do not do a follow up.' In keeping with participants' proposed strategy of improved education within primary care in our study, one review of diagnosis and treatment of gastrointestinal disease

in primary care advised updating knowledge and skills of primary care physicians via continuing medical education is the only way to improve adherence with standards and quality of care for patients.[149] Enhancing the experience, training and competence of the surgeon themselves, also became an emergent strategy to overcome challenges to quality of surgery. One participant explained how the quality of surgery performed may improve with training: 'Clearly, the quality of surgery (or, at least, the outcome of surgery) is enhanced by the experience, understanding, training and competence of the surgeon. If those factors can be enhanced then quality may be improved.' The importance of training, experience and a surgeon's learning curve and their association with patient outcomes have been extensively discussed within published literature.[150,151] Introducing a screening program in the UK for Upper GI cancer to aid early detection of disease (80.5%) and non-invasive breath testing to facilitate early detection of Upper GI cancer (75.4%) were strategies supported by most participants to overcome challenges to quality of surgery. The importance of improved investigation and screening in an effort to prevent disease progression was emphasised: 'This is badly needed as the disease can be quite advanced once diagnosed when symptoms present themselves'. Participants explained they had been impressed by the non-invasive breath test they had witnessed: 'I was in a working group on the early development of these and very impressed.' Advances in technology are bringing improved cancer screening and early diagnosis to the forefront within the fiscally stretched UK healthcare system, with one app piloted in Manchester offering smokers and ex-smokers CT scans in supermarket car parks quadrupled early diagnosis rates of lung cancer and has since been rolled out over North-Manchester.[152] Non-invasive breath testing for gastro-oesophageal cancer, currently being developed and piloted in primary care, may help to provide enhanced levels of early cancer detection.[153]

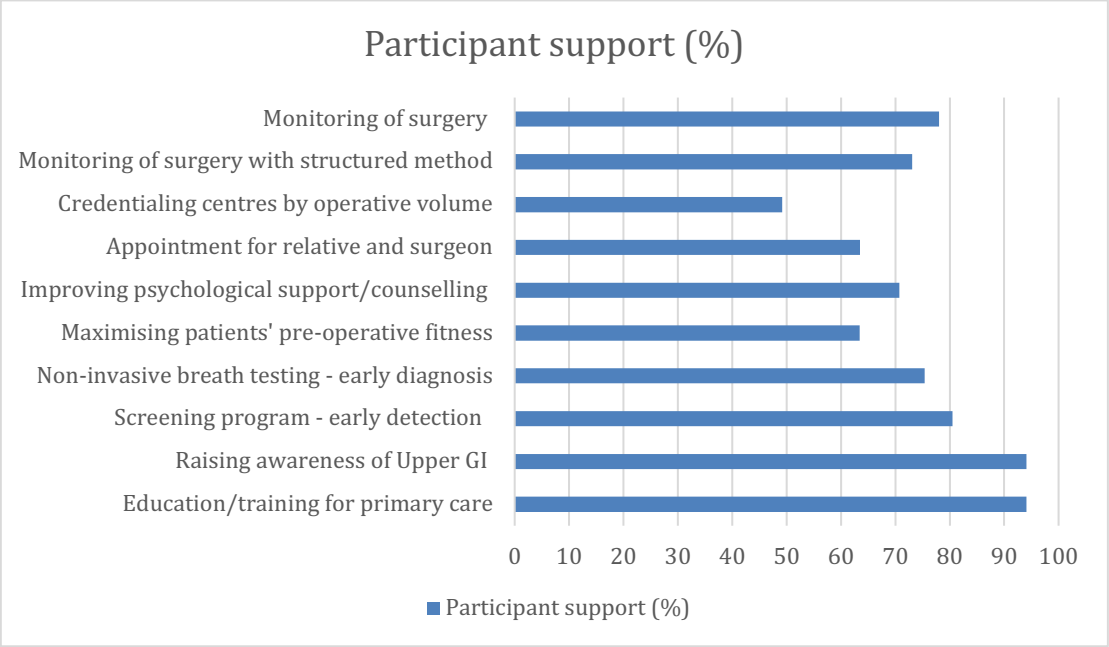

Figure 8: Participant support for strategies to overcome the challenges to quality of surgery

Credentialing Upper GI centres by specifying that they should perform a certain number of cases per year (operative volume) was supported by just less than half (49.2%) of participants to overcome challenges to quality of surgery. Limitations of using case volume alone for credentialing surgeons or centres for participation in trials has been previously recognised and noted as one of the key challenges to quality assurance of surgery in trials.[34] However, given the lack of reliable assessment tools for assessing quality of oesophagogastric cancer resections until recently, operative volume has been the most commonly employed method to credential trial centres and surgeons even in those surgical trials with the strictest SQA measures. [129] The majority of participants supported monitoring of surgery with a structured method of assessing Upper GI surgeons' skill/competencies (73.1%), and/or monitoring of surgery to check standards of operating are being met (78.0%). Although the majority agreed on the importance of operative monitoring some participants cautioned that such monitoring may affect outcomes: 'Successes rates could be affected if being monitored for good results.' Patient approval of this strategy to improve monitoring of quality of surgery using a structured method concurs with one of the expert stakeholder consensus strategies developed within chapter 4 (Section 4.4). Historically trials have been lacking in methods of objectively

monitoring surgical quality and development of such a tool is considered a critical step in improving quality assurance of surgery.[34] Participants felt that including centres in which services were centralised within Upper GI trials allows inclusion of 'Areas of Excellence', which in turn can improve quality of surgery. There may be an element of recall bias in this however as all participants are survivors of surgery in UK hospitals which are already centralised for oesophagogastric cancer surgery. A qualitative study revealed similar positive patient perceptions of centralisation of specialist cancer surgical services with important factors behind this including: highly trained staff; waiting time for cancer surgery, and; access to staff members from various disciplines with specialised skills in cancer care.[154]

5.4.1. Patient perspective on quality of surgery in context of this thesis

It is difficult to make direct comparisons between the results of Chapter 4 and those of Chapter 5. This survey focuses predominately on patient perspective of quality of surgery in general, whereas within the expert stakeholder Delphi process (Chapter 4, section 4.4), consensus was reached on mitigating strategies to overcome challenges to quality assurance of surgery in oncology trials. The expected difference in the direction of these two chapters and their subsequent results is due primarily to the difference in the research question, in which expert stakeholders were asked their opinion specifically regarding 'quality assurance of surgery in oncology trials', rather than the more generic theme of 'quality of surgery' utilised in this patient perspective study. Additionally, these two studies involved different sample populations with different levels of knowledge/expertise in the area concerned, and there will inevitably be differences in perspectives between those delivering care or designing/managing clinical trials, and those who have been the direct recipients of such care. However, despite these differences there were some similarities and comparisons worth noting. In this study factors found to influence patient perspective of quality of surgery or patient confidence in the operating surgeon included: surgical centre operative volume (number of cases operated per year), and; survival outcome statistics of that surgeon. Similarly, expert trial stakeholders reached consensus (Chapter 4, Section 4.4) that trial centre operative volume should

be considered in selection of trial centres for participation in oncology trials. They also agreed that audit of post-operative complications and outcomes is important in selecting surgeons for participation in trials. Patient perceived challenges to quality of surgery included a 'lack of funding'. Insufficient resources and financial support were also identified by expert trial stakeholders as a significant challenge both in monitoring of surgery and as a generic challenge to quality assurance of surgery in trials (Chapter 4, Section 4.3). Strategies recommended by patients to overcome challenges to quality of surgery included: use of a structured method of assessing Upper GI surgeons' skill/competencies on a regular basis, and; monitoring of surgery to check standards of operating are being met. Likewise, expert trial stakeholders agreed upon the use of a structured objective assessment tool as the optimal method for monitoring of surgery within oncology trials (Chapter 4, Section 4.4). Along with the consideration of trial centre/surgeon operative volume, patient support for monitoring of surgery demonstrates patient approval for two of the expert trial stakeholder consensus strategies (Chapter 4, section 4.4). There were also significant differences in patient perception with regard to quality of surgery when compared indirectly to the opinion of expert trial stakeholders between chapters 4 and 5. This is demonstrated by the fact that 25 (83%) of the 30 factors considered important by patients in regard to quality of surgery and/or quality of surgery in trials, were not mentioned within the expert trial stakeholder study within Chapter 4. These factors of perceived importance within the patient perspective investigation which were missing from the expert stakeholder study, included a wide variety of themes ranging from surgeons' holistic treatment of the patient and their relative including improved psychological post-operative patient support, to advocating introduction of a comprehensive screening programme for Upper GI malignancy in the UK.

5.4.2. Limitations

This study involved a convenience sample from the OPA, and as such involves only survivors of oesophagogastric cancer surgery. This was due to the accessibility of this cohort, the endeavour not to recruit patients actively involved in a trial, and the

desire to recruit patients who may have some knowledge of surgical quality or clinical trials. This inevitably leads to recall bias in that participants will be more likely to perceive the quality of their surgery positively given their own survival. 2 (20%) of focus group participants were unable to recall/specify the location of their tumour. This may indicate problems in communication from the surgical team or understanding on behalf of the patients within this cohort, and it may further indicate space for recall bias within this cohort. Additionally, within questions relating to quality of surgery in upper GI trials many participants indicated they did not understand some of the concepts described in the question. We endeavoured to mitigate against this problem through introducing and explaining key concepts and terminology at the beginning of the focus group discussion. This limitation may have been more pronounced within the survey as here there was not the opportunity to further explain the key concepts and terminology prior to commencement. The survey was piloted with three OPA members to check it for comprehensibility and feasibility and positive feedback was received indicating that the language and terminology employed was comprehensible and appropriate. Furthermore, unknown variables such as participants' educational level and/or socio-economic background, which are difficult to control for, may have influenced survey question comprehension. However, despite these challenges we still received a diverse range of both positive and negative perceptions for each survey question from the majority of participants.

5.5. Conclusion

In the first survey of patient perception of quality of surgery a high level of participant support was evident regarding patient perceived challenges to quality of surgery including: Insufficient resources and/or hospital beds for operations to be performed, insufficient screening for Upper GI cancer, and; insufficient training of trial surgeons. Importantly, this study identified patient support for consideration of trial centre/surgeons' operative volume and the use of a structured method to monitor surgery, two of the previously developed expert consensus strategies in chapter 4. Increased patient education and engagement regarding clinical research

at community and specialist centre levels are required to allow for more informed decision making at important junctures in their oncological care. Public healthcare planners and members of the surgical oncology community should heed attention regarding the importance of patients' reported challenges and their proposed strategies to improve quality of surgery, including implementation of a nationwide screening programme for Upper GI cancer, and a monitoring mechanism to ensure expert agreed operative standards are being achieved. Within chapters 4 and 5 expert and patient opinion regarding quality of surgery has been explored. Chapter 6 further affords the opportunity to assess expert opinion regarding SQA in addition to patient recruitment within an active oesophagogastric RCT. Moreover, the chapter 6 allows examination of stakeholder opinion regarding the potential feasibility of previously developed expert consensus strategies from chapter 4.

Chapter 6. Challenges to Recruitment of patients and monitoring of surgery in the ADDICT trial

6.1. Introduction

Gastrectomy with D2 lymphadenectomy is the standard treatment for curable gastric cancer.[98] An RCT designed to investigate whether there is a survival benefit from para-aortic nodal dissection (PAND) in addition to D2 lymphadenectomy for stage T2-4 tumours demonstrated no improved survival rate in curable gastric cancer.[98] Another RCT compared the feasibility of lymph node dissection in open surgery and laparoscopic surgery for advanced gastric cancer (COACT 1001).[155] Within this study, subgroup analysis of the 11p and 12a lymph node resection (routinely removed within a standard D2 gastrectomy) revealed that for stage IIB/IIIA gastric tumours they had lymph node metastases in 2.1% and 2.4-12.1% respectively. However, in stage IB/IIA gastric cancers, 0% metastases to 11 p and 12a lymph nodes were found.[155] Therefore, it was evident that 11p and 12a lymph nodes, which belong in D2 lymph nodes, need to be resected in advance gastric cancer in stage IIB or higher. However, in earlier stages of advanced gastric cancer, the probability of metastasis is very low; therefore it was hypothesised within ADDICT that resection of D1+ lymph nodes, excluding 11p and 12a, might be enough.[156] Obviously a gastrectomy involving a more limited lymphadenectomy (less lymph node removal) would be considered less invasive, potentially minimising the frequency of complications, and if survival rates were maintained, it could be considered a preferable surgical intervention for this patient cohort. Other potential advantages of a less invasive gastrectomy may include a reduction in operation time, cost, and improved quality of life.[156]

Thus the findings of the COACT 1001 RCT described above opened the route for the ADDICT trial, a multicentre randomized trial (RCT) compares D1+ versus D2 distal gastrectomy for stage IB & II advanced gastric cancer (COACT1201) with the primary aim to test non-inferiority of survival of D1+ gastrectomy versus D2 gastrectomy.[23] Recruitment of sufficient participants has been an issue for many surgical RCTs,

leading investigators to explore methods of improving recruitment.[157,158] Qualitative research can potentially improve the efficiency of trials by identifying problems with recruitment, enabling trialists to then address problems identified and increase or optimise recruitment.[158,159]

Monitoring of surgical quality within the ADDICT trial has consisted of intraoperative photographs distinguishing between D1+ and D2 procedures and a short 1-2 minute video clip on completion of lymph node dissection. Of 479 participants enrolled 397 (83%) participants have had photographs uploaded potentially indicating problems with trial centre adherence to the monitoring process. Recruitment has also been a challenge and has often lagged behind projected targets.[160] The ADDICT trial has previously extended the recruitment period in order to meet pre-specified sample-size targets due to slower than expected recruitment. In chapter 4 expert oncology trial stakeholders reached a consensus on mitigating strategies to overcome challenges to SQA in oncology trials. In order to assess their feasibility, these strategies need to be considered for implementation within an active oncology RCT, an opportunity uniquely provided by the ADDICT trial.

The aim of this study was to identify the challenges to recruitment of participants and monitoring of surgical quality within the ADDICT trial and gain stakeholder consensus on potential mitigating strategies using a Delphi process. Delphi methodology has been previously employed to support patient decision making regarding their participation in oncology trials,[141] and in several trials in order to establish standardised reporting of patient outcomes.[142,143] However, Delphi methodology has not as yet been utilised gain consensus on strategies to overcome challenges in recruitment of patients or monitoring of surgery faced by trial stakeholders.

6.1.1. Objectives

The specific study objectives were:

ii. Examination of expert opinion on challenges to SQA

iv. Investigation of SQA and recruitment within an active trial

6.2. Methods

6.2.1. Participants

A purposive sample of key stakeholders was identified within the ADDICT trial consisting of surgeons, oncologists, trial methodologists and trial managers. Stakeholders were selected on the basis of their involvement in the trial by the chief investigator (CI).

6.2.2. Study design

A qualitative research methodology consisting of four phases was employed, utilising a robust and iterative design to capture opinion of all key stakeholders. Local ethics approval was confirmed prior to commencement of study (Appendix J, IREC:18IC4347)

Phase 1.

Phase 1. Presentation of results of expert stakeholder study to ADDICT

A meeting by teleconference was held to present results of the expert trial stakeholder study (chapter 4, section 4.3) to ADDICT surgeons and the ADDICT trial management committee. Participants comprised of a purposive sample of trial committee stakeholders including the ADDICT CI, 2 trial managers (TMs), and 8 surgeon principle investigators (PIs) from different ADDICT trial centres in Korea. During this meeting stakeholders were asked their opinion regarding challenges to

monitoring of surgery and recruitment of patients, potential mitigating strategies and areas requiring further attention within the next stage of the project – the survey questionnaire. In addition, they were presented with a summary of challenges encountered previously by expert trial stakeholders (Chapter 4, Table 12) and possible mitigating strategies from the expert trial stakeholder consensus (Chapter 4, Table 13 - Full List) by the primary researcher (JB). In a subsequent teleconference meeting, utilising the feedback from the previous trial committee meeting and the relevant expert trial stakeholder reported challenges and mitigating strategies from chapter 4, the CI and a TM selected the ADDICT list of challenges and potential mitigating strategies to be included within the survey questionnaire (Table 17). Discussions were recorded, transcribed verbatim and qualitatively analysed for emergent themes using Nvivo qualitative software.

Phase 2.

Survey questionnaire (S1)

A 66 item semi-structured survey with open and closed questions was developed using the emergent themes from the aforementioned ADDICT trial stakeholder meeting (Table 18), in addition to those selected by the CI and a TM from the expert interviews and Delphi consensus project in Chapter 4 (Table 17). This on-line survey was created utilising Qualtrics XM (NPS software) which was particularly convenient as it helped to overcome the challenges of distance between the UK and Korea. The survey consisted of 54 closed questions and 12 complimentary open questions. Once completed, surveys were qualitatively analysed for emergent themes and manually coded by the main researcher (JB). A full paper version of this ADDICT on-line survey is available in Appendix L. Following review of the survey (S1) and approval by the ADDICT CI for appropriateness, comprehensibility and suitability, it was distributed to ADDICT trial stakeholders. Stakeholders within ADDICT were purposively selected by the CI for their role within the trial to request their informed consent prior to participation in the survey.

Table 17: Challenges to monitoring and recruitment, and potential mitigating strategies from previous expert interviews and Delphi consensus (Chapter 4, sections 4.3-4.4) selected for inclusion in ADDICT survey

CATEGORY	EXPERT TRIAL STAKEHOLDER PERCIEVED CHALLENGES AND RELEVANT DELPHI CONSENSUS STRATEGIES
Challenges to monitoring	**Surgeons' pride leading to difficulty in providing and receiving feedback** **Surgeons' reluctance to be monitored** **Surgeons' concerns over adverse outcomes associated with monitoring of surgery** **Surgeons' concerns over potential litigation associated with monitoring of surgery** **Potential negative impact of monitoring of surgery on generalizability of results** **Potential negative impact of monitoring of surgery on surgeon recruitment**
Challenges in recruitment	**Over-stringent selection of surgical centres** **Over-stringent standardisation of surgery**
Potential recruitment strategies	**Trial centres and surgeons increasing collaboration and working towards shared goals** **Flexible standardisation of surgical procedures focusing on those aspects affecting safety or survival** **Adjusting centre selection according to the following factors may improve recruitment:** • Operation Complexity • Disease Prevalence **Adjusting standardisation according to the following factors may improve recruitment:** • Disease process studied in the trial • Trial design (explanatory versus pragmatic)
Potential monitoring strategies	**Strategies to overcome challenges to monitoring of surgery include:** • Monitoring of surgery (photographs and/or videos) through external peer review to reinforce adherence to protocol • The process for monitoring of surgery should itself be standardised • Monitoring of surgical quality should consist of a graded monitoring system, adjusting stringency of monitoring according to performance and monitoring of initial cases • Monitoring of surgery should be kept anonymous and confidential • There should be a structured, regular and appropriately senior feedback mechanism with positive reward system for monitoring of surgery • Technology and artificial intelligence assistance in monitoring surgical quality **The optimal method for monitoring of surgery includes:** • Use of a structured objective assessment tool involving review of operative photographs, review of unedited videos or intra-operative monitoring by a visiting surgical team • Review of unedited operative videos • Review of random selection of recorded unedited trial videos • Review of operative photographs • Review of post-operative complications/outcomes and lymph node yield • Intra-operative monitoring by visiting surgical team • Histopathology assessment of quality of resected specimens

Phase 3.

Workshop

Following the survey questionnaire, the workshop was conducted with a purposive sample of ADDICT trial stakeholders identified by the chief investigator. The Korean Gastric Cancer Association (KINGCA) national conference (11th April 2019) was utilised to help recruit participants. 16 ADDICT trial stakeholders agreed to participate in the workshop. The purpose of the workshop was to strive to reach agreement, through a Delphi process amongst trial stakeholders, on strategies to improve monitoring of quality of surgery and patient recruitment to be considered for implementation within the ADDICT trial. All strategies within the written survey which received support by ADDICT stakeholders were included in the Delphi survey (See Appendix O), allowing us to acknowledge a maximal breadth of stakeholders'

opinion. The on-line Delphi survey was constructed utilising Qualtrics XM (NPS software) and results were collated and presented, along with anonymous participants' comments explaining their selection choices, to trial stakeholders between Delphi rounds (JB). Participants had the opportunity to change/alter their selection of mitigating strategies between Delphi rounds. A target agreement level of 70% between stakeholders was sought for each mitigating strategy.

Phase 4.

Recommendation of mitigating strategies for implementation within ADDICT trial

Following the Delphi process, consensus strategies to address the challenges to monitoring of surgery and recruitment in the ADDICT trial were presented to the trial committee and the CI to consider which may be suitable for implementation. The CI is currently discussing these strategies with the ADDICT PIs at two further trial meetings in Korea to decide on which are suitable for implementation. Once agreement has been reached on selected mitigating strategies, they will be implemented within a cohort of approximately 40 participants[1] within the ADDICT trial, following local ethics approval.

6.2.3. Analysis

The Phase 1 meeting for presentation of results of the expert stakeholder study to the ADDICT trial surgeons and trial management committee was recorded, transcribed verbatim and qualitatively analysed by researcher (JB) for emergent and consensus themes using Nvivo qualitative software. Surveys were manually coded and qualitatively analysed for emergent themes. The methodology utilised for qualitative analysis is thematic analysis in which data is searched for emergent themes.

1

* The number of participants and whether they are patients or surgeons participating in the trial will depend on the consensus strategies agreed for implementation by the trial committee and the rate of participant recruitment in the trial.

6.3. Results

6.3.1. Presentation of results of expert stakeholder study to ADDICT

Participants consisted of the ADDICT CI, 2 ADDICT trial managers and 8 surgeon PIs from different ADDICT trial centres (N=11). The male:female ratio was 9:2 and median age was 40 years (range 33 to 72 years). Over a meeting lasting 35 minutes conducted over Skype and facilitated by investigator (JB) a total of 23 emergent themes were identified. These themes were categorised into challenges to recruitment and monitoring, followed by stakeholder proposed strategies to overcome the identified challenges (See Table 18). For more details on qualitative analysis of emergent themes from this meeting, please see appendix K. Emergent themes were utilised directly to develop the written survey questionnaire. A diverse range of challenges to monitoring of surgery and recruitment of patients were identified, in addition to stakeholder proposed strategies to overcome these identified barriers. Key emergent perceived challenges to monitoring of surgery in ADDICT included: (i) Limitations of photographic monitoring, and; (ii) Difficulty in training trial teams to standardise the monitoring process. Prominent strategies to improve monitoring of surgery in ADDICT included: (i) Standardisation of monitoring process through utilisation of on-line resources, and; (ii) Sending regular reminders to ADDICT trial stakeholders regarding the standardisation process. Key challenges to recruitment in ADDICT included: (i) Changing epidemiology of gastric cancer, and; (ii) Trial surgeons overburdened with clinical work. Important emergent strategies to improve recruitment of patients in ADDICT included: (i) Expanding recruitment to other centres, and; (ii) Regular seminars and newsletters to maintain interest of trial investigators.

Table 18: Emergent themes from Phase 1 ADDICT trial management and surgeons meeting

CATEGORY	EMERGENT THEME
Challenges to monitoring	Limitations of photographic monitoring
	Training of trial teams to standardise monitoring process
	Poor trial stakeholder engagement with trial education website
Challenges in recruitment	Changing epidemiology of gastric cancer
	Trial surgeons overburdened with clinical work
	Concern patients may choose different hospital if they explain the trial
	Personal issues of some surgeons not wishing to help with others research
	Funding to support investigators and research nurses
	Lack of support for surgical trials/quality improvement
	Lack of clinician clinical equipoise for the ADDICT trial
	Reluctance to expand recruitment to Europe/US due to concerns over operative resections
Monitoring strategies	Standardisation of monitoring process
	Reminders of ADDICT trial and standardisation process for investigators
	Advantages of utilising video monitoring
	Increased surgeon collaboration and initiation of clinical trials
	Funding required for surgical trials
	Review of operative notes
	Review of case report forms
	Intra-operative monitoring using video recording and real-time data transfer to Trial Committee
Recruitment strategies	Expanding recruitment to other centres
	Adjusting centre selection according to Hazard ratio (e.g. for post-operative complications or 30-day mortality)
	Expanding to recruit centres within range of trial monitoring teams
	Regular seminars and newsletters to maintain interest of trial investigators

6.3.2. ADDICT Survey and Delphi process: Results and Discussion

This is the first reported survey and Delphi process performed within an active oeosphagogastric oncology trial with the aim of identifying and overcoming challenges to recruitment and monitoring of surgery. Eight out of eleven ADDICT trial stakeholders contacted completed the written survey, including 6 ADDICT surgeons and 2 trial managers, male:female ratio 6:2 and with a median age of 55 years (range 45-69). Table 19 provides a summary of challenges to recruitment and monitoring in ADDICT and potential mitigating strategies from open questions, and Table 20 outlines results from the Likert-scale questions. For full details of stakeholder responses according to the Likert-scale for closed questions please see appendix M. For a detailed qualitative analysis of each emergent theme from the written survey please see appendix N. Below, seminal findings the ADDICT survey and Delphi consensus are summarised and then discussed in the context of published literature. In Delphi Round 1, 16 ADDICT trial stakeholders participated in including 2 trial

managers and 14 surgeons (including 1 CI and 3 PIs) with a median age of 48 years (range 24-60 years). In Delphi Round 2, 11 trial stakeholders participated including 9 surgeons (including 1 CI and 3 PIs) and 2 trial managers with a median age of 50 years (range 24-59). In Delphi Round 1, we assessed agreement for 36 proposed strategies to overcome challenges to monitoring of surgery or recruitment and consensus (>70% agreement) was reached for 26 (72%) strategies. Following Delphi Round 2, trial stakeholder consensus agreement had been reached for a further 3 strategies making a total of 29 (81%) consensus strategies (See Table 21). 14 (74%) of the 19 included expert consensus stakeholder strategies from chapter 4 (Table 17) gained consensus (>70% agreement) amongst ADDICT trial stakeholders within 2 Delphi rounds (Table 21).

Five (31%) stakeholders who participated in Delphi round 1 did not respond in Delphi round 2. We attempted several reminder emails over several months, however a higher response could not be achieved. 7 strategies (19%) did not reach 70% agreement or more (See Table 22), and these strategies were excluded following Delphi Round 2 accordingly. 5 of the excluded strategies were amongst those which had previously gained agreement in the expert stakeholder consensus in Chapter 4 (see Table 22). A further Delphi Round 3 was not attempted given the attrition rate (n=5, 31%) between Delphi rounds 1 and 2, the slow rate of response of the Delphi Round 2 (2 months), and there was not a significant increase in agreement between rounds 1 and 2 for the strategies not reaching 70% agreement (See Table 21). Prominent emergent themes from the survey and strategies from the ADDICT Delphi process are summarised below and discussed within the context of published literature.

Table 19: Summary of challenges to recruitment and monitoring of surgery in ADDICT and potential mitigating strategies (open questions) for written survey.

CATEGORY	EMERGENT THEME	NUMBER OF STAKEHOLDERS (N=8)
Challenges to recruitment		n=2
	Reducing incidence of gastric cancer	
	Difficulty for clinicians explaining trial to patients	n=1
	Burden of routine clinical work on Korean surgeons	n=2
	Lack of funds for surgical trials	n=2
	Patient and trial investigator clinical equipoise	n=2
Challenges to monitoring		
	Difficulty in distinguishing between trial arm resections	n=3
	Insufficient monitoring members	n=2
Strategies to overcome challenges to recruitment		
	Additional human resources at trial centres	n=2
	Encouraging standardisation of surgery and monitoring	n=2
	Increasing surgeons' enthusiasm for clinical trials	n=2
	Early provision of trial contact details for investigators	n=1
Strategies to overcome challenges to monitoring		
	Regular seminars and monitoring manual to aid standardisation of surgery and monitoring	n=3

Table 20: Summary of challenges to recruitment and monitoring of surgery in ADDICT and potential mitigating strategies (from Likert scale questions) for written survey

CATEGORY	EMERGENT THEME	Stakeholder support (%)
Challenges to recruitment	**Stakeholder perception of factors contributing to reduced recruitment of patients in the ADDICT trial:**	
	• Insufficient research funds to support investigators including insufficient funds to hire more research nurses	62.5
	• Reduced interest amongst trial surgeons	75
	• Burden of routine clinical work on trial surgeons restricting their participation within trial	75
	• Reducing incidence and change in clinical stage of presentation of gastric cancer	62.5
	Factors leading to reluctance to expand recruitment of ADDICT centres internationally:	
	• Difference in surgeons' conception and skill to perform D2 gastrectomy internationally	87.5
	• Difficulty in assessing quality of surgery across centres internationally	87.5
Recruitment strategies	**Potential strategies to overcome challenges to recruitment include:**	
	• Expand number of centres/surgeons included in trial internationally within selection criteria (participating surgeons having previously performed 50 gastrectomies)	60
	• Increase the number of participating centres within the geographically accessible area for the surgical quality monitoring teams	75
	• Trial centres and surgeons increasing collaboration and working towards shared goals	87.5
	• Structured, regular and frequent communication between trial centres and trial committee with appropriately senior feedback mechanism and a positive reward system	100
	Stakeholder perception of adjusting centre selection according to the following factors to attempt to improve recruitment in the ADDICT trial:	
	• Operation complexity	62.5
	• Disease prevalence	62.5
	• Hazard ratio (e.g. for post-operative complications or 30-day mortality) between centres	62.5
	To maintain interest of trial investigators to improve recruitment the following strategies would be useful:	
	• Conduct frequent investigator seminars	62.5
	• Send regular updates on newsletters	100
Challenges to monitoring	**Stakeholder perception of the following potential challenges to monitoring of surgery in the ADDICT trial:**	
	• Quality of some videos is insufficient to assess quality of surgical procedure	75
	• Quality of some photographic series is insufficient to assess quality of surgical procedure	75
	• Trial teams are unfamiliar with method of taking and submitting standardised videos/photographs	62.5
Monitoring strategies	**Stakeholder perception of potential strategies to overcome challenges to monitoring surgery in ADDICT:**	
	• Trial teams utilising standardised method of recording operative videos/photographs as per instructions on the ADDICT trial website	100
	• Newsletters and messages to trial stakeholders reminding them of standardised methods for intra-operative monitoring	100
	• Monitoring of surgery (photographs and/or videos) through external peer review within the ADDICT trial to reinforce adherence to protocol	87.5
	• The process for monitoring of surgery within the ADDICT trial should be standardised	75
	• Monitoring of surgical quality should consist of a graded monitoring system, adjusting stringency of monitoring according to performance and monitoring of initial cases	75
	• Monitoring of surgery should be kept anonymous and confidential	75
	• There should be a structured, regular and appropriately senior feedback mechanism with positive reward system for monitoring of surgery	62.5
	• Improved system of funding investigator initiated surgical trials with adequate funding for quality assurance of surgical procedures	87.5
	Stakeholder perception regarding optimal method for monitoring surgery in ADDICT:	
	• Use of a structured objective assessment tool involving review of operative photographs, review of unedited videos or intra-operative monitoring by a visiting surgical team	100
	• Review of random selection of recorded unedited trial videos	62.5
	• Review of operative photographs	87.5
	• Histopathology assessment of quality of resected specimens	75
	• Review of post-operative complications/outcomes and lymph node yield	100
	• Review of operative notes	87.5
	• Review of case report forms	87.5

Table 21: ADDICT trial stakeholder expert Delphi consensus on proposed mitigating strategies to overcome challenges to recruitment of patients and monitoring of surgery

CATEGORY	ADDICT TRIAL STAKEHOLDER EXPERT CONSENSUS ON PROPOSED MITIGATING STRATEGIES	Round 1(%)	Round 2 (%)
RECRUITMENT	**Potential strategies to overcome challenges to recruitment in the ADDICT trial include:**		
	• Expand number of centres/surgeons included in trial internationally within selection criteria (participating surgeons having previously performed 50 gastrectomies)	72	
	• Increase the number of participating centres within the geographically accessible area for the surgical quality monitoring teams	94	
	• Trial centres and surgeons increasing collaboration and working towards shared goals*	100	
	• Providing contact details to trial stakeholders at the trial opening meeting and contact points during the trial to enable trial centres to overcome any challenges they encounter at the moment they arise.	89	
	• Making more clinical research assistants available to trial centres to help manage the process of the trial	100	
	• Organizing a clinical trial group which has strong enthusiasm and commitment (like the Japanese Cancer Oncology Group (JCOG))	63	72
	• Adjusting centre selection according to disease prevalence *	69	72
	• Standardisation should be adjusted according to Trial design (explanatory versus pragmatic)*	73	
	Recruitment in ADDICT may be improved through maintaining interest of trial investigators utilising the following strategies:		
	• Conducting frequent investigator seminars	83	
	• Sending regular trial updates on newsletters	94	
MONITORING	**Strategies to overcome challenges to monitoring of surgery in ADDICT include:**		
	• Trial teams utilising standardised method of recording operative videos/photographs as per instructions on the ADDICT trial website	89	
	• Educating trial surgical teams in photographic and video monitoring of surgery and giving them a manual for this process	89	
	• Frequent trial stakeholder seminars should be conducted regarding standardised surgery for the ADDICT trial	83	
	• Operative videos should be reviewed in trial stakeholder seminars to monitor the quality of surgery in the ADDICT trial	100	
	• Newsletters and messages to trial stakeholders reminding them of standardised methods for intra-operative monitoring	83	
	• Monitoring of surgery (photographs and/or videos) through external peer review within the ADDICT trial to reinforce adherence to protocol*	89	
	• The process for monitoring of surgery within the ADDICT trial should itself be standardised*	89	
	• Monitoring of surgical quality should consist of a graded monitoring system, adjusting stringency of monitoring according to performance and monitoring of initial cases*	100	
	• Monitoring of surgery should be kept anonymous and confidential*	72	
	• There should be a structured, regular and appropriately senior feedback mechanism with positive reward system for monitoring of surgery*	62.5	78
	• Improved system of funding investigator initiated surgical trials with adequate funding for quality assurance of surgical procedures	100	
	The optimal method for monitoring of surgery in the ADDICT trial includes:		
	• Use of a structured objective assessment tool involving review of operative photographs, review of unedited videos or intra-operative monitoring by a visiting surgical team*	78	
	• Review of unedited operative videos*	72	
	• Review of random selection of recorded unedited trial videos*	100	
	• Review of operative photographs*	72	
	• Histopathology assessment of quality of resected specimens*	82	
	• Review of post-operative complications/outcomes and lymph node yield*	83	
	• Review of operative notes	78	
	• Review of case report forms	72	

***Strategies incorporated from previous expert stakeholder Delphi consensus (Chapter 4, section 4.4)**

Table 22: Strategies not reaching expert consensus in ADDICT Delphi round 2

CATEGORY	STRATEGIES NOT REACHING EXPERT CONSENSUS IN DELPHI ROUND 2	Round 1 (%)	Round 2 (%)
RECRUITMENT	**Q1. Strategies to overcome challenges to recruitment in the ADDICT trial include:**		
	• Flexible standardisation of surgical procedures focusing on those aspects affecting safety or survival*	50	55
	Q2. Adjusting centre selection according to the following factors may improve recruitment in the ADDICT trial:		
	• Operation Complexity*	44	45
	• Hazard ratio (e.g. for post-operative complications or 30-day mortality) between centres	44	36
	Q3. Adjusting standardisation according to the following factors may improve recruitment in the ADDICT trial:		
	• Disease process studied in the trial*	64	64
MONITORING	**Q5. Potential strategies to overcome challenges to monitoring of surgery in ADDICT include:**		
	• Technology and artificial intelligence assistance in monitoring surgical quality*	50	55
	Q6. The optimal method for monitoring of surgery in the ADDICT trial includes		
	• Intra-operative monitoring by visiting surgical team*	31	18
	• Intra-operative monitoring using video recording and real-time data transfer to Trial Committee	25	27
	***Strategies incorporated from previous expert stakeholder Delphi consensus (Chapter 4, section 4.4)**		

Following this 66-item survey with 8 ADDICT stakeholders, 13 challenges to recruitment or monitoring of surgery and 29 strategies to overcome these challenges were identified (See Tables 19-20). The key stakeholder perceived barriers to recruitment of patients in ADDICT included: (i) Clinical burden on Korean surgeons, and; (ii) Reduced interest amongst trial surgeons. The majority of stakeholders were in agreement (80%) that burden of clinical work could reduce participation of surgeons within ADDICT. Some feeling that those in smaller volume hospitals lack the manpower and support infrastructure to engage in the trial: 'This is particularly true for small volume hospital because they lack support and human resource from the hospital.' Similar disincentives for clinicians partaking in research were identified within a previously published review of recruitment finding recruiters 'time constraints' were an important factor acting against their recruitment of participants.[161] One RCT of cardiac rehabilitation for patients with bowel cancer revealed salient reasons for under-recruitment included: over-estimation of the number of patient admissions; other reasons were (i) not assessing all patients for eligibility, (ii) not completing a screening form for eligible patients and (iii) patients who signed a screening form being lost to the study before consenting and randomisation.[162] Although these challenges have not specifically been mentioned by ADDICT stakeholders, it is possible that some of them may account for the mechanism of suboptimal recruitment within ADDICT. For example, the identified

'burden of clinical work' and 'reduced interest amongst trial surgeons' both of which had support from 75% of ADDICT stakeholders, may have led to stakeholders 'not assessing all patients for eligibility'. 75% of stakeholders felt that 'reduced interest amongst ADDOICT trial surgeons' was an important challenge to recruitment. One stakeholder explained that Korea surgeons were difficult to encourage to participate in trial recruitment: 'Maybe. It is general issue for Korean surgeons. Difficult to encourage.' Other stakeholders felt some patients who would meet criteria for the ADDICT trial would receive a D2 gastrectomy under the Japanese guidelines. This fact can reduce the clinical equipoise of ADDICT trial stakeholders making them reluctant to recruit these patients into the trial for fear of going against the Japanese guidelines. One stakeholder summarised his perceived conflict between the ADDICT trial protocol and the JCOG guidleines: 'T1N1 ~ T1. T2 and N0, N1 are currently recommended as D2 LND according to JCOG.' In one qualitative study, researchers' examined factors influencing how team working affected recruitment in oesophagogastric oncology trials. Several influential factors were identified including: the multidisciplinary team (MDT) meeting; leadership of the trial, and; the recruitment process. [163] Similarly, ADDICT stakeholders in our study perceived their clinical colleagues to have strong treatment preferences, which led to scepticism regarding whether the treatments were being allocated to patients in a balanced manner during doctor -patient consultations.

Recruitment of patients has historically been a challenge in surgical investigator initiated trials, [10,164] and the recruitment period has been extended and centres expanded within ADDICT in order to meet the recruitment target. Key ADDICT stakeholder consensus strategies to overcome challenges to recruitment in ADDICT included: (i) Allocating additional human resources for trial centres; (ii) Frequent investigator seminars and newsletters/updates, and; (iii) Sending regular trial updates on newsletters. Multiple stakeholders felt extra human resources would be useful to support investigators and aid recruitment: 'More clinical research assistants to help manage the process of this trial'. The shortage of and need for additional human resources to facilitate surgical trials is recognised internationally, and in the UK the National Cancer Research Network (NCRN) seek to identify centres

where allocation of clinical sessions, research nurses and other resources will enhance recruitment in surgical trials. [160] The NCRN states that they recognise UK surgeons are busy and in order optimise recruitment they seek to support them by providing adequate research infrastructure.[160] Consensus strategies proposed to maintain interest of investigators (Table 21) included frequent investigator seminars and newsletters/updates (supported by 83% and 94% of ADDICT stakeholders respectively). One stakeholder advocated: 'Seminars should be held frequently for training on standard surgery, and on the method of operative pictures taken. Video seminars could help the quality of surgery.' Although improved communication between trial coordinators and trial sites has been previously proposed, one review of methods to improve trial recruitment and retention finds evidence to be limited regarding effectiveness of this strategy.[157] Other ADDICT stakeholders recommended giving investigators the contact details of the trial coordinating centre at the opening trial meeting (See Table 19), in order to facilitate rapid resolution of problems faced by trial stakeholders: '…giving the trial details at the trial opening meeting and contact point during the trial to solve any challenges they encounter right at the moment.'

Key challenges to monitoring of surgery in ADDICT include: (i) Insufficient quality of video and/or photographic series; (ii) Trial teams being unfamiliar with methods of taking/submitting standardised videos/photographic series, and; (iii) Insufficient monitoring members. The majority (70%) of stakeholders felt that insufficient quality of either the photographic or video monitoring (See Table 20) posed a challenge to the monitoring of surgical quality. The difficulty in clearly distinguishing between the resections in the two trial arms was emphasised: 'Exact D1+ and D2 should be done for the success of this trial. Even though photos were taken after D1 or D2, it is hard to discriminate.' In part this was attributed to the quality of the photographs taken: 'Monitoring of surgery is done by the evaluation of operative picture which is uploaded to ADDICT trial website. However, the quality of picture is not satisfying.' Similar difficulties were encountered in a study assessing reliability of 2-stage oesophagectomy assessment tools within the ROMIO trial, in which 31 videos and 53 photographic series were rated by 3 surgeons. This research showed that a high

proportion of video and photographic data was absent or insufficient in order to rate performance. The operative time for esophagectomy and the potential intrusiveness of audio-visual recordings, given the restricted surgical access and limited operative field in open surgery, were challenges encountered in capturing the image data.[24] The majority of stakeholders (62.5%) felt that the trial teams were unfamiliar with methods of taking and submitting standardised videos/photographs. However, less trial stakeholders were convinced this was secondary to feasibility issues with either the photographic (50%) or video (37.5%) monitoring method. The difficulty in feasibility of monitoring of surgery in randomised controlled trials has been highlighted in a previously published narrative review of trial recruitment.[10] Other stakeholders felt there are an insufficient numbers of monitoring members in the trial committee to adequately monitor surgical quality. The considerable resources required in order to adequately monitor and record operative interventions has been previously recognised.[122]

Key consensus strategies to overcome challenges to monitoring of surgery in ADDICT included: use of an anonymous, graded monitoring system; structured, regular and frequent communication between trial centres and trial committee with appropriately senior feedback mechanism and a positive reward system, and; utilising optimal methods to monitor surgery in the ADDICT trial. Tailoring the quality assurance (QA) monitoring process according to individual cases, trial units and performance has been previously utilised within radiotherapy oncology trials but not as yet for surgical interventions. In a review of radiotherapy quality assurance in oncology trials the National Cancer Institute Work Group on Radiotherapy Quality Assurance recommends adjusting the intensity of QA to the clinical trial objectives (include general credentialing of trial units, trial-specific credentialing, and individual case review).[31] Structured, regular and frequent communication between trial centres and trial committee with appropriately senior feedback mechanism and a positive reward system gained universal support by trial stakeholders within the ADDICT Delphi consensus within the first round. Within the ADDICT survey the exact details of the 'positive reward system' to improve recruitment were not further delineated. Possible incentives were suggested in one review of trial recruitment

including: improved care for participants; altruism; career advancement; co-authorship of scientific outputs, and; the opportunity to keep up to date with current research.[161] Consensus optimal methods of monitoring quality of surgery in ADDICT included: review of random selection of recorded unedited trial videos; use of a structured objective assessment tool in monitoring; review of operative photographs; histopathology assessment of quality of resected specimens; review of post-operative complications/outcomes and lymph node yield; review of operative notes, and; review of case report forms. Post-operative complications were already a 'secondary end point and included in this trial' possibly accounting for its support. The majority (n=5, 62.5%) of written survey participants disagreed that intra-operative monitoring by a visiting surgical team would be an optimal monitoring strategy. Regarding intra-operative monitoring by a visiting surgical team, stakeholders expressed concerns regarding its complexity, practical difficulties and that litigation may be an issue. Video and photographic monitoring were similarly utilised within the randomised trial of open versus minimally invasive oesophagectomy (ROMIO) trial. In development of quality assessment tools for 2-stage oesophagectomy within the ROMIO study, the data from intra-operative videos and photographs was often either absent or of insufficient quality.[24] Similar to findings in this qualitative study of ADDICT trial stakeholders, authors planning the ROMIO trial recommended clear instructions for data capturing, adequate monitoring resources and strong engagement from participating surgeons are required to improve the monitoring process.[24] Although a structured objective assessment tool for monitoring 2-stage oesophagectomy has been previously developed,[24] no such tool has as yet been validated for D1+ or D2 gastrectomy as would be required for monitoring within the ADDICT trial.

The fact that 14 (74%) of the 19 included expert stakeholder consensus strategies from chapter 4 (Table 17) gained consensus (>70% agreement) amongst ADDICT trial stakeholders within 2 Delphi rounds (Table 21), implies this previously developed expert Delphi consensus is highly relevant within oesophagogastric oncology RCTs. A heightened awareness regarding the importance of quality assurance of surgery has arisen within Western oncology trials, as evidenced by the recent evaluation of the

STO3 trial and SQA planning in the ROMIO study.[22,165] However the STO3 analysis was restricted the evaluation of the trials SQA to R0 resection rate and lymph node yield.[166] Although these are essential markers of quality, there still appears to be a reluctance to utilise the other available SQA measures known to be associated with improved outcome. The fact that monitoring is often lacking in oncology trials is more concerning in the light of the known beneficial association of SQA measures on reduction in lymph node yield and in-hospital mortality.[1] Given this trend, we expect that future implementation of consensus monitoring strategies within a cohort of patients in ADDICT may affect short-term outcomes such as lymph node yield or 30-day mortality. Any differences in patient outcomes may however be difficult to detect as ADDICT already utilises multiple SQA measures including credentialing surgical centres, standardisation of surgical techniques, in addition to video and photographic monitoring.

Similar research methodologies were utilised to elicit barriers to monitoring and recruitment in the Neo-AEGIS RCT and the preliminary focus group results from this are outlined in appendix R, in addition to the direction of future research with the trial. Comparable challenges to recruitment were identified amongst Neo-AEGIS including: Surgeons within some trial centres not being fully engaged with the trial, and; a lack of infrastructure to support clinical trials within some trial centres. Barriers to monitoring differed significantly between the two trials given the extensive monitoring mechanisms already incorporated within the ADDICT trial. Challenges to monitoring in Neo-AEGIS included: Logistical difficulties of monitoring quality of surgery, and; there being no strict quality assurance measures for surgery within Neo-AEGIS. Similar to ADDICT trial stakeholders, participants in Neo-AEGIS felt that conducting frequent investigator seminars and sending regular updates on newsletters were important strategies to improve recruitment. There were significant parallels between ADDICT and Neo-AEGIS participants' perceived strategies to overcome challenges to monitoring in that both advocated: monitoring of surgery with photographic images or videos using use of a structured objective assessment tool.

The consensus strategies which have gained approval of ADDICT trial stakeholders via Delphi consensus may be utilised within ADDICT in endeavour to overcome identified challenges to recruitment and monitoring of surgery. Furthermore, these strategies may be transferable to aid in design of future oesophagogastric oncology trials, including selection of optimal credentialing and standardisation strategies according to disease prevalence, operation complexity and the phase of the trial. Additionally, the consensus strategies may be important in giving impetus to trial groups to select robust and reliable methods of monitoring surgical quality within future oesophagogastric oncology trials. Through improving SQA within trials via incorporating some of the consensus measures, which have reached agreement amongst international expert trial stakeholders and then received approval within an active RCT, recruitment of patients within future oesophagogastric trials may also be improved. This may occur either through enhanced enthusiasm of trial investigators allocating more time and/or effort to trial recruitment, or through expanded recruitment of trial centres.

6.3.3. Limitations

Limitations of this study include the fact that 5 (31%) of stakeholders who participated in Delphi round 1 did not participate in Delphi round 2. In addition, we would have ideally preferred a higher number of participants in the written survey. We suspect this was partly attributable to the distance involved between the UK and Korea in arranging these meetings and requesting study participation. Often face to face meetings facilitate participant recruitment and within this qualitative study, such meetings were limited to one visit to Korea during which the focus became Delphi Round 1 with the largest stakeholder sample attained (N=16). In addition, the participants' first language was Korean, and this language barrier may have influenced ADDICT trial stakeholders' willingness to participate in the study. It cannot always be presumed that trial stakeholders would feel comfortable answering questions in English. This factor may have also contributed to the relatively low number of stakeholders writing their comments in the open text

questions of the survey (See Table 19) (median number of participants documenting written responses: 2 (range 1-3) (25%)). We sought to mitigate this through sending the survey first to the ADDICT CI to review and check that the English was uncomplicated and comprehensible. Furthermore, there are cultural differences between the surgical professions in the UK and Korea which may have acted against participant recruitment. One of these may have been the steep hierarchy between senior surgeons (such as the CI of ADDICT) and the more junior surgeons within ADDICT. This may have potentially caused some hesitation in ADDICT stakeholders feeling able to openly provide negative feedback on aspects of the trial monitoring or recruitment. We strove to mitigate against some of these factors through being accompanied by local trial stakeholders in meeting new potential participants and answering their questions prior to survey participation in order to help avoid language difficulties. Furthermore, on recruiting new trial stakeholders into the project we emphasised that their responses would be confidential and all data would be pseudo anonymised. Despite these limitations surgeons and trial managers had input into each phase of the study and a broad range of opinions as to challenges and potential strategies became apparent. In addition, within the larger cohort of stakeholders participating in the Delphi process, consensus agreement (70% or over) was achieved by Round 2 on 29 (81%) of the initial 36 stakeholder proposed mitigating strategies.

6.4. Conclusion

A wide variety of challenges to recruitment of patients and monitoring of surgery were identified, and ADDICT trial stakeholder consensus agreement was reached on 29 strategies to overcome the outlined challenges. At present the ADDICT trial committee are reviewing consensus strategies to consider which would be suitable for implementation within a cohort of patients in the ADDICT trial (Phase 4). Unlike Chapter 4 which assessed expert trial stakeholders' perspective on challenges to SQA in oncology trials in general, this chapter has identified trial stakeholders' opinion regarding challenges to recruitment and monitoring within a real-life setting (ADDICT trial). The fact that 14 (74%) of the 19 included expert stakeholder consensus

strategies from chapter 4 gained consensus amongst ADDICT trial stakeholders within 2 Delphi rounds, indicates the previously developed expert Delphi consensus is highly relevant within oesophagogastric oncology RCTs. This also suggests the applicability and potential feasibility of these strategies within an active surgical oesophagogastric oncology RCT. The future direction of this study within the ADDICT trial (Phase 5), will be to assess stakeholder opinion regarding usability of implemented consensus strategies and to perform a statistical analysis of their impact on short-term outcomes (See Appendix P). One of the key expert consensus strategies which also reached agreement within the ADDICT Delphi consensus was the need for a structured objective assessment tool to review operative photographs and/or unedited videos. However, at the time of writing this chapter there was no such tool available for measuring the quality of an Ivor-Lewis oesophagectomy. This leads us to chapter 7 in which we will seek to test the reliability of a previously developed structured objective assessment tool for monitoring the quality of the process and end-product of an open or laparoscopic 2-stage Ivor-Lewis oesophagectomy, to potentially enable its use within oesophageal oncology trials.

Section 3.

Chapter 7. Reliability assessment of video and photographic oesophagectomy tools

7.1. Introduction

The need for proper application of robust quality assessment tools, in order to 'police our own practice', has been a long-standing goal for surgeons.[167] The current, gold standard, practice is to assess surgical outcomes clinically, radiologically and from the histopathology report of the excised specimen including the lymph node involvement/ yield. However, the SQA requirements of surgical oncology trials necessitated a shift in focus to confirm both the extent and completeness of lymphadenectomy, including an objective assessment of the amount of lymphatic tissue remaining around an anatomical landmark that should have been cleared.[24] This assessment can be achieved through use of a structured objective assessment tool. With the advantage over retrospective data collection monitoring methodology of permitting real-time monitoring in which errors can be identified and stratified according to severity, a structured objective assessment tool also allows feedback to operators in order to make adjustments and improve practice. [24] The importance of utilizing quality assurance measures within oesophagogastric oncology trials and the associated beneficial impact on short-term outcomes is well known.[125] In laparoscopic colorectal surgery such a competency assessment tool was successfully developed within the national training program and demonstrated to be reliable in assessment of surgeons' technical performance.[29] In both the international expert trial stakeholder Delphi consensus (Chapter 4, Section 4.4) and the ADDICT stakeholder Delphi consensus (Chapter 6, Section 6.3.2), agreement was reached for the use of a structured objective assessment tool to review operative photographs, unedited videos and/or intra-operative monitoring by a visiting surgical team. However, previously there had been no published reliable structured objective assessment tool for measuring quality of an Ivor-Lewis oesophagectomy.

Video and photographic esophagectomy assessment tools were developed previously by Harris et al (See Appendix Q), in order to permit independent observers to rate the safety and efficiency, as well as the quality of the end product (i.e. outcome) of the operation, either in real time or remotely at a later date. The development of this tool involved an iterative process including review of published literature and on-line digital media, combined with thematic analysis of the semi-structured interviews and structured observations to create a Heirarchical Task Analysis for 2-stage esophagectomy.[24] Although face and content validity had been demonstrated, the tool had not previously been tested for reliability.[28] The ROMIO trial allowed the opportunity to test this previously developed structured assessment tool within a an active RCT. We aimed to assess the reliability of the video and photographic assessment tools using data from the ROMIO trial.

Reliability refers to the consistency of a measure and can be categorised according to three types of consistency: over time (test-retest reliability), across items (internal consistency), and across different researchers (inter-rater reliability).[168] Inter-rater reliability is the extent to which two or more raters' agree and reflects the consistency of the implementation of a rating system.[169] Internal consistency is the consistency of an assessor's responses across the items on a multiple-item measure and the most common measure of internal consistency used by researchers is a statistic called Cronbach's α.[168] Classical reliability theory examines the relative contribution of the primary variable of interest, the performance of subjects, compared to error variance.[170] However within G-theory, various sources of error contributing to the inaccuracy of measurement can be explored and multiple reliability measures (eg, interrater, intertest, and intratest) can be included into the same model.[29] The essential difference between this and classical reliability theory, is that instead of simply dividing an observed score into true score and error, G-theory explicitly identifies multiple sources of error (facets).[170] G theory is a valuable tool in judging the methodological quality of an assessment method, improving its precision and leads to a more accurate description of the overall reliability of the assessment tool being tested.[170,171] To assess reliability of the video and

photographic esophagectomy assessment tools in this chapter we decided to employ the generalizability theory.

7.1.1. Objective

The specific study objective was:

v. Assessment of the reliability of an SQA tool

7.2. Methods

7.2.1. Participants

Three esophagogastric cancer surgeons (one based in the UK and two in Japan, male:female ratio 3:0, median age 43 (range 37-46)) were invited to assess and rate the intraoperative videos and photographs submitted by surgeons within the pilot ROMIO trial. In April 2017, 31 videos and 53 photographic series of individual patients' oesophagectomies performed within the ROMIO trial were available for assessment. [24]

7.2.2. Study design

The video and photographic oesophagectomy assessment tools (Appendix Q) were previously developed by Harris et al [27] through a standardisation process of two-stage esophagectomy, based on an iterative and robust methodology. This standardisation provided the structure for the development of the operation manual and video and photographic assessment tools. The video assessment and photographic assessment tools consisted of 35 items and 27 items respectively, divided according to operative anatomical dissection into 4 stages: diaphragmatic hiatus; abdominal lymphadenectomy; thoracic lymphadenectomy and the reconstruction.[24] In the video tool, safety and efficiency are rated in addition to completeness of lymphadenectomy, whereas the photographic tool focuses solely on the dissection. Prior to commencing data analysis, the three assessors were trained by the senior author (GH) in two videoconference meetings on the pre-defined terms for using the assessment tools and to clarify any conceived variability.

Prior to the second videoconference, each assessor was asked to independently rate two videos that had been chosen at random. These assessments formed a focal point for the discussion held during the second videoconference in order to minimize the discrepancy in assessments.[24] Videos and photographic series were procured from the ROMIO trial group manager electronically on an encrypted hard drive. These three independent blinded assessors then applied the video and photographic esophagectomy assessment tools to rate each of the 31 videos and 53 photographic series submitted to the ROMIO trial. Assessments were completed on paper forms, which were subsequently submitted as a scanned PDF file and later transcribed into Excel (Microsoft office, Redmond, USA). [24]

7.2.3. Analysis

A wide range of length of trial case videos were submitted from 4.5 to 449.4 minutes with a total length of 4106.9 minutes, giving an average of 132.5 minutes per oesophagectomy. Photographic oesophagectomy series ranged from 2 to 39 images per oesophagectomy, with a total of 482 images and an average of 9 photos per case.[24] A decision (D) study was performed to determine the combination of components that yielded the maximum generalizability. Inter-rater reliability and internal consistency were assessed with generalizability theory and Conbach's alpha utilizing G-String (version V. 2018) software and SPSS (IBM SPSS statistics (Ver.24, SPSS Inc., Chicago, IL) respectively. [24]

7.3. Results

7.3.1. Trial data

Despite a large volume of data being submitted, the three assessors identified that there was also a significant amount of data missing. After an interim review, a videoconference between the three reviewers explored the possible reasons for missing data and potential strategies to alleviate its impact on statistical analysis. The original 3-point lymphadenectomy assessment system of complete, incomplete, and not performed was deemed insufficient. [24] Alternative solutions that were considered included making missing assessment values a mean value or coding them as not performed. However, assessors had concerns that this would introduce bias and skew results. There was consensus that two further categories could be used to re-code those parts of the assessment in which assessors were unable to provide a rating. The new categories acknowledged insufficient evidence for assessors to provide a rating (i.e. videos with an obstructed field of view or blurred photographs) and absent data (i.e. no video or photograph submitted). Overall, 32.3% of video and photographic data were absent. 6.8% of video data were insufficient for assessors to provide a rating, compared with 23.4% of photographic series. [24]

7.3.2. Generalizability (G) theory results for the video assessment tool

Generalizability analyses were performed to evaluate reliability of the 35-item video assessment tool with a fully crossed design using videos (V), items (I) and assessors (A), such that (V x I x A). In total, 93 video assessment tool forms (comprising 31 oesophagectomy videos rated by 3 different oesophagogastric surgeon assessors) were used for analysis.[24] Raw scores of the 35-item video assessment tool were generalized over the assessor (A), and item (I). The overall reliability of the 3 assessors rating 31 videos each was represented by a generalizability coefficient of G(AI) = 0.744. Inter-rater reliability is usually around 0.7–0.8,[170] therefore the generalizability coefficient of G(AI) = 0.744 demonstrates the reliability of the oesophagectomy video assessment tool.[24]

D-studies were performed to examine the effect of increasing numbers of assessors (A) and video (V) oesophagectomies that they assessed (Figure 9). The critical G coefficient of 0.8 was reached with 4 assessors each rating 26 video oesophagectomies or 6 assessors each rating 16 video oesophagectomies.[24]

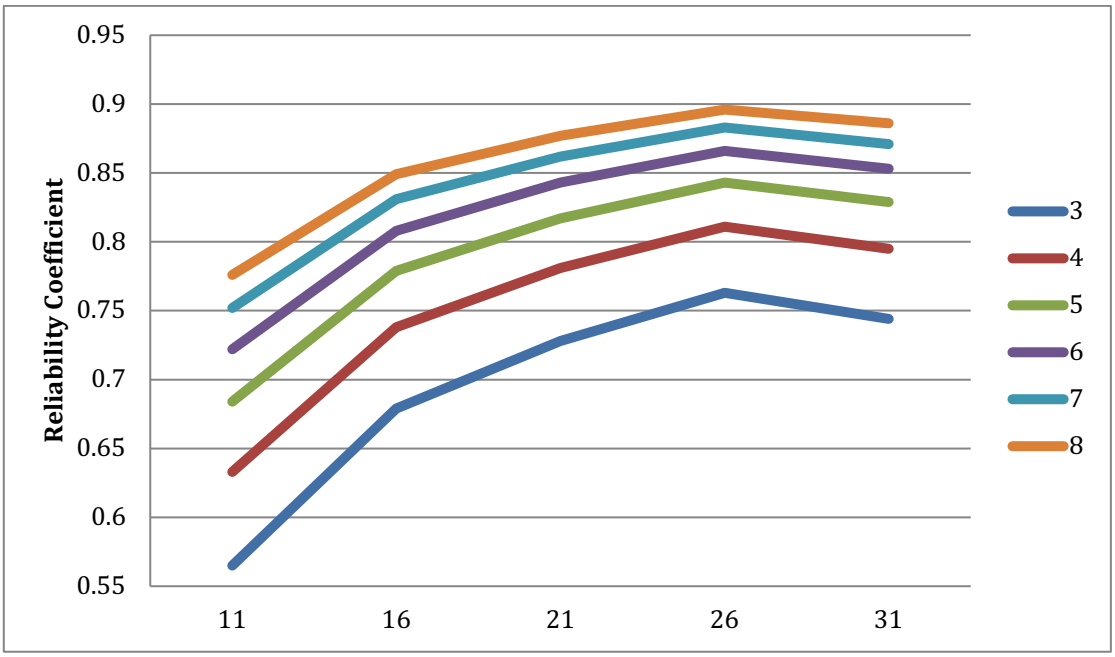

Figure 9: D-study for increasing numbers of assessors and video oesophagectomies with number of video oesophagectomies assessed along the X-axis.

7.3.3. G-theory results for the photographic assessment tool

To evaluate the reliability of the 27-item photographic assessment tool, a fully crossed design using photographs (P), items (I) and assessors (A) such that (P x I x A) was used. In total, 159 ratings (comprising 3 assessors rating 53 sets of operative photographs) of the photographic assessment tool were utilised in the analysis. Raw scores of the 27-item photographic assessment tool (P) were generalized over the assessor (A) and item (I). The overall reliability of the 3 assessors, rating 53 sets of photographs each, was represented by a generalizability coefficient of G(AI) = 0.700. As inter-rater reliability is usually around 0.7–0.8,[170] the generalizability coefficient of G(AI) = 0.700 demonstrates the reliability of the photographic oesophagectomy assessment tool.[24]

D-studies were once again performed to examine the effect of increasing numbers of assessors (A) and sets of photographs (P) of oesophagectomies that they assessed (Figure 10). The critical G coefficient of 0.8 was reached with 6 assessors each rating 38 sets of oesophagectomy photographs or 8 assessors each rating 33 sets of oesophagectomy photographs.[24]

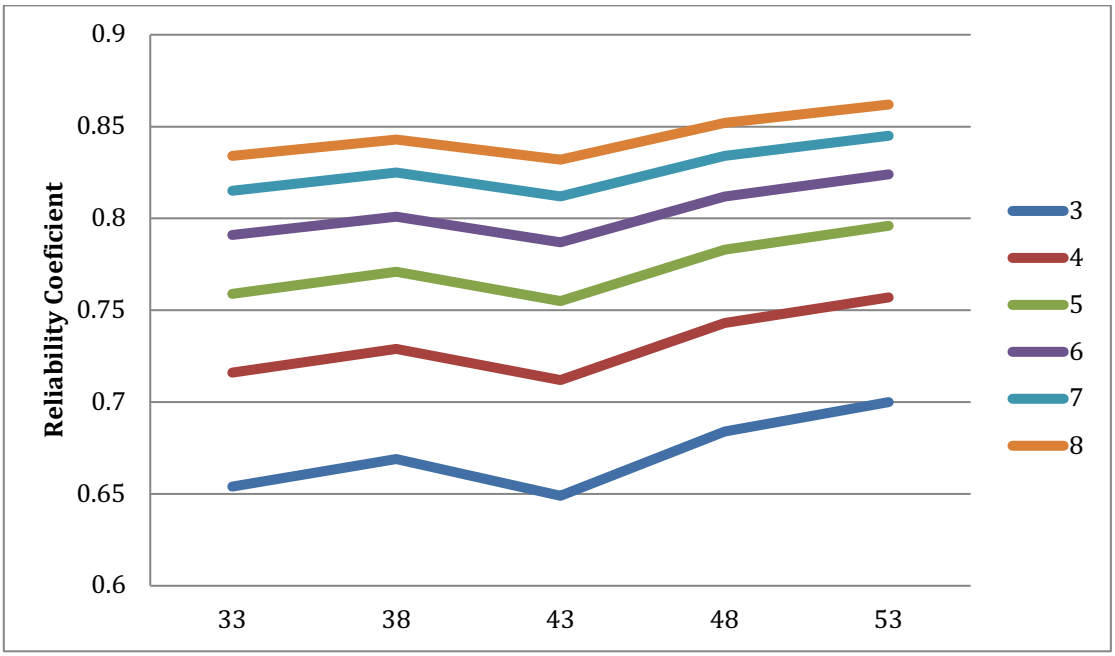

Figure 10: D study for increasing numbers of assessors and photographic oesophagectomies with number of photographic oesophagectomies assessed along the X-axis.

7.3.4. G-Coefficients by task

G-coefficients were calculated separately for each task using G-theory and G-software (See Table 23), demonstrating consistently high and similar reliability coefficients within the photographic and video assessments respectively. The video assessments had higher reliability coefficients than the photographic assessments dependably across all tasks. The highest reliability coefficient for the photographic assessment series was in the Abdomen Task 2, and the Diaphragm Task 1 for the video assessments. [24]

Table 23: G-coefficients by task [24]

G coefficient	Task 1. Diaphragm	Task 2. Abdomen	Task 3. Thorax	Task 4. Reconstruction
Photographs	0.758	0.770	0.695	0.759
Videos	0.983	0.919	0.975	0961

7.3.5. Inter-rater reliability and internal consistency

By treating one facet at a time as random, and whilst fixing the other facets, it was possible to compute the classical coefficients using G-theory. Thus, for the video assessment tool, by setting the assessor as 'random' and as 'A=1' whilst keeping the item as 'fixed', we generated the equivalent of inter-rater reliability as: Ep^2 =0.492. Then, through setting the item as random and the assessor as fixed for the video assessment tool, we generated the equivalent of internal consistency: Ep^2 =0.991, which was similar to the value calculated using SPSS with Cronbach alpha: 0.986. This indicates a moderate level of inter-rater reliability and a high level of internal consistency for the video oesophagectomy assessment tool.[24]

Through the same process, the inter-rater reliability and internal consistency for the photographic assessment tool were calculated as: Ep^2 = 0.438 and Ep^2 =0.948 respectively. Again the internal consistency was similar to Cronbach alpha calculated using SPSS: 0.942. This indicates a moderate level of inter-rater reliability and a high level of internal consistency for the photographic oesophagectomy assessment tool. [24]

7.3.6. Cronbach's alpha further analysis

Through utilizing item-total statistics in SPSS the Cronbach's alpha reliability value can be calculated if each different individual anatomical item were to be deleted.[4] On review of item-total statistics for the 31 video assessments data Cronbach's alpha remained constant at 0.986 if any anatomical item were deleted. [24]

In the 53 photographic series assessments 8 anatomical items would cause an increased value of Cronbach's alpha if deleted as shown in Table 24. [24]

Table 24: Cronbach's Alpha Item-Total Statistics for photographic series [24]

Task	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Task 1 Left Lung Assessor 1	0.067	0.943
Task 1 Aorta Assessor 2	0.222	0.943
Task 1 Pericardium Assessor 2	0.094	0.943
Task 2 Portal Vein Assessor 1	-0.02	0.943
Taks 2 Splenic Hilum Assessor 1	-0.200	0.943
Task 2 Portal Vein Assessor 2	0.010	0.943
Task 3 Carina Assessor 2	0.121	0.943
Task 4 Approximation of Sutures 1	0.226	0.943

As removal of these items would lead to a small improvement in the Cronbach's alpha and their respective 'Corrected Item-Total Correlation' values were low, this may lead one to consider removing these items from form the photographic assessment form. [24] However as a consistently high Cronbach's alpha with deletion of any anatomical item was demonstrated in the video assessment data, we suspect that this variation demonstrated with 8 items in the photographic series is secondary to the higher proportion of data that contained insufficient evidence to rate in the photographic series. [24]

7.4. Discussion

This is the first published study to assess the reliability of a structured objective assessment tool to rate the quality of the process and end-product of the 2-stage Ivor-Lewis oesophagectomy.[24] Both the photographic and video oesophagectomy tools were found to be reliable with a high reliability score using generalizability theory.

In both the previous international expert trial stakeholder Delphi consensus (Chapter 4, section 4.4) and the ADDICT stakeholder Delphi consensus (Chapter 6, section 6.3.2), agreement was reached on the importance of utilisation of a structured objective assessment tool to review operative photographs, unedited videos and/or intra-operative monitoring by a visiting surgical team. In this chapter we were able to examine the challenges and performance of a structured objective assessment tool for measuring quality of Ivor-Lewis oesophagectomy in an active RCT (ROMIO trial).

With increasing government and public interest in surgical performance some surgical specialties have made effort to assess surgical performance using proposed proxy measures such as operative time.[172] In cardiothoracic surgery various post-operative outcomes including early mortality, early re-thoracotomy for bleeding, sternal rewiring for instability, and mediastinitis were used to develop a surgical performance indicative which was shown to be associated with individual surgical quality.[173] In efforts to overcome the limitations of these retrospective attempts to assess surgical quality, authors in one review advocate Risk-Adjusted Bernoulli Cumulative Sum (RA-CUSUM), which involves prospective data collection in order to apply the cumulative sum graphical method to the surgical field and to evaluate performance changes.[174] However, the constraints of these retrospective and prospective data monitoring methods are that they inherently provide information that cannot be utilized to identify the specific area or error contributing to poor surgical quality. Thus they cannot be used to identify areas for improvement or to discriminate between environmental factors, system failure, or errors at the level of the individual operator.[24]

Structured objective assessment tools alternatively have the advantage of real-time monitoring in which errors can be identified and stratified according to severity, allowing appropriate feedback to operators in order to make adjustments and improve practice. [24] In laparoscopic colorectal surgery structured assessment tool was successfully developed within the national training program and demonstrated to be reliable in assessment of surgeons' technical performance.[29]

Capturing video or photographic data during open esophagectomy presents a challenge. This research shows that a high proportion of video and photographic data was absent or insufficient to rate performance. The operative time for oesophagectomy and the possible intrusiveness of audio-visual recordings, explains the challenge in capturing the image data. [24] Clear instructions for data capturing as well as adequate resources and strong surgeon engagement will be required to overcome these challenges in future trials. [24] An alternative approach to be explored is a short video recordings, popular with certain trial stakeholders within the expert interview study and Neo-AEGIS focus group (See Chapter 4 and Appendix R), in order to demonstrate the extent of the lymphadenectomy dissection and characteristics of the anastomotic reconstruction. The benefits of this approach would be the avoidance of long video recordings and the frequent inadequacies of photographic images, as was revealed by trial stakeholders in the ADDICT study (see chapter 6). The short recording would allow a snapshot of the operative field following lymph node dissection, permitting visualisation of anatomical structures from multiple angles (which would not be appreciated on photographic assessment from a single point of view). In addition a short video recording would provide a better assessment of conduit health and tension at the anastomosis than a photograph. [24] One limitation of such an approach would be the inability to assess the safety and efficiency of the operative tasks, as it would only show the quality of the end product. In the main ROMIO trial, the video and photographic assessment tools will also be piloted in assessment of surgical performance prior to surgeons' trial entry. [24]

This study has several limitations, including the fact that the development of the oesophagectomy assessment tool was not based on clinical outcomes, but on a hierarchical task analysis of surgical procedures and an expert consensus. This qualitative approach introduces potential bias through the subjective selection of surgical expertise. Furthermore, the validity of assessment tools to predict clinical outcomes has yet to be assessed. To assess the clinical validity an adequately large dataset will be required and this may extend beyond the scope of the ROMIO trial.[24]

7.5. Conclusion

This is the first published study to assess the reliability of a structured objective assessment tool to rate the quality of the Ivor-Lewis oesophgagectomy within an active RCT. [24] Both photographic and video oesophagectomy tools were found to be reliable with high G-coefficients and levels of internal consistency. Both tools require further validity testing in surgical oncology RCTs to assess for correlation between operative performance and clinical outcomes. [24] Within this chapter we have investigated the reliability of a structured objective assessment tool for operative assessment in an RCT, a mitigating strategy which reached consensus as a tool to overcome challenges to monitoring of surgery by both expert trial stakeholders and stakeholders within an active RCT (ADDICT). In the final chapter, we will briefly summarise the findings of this thesis, further assess strengths and limitations of each chapter, followed by an in-depth analysis of implications of this work and the future direction of research within the field.

Chapter 8. Discussion

8.1. Summary of results

Randomised controlled trials (RCTs) with surgical interventions often lack a framework to ensure surgical quality. Furthermore, expert opinion regarding challenges to surgical quality assurance (SQA) in trials and potential mitigating strategies has not been previously assessed. Although recent oncology trials, such as ADDICT comparing D1+ and D2 gastrectomy, have sought to monitor surgery there has been no demonstrably reliable tool to assess surgical quality. The aim of this thesis is to systematically investigate quality assurance of surgery in oesophagogastric oncology trials and to develop a robust and feasible framework of consensus strategies to overcome challenges to design and implementation of SQA in trials. Initially, a systematic review was conducted to review reported challenges to SQA in oncology trials involving surgery. Key challenges included: constraints of using case volume alone for credentialing surgeons; inter-centre variation in the definition and execution of interventions, and; insufficient training and deficient monitoring of surgical quality. Following this a meta-analysis of oesophagogastric RCTs with surgical interventions sought to analyse SQA measures and protocols utilised and their impact on overall survival. Public availability of oesophagogastric trial protocols and Eastern country of origin were associated with improved survival. Semi-structured interviews were conducted with trial stakeholders within the scope of a Delphi methodology aimed at examining challenges to SQA in oncology trials and then gaining expert consensus on mitigating solutions. Expert trial stakeholders' perceived challenges were examined and prominent mitigating strategies included: operative monitoring using photographs and/or videos with a structured objective assessment tool. Expert consensus was reached for 59 strategies to overcome challenges to SQA in oncology trials. Moving to explore patient opinion regarding quality of surgery, chapter 5 involved a patient focus group and survey. The patient perspective survey reinforced the importance of considering operative volume and monitoring surgery using a structured methodology to ensure surgical quality. In order to assess the feasibility of the expert consensus strategies developed in

chapter 4, a further qualitative study embedded within the ADDICT trial was conducted involving a trial stakeholder meeting, written survey and Delphi process aimed at gaining trial stakeholder consensus on mitigating strategies to improve recruitment and monitoring of surgery. 14 (74%) of the 19 included expert stakeholder consensus strategies from chapter 4 gained consensus (>70% agreement) amongst ADDICT trial stakeholders within 2 Delphi rounds, implying that the previously developed expert Delphi consensus is highly relevant within oesophagogastric oncology RCTs. Monitoring using photographs and/or videos with a structured objective assessment tool was a consensus strategy amongst expert and ADDICT trial stakeholders, however no such tool had previously been demonstrated to be reliable. Robust monitoring methods are required to achieve this and the photographic and video oesophagectomy assessment tools were demonstrated to be reliable in assessment of surgical quality using generalisability theory.

8.2. Strengths and Limitations

Chapter 2: Although published reported challenges to SQA are summarised in this review, we suspect there remain many undocumented constraints experienced by stakeholders conducting those trials. The strength of this chapter lies in its novelty being the first systematic review in this particular area and in its robust search strategy with a thorough analysis of included studies.

Chapter 3: For the first time within available literature, we were able to assess the impact of having a publicly available trial design process (protocol) on clinical outcomes. A limitation of this review is that evaluating the true impact of the SQA implementation on long-term trial outcomes was hindered due to inconsistent reporting and implementation of such measures. Stakeholders designing future trials should therefore strive to integrate SQA initiatives within their trial design and make protocols publicly available to allow greater transparency and facilitate analysis of the impact of such measures on long-term outcomes.

Chapter 4: Cultural and language barriers may have affected the interpretation of some of this data. However, by recruiting a large international sample of seventy-one expert stakeholders and utilising robust qualitative methodologies/analysis we hope to have mitigated this for and provided a relatively representative summary of international expert consensus opinion in this area. Through the international expert trial stakeholder Delphi process, consensus agreement was reached for 59 mitigating strategies, marking one of the key outputs of this thesis. A limitation of this study was that the sample of expert trial stakeholders included were predominately surgeons and oncologists, with trial managers and methodologists being under-represented. We hope that through including stakeholders from each of the four different background specialities (surgeon, oncologist, methodologist and manager) from multiple countries over Eastern and Western hemispheres, we made our sample as representative as possible of the expert oncology trial stakeholder community within the constraints of this study.

Chapter 5: Strengths of the 'patient perspective of quality of surgery' study include its novelty and ambitious nature in being the first study to endeavour to gain understanding of patient opinion within this area. This study was limited by the potential inherent recall bias of asking survivors of oesophagogastric cancer their opinion regarding quality of surgery. Additionally, it was easier to assess patient perspective regarding quality of surgery rather than quality of surgery in oncology trials. This was due to the fact that only 6 (14.6%) of the survey participants had previously taken part in a clinical trial, and we had deliberately excluded patients actively involved in oncology trials due to concerns this study may influence their willingness to continue their participation in such a trial. Additionally, for most lay people, the subject matter - quality of surgery in oncology trials - is difficult to understand or conceptualise without prior learning. Therefore, the direction of the focus group and thus the survey moved in the direction of the more accessible subject of 'quality of surgery' in general.

Chapter 6: Strengths of the ADDICT trial study include the novelty of utilising this qualitative approach to identify challenges to recruitment of patients and monitoring of surgery, and the utilisation of the Delphi process to gain expert consensus on strategies for implementation within the trial. Limitations include the potential language and cultural barriers which may possibly have influenced some responses. In addition, the time limitation of this study has meant that the implementation of consensus strategies phase and assessment of consequent stakeholder opinion of these strategies and their impact on clinical outcomes is beyond the scope of this thesis. The study within Neo-AEGIS (See Appendix R) followed a similar methodology to the ADDICT study, however it faced greater challenges in terms of time taken to receive ethical approval and to recruit sufficient stakeholders. Therefore the final results of the Neo-AEGIS qualitative study are beyond the scope of this thesis.

Chapter 7: This is the first study to demonstrate reliability of an oesophagectomy assessment tool from data in an Upper GI trial and this may prove itself valuable for future oesophageal oncology trials in facilitating the objective assessment of surgical quality. One limitation of this study is that the validity of assessment tools to predict clinical outcomes has yet to be assessed. In order to assess its clinical validity, an adequately large dataset would be required which may extend beyond the scope of the ROMIO trial.

Further limitations: Three chapters within this thesis utilised qualitative methodologies, which some may argue can be inherently prone to bias due to the fact that one is reliant on the subjective opinion of other people rather than objective data. However, given the lack of published objective scientific data within this subject area and the diverse range of opinion regarding oncology trial SQA, we propose that our qualitative approach was the optimum method to gain understanding of the challenges faced and to develop plausible strategies to overcome these barriers.

8.3. Implications of this thesis

This thesis reports on the use of a systematic and robust methodology to investigate key stakeholders' perception of challenges to SQA within trials and Delphi methodology to develop consensus mitigating strategies. Furthermore, through aligned qualitative and quantitative methods some of these consensus measures were demonstrated to be reliable, and others to have approval of patients and trial stakeholders within an active RCT (ADDICT). This thesis outlines a dependable approach for future academic trialist within surgical oncology RCTs to explore challenges and develop suitable mitigating strategies. Moreover, the consensus framework outlined below sets a blueprint for the design of SQA within future surgical oncology RCTs which may be considered by RCT trial steering groups and surgeons, as a benchmark for future trials.

8.3.1. The FOSQAT consensus

Through summarising and condensing the expert stakeholder Delphi consensus in chapter 4, we have developed a '**F**ramework of strategies to overcome challenges to implementation of **Q**uality **A**ssurance of **S**urgical Interventions in **O**ncology **T**rials' (FOSQAT consensus) (Table 25). Multiple strategies within the expert consensus framework developed in this thesis have been approved by trial stakeholders within an active RCT (ADDICT) and proved popular with the most important service stakeholder – the patient. Moreover, 14 (74%) of the 19 included expert stakeholder consensus strategies from chapter 4 (Table 17) gained consensus (>70% agreement) amongst ADDICT trial stakeholders within 2 Delphi rounds (Table 21), indicating that the previously developed expert Delphi consensus is highly relevant within oesophagogastric oncology RCTs. This also suggests the applicability and potential feasibility of these strategies. Therefore, in future studies we recommend that trial committees and surgeons will not be required to do a Delphi process each time to identify relevant strategies for implementation, but rather they can select relevant strategies from the FOSQAT consensus developed in this thesis. This framework could both aid in the design of future surgical oncology trials and be utilised to

overcome challenges to implementation of SQA within active trials. In practice we aspire for the FOSQAT consensus to be published and then included within the Enhancing the QUAlity and Transparency Of health Research (EQUATOR) database.[175] The EQUATOR database is a well-established online resource which trialists utilise in order to facilitate improving trial design and processes, currently including other relevant guidelines such as the IDEAL, CONSORT NPT and SPIRIT checklist.[175] Following the uptake and implementation of FOSQAT consensus measures within trials internationally, there will be multiple opportunities to analyse their validity through assessment of their impact on short and long-term post-operative outcomes, recruitment. Furthermore, the relevance, applicability and feasibility of FOSQAT measures can be further assessed through analysis of patient satisfaction and expert trial stakeholder opinion using relevant qualitative research methodologies. Although the FOSQAT consensus has been primarily designed based upon the opinion of oesophagogastric oncology trial stakeholders (Chapter 4), we believe the underlying principles are generically transferable to other surgical oncology trials and thus this framework may be utilised within oncology trials involving other specialities.

Table 25: Framework of Strategies to overcome challenges to implementation of Quality Assurance of Surgical interventions in Oncology Trials (FOSQAT)

	GENERIC MITIGATING STRATEGIES
1	Developing shared goals is required to facilitate the development of research relationships between trial centres
2	Developing surgeon consensus on acceptable SQA measures prior to protocol development is required for each oncology trial involving surgical interventions
3	A centralised SQA initiative with strong trial leadership is required to aid SQA
4	Adequate training and education is required to improve surgeons' understanding of research methodology
5	Trial group learning with regular meetings is important to facilitate learning regarding quality assurance from other trial groups
6	The SQA process within oncology trials should itself be standardised and monitored
7	Robotics, technology and virtual training systems may be useful for standardisation and monitoring of surgery in oncology trials
	CREDENTIALING
8	Stringency of selection of trial centres should be adjusted according to: Trial design; phase of trial; disease prevalence, and; trial operation complexity
9	A national database of operative centres is required to aid selection surgical centres for participation in oncology trials
10	Trial centre operative volume should be considered in selection of participating centres
11	Training, education and mentoring of surgeons is required to ensure surgeons have required level of trial operative skills
12	The following factors should be considered in selecting surgeons for participation in oncology trials: Operative case volume; position of surgeons on their learning curve, and; audit of histopathology of resected specimens/intraoperative bloods loss/postoperative complications and outcomes
13	Surgeons should undertake an accreditation/qualification processes prior to participation in oncology trials
14	Stringency of surgeon selection should be adjusted according to trial operation complexity and trial design
15	The optimal method of selecting surgeons for participation in oncology trials includes: Structured objective assessment tools, and; review of unedited operative videos
	STANDARDISATION
16	Trial stakeholder consensus on definition of quality of surgery is required prior to standardisation of surgery within oncology trials
17	Prospective data collection is required to provide evidence to facilitate gaining consensus on definition of quality of surgery
18	Regular trial stakeholder meetings should be conducted to aid standardisation.
19	Trial centres should be given time to adequately prepare before trials to aid standardisation of surgical procedures
20	Stringency of standardisation of surgery should be adjusted according to the following factors: Study question; disease process studied in the trial, and; trial design (explanatory versus pragmatic)
21	Standardisation of surgery should focus on operative steps affecting primary end-points, safety or survival
22	Regular feedback on quality assurance of surgery within trial centres should be provided to trial centres by the trial committee to facilitate standardisation of surgery in oncology trials
23	Inter-institutional operative team visit exchanges are useful to facilitate standardisation
24	Trial launch days in which trial surgeons and stakeholders are invited to an event to discuss the trial operative protocol should be organised to reinforce standardisation of surgery in oncology trials.
25	An operation manual should be utilised to aid standardisation of surgery in oncology trials
26	The following modalities as adjuncts to an operation manual would be useful to reinforce standardisation of surgery: Operative videos; operative photographs, and; webinars
27	Adherence to the trial operation manual should be assessed to aid standardisation within oncology trials.
28	To facilitate standardisation of surgery in oncology trials the following methods are recommended: Histopathology assessment of resected specimens; standardisation of radiology, diagnostic processes and peri-operative care by the clinical team, and; display of operation room boards detailing required steps for each specific trial operation.
	MONITORING
29	The process for monitoring of surgery within oncology trials itself should be standardised
30	Stringency and methods of monitoring of surgery in trials should be adjusted according to the following factors: Trial design and research question, and; surgical procedure involved within the trial
31	The optimal method for monitoring of surgery in oncology trials includes: Use of a structured objective assessment tool involving review of a random selection of operative photographs, unedited videos, and/ or intraoperative assessment; histopathology assessment of quality of resected specimens, and/or; review of post-operative complications/outcomes and lymph node yield.
32	Monitoring of surgical quality should be anonymous, confidential, involve external peer review and consist of a graded monitoring system, adjusting stringency of monitoring according to performance of initial cases
33	There should be a structured, regular and appropriately senior feedback mechanism with positive reward system for monitoring of surgery in oncology trials

8.3.2. Importance of trial design and protocol availability

The meta-analysis in Chapter 3 revealed trials with publicly available protocols and from an Eastern country of origin had improved overall long-term survival. This may indicate that improved planning and design within oesophagogastric oncology trials leads to higher quality oncological care for patients within the trial. Similarly, a previous systematic review demonstrated improved 5-year survival following gastrectomy in Eastern versus Western centres in which a more radical lymphadenectomy is performed. [117] We advocate those designing future oeosphagogastric oncology trials should either endeavour to publish their protocol or make it publicly available. Furthermore, we advocate improved collaboration, fellowship availability and training program establishment between Western and Eastern hemisphere surgical institutions, to enable trainees to develop and improve their oesophagogastric resection skills through learning from the leading centres in an endeavour to optimise surgical outcomes.

8.3.3. Trial stakeholder opinion within active RCTs

Through the qualitative study in chapter 6, challenges to monitoring of surgery and recruitment of patients were successfully identified within an active RCT (ADDICT). A similar methodology was utilised to elicit barriers to monitoring and recruitment in the Neo-AEGIS RCT and the preliminary focus group results from this are outlined in appendix R, in addition to the direction of future research with the trial. Key challenges to recruitment in Neo-AEGIS included: Clinical equipoise lacking amongst investigators, and; surgeons within some trial centres not being fully engaged with the trial. Key barriers to monitoring included: Logistical difficulties of monitoring quality of surgery, and; there being no strict quality assurance measures for surgery within Neo-AEGIS. Seminal strategies to overcome identified challenges to recruitment included: Trial leaders (Chief or Principle investigators) visiting other centres, and; conducting frequent investigator seminars, phone calls and steering group meetings. Prominent monitoring strategies included: monitoring of surgery

with photographic images or short videos following operative resection, and; use of a structured objective assessment tool for monitoring of surgery. In addition, trial stakeholder consensus was gained on mitigating strategies to overcome these challenges. This lends further support to the employment of qualitative studies embedded within RCTs in order to improve recruitment as has been previously shown to be beneficial in other qualitative studies. [158,159] Furthermore, the robust qualitative methodology outlined in chapter 6 provides a guide for use by future trialists involved in active surgical oncology trials to explore challenges to monitoring of surgery and recruitment allowing development of suitable mitigating strategies.

8.3.4. A reliable tool for assessment of surgical quality

Prior to the research conducted in Chapter 7 of this thesis, there was no previously published reliable tool to assess the quality of the process and/or end-product of an open or minimally invasive 2-stage oesophagectomy. Historically, when such robust methods of credentialing, standardisation and monitoring were not possible within oesophagogastric oncology trials, outcomes have been difficult to interpret due to homogenisation of the trial arm lymphadenectomies and higher than expected morbidity and mortality rates. .[20,52] Both photographic and video assessment tools were found to be reliable. [24] Widespread adoption and utilisation of this tool within oesophageal oncology RCTs would permit improved standardisation of the processes of both monitoring of the surgery within trials and credentialing of surgeons prior to recruitment. This may in turn enhance interpretation and comparison of patient outcomes both within individual and between different oesophageal oncology RCTs.

8.3.5. Patient perspective on quality of surgery

Within this thesis we also sought the opinion of the most important and often underrepresented stakeholder within oesophagogastric oncology trials – the patient. Apart from demonstrating that patient stakeholder opinion supports two of the FOSQAT strategies, this qualitative study revealed a diverse range of challenges to quality of surgery and potential mitigating solutions. This has potentially important implications for public healthcare planners and members of the surgical oncology community, who should recognise the importance these reported challenges and proposed solutions to improve quality of surgery. In doing so they will likely appreciate the imperative need for development and implementation of a nationwide, effective screening programme for Upper GI cancer in the UK, in addition to a monitoring mechanism to ensure expert agreed operative standards are being achieved within oesophagogastric cancer surgery.

8.3.6. Paradigm shift within surgical oncology community

There is a tendency within the UK to consider 'pragmatic trials' as the most sensible model upon which to base the design of surgical oncology trials with the primary focus on generalisability of results rather than surgical quality, [133] despite the previous challenges faced by trials without sufficient implementation of SQA. [20,34,52] Others advocate an 'expertise-based trial' to address the inherent problems of conventional surgical RCTs. This involves a trial in which health professionals only deliver an intervention in which they have expertise, rather than the conventional two-arm RCT design in which participants are assigned to an intervention and participating surgeons are expected to deliver both interventions.[176] However, we propose that even an expertise-based trial would face similar challenges and limitations without careful design and implementation of FOSQAT consensus items within the trial. For example - how can one ensure surgeons with adequate surgical expertise are recruited to participate and how would the trial team adequately monitor the quality of the surgery or adherence to the protocol by the 'experts' involved without a reliable assessment tool? We aspire through this thesis and its related projects to have raised awareness of the importance of design and implementation of SQA within clinical trials amongst the surgical oncology community. We also aim to encourage trial stakeholders to increase implementation of such SQA initiatives within trials internationally. Additionally, we endeavour through the patient perspective study to raise awareness within the surgical oncology and public health community of patient perceived challenges to surgical quality and their proposed strategies to overcome these barriers. Through the accumulative efforts of this thesis we hope to inspire a paradigm shift within the surgical oncology community in which the implementation of FOSQAT consensus measures are considered as complementary to all surgical oncology trials. A cultural shift in which trialists, whether involved in explanatory, expertise based or pragmatic RCTs, may be permitted to utilise items from the FOSQAT consensus in efforts to reduce variation of surgical interventions, improve surgical quality and improve comparability of trial outcomes across different studies.

8.4. Future research

Future work should include implementation of consensus stakeholder strategies within ADDICT followed by analysis of stakeholder opinion and effect on clinical outcomes (Appendix P). This will permit clinical validation of certain FOSQAT consensus measures through assessing impact of implemented strategies on short and long-term outcomes. A qualitative project with similar methodology is currently on-going within another RCT involving oesophageal cancer (Neo-AEGIS) and evidence of current progress is available in appendix R. We hope that with the uptake and implementation of FOSQAT consensus measures within trials there will be ample opportunity to analyse the impact of these upon post-operative outcomes, recruitment, patient satisfaction and expert trial stakeholder opinion. The further analysis within ADDICT and Neo-AEGIS is however beyond the scope of this thesis.

In future when SQA measures are incorporated into publicly available trial protocols and utilised within surgical oncology trials, it will be possible to formally assess the impact of adherence to protocol and SQA implementation on long-term outcomes including overall survival and disease recurrence.

Surgical education can draw on strengths of trial designs and assessment tools. A clear example of this is the national laparoscopic colorectal surgery "training the trainer" (Lapco TT) curriculum with its use of a structured objective assessment tool, which was demonstrated to improve training performance and enhance the learning curve of delegates.[30] Incorporation of the assessment tools for 2-stage oesophagectomy, demonstrated to be reliable within chapter 7, within the national training programme may contribute to similar improvements in trainee performance and could provide multiple educational and academic opportunities. The oesophagectomy assessment tool requires clinical validation within a large oesophagogastric RCT with sufficient volume to statistically assess the clinical significance of achieving higher scores versus lower scores.

Patient proposed strategies to improve quality of surgery should be considered by public healthcare planners, trialists and members of the surgical oncology community for implementation within UK oncology centres and within upper GI oncology RCTs. Following implementation of these strategies, clinical outcomes should be carefully audited, and patient experience/opinion further surveyed to assess the utility and efficacy of such initiatives.

In the long-term through the cumulative effect of the aforementioned efforts, and through integration of the FOSQAT consensus within trial protocols, design and implementation of surgical interventions, we aspire to improve the quality of surgery in oesophagogastric oncology trials.

“Quality is not an act, it is a habit”

Aristotle

“Quality is never an accident; it is always the result of high intention, sincere effort, intelligent direction and skilful execution; it represents the wise choice of many alternatives.”

William A. Foster

REFERENCES

1. Markar SR WT, Ni M, Steyerberg EW, Van Lanschot JJ B, Sasako M, Hanna GB. Assessment of the quality of surgery within randomised controlled trials for the treatment of gastro-oesophageal cancer: a systematic review. *Lancet Oncology* 2015; **16**(1): 23-31.
2. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; **340**: c869.
3. Ergina PLC, Jonathan A ; Blazeby, Jane M ; Boutron, Isabelle ; Clavien, Pierre-Alain ; Reeves, Barnaby C ; Seiler, Christoph M. Challenges in evaluating surgical innovation. *The Lancet,* 2009; **374(9695)**: 1097-104.
4. Fielding LP S-BS, Dudley HA. Surgeon-related variables and the clinical trial. *The Lancet* 1978 **7**(2(8093)): 778-9.
5. Excellence NIoC. Oesophago-gastric cancer: assessment and management in adults. 2018 https://www.nice.org.uk/guidance/ng83/chapter/Recommendations#radical-treatment (accessed 16.2.21 2021).
6. Powell L HG, Wyman A Gastrectomy for adenocarcinoma. *Continuing Education in Anaesthesia, Critical Care & Pain* 2009; **9**.
7. T D. Radical D2 Gastrectomy for Cancer. *Ann R Coll Surg Engl* 2007; **89**(7): 672–6.
8. Booka ET SK, Fukuda K, Nakamura R, Wada N, Kawakubo H, Kitagawa Y. Meta‐analysis of the impact of postoperative complications on survival after oesophagectomy for cancer. *BJS Open* 2018 **2**(5): 276-84.
9. Chen D MY, Xue Y, Sang Y, Liu D, Chen Y. Does the lymph node yield affect survival in patients with esophageal cancer receiving neoadjuvant therapy plus esophagectomy? A systematic review and updated meta-analysis. *EClinicalMedicine* 2020; **25**.
10. Cook JA. The challenges faced in the design, conduct and analysis of surgical randomised controlled trials. *Trials* 2009; **10**: 9.
11. Foster JD, Mackenzie H, Nelson H, Hanna GB, Francis NK. Methods of quality assurance in multicenter trials in laparoscopic colorectal surgery: a systematic review. *Ann Surg* 2014; **260**(2): 220-9.
12. Boutron IM, David ; Altman, Douglas G ; Schulz, Kenneth F ; Ravaud, Philippe. Extending the CONSORT Statement to Randomized Trials of Nonpharmacologic Treatment- Explanation and Elaboration. *Annals of internal medicine,* 2008; **148**(4): 295-309.
13. An-Wen Chan JMT, Douglas G. Altman, Andreas Laupacis. SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. . 2013. at:http://annals.org/aim/fullarticle/1556168/spirit-2013-statement-defining-standard-protocol-items-clinical-trials (accessed 30.01 2018).
14. Tanis E, Caballero C, Collette L, et al. The European Organization for Research and Treatment for Cancer (EORTC) strategy for quality assurance in surgical clinical research: Assessment of the past and moving towards the future. *European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology* 2016; **42**(8): 1115-22.

15. McCulloch M AD, Campbell WB, et al. 2009; 374(9695): 1105-12. No surgical innovation without evaluation- the IDEAL recommendations. . *The Lancet* 2009; 374(9695): 1105-12.; **374**: 1105-12.
16. Macdonald JS, Smalley SR, Benedetti J, et al. Chemoradiotherapy after surgery compared with surgery alone for adenocarcinoma of the stomach or gastroesophageal junction. *New England Journal of Medicine* 2001; **345**(10): 725-30.
17. Bonenkamp JJ, van de Velde CJ, Sasako M, Hermans J. R2 compared with R1 resection for gastric cancer: morbidity and mortality in a prospective, randomised trial. *The European journal of surgery = Acta chirurgica* 1992; **158**(8): 413-8.
18. Cuschieri A. Does chemoradiotherapy after intended curative surgery increase survival of gastric cancer patients? *Gut* 2002; **50**(6): 751.
19. Cuschieri AJ, V ; Fayers, P ; Cook, P ; Fielding, J ; Craven, J ; Bancewicz, J, . Postoperative morbidity and mortality after D1 and D2 resections for gastric cancer- preliminary results of the MRC randomised controlled surgical trial. *The Lancet,* 1996,; **347(9007)**: 995-9.
20. Songun I, Putter H, Kranenbarg EM, Sasako M, van de Velde CJ. Surgical treatment of gastric cancer: 15-year follow-up results of the randomised nationwide Dutch D1D2 trial. *Lancet Oncology* 2010; **11**(5): 439-49.
21. Blencowe NS BA, Harris A, Hanna T, Whiting A, Cook JA, Blazeby JM. Systematic review of intervention design and delivery in pragmatic and explanatory surgical randomized clinical trials. *Br J Surg* 2015; **102**(9): 1037-47.
22. Blencowe NS, Skilton A, Gaunt D, et al. Protocol for developing quality assurance measures to use in surgical trials: an example from the ROMIO study. *BMJ open* 2019; **9**(3): e026209.
23. Park YK, Yoon HM, Kim YW, et al. Laparoscopy-Assisted versus Open D2 Distal Gastrectomy for Advanced Gastric Cancer: Results from a Randomized Phase II Multicenter Clinical Trial (COACT 1001). *Annals of Surgery* 2017; **09**.
24. Harris A BJ, Boshier PR, Mackenzie H,Tokunaga M, Sunagawa H, Mavroveli S, Ni M, Mikhail S, Yeh C-C, Blencowe NS, Avery KNL, Hardwick R,Hoelscher A, Pera M, Zaninotto G, Law S, Low DE, Lanschot JJB, Berrisford R, Barham CP, Blazeby JM, Hanna GB. Development of a Reliable Surgical Quality Assurance System for 2-stage Esophagectomy in Randomized Controlled Trials. *Annals of surgery* 2020: 1528-140.
25. Miskovic D NM, Wyles SM, Tekkis P, Hanna GB. . Learning curve and case selection in laparoscopic colorectal surgery: systematic review and international multicenter analysis of 4852 cases. *Dis Colon Rectum* 2012; **55**(12): 1300-10.
26. Deijen CL, Velthuis S, Tsai A, et al. COLOR III: a multicentre randomised clinical trial comparing transanal TME versus laparoscopic TME for mid and low rectal cancer. *Surg Endosc* 2016; **30**(8): 3210-5.
27. Harris A MS, Hanna GB. . Oesophagectomy Photographic Assessment Tool. 2015.
28. Harris A. Mavroveli S HG. Oesophagectomy Video Assessment Tool. 2015.
29. Miskovic MD; Ni H M; Wyles K, Susannah ; Kennedy, A., Robin ; Francis, M., Nader ; Parvaiz, G., Amjad ; Cunningham, B., Chris ; Rockall, B., Timothy ; Gudgeon, B., Andrew ; Coleman, B., Mark ; Hanna, B., George. Is Competency Assessment at the Specialist Level Achievable? A Study for the National Training

Programme in Laparoscopic Colorectal Surgery in England. *Annals of Surgery* 2013; **257**(3): 476-82.
30. Mackenzie H CT, Miskovic D, Wyles SM, Langsford L, Anderson J, Thomas-Gibson S, Valori R, Hanna GB, Mark G Coleman, Francis N. Design, delivery, and validation of a trainer curriculum for the national laparoscopic colorectal training program in England. 2015; **261**(1): 149-56.
31. Bekelman JE DJ, Vikram B, Bentzen SM, Bruner D, Curran WJ, et al. Redesigning Radiotherapy Quality Assurance: Opportunities to Develop an Efficient, Evidence-Based System to Support Clinical Trials--Report of the National Cancer Institute Work Group on Radiotherapy Quality Assurance. *Int J Radiat Oncol Biol Phys* 2012; **83**(3): 782-90.
32. Dixon JLTMMW-JHS, J ; Papaconstantinou H T. Patients' Perspectives of Surgical Safety: Do They Feel Safe? *The Ochsner journal* 2015; **15**(2): 143-8.
33. Regula CM, J J ; Mauger, DT ; Marks, J G. Quality of Care From a Patient's Perspective. *Archives of Dermatology,* 2007; **143**(12): 1589-603.
34. Butterworth JW BP, Mavroveli S, Lanschot JBBV, Sasako M, Reynolds JV, Hanna GB. Challenges to quality assurance of surgical interventions in clinical oncology trials: A systematic review. *EJSO* 2020.
35. Boutron IM, David ; Altman, Douglas G ; Schulz, Kenneth F ; Ravaud, Philippe. Methods and Processes of the CONSORT Group- Example of an Extension for Trials Assessing Nonpharmacologic Treatments.. *Annals of internal medicine,* 2008; **148**(4): 60-6.
36. Remes O, Brayne C, Lafortune L. The prevalence of anxiety disorders across the life course: a systematic review of reviews. *The Lancet* 2014; **384**.
37. Krishnaratne S, Hensen B, Cordes J, Enstone J, Hargreaves JR. Interventions to strengthen the HIV prevention cascade: a systematic review of reviews. *The Lancet HIV* 2016; **3**(7): e307-e17.
38. Higgins JP1 AD, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L, Sterne JA; Cochrane Bias Methods Group; Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011; **18**(343).
39. Liberati A AD, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. . *PLoS Med* 2009; **6**(7): e1000100.
40. Balch CM SR BA. A randomized prospective clinical trial of adjuvant C. parvum immunotherapy in 260 patients with clinically localized melanoma (Stage I)- prognostic factors analysis and preliminary results of immunotherapy. *Cancer* 1982; **49**: 1079-84.
41. Balch CM DJ BA. The impact of surgical quality control in multi-institutional group trials involving adjuvant cancer treatments. *Annals of Surgery* 1983; **198**: 164-7.
42. Gérard A1 BM NB, Loygue J, Pène F, Kempf P, Bosset JF, Gignoux M, Arnaud JP, Desaive C. . Preoperative radiotherapy as adjuvant treatment in rectal cancer. Final results of a randomized study of the European Organization for Research and Treatment of Cancer (EORTC). *Annals of Surgery* 1988; **208**: 606-14.
43. Krook JE MC GL, Wieand HS, Collins RT, Beart RW, Kubista TP, Poon MA, Meyers WC, Mailliard JA. Effective surgical adjuvant therapy for high-risk rectal carcinoma. *New England Journal of Medicine* 1991; **324**: 709-15.

44. Sasako M MK KT, Bonenkamp JJ, Cornells J. H. van de Velde,Hermans J and Cooperating Investigators. . Quality control of surgical technique in a multicenter, prospective, randomized, controlled study on the surgical treatment of gastric cancer. *Jpn J Clin Oncol* 1992; **22**.
45. Bonenkamp JJ SI HJ. Randomised comparison of morbidity after D1 and D2 dissection for gastric cancer in 996 Dutch patients. *Lancet* 1995; **345**: 745-8.
46. Bunt AM HJ vdVC. Lymph node retrieval in a randomized trial on western-type versus Japanese-type surgery in gastric cancer. *Journal of Clinical Oncology* 1996; **14**: 2289-94.
47. Sasako M KT, Bonenkamp JJ, Cornells J. H. van de Velde,Hermans J and Cooperating Investigators. . Risk factors for surgical treatment in the Dutch Gastric Cancer Trial. *British Journal of Surgery* 1997; **84**: 1567-71.
48. Bonenkamp JJ1 HJ SM-. Quality control of lymph node dissection in the Dutch randomized trial of D1 and D2 lymph node dissection for gastric cancer. *Gastric Cancer* 1998; **1**: 152-59.
49. Bonenkamp JJ HJ SM. Extended lymph-node dissection for gastric cancer. *New England Journal of Medicine* 1999; **340**: 908-14.
50. Hartgrink HH, van de Velde CJ, Putter H, et al. Extended lymph node dissection for gastric cancer: who may benefit? Final results of the randomized Dutch gastric cancer group trial. *Journal of Clinical Oncology* 2004; **22**(11): 2069-77.
51. Cuschieri A FP FJ. Postoperative morbidity and mortality after D1 and D2 resections for gastric cancer- preliminary results of the MRC randomised controlled surgical trial. *Lancet* 1996; **347**: 995-99.
52. Cuschieri A WS FJ. Patient survival after D1 and D2 resections for gastric cancer- long-term results of the MRC randomized surgical trial. Surgical Co-operative Group. *British Journal of Surgery* 1999; **79**: 1522-30.
53. Gouillou PJ QP BN. MRC CLASSIC Trial Protocol. 1996 (accessed accessed 1st June 2017 2017).
54. Tveit KM GI HS, Trondsen E, Harbitz T, Nygaard K, Nilsen JB, Wist E, Hannisdal E. . Randomized controlled trial of postoperative radiotherapy and short-term time-scheduled 5-fluorouracil against surgery alone in the treatment of Dukes B and C rectal cancer. *British Journal of Surgery* 1997; **84**: 1130-5.
55. Holm T JH CB, Ekelund G, Rutqvist LE. Influence of hospital- and surgeon-related factors on outcome after treatment of rectal cancer with or without preoperative radiotherapy. . *British Journal of Surgery* 1997; **84**: 657-63.
56. Hundahl SA, Macdonald JS, Benedetti J, Fitzsimmons T. Surgical treatment variation in a prospective, randomized trial of chemoradiotherapy in gastric cancer: The effect of undertreatment. *Annals of Surgical Oncology* 2002; **9**(3): 278-86.
57. Hazebroek EJCSG. COLOR- a randomized clinical trial comparing laparoscopic and open resection for colon cancer. *Surgical Endoscopy* 2002; **16**(6): 949-53.
58. Sano T, Sasako M, Yamamoto S, et al. Gastric cancer surgery: morbidity and mortality results from a prospective randomized controlled trial comparing D2 and extended para-aortic lymphadenectomy--Japan Clinical Oncology Group study 9501. *Journal of Clinical Oncology* 2004; **22**(14): 2767-73.
59. Veldkamp R KE HW, Jeekel J, Kazemier G, Bonjer HJ, Haglind E, Påhlman L, Cuesta MA, Msika S, Morino M, Lacy AM. . Laparoscopic surgery versus open

surgery for colon cancer: short-term outcomes of a randomised trial. *The Lancet Oncology* 2005; **6**(7): 477-84.
60. Wind J HJ PB, Hollmann MW. 2006; 6: 16. Perioperative strategy in colonic surgery; LAparoscopy and:or FAst track multimodal management versus standard care (LAFA trial). *BMC Anesthesiology* 2006; (6): 16.
61. Fleshman J SD, Green E, et al. Laparoscopic colectomy for cancer is not inferior to open surgery based on 5-year data from the COST Study Group trial. *Ann Surg* 2007; **246**(4): 655-62; discussion 62-64.
62. Wright FC GA LC, Last LD, Klevan AE, Hongjinda S, Stitt LW, Klar N, Ryan DP, Smith AJ. . A randomized controlled trial to improve lymph node assessment in stage II colon cancer. *Archives of Surgery* 2008.
63. Hewett PJ, Allardyce RA, Bagshaw PF, et al. Short-term outcomes of the Australasian randomized clinical study comparing laparoscopic and conventional open surgical treatments for colon cancer: the ALCCaS trial. *Annals of Surgery* 2008; **248**(5): 728-38.
64. Allardyce RA, Bagshaw PF, Frampton CM, et al. Australian and New Zealand study comparing laparoscopic and open surgeries for colon cancer in adults: organization and conduct. *ANZ Journal of Surgery* 2008; **78**(10): 840-7.
65. Kitchener H SA QQ, Amos C, Parmar MK, ASTEC study group. . Efficacy of systematic pelvic lymphadenectomy in endometrial cancer (MRC ASTEC trial)- a randomised study. . *Lancet* 2009; **373**: 125-36.
66. Neudecker J KF, Bittner R, et al. . Short-term outcomes from a prospective randomized trial comparing laparoscopic and open surgery for colorectal cancer. *Br J Surg* 2009; **96**(12): 1458-67.
67. Simunovic M CA, Goldsmith CH, et al. . The cluster-randomized Quality Initiative in Rectal Cancer trial: evaluating a quality-improvement strategy in surgery. . *CMAJ* 2010; **182**(12): 1301-6.
68. Kennedy RH FA, Dutton S, Love S, Pearson S, Blazeby JM, Quirke P, Franks PJ, Kerr DJ. A multicentre randomised trial of conventional versus laparoscopic surgery for colorectal cancer within an enhanced recovery programme. *BMC Cancer,* 2012; **12**(1): 181.
69. van der Pas MHGM HE, Cuesta MA, et al. . Laparoscopic versus open surgery for rectal cancer (COLOR II): short-term outcomes of a randomised, phase 3 trial. *The Lancet Oncology* 2013; **14**(3): 210-8.
70. Degiuli M SM, Ponti A, et al. Randomized clinical trial comparing survival after D1 or D2 gastrectomy for gastric cancer. . Randomized clinical trial comparing survival after D1 or D2 gastrectomy for gastric cancer. . *Br J Surg* 2014; **101**(2): 23-31.
71. Primrose J FS, Finch-Jones M, et al. Systemic chemotherapy with or without cetuximab in patients with resectable colorectal liver metastasis: the New EPOC randomised controlled trial. . *The Lancet Oncology* 2014; **15**: 601-11.
72. Macdonald JS BN, Estes N, et al. ASCO-GI, 2004. January 22-24, 2004; San Francisco, CA. Post- operative combined radiation and chemotherapy improves disease-free survival (DFS) and overall survival (OS) in resected adenocarcinoma of the stomach and gastroesophageal junction: update of the results of intergroup study INT-0116 (SWOG 9008). American Society of Clinical Oncology Gastrointestinal Cancers Symposium;; 2004.
73. Cuschieri SA HG. Meta-analysis of D1 versus D2 gastrectomy for gastric adenocarcinoma: let us move on to another era. *Ann Surg* 2014; **259**(6): e90.

74. Forster AJ DG, Martin C, Papp S Improving patient safety through the systematic evaluation of patient outcomes. *Can J Surg* 2012; **55**(6): 418-25.
75. Neumayer L MM, Vanderhoof L, Hinson D. Using the Veterans Administration National Surgical Quality Improvement Program to improve patient outcomes. *J Surg Res* 2000; **88**(1): 58-61.
76. Evrard S AR, Poston G, Caballero C, Kataoka K, Fontein D, et al. . From a comic opera to surcare an open letter to whom clinical research in surgery is a concern: announcing the launch of SURCARE. . *Ann Surg* 2016; **264**(6): 911.
77. Nakajima K IM, Akagi T, Etoh T, Sugihara K, Watanabe M, et al. Quality control by photo documentation for evaluation of laparoscopic and open colectomy with D3 resection for stage II/III colorectal cancer: Japan clinical oncology group study JCOG 0404. . *Jpn J Clin Oncol* 2014; **44**(9): 799-806.
78. de Balincourt C LD, Coens C, den Dulk M, Machiels J-P, Weber D, et al. Multidisciplinary quality assurance and control in oncological trials: perspectives from European organisation for research and treatment of cancer (EORTC). *European journal of cancer (Oxford, England : 1990)* 2017; **86**: 91-100.
79. Barger S SS, Bell-Brown A, Bott B, Ciccarella A-M, Golenski J, et al. Effective stakeholder engagement: design and implementation of a clinical trial (SWOG S1415CD) to improve cancer care. *BMC Med Res Methodol* 2019; **19**: 119.
80. Snyder C SK, Tolbert E, Bantug E. Brundage M and the PRO data presentation stakeholder advisory board. . Partnering with stakeholders using an example patient-reported outcomes project *Journal of supportive and community oncology* 2017.
81. Smalley SR, Benedetti JK, Haller DG, et al. Updated analysis of SWOG-directed intergroup study 0116: a phase III trial of adjuvant radiochemotherapy versus observation after curative gastric cancer resection. *Journal of Clinical Oncology* 2012; **30**(19): 2327-33.
82. Filters. SIGNSS. Scottish Intercollegiate Guidelines Network (SIGN) Search Filters. . http://www.sign.ac.uk/search-filters.htm (accessed 07.06 2018).
83. Arends LR, Hunink MG, Stijnen T. Meta-analysis of summary survival curve data. *Stat Med* 2008; **27**(22): 4381-96.
84. Digitaliser WP. Web Plot Digitaliser 2018.
85. Bajetta E, Buzzoni R, Mariani L, et al. Adjuvant chemotherapy in gastric cancer: 5-year results of a randomised study by the Italian Trials in Medical Oncology (ITMO) Group. *Ann Oncol* 2002; **13**(2): 299-307.
86. Nashimoto A, Nakajima T, Furukawa H, et al. Randomized trial of adjuvant chemotherapy with mitomycin, Fluorouracil, and Cytosine arabinoside followed by oral Fluorouracil in serosa-negative gastric cancer: Japan Clinical Oncology Group 9206-1. *J Clin Oncol* 2003; **21**(12): 2282-7.
87. Xiao ZF, Yang ZY, Liang J, et al. Value of radiotherapy after radical surgery for esophageal carcinoma: a report of 495 patients. *Annals of Thoracic Surgery* 2003; **75**(2): 331-6.
88. Chipponi J, Huguier M, Pezet D, et al. Randomized trial of adjuvant chemotherapy after curative resection for gastric cancer. *American Journal of Surgery* 2004; **187**(3): 440-5.
89. Bouche O, Ychou M, Burtin P, et al. Adjuvant chemotherapy with 5-fluorouracil and cisplatin compared with surgery alone for gastric cancer: 7-year results of the FFCD randomized phase III trial (8801). *Ann Oncol* 2005; **16**(9): 1488-97.

90. Burmeister BH, Smithers BM, Gebski V, et al. Surgery alone versus chemoradiotherapy followed by surgery for resectable cancer of the oesophagus: a randomised controlled phase III trial. *The Lancet Oncology* 2005; **6**(9): 659-68.
91. Yu W, Choi GS, Chung HY. Randomized clinical trial of splenectomy versus splenic preservation in patients with proximal gastric cancer. *British Journal of Surgery* 2006; **93**(5): 559-63.
92. Cunningham D, Allum WH, Stenning SP, et al. Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. *New England Journal of Medicine* 2006; **355**(1): 11-20.
93. Wu CW, Hsiung CA, Lo SS, et al. Nodal dissection for patients with gastric cancer: a randomised controlled trial. *Lancet Oncology* 2006; **7**(4): 309-15.
94. De Vita F, Giuliani F, Orditura M, et al. Adjuvant chemotherapy with epirubicin, leucovorin, 5-fluorouracil and etoposide regimen in resected gastric cancer patients: A randomized phase III trial by the Gruppo Oncologico Italia Meridionale (GOIM 9602 Study). *Annals of Oncology* 2007; **18**(8): 1354-8.
95. Kelsen DP, Winter KA, Gunderson LL, et al. Long-term results of RTOG trial 8911 (USA Intergroup 113): a random assignment trial comparison of chemotherapy followed by surgery compared with surgery alone for esophageal cancer. *Journal of Clinical Oncology* 2007; **25**(24): 3719-25.
96. Omloo JM, Lagarde SM, Hulscher JB, et al. Extended transthoracic resection compared with limited transhiatal resection for adenocarcinoma of the mid/distal esophagus: five-year survival of a randomized clinical trial. *Annals of Surgery* 2007; **246**(6): 992-1000; discussion -1.
97. Di Costanzo F, Gasperoni S, Manzione L, et al. Adjuvant chemotherapy in completely resected gastric cancer: a randomized phase III trial conducted by GOIRC. *Journal of the National Cancer Institute* 2008; **100**(6): 388-98.
98. Sasako MS, Takeshi ; Yamamoto, Seiichiro ; Kurokawa, Yukinori ; Nashimoto, Atsushi ; Kurita, Akira ; Hiratsuka, Masahiro ; Tsujinaka, Toshimasa ; Kinoshita, Taira ; Arai, Kuniyoshi ; Yamamura, Yoshitaka ; Okajima, Kunio. D2 lymphadenectomy alone or with para-aortic nodal dissection for gastric cancer. *New England Journal of Medicine* 2008; **359**(5): 453-62
99. Allum WH, Stenning SP, Bancewicz J, Clark PI, Langley RE. Long-term results of a randomized trial of surgery with or without preoperative chemotherapy in esophageal cancer. *Journal of Clinical Oncology* 2009; **27**(30): 5062-7.
100. Kulig J, Kolodziejczyk P, Sierzega M, et al. Adjuvant chemotherapy with etoposide, adriamycin and cisplatin compared with surgery alone in the treatment of gastric cancer: a phase III randomized, multicenter, clinical trial. *Oncology* 2010; **78**(1): 54-61.
101. Songun I, Putter H, Kranenbarg EM, Sasako M, van de Velde CJ. Surgical treatment of gastric cancer: 15-year follow-up results of the randomised nationwide Dutch D1D2 trial. *Lancet Oncol* 2010; **11**(5): 439-49.
102. Miyashiro I, Furukawa H, Sasako M, et al. Randomized clinical trial of adjuvant chemotherapy with intraperitoneal and intravenous cisplatin followed by oral fluorouracil (UFT) in serosa-positive gastric cancer versus curative resection alone: final results of the Japan Clinical Oncology Group trial JCOG9206-2. *Gastric Cancer* 2011; **14**(3): 212-8.
103. Sasako M, Sakuramoto S, Katai H, et al. Five-year outcomes of a randomized phase III trial comparing adjuvant chemotherapy with S-1 versus

surgery alone in stage II or III gastric cancer. *Journal of Clinical Oncology* 2011; **29**(33): 4387-93.
104. Van Hagen PH, M. C. C. M ; Van Lanschot, J. J. B ; Steyerberg, E. W ; Henegouwen, M. I. Van Berge ; Wijnhoven, B. P. L ; Richel, D. J ; Nieuwenhuijzen, G. A. P ; Hospers, G. A. P ; Bonenkamp, J. J ; Cuesta, M. A ; Blaisse, R. J. B ; Busch, O. R. C ; Ten Kate, F. J. W ; Creemers, G. - J ; Punt, C. J. A ; Plukker, J. T. M ; Verheul, H. M. W ; Bilgen, E. J. Spillenaar ; Van Dekken, H ; Van Der Sangen, M. J. C ; Rozema, T ; Biermann, K ; Beukema, J. C ; Piet, A. H. M ; Van Rij, C. M ; Reinders, J. G ; Tilanus, H. W ; Van Der Gaast, A. Preoperative Chemoradiotherapy for Esophageal or Junctional Cancer. *New England Journal of Medicine* 2012; **366**(22): pp.2074-84.
105. Bass GA, Furlong H, O'Sullivan KE, Hennessy TP, Walsh TN. Chemoradiotherapy, with adjuvant surgery for local control, confers a durable survival advantage in adenocarcinoma and squamous cell carcinoma of the oesophagus. *European journal of cancer (Oxford, England : 1990)* 2014; **50**(6): 1065-75.
106. Noh SH, Park SR, Yang HK, et al. Adjuvant capecitabine plus oxaliplatin for gastric cancer after D2 gastrectomy (CLASSIC): 5-year follow-up of an open-label, randomised phase 3 trial. *Lancet Oncology* 2014; **15**(12): 1389-96.
107. Degiuli M, Sasako M, Ponti A, et al. Randomized clinical trial comparing survival after D1 or D2 gastrectomy for gastric cancer. *British Journal of Surgery* 2014; **101**(2): 23-31.
108. Mariette C, Dahan L, Mornex F, et al. Surgery alone versus chemoradiotherapy followed by surgery for stage I and II esophageal cancer: final analysis of randomized controlled phase III trial FFCD 9901. *Journal of Clinical Oncology* 2014; **32**(23): 2416-22.
109. Hirao M, Kurokawa Y, Fujita J, et al. Long-term outcomes after prophylactic bursectomy in patients with resectable gastric cancer: Final analysis of a multicenter randomized controlled trial. *Surgery* 2015; **157**(6): 1099-105.
110. Yang ZQ, Lu HX, Zhang JH, Wang J. Comparative study on long-term survival results between minimally invasive surgery and traditional resection for esophageal squamous cell carcinoma. *European Review for Medical & Pharmacological Sciences* 2016; **20**(16): 3368-72.
111. Sano T, Sasako M, Mizusawa J, et al. Randomized Controlled Trial to Evaluate Splenectomy in Total Gastrectomy for Proximal Gastric Carcinoma. *Annals of Surgery* 2017; **265**(2): 277-83.
112. van Heijl M, van Lanschot JJ, Koppert LB, et al. Neoadjuvant chemoradiation followed by surgery versus surgery alone for patients with adenocarcinoma or squamous cell carcinoma of the esophagus (CROSS). *BMC Surgery* 2008; **8**: 21.
113. Van Lanschot JJB OH, Tilanus HW, de Graaf PW, Lerut T. Tranhiatal oesophageal resection without thoracotomy verus transthoracic oesophageal resection with two-fields lymphadenectomy: a randomised study with adenocarcinoma of the oesophagus and/or gastro-oesophageal junction. In: Academic Medical Centre A, editor. Protocol ed; 1993.
114. van Hagen P, Hulshof MC, van Lanschot JJ, et al. Preoperative chemoradiotherapy for esophageal or junctional cancer. *New England Journal of Medicine* 2012; **366**(22): 2074-84.

115. Sasako M, Sano T, Yamamoto S, et al. D2 lymphadenectomy alone or with para-aortic nodal dissection for gastric cancer. *New England Journal of Medicine* 2008; **359**(5): 453-62.
116. Jamieson GG MG, Ludemann R, Wayman J, Myers JC, Devitt PG. . Postoperative mortality following oesophagectomy and problems in reporting its rate. *Br J Surg* 2004; **91**(8): 943-71.
117. Markar SR KA, Jackson D, Hanna GB. Long-Term Survival After Gastrectomy for Cancer in Randomized, Controlled Oncological Trials: Comparison between West and East. *Annals of Surgical Oncology* 2013; **20**(7): 2328-38.
118. Noguchi Y, Yoshikawa T, Tsuburaya A, Motohashi H, Karpeh MS, Brennan MF. Is gastric carcinoma different between Japan and the United States? *Cancer* 2000; **89**(11): 2237-46.
119. Bollschweiler E BK, Hoelscher AH, et al. . Is the prognosis for Japanese and German patients with gastric cancer really different? *Cancer* 1993; **71**(10): 2918-25.
120. Hanna GB BP, Knaggs A GR, Sasako M, . Improving Outcomes After Gastroesophageal Cancer Resection Can Japanese Results Be Reproduced in Western Centers? *Archives of Surgery* 2012; **147**(8): 738-45.
121. Hirst A, Philippou Y, Blazeby J, et al. No Surgical Innovation Without Evaluation: Evolution and Further Development of the IDEAL Framework and Recommendations. *Ann Surg* 2018.
122. McCulloch P TI, Sasako M, Lovett B, Griffin D. . Randomised trials in surgery: problems and possible solutions. *BMJ* 2002; **324**: 1448-51.
123. Unger CA BM. Studying surgical innovations: challenges of the randomized controlled trial. *J Minim Invasive Gynecol* 2015; **22**: 573-82.
124. Huddy JR NM, Barlow J et al. . Point-of-care C reactive protein for the diagnosis of lower respiratory tract infection in NHS primary care: a qualitative study of barriers and facilitators to adoption. *BMJ Open* 2016.
125. Markar SR, Wiggins T, Ni M, et al. Assessment of the quality of surgery within randomised controlled trials for the treatment of gastro-oesophageal cancer: a systematic review. *The Lancet Oncology* 2015; **16**(1): e23-e31.
126. Fischer C LH, Klazinga N, Hardwick R, Cromwell D, Steyerberg E, Groene O. Volume-outcome revisited: The effect of hospital and surgeon volumes on multiple outcome measures in oesophago-gastric cancer surgery. *PLoS One* 2017; **12**(10).
127. Keller DS HR. The learning curve in TaTME—considerations to guide safe implementation. *Annals of Laparoscopic and Endoscopic Surgery* 2019; **4**.
128. Blencowe NS MN, Cook AJ, Donovan JL, Rogers CA, Whiting P, et al. Standardizing and monitoring the delivery of surgical interventions in randomized clinical trials. *Br J Surg* 2016; **103**(10): 1377-84.
129. Kim J BW, Eom BW, Nam BH, Yoon HM, Kim YW. ADDICT – A multicenter randomized clinical trial of D1+ versus D2 distal gastrectomy for stage IB & II advanced gastric cancer: Design and rationale. *European Journal of Surgical Oncology* 2014; **40**(11).
130. Ibbott GS HA, Followill DS. Quality Assurance for Clinical Trials. *Front Oncol* 2013; **3**(31).

131. Williams L AK, Plahe G. Restricted capital spending in the English NHS: a qualitative enquiry and analysis of implications. In: Birmingham Uo, editor.; 2018.
132. Ke KM HW, Ness AR. The costs of centralisation: a systematic review of the economic impact of the centralisation of cancer services. *Eur J Cancer Care (Engl)* 2012; **21**(2): 158-68.
133. DemangeI MK FF. Limits to clinical trials in surgical areas. *Clinics (Sao Paulo)* 2011; **66**(1): 159–61.
134. Sideris M HJ, Staikoglou N, Pantelidis P, et al. Optimizing engagement of undergraduate students in medical education research: The eMERG training network. *Ann Med Surg (Lond)* 2018; **26**(31): 6-10.
135. Khan M. Awareness of Post Graduate Residents regarding Medical Research: A Tertiary Care Hospital Experience. *Review of Global Medicine and Healthcare Research* 2012; **3**(41).
136. Chu LF UA, Kadry B, Kucharski SE, Campos H, Crockett J, Dawson N, Clauson KA. "Nothing about us without us"-patient partnership in medical conferences. *BMJ* 2016; **14**(354): 3883.
137. England N. Involving people in their own care. 2017. https://www.england.nhs.uk/ourwork/patient-participation/ (accessed 10/08/2019 2019).
138. Wise J. Disruptive technologies making cancer care more patient centred. *BMJ* 2017; **359**: 1756-833.
139. Kudsi OY CA, Kaza S, McCarty J, Dickens E, Martin D, Tiesenga FM, Konstantinidis K, Hirides P, Mehendale S, Gonzalez A. Cosmesis, patient satisfaction, and quality of life after da Vinci Single-Site cholecystectomy and multiport laparoscopic cholecystectomy: short-term results from a prospective, multicenter, randomized, controlled trial. . *Surg Endosc* 2017; **81**(8): 3242-50.
140. Blazeby JM AD, Farndon JR. . Quality of life in patients with oesophageal cancer. . *Recent Results Cancer Res* 2000; **155**: 193-204.
141. Kao J, Genden EM, Gupta V, et al. Phase 2 trial of concurrent 5-fluorouracil, hydroxyurea, cetuximab, and hyperfractionated intensity-modulated radiation therapy for locally advanced head and neck cancer. *Cancer* 2011; **117**(2): 318-26.
142. Schlessinger DI IS, Yanes AF, Lazaroff JM, Godinez-Puig V, Chen BR, Kurta AO, Henley JK, Chiren SG, Furlan KC, Schmitt J, Deckert S, Poon E, Sobanko JF, Cartee TV, Alam M, Maher IA. Development of a core outcome set for clinical trials in basal cell carcinoma: study protocol for a systematic review of the literature and identification of a core outcome set using a Delphi survey. *Trials* 2017; **18**(1): 490.
143. Alkhaffaf B GA, Blazeby JM, Williamson P, Bruce IA. Standardising the reporting of outcomes in gastric cancer surgery trials: protocol for the development of a core outcome set and accompanying outcome measurement instrument set (the GASTROS study). *Trials* 2017 **18**(1): 370.
144. Bailo L GP, Vergani L, Marton G, Pravettoni G. The patient perspective: investigating patient empowerment enablers and barriers within the oncological care process. *Ecancermedicalscience* 2019; (13): 912.
145. Woods P GR, King M A. Generating or developing grounded theory: methods to understand health and illness. *Int J Clin Pharm* 2016; **38**(3): 663-70.

146. DM E. Have We Created a Crisis in Confidence for General Surgery Residents? A Systematic Review and Qualitative Discourse Analysis. *JAMA Surg* 2016; **151**(12): 1166–75.
147. Choi KS JJ, Suh M, Park B, Noh DK, Song SH, Jung JW, et al. Effect of endoscopy screening on stage at gastric cancer diagnosis: results of the National Cancer Screening Programme in Korea. *Br J Cancer* 2015; **112**(3): 608–12.
148. Shalowitz DI MF. Communicating the results of clinical research to participants: attitudes, practices, and future directions. *PLoS Med* 2008; **5**(5): e91.
149. Gikas A TJ. The role of primary care physicians in early diagnosis and treatment of chronic gastrointestinal diseases. *Int J Gen Med* 2014; **7**: 159–73.
150. Chioreso C VN, Schweizer ML, Schlichting J, Gribovskaja-Rupp I, Charlton ME. The association between hospital and surgeon volume and rectal cancer surgery outcomes in rectal cancer patients treated since 2000: systematic literature review and meta-analysis. *Dis Colon Rectum* 2018; **61**(11): 1320–32.
151. Morche J MT, Pieper D. Relationship between surgeon volume and outcomes: a systematic review of systematic reviews. *Syst Rev* 2016; **5**: 204.
152. J. W. Mobile lung cancer testing in supermarket car parks is to be expanded. *BMJ* 2017; **359**: 5450.
153. Markar SR LJ, Hanna GB Research protocol for a diagnostic study of non-invasive exhaled breath analysis for the prediction of oesophago-gastric cancer. *BMJ open* 2016; **6**: e009139.
154. Melnychuk M V-PC, Aitchison M, Clarke CS, Fulop NJ, Levermore C, Maddineni SB, Moore CM, Mughal MM, Perry C, Pritchard-Jones K, Ramsay AIG, Shackley D, Vickers J, Morris S. Centralising specialist cancer surgery services in England: survey of factors that matter to patients and carers and health professionals. *BMC Cancer* 2018; **18**(1): 226.
155. Park YK YH, Kim YW, Park JY, Ryu KW, Lee YJ, Jeong O, Yoon KY, Lee JH, Lee SE, Yu W, Jeong SH, Kim T, Kim S, Nam BH. COACT group. Laparoscopy-Assisted versus Open D2 Distal Gastrectomy for Advanced Gastric Cancer: Results from a Randomized Phase II Multicenter Clinical Trial (COACT 1001). . *Ann Surg* 2017.
156. Kim Y-W. A Clinical Trial of D1+ Versus D2 Distal Gastrectomy for Stage IB & II Advanced Gastric Cancer (ADDICT). 2016. https://clinicaltrials.gov/ct2/show/NCT02144727.
157. Bower P BV, Gamble C, Treweek S, Smith TC, Young B, Williamson P. Interventions to improve recruitment and retention in clinical trials: a survey and workshop to assess current practice and future priorities. *Trials* 2014; **15**(399).
158. Fletcher B GA, Moore D, Wilson S, Damery S. . Improving the recruitment activity of clinicians in randomised controlled trials: a systematic review. *BMJ open* 2012; **2**(1): e000496.
159. Hennessy M HA, Healy P, Galvin S, Houghton C. Improving trial recruitment processes: how qualitative methodologies can be used to address the top 10 research priorities identified within the PRioRiTy study. *Trials* 2018; **19**(584).
160. Kaur G HI, Mehanna H, Williamson P, Shaw R, Smith CT. Barriers to recruitment for surgical trials in head and neck oncology: a survey of trial investigators. *BMJ open* 2013; **11**(3).

161. Parkinson B MR, Sutton M, Fichera E, Mills N, Shorter GW, Treweek S, Harman NL, Brown RCH, Gillies K, Bower P. Designing and using incentives to support recruitment and retention in clinical trials: a scoping review and a checklist for design. *Trials* 2019; **20**(624).
162. Hubbard G CA, Davies Z, Munro J, Ireland AV, Leslie S, Watson AJ, Treweek S. Experiences of recruiting to a pilot trial of Cardiac Rehabilitation In patients with Bowel cancer (CRIB) with an embedded process evaluation: lessons learned to improve recruitment. *Pilot Feasibility Stud* 2015; **14**(1): 15.
163. Strong S, Paramasivan S, Mills N, Wilson C, Donovan JL, Blazeby JM. 'The trial is owned by the team, not by an individual': a qualitative study exploring the role of teamwork in recruitment to randomised controlled trials in surgical oncology. *Trials [Electronic Resource]* 2016; **17**(1): 212.
164. Avery KN WP, Gamble C, O'Connell Francischetto E, Metcalfe C, Davidson P, Williams H, Blazeby JM; members of the Internal Pilot Trials Workshop supported by the Hubs for Trials Methodology Research. Informing efficient randomised controlled trials: exploration of challenges in developing progression criteria for internal pilot studies. *BMJ Open* 2017; **7**(2).
165. Metcalfe C, Avery K, Berrisford R, et al. Comparing open and minimally invasive surgical procedures for oesophagectomy in the treatment of cancer: The romio (randomised oesophagectomy: Minimally invasive or open) feasibility study and pilot trial. *Health Technology Assessment* 2016; **20**(48): 1-100.
166. Allum WH SE, Blazeby JM, Grabsch HI, Griffin SM, Rowley S, Cafferty FH, Langley RE, Cunningham D. Quality assurance of surgery in the randomized ST03 trial of perioperative chemotherapy in carcinoma of the stomach and gastro-oesophageal junction. *Br J Surg* 2019; **106**(9): 1204-15.
167. Dindo D CP. Quality assessment in surgery: mission impossible? *Patient Saf Surg* 2010; **4**(1).
168. Price PC JR, I-Chant A, Chiang, DC, Leighton, Cuttler C Reliability and Validity of Measurement 2009. https://opentext.wsu.edu/carriecuttler/chapter/reliability-and-validity-of-measurement/.
169. RT L. Inter-rater reliability 2011. https://link.springer.com/referenceworkentry/10.1007%2F978-0-387-79948-3_1203.
170. Bloch R NG. Generalizability theory for the perplexed: A practical introduction and guide. *Medical Teacher* 2012; **34**(11): 960-92.
171. Streiner N NG. Health Measurement Scales. A Practical Guide to Their Development and Use. 4th Edition. Oxford University Press; 2008.
172. Maruthappu M DA, Orgill D, Carty MJ. A monitoring tool for performance improvement in plastic surgery at the individual level. *Plast Reconstr Surg* 2013; **131**(5): 702-10.
173. Hartrumpf M CT, Erb M, Albes JM. Surgeon performance index: tool for assessment of individual surgical quality in total quality management *European Journal of Cardio-Thoracic Surgery* 2009; **35**(5): 751-8.
174. Agostino PA LFC, Francesca D S, Alessandra S, Marco B and Fabio V. Monitoring Surgical Performance: Current Models and Limitations. *SOJ Surgery* 2017; **4**(1): 1-3.
175. Centre UE. Enhancing the QUAlity and Transparency Of health Research. 2021. https://www.equator-network.org/ (accessed 19.2.21 2021).

176. Cook JA EA, Boachie C, Bassinga T, Fraser C, Altman DG, Boutron I, Ramsay CR, MacLennan GS. A systematic review of the use of an expertise-based randomised controlled trial design. *Trials* 2015; **16**: 241.
177. Junor EJ HD, Gillis CR. Management of ovarian cancer: referral to a multidisciplinary team matters. *Br J Cancer* 1994; **70**: 363-70.
178. Forrest LM MD, McArdle CS,. 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* 2005; **93**: 977–8.
179. Evans L DB, Liu Y, Shaw T, Harnett P. A tool to improve the performance of multidisciplinary teams in cancer care. *BMJ Open Qual* 2019; **8**(2).
180. Anderson BO. What Outcomes Matter to Patients: A Surgeon's Perspective. *MEDICAL CARE* 2002; **40**(6).
181. Sheetz KH WS, Girotti ME, Campbell DA Jr, Englesbe MJ. Patients' perspectives of care and surgical outcomes in Michigan: an analysis using the CAHPS hospital survey. *Ann Surg* 2014; **260**(1): 5-9.
182. Macleod U MED, Burgess C, Macdonald S, Ramirez AJ. Risk factors for delayed presentation and referral of symptomatic cancer: evidence for common cancers. *Br J Cancer* 2009; **101**(2): S92–S101.
183. Committee on Improving the Quality of Cancer Care: Addressing the Challenges of an Aging Population; Board on Health Care Services; Institute of Medicine; Levit L BE, Nass S, et al., editors. Delivering High-Quality Cancer Care: Charting a New Course for a System in Crisis.
Show details. *National Academies Press (US);* 2013.
184. Niedzwiedz CL KL, Robb KA, Katikireddi SV, Smith DJ. Depression and anxiety among people living with and beyond cancer: a growing clinical and research priority. *BMC Cancer* 2019; **19**: 943.
185. Fregene TW, S ; Venkat R, Vi; Edmond, H ; Rizvi, S. Making the experience of elective surgery better. *BMJ Open Quality* 2017; **6**(2).
186. Ke Chen YP, Shu-ting Zhai, Jia-qin Cai, Qi-long Chen, Ding-wei Chen, Yi-ping Zhu, Yu Zhang, Ya-ping Zhang, Hendi Maher, Xian-fa Wang. Laparoscopic gastrectomy in obese gastric cancer patients: a comparative study with non-obese patients and evaluation of difference in laparoscopic methods. *BMC Gastroenterol* 2017; **17**(78).
187. JG M. The Problem Surgical Colleague. *Ann R Coll Surg Engl* 2007 **89**(5): 457–63.
188. McCall MD GP, Bathe OF. Quality of life: A critical outcome for all surgical treatments of gastric cancer. *World J Gastroenterol* 2016; **22**(3): 1101–13.
189. Cook AJ. Statistical evaluation of learning effects in surgical trials. *Clinical Trials* 2004; **1**(5): 421-7.
190. Chou E AH, Khan MS, Dasgupta P, Ahmed K. Should surgical outcomes be published? *J R Soc Med* 2015; **108**(4): 127–35.
191. Jones IR GJ, Exworthy M. . Will publishingsurgeon performance improve outcomes? 2013. http://www.hsj.co.uk/comment/will-publishing-surgeon-performance-improve-outcomes/5060344.article#.VQH9BiiGtbw (accessed 13.10.2020 2020).
192. Richardson K LD, Jack S, Grocott MPW. Fit for surgery? Perspectives on preoperative exercise testing and training. *Review Br J Anaesth* 2017; **119**: i34-i43.

193. D'Angelica M HK, Ross H, Passik S, Brennan MF. Surgeon-patient communication in the treatment of pancreatic cancer. *Arch Surg* 1998; **133**(9): 962-6.
194. Fund TK. Closing the gap. Key areas for action on the health and care workforce. 2019.
195. England RCoSo. NHS patients will benefit from £20 million investment into surgical research. 2014. https://www.rcseng.ac.uk/news-and-events/media-centre/press-releases/nhs-patients-will-benefit-from-investment-into-surgical-research/ (accessed 13.10.2020 2020).
196. Graham DY. History of Helicobacter pylori, duodenal ulcer, gastric ulcer and gastric cancer. *World Journal of Gastroenterology* 2014; **14**(20): 5191–204.
197. LH A. The coeliac axis in the expansion of the operation for gastric carcinoma. *Cancer* 1953; **6**: 704-7.
198. Sue-Ling HM JD. D1 Versus D2 Dissection for Gastric Cancer. *Lancet* 1995; **10**(345): 1515-6.
199. J M. Doctors and Medical Science. *Australas Med J* 2012; **5**: 462–7.
200. Steihaug S JA, Ådnanes M, Paulsen B, Mannion R. Challenges in Achieving Collaboration in Clinical Practice: The Case of Norwegian Health Care. *Int J Integr Care* 2016; **16**(3).
201. Bertram WA. Moore WM, Gooberman-Hill R Optimising recruitment into trials using an internal pilot. *Trials* 2019; **207**.
202. Uozumi R YS, Kawaguchi A Patient recruitment strategies for adaptive enrichment designs with time-to-event endpoints. *BMC Medical Research Methodology* 2019; **19**.
203. Kitto SC GR, Peller J, Moulton C-A, Gallinger S What's in a name? Tensions between formal and informal communities of practice among regional subspecialty cancer surgeons. *Advances in Health Sciences Education* **23**: 95-113.
204. Ergina. Challenges in evaluating surgical innovation.
205. Roberts DJ ZD, Ball, CG, Kirkpatrick AW, et al. . Challenges and potential solutions to the evaluation, monitoring, and regulation of surgical innovations. *BMC Surg* 2019; **19**(119).
206. Wartolowska K CG, Hopewell S, et al Feasibility of surgical randomised controlled trials with a placebo arm: a systematic review *BMJ Open* 2016; **6**: 010194.
207. D'Lima D AG, Brett SJ, Bottle A, Smith A, Benn J. Continuous monitoring and feedback of quality of recovery indicators for anaesthetists: a qualitative investigation of reported effects on professional behaviour. *Br J Anaesth* 2017; **119**(1): 115-24.
208. Hashimoto DA RG, Rus D, Meireles OR, . Artificial Intelligence in Surgery: Promises and Perils. *Ann Surg* 2018; **268**(1): 70-6.
209. Reynolds JV, Preston SR, O'Neill B, et al. ICORG 10-14: NEOadjuvant trial in Adenocarcinoma of the oEsophagus and oesophagoGastric junction International Study (Neo-AEGIS). *BMC Cancer* 2017; **17 (1) (no pagination)**(401).

APPENDICES

Appendix A: Embase online search strategy for 'Systematic Review of Challenges to quality assurance of surgery in oncology trials' (Chapter 2)

	Searches	Results
1	Quality Control/	180310
2	Quality.mp. and (assurance or improvement*).ab,ti. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	274431
3	1 or 2	421505
4	exp surgery/	5155294
5	(surgical or surgery).ab,ti.	2464239
6	4 or 5	5765029
7	"Trial*".ab,ti.	1442060
8	"Barrier*".ab,ti.	359245
9	"Challenge*".ab,ti.	799156
10	"difficult*".ab,ti.	899495
11	8 or 9 or 10	1954996
12	limitation.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	131078
13	11 or 12	2071694
14	3 and 6 and 7 and 13	1556

Appendix B: Risk of bias assessment of randomised trials for Systematic Review of Challenges to Quality assurance of Surgery in Oncology Trials (Chapter 2)

	Bonenkamp [45]	Cuschieri A [51]	Degiuli M [70]	Fleshman J [61]	Neudecker J [66]	Van der Pas MH [69]	Hewett PJ [63]	Simunovic M, [67]	Sano T [58]	Veldkamp R [59]	Kitchener H [65]	Primrose J [71]	Wright FC [62]	Macdonald JS [16]	Krook JE [43]	Gerard A [42]	Tveit KM [54]	Balch CM [40]	Holm T [55]
Sequence generation (1a)	?	?	+	+	?	+	?	?	?	+	+	+	?	?	?	?	?	?	?
Allocation concealment (1b)	?	+	+	+	+	+	?	?	+	+	+	+	?	?	?	+	?	?	?
Performance bias (2)	?	?	+	?	?	-	-	?	?	-	?	+	+	?	?	?	?	?	?
Detection bias (3)	?	?	+	?	?	-	-	+	?	-	?	+	-	?	?	?	?	?	?
Reporting bias (4)	+	+	+	+	+	+	+	+	+	+	+	+	+	?	+	?	+	-	+
Attrition bias (5)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Other bias (6)	+	+	+	?	?	+	?	+	+	+	+	+	+	-	-	+	+	-	+
Overall	?	?	+	?	?	-	-	?	?	-	?	+	-	-	-	?	?	-	?

Appendix C: Embase online search strategy for 'Systematic Review of Oesophagogastric Randomised Controlled Trials with surgical interventions' (Chapter 3)

No	Search	Results
1	gastric cancer.mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word]	87276
2	oesophageal cancer.mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word]	4991
3	esophagogastric cancer.mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word]	481
4	esophageal cancer.mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word]	29796
5	squamous cell carcinoma/	129665
6	stomach tumour/	35165
7	esophageal tumour/ or oesophageal cancer/	51003
8	gastroesophageal junction/	3922
9	adenocarcinoma/	106112
10	neoplasm	484258
11	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10	823494
12	General surgery/	16496
13	esophagus resection/ or esophagus surgery/	25138
14	exp gastrectomy/	67613
15	(resection or gastrectomy or oesophagectomy or esophagectomy or surgery).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word]	3877486
16	12 or 13 or 14 or 15	3877486
17	Clinical Trial/	987728
18	Randomized Controlled Trial/	596195
19	controlled clinical trial/	463931
20	multicenter study/	244899
21	Phase 3 clinical trial/	45571
22	Phase 4 clinical trial/	3797
23	exp RANDOMIZATION/	86676
24	Single Blind Procedure/	38229
25	Double Blind Procedure/	172830
26	Crossover Procedure/	62765
27	PLACEBO/	357905
28	randomi?ed controlled trial$.tw.	223001
29	rct.tw.	36123
30	(random$ adj2 allocat$).tw.	42431
31	single blind$.tw.	24616
32	double blind$.tw.	212365
33	((treble or triple) adj blind$).tw.	1159
34	placebo$.tw.	309465
35	Prospective Study/	588511

36	or/17-25	1531201
37	Case Study/	76703
38	case report.tw.	432361
39	abstract report/ or letter/	1133798
40	Conference proceeding.pt.	0
41	Conference abstract.pt.	3721324
42	Editorial.pt.	644785
43	Letter.pt.	1103784
44	Note.pt.	787716
45	or/37-44	6733532
46	36 not 45	1164051
47	11 and 16 and 46	7340
48	limit 47 to yr="2000 - 2018"	5437
49	limit 48 to randomized controlled trial	1576

Appendix D: Risk of bias in oesophagogastric RCTs using Higgins tool (Chapter 3)

Higgins	Allum[99]	Bajetta[85]	Bouche[89]	Burmeister[90]	Bass[105]	Cunningham[92]	Degiuli[107]	De Vita[94]	Di Costanzo[97]	Hiraro[109]	Kelson[95]	Kulig[100]	Marriette[108]	Miyashiro[102]	Nashimoto[86]	Noh[106]	Sano T[111]	Sasako[98]	Sasako[103]	Smalley[4]	Songun[101]	Van Hagen[104] (CROSS)	Omloo JM[96]	WU CW[93]	Yang ZQ[110]	Yu W[91]
Selection bias - random sequence	?	?	?	+	?	+	+	?	+	+	?	+	+	+	?	+	+	+	+	?	?	+	?	+	?	+
Selection bias - allocation concealment	?	+	?	+	?	?	+	?	?	+	?	?	?	+	?	+	+	+	?	?	+	?	?	+	?	+
Performance bias.	?	+	?	+	?	?	+	?	?	?	?	?	?	?	?	+	?	?	?	?	+	?	?	+	?	?
Detection bias.	?	+	?	+	?	?	+	?	?	?	?	?	?	?	?	-	?	?	?	?	+	?	?	+	?	?
Attrition bias.	+	-	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Reporting bias.	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+
Other bias.	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	+	+	-	-	+	+	+	+	-	+
Overall	?	-	?	+	?	?	+	?	?	?	-	-	?	?	?	-	?	?	-	-	?	?	?	+	+	?

Appendix E: Qualitative analysis of emergent themes within expert interview study (Chapter 4)

4.3.1. Credentialing

4.3.1.1. Challenges to credentialing of surgeons and centres in oncology trials

Limitations of surgical volume in selecting centres for participation (n=14)

The variability of surgical quality irrespective of a centres operative volume became a prominent concern amongst study participants. Many stakeholders felt operative quality could be low regardless of the total number of procedures performed within a centre.

'And there is another problem with threshold, is that you could do 100 bad operations a year and 20 good ones. And so, in isolation, a threshold is not sufficient to make an assessment of quality.' (Quote, Surgeon 25)

'You need the expertise. I know that there are some centres that have enough cases, but they have a very, very bad quality. So this is, it's both.' (Quote, Surgeon 33)

Other emergent sub-themes involving limitations in centre selection according to operative volume included:

- *High volume operating may lead to constraints on other resources*
- *Association between high volume centres and adverse outcomes*
- *Credentialing according to centre operative volume not accounting for complexity of the cases performed*

Many stakeholders were concerned that in some high volume centres, the operative volume itself may lead to constraints on supportive services within the centre such as intensive care bed capacity, which may in turn lead to adverse outcomes. Others

expressed caution that operative volume may not account for the complexity/difficulty of cases and thus may be misleading.

Personal factors and bias influencing surgical trial centre selection (n=11)

The perception that trial centres are selected according to expectation of collaboration or on the basis of a 'friendly relationship' rather than utilising objective criteria has became one prominent emergent theme.

'Well the factors that people use when selecting surgical centres generally are affected by the expectation of cooperation as their primary aim, which may not be the best criteria for selecting a centre.' (Quote, Surgeon25)

'I think the sense of people trying to deliver a trial quickly and efficiently, they choose the people that they think they know because they perceive that a friendly relationship is as good as a good high-quality centre. I think it could be improve.' (Quote, Surgeon 37)

The difficulty of maintaining quality of surgery with younger (trainee) surgeons operating within trials under supervision of participating trial surgeons was highlighted. On the other hand, concerns regarding newer consultants commencing work within centres already included in a trial were emphasised as an important challenge to ensuring surgical quality within trials. This was due to the fact that they would likely be less experienced, and would not likely be supervised whilst operating, making it more difficult to maintain the quality of trial operations.

'….a younger guy, gets trained by the participating surgeon and we struggle with how to deal with it, because yeah you're not completely sure that…' (Quote, Surgeon 34)

‘Within ROMIO we have trainees doing parts of the procedure and that's entirely reasonable given that they will be supervised, and a new consultant may not be, and so I think you've got a difficult area of new consultants.’ (Quote, Surgeon 23)

Excessive focus on surgeons/selection bias and difficulty in surgeon recruitment (n=5)

It was felt that excessive focus and scrutiny on selection of surgeons for participation in oncology trials could cause potential problems in recruitment by inadvertently causing offence to the potential trial surgeons.

‘I think if the criteria, the selection criteria is very strict and very specific maybe the surgeons won’t agree to participate in the trial if they feel like it is too much like digging deep into the background.’ (Quote, Oncologist 52)

‘…some of surgeons would be insulted by the idea of somebody coming along and saying we want to do a trial but we need to assess your competency to do that operation, it might … you know some surgeons might just say well thanks but I won’t bother. (Quote, Oncologist 55)

Some stakeholders argued that junior surgeons would be more likely to be selected for trial participation as they would be more likely to tolerate such increased levels of scrutiny and monitoring, leading to the sub-theme below:

- *Increased focus on surgeon selection leading to selection bias*

Difficulty in determining surgeon experience or quality of operating and consequent restrospective case registration (n=4)

The inherent difficulty of quantifiably evaluating the quality of surgery or of a surgeon was also highlighted.

‘But it's difficult enough to know who's good and who isn't in surgery anyway….But I don't know how you can evaluate quality in surgery in general.’ (Quote, Oncologist 69)

This difficulty in credentialing surgeons for trials was reflected in the historical example of the IT 116 SWOG trial in which there was retrospective registration. Surgeons were selected to participate in the trial following trial committee receipt of the patient pathology report emerged as stakeholder perceived challenges to surgeon selection. This was postulated to be due to the size of the IT 116 SWOG trial in the United States and insufficient resources for SQA.

‘I suppose because that trial was so big it just didn’t have the resources to look at surgical quality and also that trial was done after the event. So patients only participated after they had the pathology report.’ (Quote, Oncologist 58)

4.3.1.2. Strategies to overcome challenges to implementation of credentialing within oncology trials

Case volume should be considered in selecting surgical centres for participation in oncology trials (n=31)

A common strategy expressed by stakeholders (n=31) to overcome challenges to credentialing of centres in oncology trials included review of case volume of the operative centres. Multiple stakeholders supported the concepts that centre volume is associated with clinical outcomes, and that it should not be the only factor considered in selecting centres for trial participation.

‘Certainly a high volume centre with good outcomes. If you’re trying to investigate a new treatment or a new technique the last thing you want is poor surgery or other factors before surgery other factors to start diluting or biasing the results that you’re getting.’ (Quote, Oncologist 61)

‘You can’t do it on one marker. But a low volume hospital we know is less likely to have good outcomes than a high-volume unit so it is important.’ (Quote, Surgeon 21)

Case volume used as a minimal threshold criterion

Utilising trial centre operative volume as a minimum inclusion criterion became a popular emergent theme. Some stakeholders felt that those selecting trial surgeons should be required to adopt a broader appreciation surgeons abilities beyond this minimal operative volume.

‘So, we have to be very, very objective and transparent and then surgical volume is, of course, a good start but it doesn’t tell us very much. It’s the minimum requirement to have a decent volume but then you have to go into another type of assessment of the surgical technique and the surgical proficiency.’ (Quote, Surgeon 27)

Volume differs according to geographical region

The geographical variability in centre operative volume and the perception of what should be considered ‘high volume’ became an emergent theme and was conveyed within the following participants’ quote:

‘My speciality is gastric cancer surgery. But in Japan, high volume centre means the hospital that has maybe between 200 and 600 cases. In Japan, but in Korea or in China, the circumstances are different. Maybe in Korea, the high volume centre means the hospital has maybe more than one thousand.’ (Quote, Surgeon 9)

Surgical expertise and quality of surgery associated with centre operative volume

Another prominent theme amongst surgeon stakeholders was that high centre volume is associated with certain surgeon characteristics including surgical expertise

and even clinical equipoise. The example of the CROSS trial which demonstrated the efficacy of neoadjuvant chemoradiotherapy in oesophageal adenocarcinoma was utilised by one stakeholder to attribute the 'high quality of this trial' to its selection of high volume academic centres.

'I think, surgical expertise, that is highly related with volume, hospital volume' (Quote, Surgeon 1)

'I think the CROSS trial is interesting because it established a very high quality trial with high standards and basically its quite a select number of very high volume academic centres.' (Quote, Surgeon 345)

Case volume should be considered in selecting surgeons for participation in oncology trials (n=18)

The importance of overall case volume of potential participating trial surgeons was highlighted by multiple stakeholders as a consensus theme.

'I think that, well the number of cases performed is important, and it is easy to assess and to ask the people.' (Quote, Surgeon 22)

Specific case volume required

Many participants feel that a specific case volume is required and give examples of specific volumes they consider the acceptable for surgeon selection within a trial.

'So, each surgeon who perform 350 or more cases should be included into the trial every year.' (Quote, Surgeon 3)

'Of course they should have a decent experience and I don't know the exact cut off point but if you do a new technology they should have done at least from 50 or 100 or… It always used to be 20 in most trials, but I think 20 is way below a reasonable learning curve, I think the number should be higher.' (Quote, Surgeon 34)

Surgeon case volume and post-operative complication rates

The fact that ones post-operative complication rates will appear more significant within a junior consultants smaller overall case volume was highlighted. This would be relevant in the case below if consultants were required to have a percentage complications rate (e.g. strokes following carotid endarterectomise) below a certain threshold

'And it is just bad luck that the cohort patients are asymptomatic, and they could have a stroke at any minute and they had it peri-procedurally. Like it is just one of

those things but on paper you are like, 'Oh no he has got a whatever, 20% stroke factor.' (Quote, Trial Manager 40)

A surgeon's learning curve should be considered in selecting surgeons for participation in oncology trials assessing surgical interventions (n=34)

Importance of assessing learning curves of surgeons in trials assessing new techniques

The particular importance of this with trials comparing new to more established techniques was reinforced by several participants.

'If this is concerned about a very new technique like laparoscopy or something, then if they are in the learning curve then this video check examination is very useful and important. (Quote, Surgeon 2)

'If it's a brand new procedure then you have to accommodate the learning curve within your study…..Whereas if it's a well-established procedure, such as let's just say, esophagectomy that's something where there shouldn't be a learning curve.' (Quote, Surgeon 25)

Other emergent sub-themes included:

- *Generic importance of assessing the learning curve and avoiding risks to patients*
- *Surgeons' learning curve plateaux*
- *Variability of the learning curve*
- *Accommodate statistically for the learning curve*
- *Surrogate operations to measure experience*

The importance of assessing the learning curve comprehensively including review of several 'quality indicators' over time is reinforced. The importance of trial surgeons

reaching a plateaux and trial patients not receiving operations within a surgeons 'learning curve phase' have also been emphasised in endeavour to minimise risks to trial patients. The number of cases required for a surgeon to develop operative skills to a pre-specified level may vary depending on a number of factors including the aptitude of the surgeon for learning new skills or the difficulty of operative cases. With the advantage of avoiding the logistical problems of assessing surgeons learning curves prior to recruitment thereby potentially limiting surgeon and patient recruitment to the trial, many stakeholders advocate accommodating statistically for the learning curve in oncology trials assessing novel interventions. The concept of utilising other surrogate skills to credential surgeons for a trial involving novel techniques was also proposed, on the premise that other relevant operative experience would permit a shorter required learning curve for the operative procedure within the trial.

Training, education and mentoring of surgeons to gain required trial operative skills to allow accreditation/qualification process prior to trials participation (n=5)

Involving surgeon trainees in the trial SQA process was felt to be important in addition to specific mentoring of trainees in trial specific operative skills allowing them to gain trial operation specific accreditation.

'…there should be an accreditation process at some point assessing the surgeons on maybe how many of those operations they have done, and their skill level, how long they have been practicing and things like that' (Quote, Trial Manager 38)

'And mentorship might be important for this, that you have done the procedure with somebody who is doing it routinely and has demonstrated that they are competent and they do know what they are doing.' (Quote, Oncologist 54)

Education to improve surgeons' understanding of research methodology (n=5)

Some stakeholders proposed a specific course integrated within the UK surgical training system to ensure that on gaining CCT consultants have a required understanding of research methodology.

'I think the other thing we should do though, in surgical training exposure to research methods and an understanding of how to read a scientific paper or how to appraise a scientific work, I think they should be compulsory elements of training. 'You can't become a consultant until you have got your badge and ticked a box. ' (Quote, Surgeon18)

Credentialing of centres prior to centre selection for trial participation according to recruitment target (n=5)

One participant made clear the importance of the timing of credentialing being before commencing the trial and before selecting trial centres. Stakeholders perceived centre volume is considered essential by many for meeting trial recruitment targets in order to achieve statistically required sample sizes.

'Before doing clinical trial you should check the surgical skill of all candidate hospitals.' (Quote, Surgeon 12)

'We certainly would want a high caseload when looking at sites because we need to know that they are going to be able to recruit.' (Quote, Trial Manager 40)

Adjust stringency of selection of trial centres according to trial design, phase of trial, disease prevalence or operation complexity (n=13)

Adjust stringency of centre selection according to prevalence of disease and complexity of surgical procedure

Multiple stakeholders suggested that a trial committee should relax stringency of centre selection with increased prevalence of the trial disease.

'So if its common disease like colon cancer, then the importance of, of course some range of quality should be counted, but its less important considering the future application, because many people do it. But less common disease like EG junction tumour or pancreas cancer, the more important is quality in surgery and selection of surgeons, because these patients go to some centres.' (Quote, Surgeon 1)

'For more common operations you need to understand what is going on in the real world.' (Quote, Surgeon 15)

Another common theme was to increase stringency of selection of centres with increased complexity of surgical procedures performed within the trial.

'It also, obviously, depends a little bit on the complexity of the surgical procedure or the issue that's being assessed. The aspects of assessing outcomes in inguinal hernias is different than assessing outcomes in Whipple procedures.' (Quote, Surgeon 24)

Other emergent sub-themes included:

- *Adjust stringency according to Hazard Ratio*
- *Adjust stringency according to trial design and purpose*
- *Adjust stringency according to phase of trial*

Another emergent sub-theme important in selecting centres was that of the Hazard ratio. In trials when the Hazard ratio between centres remained constant for a given surgical procedure then it was advocated to relax the stringency of their centre selection criteria. Stakeholders argue that explanatory trials should be selecting centres or those renowned as 'centres of excellence' in that specific procedure or teaching hospitals, whereas one can broaden the selection remit if the trial question focuses on adoption rather than delineating the most effective surgical technique for

a given disease process. The concept of adjusting trial centre selection stringency according to the phase of the trial was also a popular theme. In phase 2 trials stakeholders felt more stringent centre selection was required to ensure "best possible circumstances" (Quote, Trial Methodologist 50), whereas for phase 3 trials stakeholders suggested building "flexibility into your trial" (Quote, Oncologist 66).

Trial centres required to meet certain criteria for consideration of inclusion in oncology trials including: being specialist centres; part of national audit; having published expertise; not have outlying results (n=29)

Required Infrastructure, resources and experience

The required research infrastructure and resources was a prominent theme amongst stakeholders.

'So I would say not only having those engaged surgeons, but also who can be supported by adequately resourced staff, trained research staff.' (Quote, Trial Manager 46)

Other emergent sub-themes included:

- *National specialist/centralised centres and audit/monitor their outcomes* with published expertise
- *Commitment of participating centre's surgeons*
- *Documentation and management of post-operative complications*
- *Centres with teams with clinical equipoise*
- *Qualifications and accreditation of surgeons within centres for selection*

The suggestion of selecting national, specialist centres with published expertise which audit their outcome data became a prominent theme. Others felt the commitment of surgeons within specific centres should be considered as a factor in

centre selection. Identifying centres with surgeons whom are willing to be assessed within the trial and potentially make changes to their practice emerged as another important factor in centre selection. Thorough documentation and operative team detection and appropriate management of post-operative complications became an important theme in centre selection. Trial surgical teams having the required clinical equipoise regarding the trial questions emerged as an important factor in trial centre selection. Another emergent theme indicated the importance of trial centre and their surgeons' qualifications/certifications that they can adequately perform trial operations, in order to adequately answer the trial question.

Impartial committee required in selecting surgeons for participation in oncology trials (n=3)

The importance of impartiality and objectivity in surgeons became an emergent theme with some advocating that surgeons from different specialities and without trial involvement should be involved in selection process to minimise potential bias.

'They are not involved in the trial and they do not know each other so there is no conflict of interest.' (Quote, Surgeon 21)

Adjust stringency of surgeon selection according to trial design, phase of trial, disease, operation complexity and other specific factors (n=8)

Adjust stringency depending on the trial surgical procedures

The concept of adjusting surgeon selection stringency according to how commonly performed the trial surgical procedure is and its complexity became a prominent theme.

'I mean again I think it depends on how commonly performed that procedure is. I mean obviously if it is a bread and butter procedure like a laparoscopic anterior

resection or abdomino-perineal resection then in theory any surgeon in that centre should be more than capable of performing that procedure.' (Quote, Oncologist 58)

'So it depends if the intervention is dependent on the skill of the surgeon then yes it is important to again see whether you should be selective or not of the surgeon for your results.' (Quote, Trial Methodologist 59)

Altering the stringency of surgeon selection, rather than centre selection as outlined previously, according to trial design became also became a prominent theme (as shown below). Stakeholders advocated increasing stringency of surgeon selection in explanatory designs and reducing it for pragmatic trials.

- *Adjust stringency depending on the trial design*

Graded method of surgeon selection utilising structured objective assessment tools (n=3)

An emergent theme particularly amongst oncologists and methodologists with experience in radiotherapy oncology trials was to select surgeons using an individualised, staged process in which surgeons are initially screened and then a more detailed assessment of their skills initiated if they do not meet initial standards.

'Whether it is possible that in a feature that you have different layers of assessments, so you use a picture first, so basically a screen of the surgeon. If you have a lot of surgeons for recruitment, then you use more detailed video assessment.' (Quote, Trial Methodologist 65)

Selection of surgeons according to postoperative outcomes through audit of: Histopathology; intraoperative bloods loss, postoperative complications and outcomes (n=15)

Various post-operative outcomes including operative duration, lymph node yield, circumferential and longitudinal resection involvement rates and post-operative complications rates within longitudinal series were recommended by multiple expert stakeholders.

'…for oncological trials, the main determinator of outcomes, oncologically, would be the radicality of the lymph node dissection, the combination with the all zero rates and all zero rates are easily assessed, pathologically' (Quote, Surgeon 27)

'Outcomes, re-occurrences, complications. We can all find for any one individual case that went badly wrong or you can possibly think of mitigating circumstances but hence why you need data of a long enough period of time.' (Quote, Oncologist 60)

4.3.2. Standardisation

4.3.2.1 Challenges to implementation of standardisation of surgery in oncology trials

Challenges to surgeon adherence to protocol (n=7)

When asked about challenges to standardisation of surgery in oncology trials stakeholders felt that adherence to trial operative protocol is a significant challenge, particularly for complex interventions in multi-centre and international trials.

'So being that difficult I can imagine that other factors which are very strict but not as obvious like for example, the way you do your lymph node dissection in the spleen area will become also a point which will be easily neglected by the participating centres.' (Quote, Surgeon 14)

'If I were directing a surgical clinical trial comparing techniques, it would be extremely difficult to keep the consistency across one unit never mind multi-units, never mind multi-units across national divides.' (Quote, Surgeon 19)

Differing oncological beliefs and resistance to change adoption leading to inter-surgeon and inter-institutional operative variation (n=18)

The difficulty of gaining national or international agreement within oesophagogastric surgery were highlighted and stakeholder perceived reasons behind this explored.

'You know I think particularly with oesophageal surgery it's very, very difficult to even get two separate surgeons to agree what's the best operation. If you open that to different countries and different units, that will be very difficult to get agreement on.' (Quote, Surgeon 23)

'…the surgeons do not like to change the technique and surgeons being uncomfortable doing it, might perform worse, doing it. So, you have to be very, very, very selective in what aspects of the procedure that you stand by, that you put your focus on.' (Quote, Surgeon 27)

Resistance to change adoption

The difficulty in persuading experienced surgeons who have performed thousands of operations using a certain method to follow a different protocol is highlighted and explained through the bond each individual surgeon develops with their own techniques. From a radiotherapy perspective this struggle is explained that in endeavouring to change surgeons techniques to comply with protocols you are venturing outside the general thought process.

'Well it is difficult to agree on standardisation of techniques, because we all have our own tips and tricks and we believe they are fundamental to the substance of our surgery. I don't think this is true, but I don't think you can force any surgeon to change his technique too much.' (Quote, Surgeon 32)

Difficulty in consensus building, following guidelines and providing evidence

Examples are provided from stakeholders' experience of the difficulty experienced gaining consensus on standardisation of surgery within a trial and a lack of evidence can contributes to this. The importance of bringing all national societies together to help formulate a consensus is also highlighted.

Overly prescriptive protocols and standardisation causing difficulty in recruitment of surgeons (n=12)

Cautions are raised that surgeon recruitment can be compromised in addition to patient care if protocols are too strict, which is in due to the inherent unpredictable

nature of general surgery and the surgeons duty to perform the procedure in the patients best interests.

‘And the operation procedure itself is very, very different each by each. And then also if you make too much standardisation for the protocol nobody can join you. Too strict. ‘I want to do it this way.’ ‘I want to do it this way.’ You know?’ (Quote, Surgeon 10)

‘Again you can't if it is a very --- you can't have very, very strict guidelines either, because the surgeon has to do what's best for the patient's care. And sometimes that means a deviation from the trial protocol as well.’ (Quote, Trial Manager 43)

International concerns regarding consequences of standardisation and operative variation

Stakeholders highlight the difficulty of standardising procedures across continents within multi-centre trials trials comparing curgical interventions.

‘Yes, regarding oesophagogastric and gastric cancer. But one big problem is standardisation between different continents, different countries.’ (Quote, Surgeon 09)

Regional differences in perception of surgical quality and insufficient quality of data regarding outcomes

Regional differences in perception of surgical quality are highlighted and the fact that data regarding surgical quality and outcomes within routine surgical practice is insufficient was also highlighted.

‘I think it is important to identify that what we look on as a quality approach to staging treatment, operative approach and recovery in Europe and North America may be a very appropriate approach from our standpoint, but it does not necessarily

match resources, perception or historical experience in other areas.' (Quote, Surgeon24)

'One of the problems is I think if you then want to go way back to the beginning is about quality of data. Unless we as centres are reliably and reproducibly recording what we do – in routine practice.' (Quote, Oncologist 60)

4.3.2.2. Strategies to overcome challenges to standardisation of surgery in oncology trials

Use of operation manual to aid standardisation in oncology trials (n=34)

The importance of an operation manual as an adjunct to aid standardisation of operating in trials became a dominant theme.

'Yes, I think if you are working to a standard, then a manual and a video would be helpful. (Quote, Surgeon 20)

'In the European Group of the Gastric Cancer, we have asked the real experts, international experts, what is the true lymphadenectomy and there were very interesting interpretations, you know. Everyone is saying yes, I know what it is, but the reality is sometimes different. Therefore, it's really important to have a manual to say if you do an oesophagectomy for this trial…' (Quote, Surgeon 33)

Developing a consensus on the protocol

The need to develop the consensus on standardisation of the operation manual across multiple centres by a certified team was also reinforced.

'Yes I think depending on how the manual was developed and who it was developed by. I think you would need to have a certified team of people to develop something like that.' (Quote, Oncologist 54)

Other emergent sub-themes included:

- *According to the trial design – useful for multicentre and explanatory trials*
- *Learning from experience of radiotherapy trials with procedure manuals*
- *In favour of operation manual with flexibility for compulsory and optional steps*
- *In favour of operation manual without optional steps*

Use of an operation manual was felt to be particularly important in explanatory, multi-centre trials rather than single centre, pragmatic trials. Clinical oncologist expert stakeholders shared their experience from radiotherapy trials highlighting the importance of procedure manuals. Due to the unique nature of each operation determined by factors such as anatomical variation and unpredictable physiological responses or events, multiple stakeholders emphasise the importance of maintaining optional steps within a protocol. However, some stakeholders felt a more prescriptive and detailed approach to protocol development was more appropriate, and would lead to less problems potentiated by leaving space for optional steps.

Additional methods to aid standardisation including: trial launch days, operative room boards explaining trial intervention, standardisation of histopathological processes and radiology diagnostics and use of videos, photographs, virtual reality and webinars (n=18)

Operating room board

Utilising an intra-operative standardisation aid was thought be beneficial for trial surgeons.

'So we made a board of summary of the two types of surgery which should be placed in the operation room. And when a patient is allocated to one arm, before that they

quickly review which lymph nodes should be removed and so on, so on site.' (Quote, Surgeon 2)

'It can be lots of paper. Something that can be with them in theatre that they can be following could be really helpful. It certainly would mirror what we do in radiotherapy' (Quote, Oncologist 61)

Pre-trial Training

The importance of pre-trial operative training to facilitate oncology trial standardisation was reinforced by multiple stakeholders.

'With one of the criteria to enter the trial, they also can follow our course, we have a basic course of minimum invasive gastrectomy and open and there we show step by step with a video, like you said, the whole procedure, so all surgeons have to follow the same course. I think it was very useful because everybody is doing more or less the same.' (Quote, Surgeon 16)

'

Other emergent sub-themes included:

- *Standardisation of histopathological dissection*
- *Standardisation of radiology diagnostics*
- *Use of adjunctive methods to aid standardisation*
- *Virtual reality*
- *Launch days*

Variance of lymph node yield depending on method of histopathological assessment was explained to justify standardisation of pathological assessment post-operatively. Variability of pre-operative radiological assessment and staging and the consequent potential effect of this on the ensuing operation was also expressed. The use of adjunctive methods in addition to an operation manual to reinforce standardisation

including photographs, videos and webinars became a dominant theme. Employment of virtual reality technology was another popular strategy recommended to facilitate standardisation. As an opportunity for the trial committee to introduce novel techniques, and for trial stakeholders to ask questions and practice before the trial commences, the use of 'launch days' was suggested.

Flexible standardisation depending on trial design and study question, focussing on standardisation of steps affecting primary end-points, safety or survival (n=22)

The concept of certain operative steps being pre-requisite became popular, in particular those steps affecting outcomes, safety and survival.

'there are some steps of an operation that are absolutely prerequisite and have to be done. And there are others that are not in my opinion. And there has to be agreement on what the absolute prerequisite steps of the operation are (Quote, Southampton, Surgeon 15)

'So I think that if the points which may affect the results, so safety and survival.' (Quote, Surgeon 1)

Permissible variation

Stakeholders felt variation should be possible for surgeons when not affecting outcomes and to maintain individual surgeon choice within these areas.

'I think it is important. I think that there are some essential standards that are easy to describe. I think there are some degrees of variation which are permissible because they are not going to alter some outcomes.' (Quote, Surgeon 18)

Other emergent sub-themes included:

- *Standardisation depending on the trial design and study question*

- *Learning from other oncology trial stakeholders*

In addition it was proposed these prerequisite steps could vary depending on the trial question and depending on the trial design (e.g pragmatic versus explanatory, or the phase of the trial). The importance of standardisation in trials comparing surgical techniques is also highlighted. Radiation oncologists' described how a standard deviation around compliance with a protocol can be developed within their trials and recommends such a process to be developed for surgical trials.

Virtual training system, technology, artificial intelligence and robotics to aid credentialing, standardisation and monitoring of surgery in trials (n=10)

Use of technology in standardisation of surgical procedures

Stakeholders proposed with time technology may help with standardisation through devices to access operation manuals to superimposition of patient specific diagnostic images onto robotic views in real-time.

'Now maybe with better tools. Maybe we can make a very accurate, precise manual with steps. Or we can use a…., maybe an ipad or such a tools we could use?' (Quote, Surgeon 9)

'And this is a bit far-fetched for oesophageal-gastric surgery but you might have seen this thing for liver surgery where people will overlay an MRI or a CT projection onto their robotic picture, their robotic image of the liver. Yes? So you can as it were make the CT scan be the picture that you are looking at.' (Quote, Surgeon 18)

Use of technology in credentialing of surgeons for participation in oncology trials and training

Technology, virtual reality and simulation were considered important in assessment surgeons skills prior to commencement of trial participation, particularly for those trials introducing novel surgical techniques

'At least you have to clear this level. This is good to introduce some new technique for surgeons. You just have to just clear this programme by the virtual training system. This is one of the ways.' (Quote, Surgeon 10)

Use of technology in monitoring of surgical procedures

Artificial intelligence was considered important in monitoring of surgery for identifying 'dangerous procedures' and assisting in video monitoring by helping to identify the better images for further review and measuring other ergonomic factors.

Inter-institutional operative team visit exchanges to facilitate standardisation

Operative team exchanges between trial centres was considered a useful in facilitating standardisation.

'I think what's sometimes useful is the whole team to go and visit another institution and actually follow a particular surgical pathway, even from the theatre set-up or recovery set-up, of how things flow, that might be helpful in developing a particular trial.' (Quote, Oncologist 62)

'Also, for example, The Pelican Centre are very effective at training MDTs to do locally advanced rectal cancer, high quality teaching, high quality materials, and a clear beneficial outcome.' (Quote, Oncologist 66)

4.3.3. Monitoring

4.3.3.1. Challenges to monitoring of surgery in oncology trials

Surgeons pride, fame and national culture leading to difficulty in providing and receiving feedback (n=14)

Stakeholders identified some famous and influential surgeons' pride as producing challenges in surgeons providing or receiving authentic feedback.

'I think the most problem is the surgeon's pride, it's what I think. That's, all surgeons think, "I'm the best." …… So, even if I have a chance to rate some surgeon's surgical skill it's difficult to rate, to give bad score to him. Most surgeons participate in the study are very famous surgeons, very powerful surgeons and they have pride so it's very difficult to rate them……' (Quote, Surgeon 3)

'I think the biggest challenge is to make surgeons comfortable with this. To make surgeons really understand the benefits and to make them happy in accepting being assessed and making changes in their practice.' (Quote, Surgeon 27)

Cultural factors

The role of surgeons' national culture was felt to contribute to difficulties in providing and receiving feedback within Eastern and Western Societies.

'You know, if we ask, it maybe could be very embarrassing if you ask somebody to participate in a trial then they send you a video of this operation, you assess it and say well, your quality of surgery is poor. I mean I don't see… At least in the Western world, I don't see that it's a very practical way to go.' (Quote, Surgeon 22)

Concern over exposure, judgement and human factors

Stakeholders had concerns over exposure and potential judgement within the surgical monitoring process. These concerns extended to providing feedback to experienced surgeons and to addressing human factors such as surgeon fallibility.

Limitation of operative videos to monitor surgery in trials (n=17)

Feasibility issues including financial resources and time considerations

Feasibility issues in terms of time to monitor videos and financial resources associated with this process was the most frequently cited limitation of the use operative video monitoring (n=8)

'if we can video, all videos it's the best quality assurance of, but so far, it's very difficult because it's time consuming, yes.' (Quote, Surgeon 3)

'So if you were setting up a clinical trial up using video review, you have to have a fund of thousands, tens of thousands of dollars to pay somebody to review, because you can't do it.' (Quote, Surgeon 25)

Quality of videos in open procedures

Difficulty in producing quality videos of open procedures was considered an important challenge to video monitoring in open procedures, largely secondary to ergonomic considerations.

'Easily done, not so costly and as we, it would be challenging if you were in the, if you were studying open procedures, that might be more difficult, even though there are new cameras that have good quality, but that's the problem with open surgery.' (Quote, Surgeon 30)

Other emergent sub-themes included:

- *Variation in Surgical Practice*

- *Video monitoring affecting recruitment*
- *Video monitoring affecting generalizability of results*
- *Subjectivity of interpretation of video assessors*

Inter-surgeon operative variation was also considered to pose a significant challenge to video monitoring Concerns that some centres and surgeons may not want to partake in trials with video monitoring which would impact negatively on recruitment was emphasised. The concept that video monitoring may in itself affect the operating practices of trial surgeons was proposed. This may inherently affect the generalizability of the trial operative procedures performed and the associated outcome variables. Concern over the potential subjectivity of the video monitoring process was also expressed.

Limitations of photographs to monitor operative elements (n=12)

Inadequate reflection of quality of surgical procedure

Stakeholders comment that even a nice photograph does not necessarily indicate a good procedure and that they cannot inform operative assessors regarding several factors including operative efficiency or tissue handling.

'The quality is mostly poor. And what it doesn't demonstrate, for example, when you have a gastric tube reconstruction, how you handle the gastric tube or it doesn't show you the tissue handling. It is just a static picture.' (Quote, Surgeon 14)

'…but if you're looking at how efficient you were at performing that procedure, then photographs wouldn't suffice.' (Quote, Surgeon 20)

Insufficient to represent complexity/difficulty of operation

Other stakeholders felt photographic monitoring was unable to capture the depth or complexity of surgical procedures.

So, I would say photographic is a snap-shot, well my personal opinion is, humans are 3D. We have depth and there's positioning…. But I would have thought surgical technique is more complex than a photograph, I would have thought it's a video. (Quote, Trial Manager 37)

Feasibility and costs

Costs, coordination and anonymisation of images are listed as some of the feasibility concerns with photographic monitoring and the potential consequences of such burdensome tasks were outlined.

Feasibility and other concerns with intra-operative monitoring (n=6)

Resources concerns particularly with open surgery

Concerns regarding intra-operative monitoring were expressed particularly the extensive man-power that would be required to consistently employ intra-operative monitoring within a trial.

'I suspect I've answered that in that it's just consistency is almost impossible to maintain unless you've got the manpower to be constantly there, constantly mentoring or supervising or, of course it's easier in the laparoscopic cases where you can have video evidence. But with open surgery it's very difficult.' (Quote, Surgeon 19)

'I have no more ideas. In case of Dutch study, if they performed D1 study, research fellow went to that hospital every time, but actually its not feasible…..But during that time we had about 60 cases randomised, and I could participate in 34 D2 during my stay and the others with D1. But if the period is a little bit longer.' (Quote, Surgeon 1)

Observer bias concerns

Concerns that intraoperative monitoring itself may change the operators practice leading to a substantial source of bias was also expressed.

'I suppose someone being in the room you will naturally you know, demonstrate your best practice. I suppose video or the photographs are less invasive in a way…. Yes, and less prone to bias. I don't know, I think if someone was in the room it would cause some sort of bias.' (Quote, Trial Manager 39)

Stakeholder perceived limitations of use of other methods of monitoring of surgery (including post-operative complications/complications/review of operative notes)

Problems with review of operative notes or case report forms are highlighted particularly that of missing information and risk of bias indicating it is not a reliable methods to assess surgical quality within trials.

Referring to use of case report forms for monitoring od quality of surgery stakeholder 49 explained: 'So I would regard those as considerably more at risk of bias than some of the other things that we've been talking about.'(Quote, Trial Methodologist 49)

'…looking at the operating notes which is a way of assessing the quality, in the emergency settings, we have lots of issues with the notes might be incomplete, things might go missing because the patient might not belong to that hospital.' (Quote, Trial Manager 44)

Consequences of monitoring of surgery and providing feedback (n=7)

Exclusion of centres or surgeons

Several stakeholders explain the serious potential consequences of monitoring surgery when protocol adherence is not met or outcomes are unsatisfactory.

'So if we see some hospitals where the morbidity, mortality is outside of 95%, then eh maybe, we will kick them out. That kind of thing is very important.' (Quote, Surgeon 1)

'The same operation being observed by an external observer that you've seen, if the things are not good you have to stop the trial.' (Quote, Surgeon 26)

Managing unsatisfactory centre/surgeon performance

Rather than centre or surgeon exclusion other stakeholders explain their preference for other methods including conducting audits, mentoring programs or Corrective Preventative Action at centres in which unsatisfactory performance was detected. Stakeholders expressed the difficulty in deciding on the most appropriate method of managing sub-satisfactory performance.

'But whether you monitor when things go wrong, is important, it's important that you monitor when things go right and also that when things go wrong, and you need that again, that feedback mechanism, that CPA, that corrective preventative action to make sure that, "What can you do? Do you require more training? Can you get the expert in to shadow you for the next one?"' (Quote, Trial Methodologist 42)

Insufficient monitoring, surgeon reluctance to be monitored and concerns over adverse outcomes/litigation due to monitoring (n=8)

Litigation

Concerns regarding potential litigation in event of a recurrence and existence of an operative video have raised concerns

'Obviously you know, you can pedantic about it and say you know, well if somebody does have a recurrence and you though they want to look at this, its imperfect. That's the challenge that needs to be discussed and some advice would need to be taken. I can't see it as being a big big issue in reality.' (Quote, Surgeon 345)

Adverse outcomes

Other stakeholders were concerned the actual process of recording images may increase operative risks due to increased operative duration.

'Yeah, yeah, and also they would perceive that certain photographs looking at the specimen might be creating risk of necrosis, you know, vascular things because the operation takes longer.' (Quote, Oncologist 61)

Other emergent sub-themes included:

- *Insufficient monitoring*
- *Reluctance to be monitored*

The general lack of monitoring and perhaps the futility of such an end endeavor is suggested by several stakeholders, and this is justified by the trial design (pragmatic) and their aim to produce generalizable results. Surgeon reluctance to be monitored is expressed as a significant challenge with some explaining this through reluctance to undergo scrutiny and others in lack of perceived personal benefit from participation. Potential impact on a surgeons practice due to monitoring was another emerging concern.

Challenges to monitoring of surgery in trials including insufficient resources and finances, and a potential negative impact on generalizability and surgeon recruitment (n=5)

Generalizability and trial design

Stakeholders explain the inherent tension in the UK between generalizability of pragmatic trials and trial validity.

'I am always wary of things that interfere with how care is delivered within the NHS. There is tension between the internal validity of the trial and generalisability of the finding.' (Quote, Trial Methodologist 47)

'And actually, if they don't do it you intervene and give them more training but you are actually trying to design a different sort of trial. So most of the ones I do are definitely more at the pragmatic end than the explanatory end.' (Quote, Trial Methodologist 41)

Resources and finances

The burden and expense of monitoring adherence to protocol is emphasized by trial stakeholders.

'Another barrier would be because I think monitoring and auditing is so important, might be financial because it takes, for example, I've done monitoring, it takes sometimes a whole day and you end up looking at about 15 patients or something. It's very time consuming.' (Quote, Trial Manager 44)

Other emergent sub-themes included:

- *Surgeon recruitment*

The burden of the monitoring process and the scrutiny involved in monitoring surgeons with their potential adverse affect on surgeon recruitment became another emergent theme.

4.3.3.2. Strategies to overcome challenges to monitoring of surgery in oncology trials

Surgery in oncology trials should be monitored (11)

The overall importance of monitoring of consistency of surgical techniques within oncology trials became a popular emergent theme, particularly in trials in which the surgical technique is central within the trial.

'If the very meticulous procedure is the goal, if surgical technique is the goal of the trial then it should be recorded.' (Quote, Surgeon 2)

'I would say that I think that it's an obvious area of weakness in a poorly designed trial, if you haven't considered how you are ensuring that the activities by the surgeons are being undertaken in a consistent way to give the same quality.' (Quote, Trial Manager 37)

Surgery in oncology trials should be monitored with video recording (n=22)

Feasibility for laparoscopic surgery

Video assessment was regarded by many as the optimal form of monitoring, particularly for laparoscopic operating, due to the clear visualisation of operative events.

'Yes, well that is the ideal way. For something that you can video that would be absolutely the ideal thing. If you had a series of people scoring blindly loads of videos. Yes, that would be ideal.' (Quote, Surgeon 21)

Random selection of videos for monitoring

Stakeholders often advocated random selection of such videos for review by a monitoring group. The random selection was felt to help overcome the time-consuming/resource intensive concerns some stakeholders had whilst minimising potential bias within the video assessment process.

'Well the best or preferred aspect for ensuring a standardised approach would be the aspect of videoing the operations and reviewing them randomly.' (Quote, Surgeon 24)

Other emergent sub-themes included:

- *Video monitoring perceived as less susceptible to bias*
- *Videos are more representative of operative events*
- *Educational value*

Other stakeholders regarded video has less susceptible to bias than photographic monitoring and that it would be more representative of operative events. The concept that videos are more demonstrative of operative events and the surgical technique became a popular emergent theme. The educational value of video monitoring was also thought to be important for the surgeons involved.

Surgery in oncology trials should be monitored with photographs (n=11)

Photographs following lymph node dissection

Stakeholders felt photographs were particularly useful for demonstrating satisfactory completion of lymph node dissection.

'….I think, do some still images of the lymph node dissection is the best way to assess.' (Quote, Surgeon 3)

'OK the essential criteria might be that you remove these lymph nodes, you prove that you have removed those lymph nodes, you give me a picture or photograph to show me what it is like afterwards and you show me the pathology that the pathologist says that was removed.' (Quote, Surgeon 18)

Photographs are useful for open surgery

The utility of photographic monitoring of surgery for open surgery was a popular emergent theme.

'Speaking of video monitoring stakeholder 02 explained: 'But with open surgery we don't do that but the pictures is useful.' (Quote, Surgeon 2)

'…in case of laparoscopic surgery, video is feasible, but in the case of open surgery it is not feasible, so instead of that we use pictures and photographs. So photographs after dissection, lymph node dissection.' (Quote, Surgeon 1)

Other emergent sub-themes include:

- *Improved feasibility of photographic monitoring*

- *Standardisation of photographic monitoring*
- *Photographic monitoring reported as useful in previous trials*

Multiple surgeon stakeholders considered simplicity and efficiency of photographic monitoring as appealing attributes of this monitoring method. The importance of clear specifications as to the exact anatomical and operative stage to take the photograph was reinforced. Stakeholders' positive experience of photographic monitoring within colorectal trials, such as the ARISTOTLE trial (comparing standard versus novel chemoradiation treatment (CRT) as pre-operative treatment for locally advanced rectal cancer), became a popular emergent theme.

Surgery in oncology trials should be monitored using intra-operative monitoring (n=15)

Reliability of intra-operative monitoring

Multiple stakeholders felt this was the most reliable and less susceptible to bias than other methods of monitoring of surgical quality.

'One possible solution is visiting the hospital and observation. But it takes time and it takes cost. It is very very reliable. Visit of one professor or one reviewer, visit the hospital and watch' (Quote, Surgeon 9)

'…you cannot just say, "Oh he did 100, you need to be sure that the person who is involved has experience in and if you don't know he has to show, you have to go there and he has to show because a video can be a fake.' (Quote, Surgeon 31)

Relative superiority of intra-operative monitoring

Multiple stakeholders explain the superiority of intra-operative monitoring through comparison with other available methods.

'At least a video, but from what I've learnt from Kristoff Marriett, is because I asked him, "How did you do that?" He said, "You know, the video is not enough, you have to visit every centre and you have to see the surgeon who is performing the operation otherwise you don't know."' (Quote, Surgeon 31)

Other emergent sub-themes included:

- *Intra-operative monitoring within graded monitoring system*

- *Intra-operative monitoring in the form of mentoring*

The concept of intra-operative monitoring being the next step if a deviation from protocol is detected on video monitoring emerged. Intra-operative monitoring in the form of mentorship was another popular theme taking emphasis away from scrutinising surgeons' practice to helping surgeons to develop skills and improve where required.

Surgery in oncology trials should be monitored according to histopathological assessment of trial resection specimens (n=10)

Standardisation of the histopathological process

In reinforcing the critical importance of monitoring histopathological assessment of resected trial specimens, stakeholders also advocated standardising this process.

'Yes we have got to have standardised assessment of resections. That is the critical element of any trial. And so there has to be agreement between pathologists before you set something up as to how they are going to process the specimen and then analyse it' (Quote, Surgeon 21)

'So they are actually, the results should, the most useful checking system is the pathology.' (Quote, Surgeon 6)

Pragmatism of monitoring surgery through histopathological assessment

The pragmatism of monitoring surgery via histopathological assessment of resected specimens was reinforced.

'Well I would probably prefer the pathology style route. Just from a pragmatic viewpoint.' (Quote, Trial Methodologist 47)

Other emergent sub-themes included:

- *Established tools for assessment of surgical quality*

Utilising established tools, such as histopathological assessment, for monitoring surgery became a popular emergent theme and was considered a well recognised measure of quality of surgery through its historical utility within colorectal trials.

Other suggested methods of monitoring surgical quality in trials including post-operative complications/outcomes, lymph node yield, operative notes, case report forms, real time data monitoring (n=32)

Blood loss/operative duration

Monitoring operative duration and blood loss emerged as popular monitoring strategies.

'There are all the standard things we would all immediately say, blood loss, how many hours in theatre. They are all small elements which you can monitor which are quantifiable.' (Quote, Surgeon 21)

Lymph node count

The importance of counting the number of lymph nodes within the resected upper gastrointestinal specimen and reaching a certain minimum threshold, commonly percieved to be 25, was emphasised by multiple stakeholders.

'But it is important to say 25 lymph nodes is a minimum and if you, the pathologist want to and it does work. It's to look again and again and again. It's one point it's not the most important point, but it's one point.' (Quote, Surgeon 33)

Conducting full analyses of adverse events evaluating root causes behind such incidents became an emergent theme. Use of novel plasma markers including pre and post-operative DNA load levels were suggested to be potentially useful monitoring strategies once developed. Review of pre-operative investigations, pre-operative briefs and completion of operation notes were proposed as potential surrogate markers of surgical quality. Remote central monitoring of trial surgeon's performance was also emphasised, including making a comparison between ablated tissue with pre-operative diagnostic imaging. In addition to traditional monitoring of

post-operative complications the need to assess patient quality of life was re-enforced. Assessing short-term outcomes including length of stay, morbidity and mortality were considered important according to multiple stakeholders. Monitoring of local recurrence rates, progression-free and overall survival were considered essential monitoring strategies by multiple stakeholders. Some stakeholders advised audit of a centres' general practice in addition to their trial patient outcomes and comparing these to international (US and European) standards for benchmarking. Some stakeholders advocated auditing a certain percentage of recorded procedures whereas other advocated a random selection of procedures should be audited. Electronic case report forms have also been suggested allowing more frequent, surgeon specific recruitment monitoring. Other stakeholders maintain belief that review of operative or case notes remains an important method of monitoring of quality of surgery. Standardisation of operative notes documentation was also proposed.

Adjusting stringency and methods of monitoring depending on trial design and research question/surgical procedure

Adjusting stringency to trial design/question

Adjusting according what is appropriate for that particular trial was recommended including particular steps that are required to be monitored and the feasibility of such monitoring with available tools.

'I think it is entirely dependent on the trial and what makes most sense.' (Quote, Trial Methodologist 41)

'If they're really simple things like stitching or some cut or some removal, if they're things that could be easily captured in a photograph you're going to be better off with a photograph but if it's a bit subtler that that you're better with video' (Quote, Trial Methodologist 48)

Traditional technique versus new innovation

Adjusting the intensity of monitoring according to surgeons familiarity with the procedure was proposed.

'So, if this surgical intervention is a very well-known surgical procedure that's been done, many times, but this is just being done for a slight variation of a theme, then I would say that it needs less monitoring than perhaps an innovative, new in-development, surgical intervention.' (Quote, Trial Manager 37)

Standardised monitoring process (n=10)

Multiple stakeholders emphasised the importance of standardising the monitoring process.

‘It should be standardised, we haven’t done that, but we should actually because they loose focus and they blame the nurse, they say, oh, the nurse was not taking a good photo, come on ... not their responsibility.’ (Quote, Surgeon, 16)

‘But I think obviously external audits of the procedure is key in those trials, but that’s quite difficult to get organised.’ (Quote, Surgeon 35)

Selective reviewing of recorded trial videos by trial committee assessors

Rather than full review of unedited videos during surgical quality monitoring some stakeholders advocated fast forwarding to key operative steps. Other stakeholders specify the operative stages at which monitoring would be required indicating, during a three phase oesophagectomy, four short videos of each the abdominal compartment, hiatus, carina and anastomosis following lymphadenectomy would be required.

‘In the precision of bursectomy, only that part, we can quickly fast forward to review to see, ‘OK that has been done or no they did not touch one.’ These can be checked, not all three or four hours of the surgery.‘(Quote, Surgeon 2)

‘I don't think it is feasible to review long videos. So, find a compromise just short video showing the field after the dissection and showing the main few key points of the reconstruction for example, probably be sufficient and more.’ (Quote, Surgeon 32)

Graded monitoring system, adjusting stringency of monitoring according to performance and monitoring of initial cases (n=9)

A graded monitoring system emerged as a popular theme, particularly amongst oncologists where it appears to have been extensively utilised within oncology trials involving radiotherapy. Stakeholders have advocated a risk-based monitoring system

in which the initial few cases would be monitored in their entirety, followed by a random selection of cases performed in each centre.

'We do everyone's first case in real time and then if there's been an issue with that when we do the next one or subsequent ones, however long it takes, but we then have a random selection so we've looking at overall 10% of cases of the study will be QA'd.' (Quote, Oncologist 61)

'That could be risk based or it could be you go in and film the first three or four and after someone has demonstrated they are doing it properly then you don't film it anymore.' (Quote, Trial Methodologist 47)

Graded monitoring according to quality monitoring in previous trial

Another proposed variant of a graded monitoring system involves adjusting intensity of the monitoring process at a centre according to their quality assurance performance in previous trials.

'…they used to accept that if there was a good radiotherapy QA process in another trial in the same area then they'd do a lighter touch radiotherapy QA for the next subsequent trial.' (Quote, Trial Methodologist 71)

Balance between comprehensive and pragmatic monitoring

Initial of screening trial operations using photographic monitoring followed by video monitoring if standards fell below an agreed guideline was another emergent concept.

Anonymity and confidentiality of monitoring assessments of surgeons

The importance of anonymity and confidentiality within the monitoring process and in providing feedback to trial surgeons became an emergent theme, particularly prominent amongst surgeon stakeholders.

'I think there has to be anonymity, that's a given really. The difficulty is that if there are issues that you want to feedback then that is very difficult, but it needs to be anonymous and it needs to be benchmarked.' (Quote, Surgeon 23)

Structured, regular and appropriately senior feedback mechanism with positive reward system and external peer review

Providing positive stimulation to trial centres was felt to aid in maintaining centre adherence to protocol. Through such a structured feedback process stakeholders advocate one should aim for a 'win-win' approach and focus on adherence to protocol rather than sounding judgemental regarding quality of a surgeon's operative performance.

'I believe more in monitoring of course but also in rewards, in giving rewards, giving positive stimulation towards participating centres when they keep up the good work even after one year of participating in the trial, or even two years after.' (Quote, Surgeon 14)

'…you must create a win-win situation and then you will be successful. And then it is all about quality, you want to increase quality, you do not want to punish people and you want the people getting better that's what you want, that is my opinion.' (Quote, Surgeon 31)

Senior feedback

The importance of this feedback being from a senior surgeon was also reinforced.

‘Have your senior PI responsible, not the guy at the same level in training’ (Quote, Surgeon 16)

4.3.4. Generic challenges and mitigating strategies

4.3.4.1. Generic challenges to SQA in oncology trials

Lack of clinical equipoise for trial surgical interventions amongst surgeons and a lack of trust in trial data/design

Insufficient clinical equipoise sometimes due to competing interests amongst surgeons was reported as an important challenge to SQA. Historical trials with poor SQA have led some surgeon stakeholders to distrust trial data and the validity of SQA measures within trials.

'So I think that's some of the complexities that we'll come across with the surgical interventions: the equipoise of randomising to something that everyone says, "Well, this doesn't work because we've got this really fancy machine that we bought for £10 million…"' (Quote, Trial Methodologist 71)

'Most of the time it is insufficient. It is not really quality assurance of surgery in trials and sometimes also you try to put in place a strategy it is only on papers, there's no control in most of the trials. Not real control.' (Quote, Surgeon 32)

Insufficient resources and implementation of SQA measures (n=13)

The need for substantial resources to effectively conduct SQA within oncology trials was recognised in addition to the difficulty of providing this within a resource-limited environment.

'So I think it really comes down, as usual, to commitment and resources.' (Quote, Surgeon 24)

Time and cost of staff for SQA

The time for adequately trained staff to monitor surgery emerged as a prominent challenge amongst trial stakeholders.

Regional variability in operative standards, diagnostics, patient factors, concerns over patient access to surgical interventions and insufficient generalizability of trial results (n=9)

Regional differences in perception of operative approach and perioperative care were considered a significant challenge.

'I think it is important to identify that what we look on as a quality approach to staging treatment, operative approach and recovery in Europe and North America may be a very appropriate approach from our standpoint but it does not necessarily match resources, perception or historical experience in other areas.' (Quote, Surgeon 24)

'Yes, regarding oesophagogastric and gastric cancer. But one big problem is standardisation between different continents, different countries.' (Quote, Surgeon 9)

Patient factors

Differences in patient demographics regionally including a higher proportion of patients with late clinical presentations and obesity was presented as one of the problems in conducting multi-centre, international surgical trials.

'So in Europe surgery is more difficult because you have more advanced cancer and obese patients. So the difficulty is different I think. So maybe during some surgical trials morbidity rate may be a little bit different, this would be a challenge.' (Quote, Surgeon 9)

Other emergent sub-themes included:

- *Generalisability of trial data*
- *Patient access to care*
- *Representative surgical interventions and pragmatic trials*

Concern that SQA in oncology trials may impact on the generalisability of trial results became an emergent theme. The potential impact of the cost of SQA and centralisation of services on patient access to care also became an emergent concern. The concept that surgical interventions within trials should be representative to how surgical procedures are performed in routine care was felt to be an important challenge to SQA in oncology trials by trial stakeholders.

4.3.4.2. Strategies to overcome generic challenges to SQA of surgery in oncology trials

National, centralised SQA initiative and the importance of strong trial leadership shared between trial centres

Centralised SQA initiative with funding organisations

Stakeholders highlighted the importance of a national SQA effort with coordination by the royal colleges and in association with several research and funding bodies including NCRI, CRUK and the NIHR.

'Centralised surgical QA group under the NCRI which I think will be very beneficial because the problem I found, one of the limitations I found is that, as a clinical trials unit, you know that Quality Assurance of the treatment is very, very important, but it can be a massive burden in terms of time to ensure it.' (Quote, Oncologist 71)

'I know the Royal College of Surgeons has funded some clinical trials units around the UK, …..I certainly think that group, if it isn't, should be involved in a national initiative sort of set-up for surgical QA.' (Quote, Oncologist 71)

Shared trial leadership between trial centres

The concept of sharing trial leadership or at least making it representative of the multiple involved centres became a prominent theme.

'So I believe that the quality assurance would be better done by a collaborative group that represents the wider group, who have already met and have a consensus, and then are chosen from that group.' (Quote, Surgeon 25)

Strong trial leadership

The idea of having a strong trial leader with experience and influence in order to maintain standards with other trial stakeholders was proposed.

Developing shared goals and surgeon consensus on acceptable SQA measures prior protocol development, facilitating the development of research relationships between trial centres (n=6)

Working together in a transparent way towards shared goals and aiming for 'win-win' solutions became a prominent emergent mitigating strategy

'So I think that's the key way to sell it that actually it's you know win-win for everybody …' (Quote, Oncologist 70)

'And I think as long as you approach it in an open, transparent way and you see it as working together with the sites... it wasn't a problem in radiotherapy, but obviously a lot of it is building up relationships' (Quote, Trial Methodologist 71)

Surgeon consensus on acceptable quality assurance strategies

Initially establishing a consensus on what are acceptable SQA measures within trials according to trial surgeons, being flexible where evidence is lacking. Reinforcing this process through prospective auditing of outcomes in order to generate data required for such a consensus was also highlighted.

'So I think there's something about getting together in a workshop type environment whereby you determine what is mindful to quality assure the surgeons or protocol, and where you can be flexible where there's no evidence it affects outcomes and therefore, be a little bit more open to variations..' (Quote, Oncologist 53)

Adequate design and costing/funding of SQA in trials to facilitate surgeon commitment and SQA (n=7)

Adequate design and costing of SQA within trials to facilitae the SQA process leading to minimum interruption to a surgeons usual practice was considered important.

'…just to engage with funders and convince them about how it's important to fund that type of activity within the trial and ensure that they see it as value for money.' (Quote, Trial Methodologist 64)

'The trial has to be designed and costed to make it easy for the surgeon to commit to doing the standardisation and having it recorded. So, the surgeon in my view needs to do nothing different to the normal surgery, but it is all looked after.' (Quote, Surgeon 15)

Embed SQA early within trial

Multiple stakeholders express the importance of early integration of SQA within the trial design and protocol development process. Centralised review and feedback on the design including the SQA was also suggested to complement this process.

'Your best bet would be to embed these principles very, very early on in the trial design process and that may require engagement with pharmaceutical companies that are doing commercial trials as well as NCRI.' (Quote, Oncologist 57)

Ensuring centre preparedness prior to trial commencement and conducting pre-trial feasibility studies

Centre preparedness

Ensuring the practicalities required in order to commence the study effectively, recording trial data and monitoring are in place prior to trial commencement was an emergent theme.

'I think the study group has to make it as easy as possible for the teams at the centres to be able to do what they're wanting them to do. So, thinks like data sticks and reminders about missing information, and having a database open and waiting at the beginning of the study…' (Quote, Surgeon 23)

Feasibility studies

The importance of pre-trial feasibility studies in order to assess surgical skills prior to centre recruitment was emphasised.

'We should have phase II feasibility study using eh 20 hospitals.' (Quote, Surgeon 1)

'Or before doing feasibility study you should check the surgical skill. After that you can select the hospital.' (Quote, Surgeon 12)

Team work and improving quality of clinical practice through skill development and introduction of new technologies

Team work and engagement

The important of teamwork and trust within a team are underlined by stakeholders.

‘But at the end of the day you have to essentially work with teams and people that you trust. You have trust in their experience and their ability to deliver a quality product.’ (Quote, Surgeon 21)

‘So, it’s about team working and engagement, so it’s about actually understanding that this is part of quality initiative, as much as a research initiative, can actually drive up the care and quality and safety of patients’ (Quote, Oncologist 66)

Improving expertise and routine clinical practice

Particularly through learning from the experience of clinical oncologists, stakeholders explained that oncology trials can be used to improve routine clinical practice and enhance staff expertise.

‘So new technologies should be implemented through trials and trials also provide us with training and the uplift and expertise of the staff. So there's a big incentive to do it.’ (Quote, Oncologist 67)

Trial group learning, regular meetings and focus groups facilitating learning of QA from radiotherapy community/other trial groups (n=7)

Stakeholders highlight that regular trial group meetings allow experiences to be shared and lessons learnt from other centres, development of shared trial ownership and allows review of operative variation. In the first quotation below, one stakeholder explains how one group who have hypothetically overcome a problem with patient recruitment, would be able to explain how this was accomplished to another group within a trial workshop.

‘…And then when you're a group it's okay, we had the same problem with this centre and they have improved, can you explain to them how you did it?’ (Quote, Surgeon 16)

‘Okay, one thing I've not mentioned is ownership of the problem by the people who you're trying to standardise. And so, therefore, having group workshops is really useful, to get the surgeons, who are going to be involved in a trial, together to discuss the variation and how they see that variation.’ (Quote, Surgeon 25)

Facilitating stakeholder conviction towards trial protocol

Multiple stakeholders felt trial meetings could help facilitate stakeholder comittment towards the trial protocol.

‘Bring people on board to try and explain as far as possible this is why we are doing these things. That can help. If people believe in your protocol they’re more likely to comply with it.’ (Quote, Oncologist 61)

Sharing radiotherapy oncology trial experience within trial meeting’s

Some stakeholders felt such meetings could act as a good platform to share experiences from previous radiotherapy trials and for trial stakeholders to share their experience of their learning curve.

‘The only thing that was mentioned in our previous discussion was learning from the radiotherapy trials assurance group and the learning curve that they have gone through.’ (Quote, Oncologist 68)

Quality assure the SQA process within oncology trials (n=6)

Assessing the reliability of the trial monitoring team and monitoring the assessment process was felt to be important amongst trial stakeholders.

‘It's very difficult to find people that do that and you know, at the end of the day you should also assess the quality of the assessor because I mean after the second video, you know, you say well it's okay’ (Quote, Surgeon 22)

‘I mean, whatever you, whatever type you’re doing, you have to validate, is that is it trustworthy, I mean even if you have human beings studying the videos, you should have kind of control of the controllers really….’ (Quote, Surgeon 30)

Appendix F: Strategies not reaching consensus within Delph process for expert trial stakeholders (Chapter 4)

STRATEGIES NOT REACHING CONSENSUS	Delphi Round 1	Delphi Round 2	Delphi Round 3
Robotics, technology and virtual training systems may be useful for:			
• **Credentialing surgeons**	**43**	**55**	**55**
The following factors should be considered in selection of trial centres for participation in surgical oncology trials:			
• **National specialist centre status**	**33**	**65**	**60**
• **An impartial committee is required in order to select surgeons for participation in oncology trials**	**48**	**45**	**40**
Stringency of surgeon selection should be adjusted according to the following factors:			
• **Phase of trial**	**52**	**55**	**50**
• **Disease process being studied (e.g colorectal versus upper GI)**	**57**	**65**	**65**
The optimal method of selecting surgeons for participation in oncology trials includes:			
• **Intraoperative assessment by a visiting surgical team from the trial committee**	**66**	**65**	**60**
The following modalities as adjuncts to an operation manual would be useful to reinforce standardisation in oncology trials:			
• **Virtual reality**	**43**	**45**	**30**
The optimal method for monitoring of surgery in oncology trials includes:			
• **Intra-operative monitoring using video recording and real-time data transfer to Trial Committee**	**43**	**45**	**60**
• **Review of operative notes**	**24**	**30**	**30**
• **Review of case report forms**	**62**	**55**	**65**
• **Real time data monitoring with images/videos sent from operating room to trial monitoring committee in real-time**	**43**	**45**	**25**

Appendix G: Ethics approval for study 'Exploring patient perception of quality of surgery in clinical trials

Imperial College
London

Imperial College Research Ethics Committee
Imperial College London
Room 221
Medical School Building
St Marys Campus
London
W2 1PG
Tel: +44 (0)207 594 9484

researchethicscommittee@imperial.ac.uk

Professor GB Hanna
Queen Elizabeth the Queen Mother Wing (QEQM)
St Mary's Campus

10th December 2018

Dear Professor Hanna

Study Title:. Exploring patient perception of Quality of Surgery in Clinical trials

ICREC reference: 18IC4857

The above study was approved by your Head of Department on 22/10/18 and by the Joint Research Compliance Office on 10/12/18.

Under the Imperial College Research Ethics Committee process, a study that has been reviewed by the Joint Research Compliance Office and Head of Division/Department (or Principal), where no significant ethical issues have been identified in the protocol or ethics application, can be approved without requiring it to go to full committee.

Documents
The documents reviewed were:

- ICREC Application form
- Protocol (v4 08/12/18)
- Participant Information Sheet – Focus Group (v3 07/12/18)
- Participant Information Sheet – Survey (v3 07/12/18)
- Consent Form - Focus Group (v2 29/11/18)
- Consent Form – Survey (v2 29/11/18)
- Focus Group Plan (v2 29/11/18)
- Survey (v2 29/11/18)

Yours sincerely,

Ruth Nicholson,
Research Governance Manager,
Imperial College London

Patient Perspective on Quality of Surgery

Name: …… Date of Birth: ……. Institution in which you received treatment: …….

Thank you for agreeing to participate in this survey studying patient perspective on 'Quality of Surgery'. Please complete all sections of the survey.
It can take up to 20 minutes to complete (though may be longer depending on your answers to the open questions)

Definition of terms:

Upper GI surgeon - surgeon who performs operations for cancer of the stomach or oesophagus

Upper GI centre - Centre in the UK which has surgeons performing operations for cancer of the stomach or oesophagus.

Surgery - In this survey we are referring to the whole process of your surgical care including the care you or other patients receive before, during and/or after surgery.

Trial - Form of experiment in which patients are entered randomly entered into different treatment groups to assess impact on patient outcomes.

Demographics

Date of treatment for oesophageal or gastric disease:

Name of gastric/oesophageal condition you received surgery for (E.g gastric cancer/oesophageal cancer):

Did you receive chemotherapy or radiotherapy before or after surgery?

Were you previously involved in a trial for the treatment of your gastric/oesophageal condition? If you were involved in a trial please describe the type of trial or its name.

Were you seen in clinic by your surgeon or other members of the multi-disciplinary team after the operation?

If so please specify for how long you were followed up (e.g 4 years), how frequently (e.g every 6 months), and by which member of the surgical team (e.g. surgeon/nurse/oncologist)?

Q1. What are your thoughts regarding the quality of surgery?						
Q2. What are your thoughts on quality of surgery in trials?						
Q3. Please indicate whether you agree or disagree that the following factors are important in your perception of quality of surgery:	**Strongly Disagree**	**Disagree**	**Undecided**	**Agree**	**Strongly Agree**	
• Confidence in operating surgeon	☐	☐	☐	☐	☐	
• Care received by the multi-disciplinary team including nurses, specialist nurses, oncologist, radiotherapists, physiotherapists and dieticians.	☐	☐	☐	☐	☐	
• Care received following the operation in Intensive Care Unit	☐	☐	☐	☐	☐	
• Care received following the operation on the ward	☐	☐	☐	☐	☐	

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	Comments
• Follow-up care by surgeon following leaving hospital	☐	☐	☐	☐	☐	
• Follow-up care by other members of multi-disciplinary team including oncologists, nurses, and physiotherapists	☐	☐	☐	☐	☐	
• The patient recovery following the operation (time taken for recovery and how close to ones normal self one can feel following the operation)	☐	☐	☐	☐	☐	
• Quality of life patient experiences following the operation including patient experience of discomfort and other side-effects following the operation.	☐	☐	☐	☐	☐	
Any other comments?						
Q4. Please indicate whether you disagree or agree that the following factors may influence your perception of the quality of surgery performed by a surgeon or a surgical centre:						
• Surgeon annual operative volume (number of cases operated per year)	☐	☐	☐	☐	☐	
• Surgical centre operative volume (number of cases operated per year)	☐	☐	☐	☐	☐	**Comments**
• Size of operative team within surgical centre (number of operating surgeons within operative team)	☐	☐	☐	☐	☐	
• Standardisation of surgery (surgeons within centres performing the same pre-specified key steps of operation)	☐	☐	☐	☐	☐	
Any other comments?						

Q5. Please state your disagreement or agreement as to whether the following factors may contribute towards patient confidence in their operating surgeon:	**Strongly Disagree**	**Disagree**	**Undecided**	**Agree**	**Strongly Agree**	**Comments**
• **Perceived self-confidence** of the surgeon	☐	☐	☐	☐	☐	
• Surgeon's **perceived character**	☐	☐	☐	☐	☐	
• A surgeon's explanation of diagnosis and management plan	☐	☐	☐	☐	☐	
• A **recommendation to see a specific surgeon** by another healthcare professional	☐	☐	☐	☐	☐	
• **A surgeons operative volume** (number of cases they perform per year)	☐	☐	☐	☐	☐	
• The **operative volume of the centre** in which the surgeon works (number of cases performed in that centre per year)	☐	☐	☐	☐	☐	
• **Survival outcome statistics** of that surgeon	☐	☐	☐	☐	☐	
• A surgeons **holistic treatment of the patient** (treating 'the whole patient')	☐	☐	☐	☐	☐	
• A surgeons **treatment of a patient's relative** (e.g. allowing relatives to ask questions or setting expectations)	☐	☐	☐	☐	☐	
Q6. What do you feel are the challenges to quality of surgery?						

Q7. What do you feel are the challenges to quality of surgery in trials?						
Q8. Please indicate your level of disagreement or agreement as to whether you would consider the following factors challenges to quality of surgery:	**Strongly Disagree**	**Disagree**	**Undecided**	**Agree**	**Strongly Agree**	
• **Lack of funding**	☐	☐	☐	☐	☐	
• **Lack of sufficient beds** for operations to be performed (e.g. Intensive Care bed spaces)	☐	☐	☐	☐	☐	
• **Slow investigation of Upper GI cancer** due to slow pace of investigations from GPs in NHS in UK.	☐	☐	☐	☐	☐	
• **Lack of awareness of Upper GI cancer** within other specialties (e.g. radiographers not having wedges for neck support during scans), and non specialist centres	☐	☐	☐	☐	☐	
• **Lack of contact with specific members** of multi-disciplinary team (e.g. specialist nurse/oncologist)	☐	☐	☐	☐	☐	
• **Lack of appropriate clinical cover** when operating surgeon away on holiday	☐	☐	☐	☐	☐	
• **Poor communication of surgical team** with the patient's relative	☐	☐	☐	☐	☐	
• Patient concern and anxiety on morning of surgery due to uncertainty over ITU bed status	☐	☐	☐	☐	☐	

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	
• **Time of diagnostic uncertainty** (e.g time waiting for diagnosis following endoscopy when waiting for histology results)	☐	☐	☐	☐	☐	
• **Limitations of using operative volume** to select surgical centres to partake in trials (e.g. statistics may not represent complexity of cases)	☐	☐	☐	☐	☐	
• **Lack of sufficient imaging** in follow-up appointments following surgery	☐	☐	☐	☐	☐	
• **Long wait for counselling** for patients following surgery and insufficient of mental health support	☐	☐	☐	☐	☐	
Any other comments?						
Q9. How did you feel on the night before surgery? If relevant, please outline any concerns you may have had regarding the surgery itself or the quality of the surgery						
Q10. How would you have felt regarding participating in an oncology trial comparing two different treatments when you were first diagnosed with your condition? (Please note that the assumption in this question would be there is no known difference between to the two known treatments being compared in the trial in terms of outcomes/survival)						

Q11. Please indicate your level of disagreement or agreement on whether the following factors would have contributed to your decision on whether or not to participate in an oncology trial:	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	
• A perception that there may not be genuine clinical uncertainty between the different treatments being offered in the trial.	☐	☐	☐	☐	☐	
• General patient scepticism of clinical trials	☐	☐	☐	☐	☐	
• A perception that there must be a 'gold standard' for all types of surgery	☐	☐	☐	☐	☐	
• A perception that to participate in a clinical trial would be to put one's 'life at risk'	☐	☐	☐	☐	☐	
Any other comments?						

Q12. Which strategies would you recommend to overcome the challenges to quality of surgery?

Q13. Which strategies would you recommend to overcome the challenges to quality of surgery in trials? (If you are unable to comment for this question please move on to the next questions)

Q14. Please indicate your level of disagreement or agreement regarding the following potential mitigating strategies to overcome challenges to quality of surgery:	**Strongly Disagree**	**Disagree**	**Undecided**	**Agree**	**Strongly Agree**	
• **Further education or training for community health services and GPs** in recognising symptoms of Upper GI cancer to aid in early diagnosis	☐	☐	☐	☐	☐	
• **Raising awareness of Upper GI cancer and potential postoperative complications** amongst other health care professionals and non-specialist centres	☐	☐	☐	☐	☐	
• **Non-invasive breath testing to be used as screening tool i**n UK to aid early detection of Upper GI cancer	☐	☐	☐	☐	☐	
• **Patients receiving physiotherapy and cardiovascular training prior to operation to enhance fitness**	☐	☐	☐	☐	☐	
• Provision of **counselling service to patient and relative/next of kin** before and following operation	☐	☐	☐	☐	☐	
• **Provision of personal appointment for patient relative with operating surgeon before** and after surgery in order for them to ask questions/discuss concerns/issues (providing patient consents to this process)	☐	☐	☐	☐	☐	
• **Upper GI centres having to perform a pre-specified number of cases per year (operative volume)**	☐	☐	☐	☐	☐	
• **Use of a structured method of assessing Upper GI surgeons' skill/competencies** on a regular basis	☐	☐	☐	☐	☐	
• **Monitoring of surgery in Upper GI centres** to check standards of operating are being met	☐	☐	☐	☐	☐	

• **Patient contact with anaesthetic team prior to operation** for patients to discuss options and express preferences (e.g. music playing in anaesthetic/operating room)	☐	☐	☐	☐	☐	
• **Spiritual support to be offered to patients and relatives before and after operation**	☐	☐	☐	☐	☐	
Any other comments?						

Appendix I: Detailed qualitative analysis of emergent themes for the patient perspective study (Chapter 5)

Patients' perspective on quality of surgery and quality of surgery in Upper GI cancer trials

Personal experience

The majority of respondents (90%) felt positively regarding the quality of their surgery often referencing their outcome as evidence of the quality of their surgery. Others felt the quality of their surgery was related to the reputation of the centre in which they received their care.

'I feel the quality of my surgery must have been excellent as far as I can tell.'

Quality of surgery in trials

8 (19.5%) participants responded with to this question indicating their generic opinion and positive perception of quality in trials. The importance of trials as a necessary component of research and a primary method of improving treatment were also expressed: 'Trials are a way of continually improving treatment, It has to be correct to encourage people to participate'. Another participant explained that trials are '…the only way really advances will be made'. However, many of participants felt unable to comment adequately to this questions, responding with 'NA' indicating they felt unable to comment for this question (n=15, 36.6%). 11 (26.8%) participants explain they either 'don't know' or have not been involved in a trial previously.

Factors important in patients' perception of quality of surgery

Participants' confidence in their operating surgeon, outpatient follow-up experiences, care experienced by each member of Multidisciplinary Disciplinary

team, and their experience of post-operative ward and ITU care were the most prominent factors shaping patient perception of quality of surgery.

- ***Trust and Confidence in the operating surgeon***

The majority of participants (95.2%) agreed that their 'trust' and/or 'confidence in their surgeon' influenced their perception of quality of surgery, explaining it is 'vital before surgery'. Multiple participants explained this trust developed through reference to the 'thorough explanations' and reassurance they had received from their surgeons pre-operatively: 'He instilled confidence that he would do a good job, very thorough in explaining procedure, and very sympathetic and professional in his approach'. The prominence of the emergent concepts of patient trust and confidence in the operating surgeon within this study was in keeping with surveys of patient perception of safety of surgery in another qualitative study, in which physician-patient interactions, relationships and trust were the most positive factors influencing their perception of the safety environment.[32]

The holistic approach of surgeons to their patients became an important patient perceived characteristic (supported by 78.6% of participants) contributing to patient trust. This was in keeping with a study assessing factors affecting patient perception of quality of care involving 174 surveys in which 'concern/caring of physician (15.7%), and a physician who listens (10.1%)' were amongst the most important factors.[33]

- ***Factors contributing towards patients' confidence in the operating surgeon***

As confidence in the operating surgeon emerged as an important factor contributing to patients' perception of quality of surgery within the focus group, elements contributing to this were separately explored within the survey. These factors fell under two main sub-categories: operative volume and outcomes, and; character traits and communication skills.

- ***Operative volume and outcomes***

Most participants agreed that the following factors contributed to their confidence in the operating surgeon: a surgeons' operative volume (number of cases they perform per year) (78%); the operative volume of the centre in which the surgeon works (number of cases performed in that centre per year) (75.6%), and; the survival outcome statistics of that surgeon (85.4%). Although participants supported patient access to survival and other outcome statistics, they advised caution in its interpretation without sufficient understanding of the complexity of the cases: 'But probably should not without knowledge of difficulty of cases treated.' Although the surgeons' operative volume may significantly contribute to a patient's confidence in their surgeon, it may not be as closely linked to the confidence the surgeon has in their own ability. Operative volume was recognised as a contributing factor to graduating surgeons' low self-confidence withing qualitative studies, however it was not found to be a pivotal factor for some residents expressing concerns regarding their readiness to practice.[146]

- ***Character traits and communication skills***

A surgeons' character traits, treatment approach and communications skills also appeared to play a significant role, with the following factors contributing to patient confidence: perceived self-confidence of the surgeon (95.1%); surgeon's perceived character (92.7%); a surgeon's holistic treatment of the patient (80.6%); a surgeon's explanation of diagnosis and management plan (100%), and; a surgeon's treatment of a patient's relative (e.g. allowing relatives to ask questions or setting expectations). The importance of communication skills and the approach of the surgeon are expressed within these participant comments: 'Their people skills are vital to inspire confidence'. Despite the avowed importance to patients' of a surgeons self-confidence, this character-trait has been reported to be lacking surgical graduates. [146] In a review of fifteen qualitative studies, ten survey studies reported low confidence in general surgery graduates, a phenomenon found to be

mainly attributable to complex social and cultural factors. [146] Confidence in the operating surgeon due to a patient having a 'recommendation to see a specific surgeon by another healthcare professional' was less popular than the other themes, but still gained support of the majority of participants (56.1%).

'My specialist nurse recommended a specific surgeon and I was happy to accept this as I felt she knew what she was talking about.'

- ***Care received by different members of the care team***

Participants universally agreed (100%) pre-operative care provide by the multi-discipliniary team (MDT) and care received following the operation in Intensive Care Unit were important factors contributing in their perception of quality of surgery. The majority of stakeholders similarly agreed that post-operative care provided on the ward (97.6%), by the MDT including oncologists, nurses and physiotherapists (90.2%%), and by the operative surgeon and as an outpatient (97.6%) were important factors. Participants additionally expressed their belief in the importance of a patient feeling they are receiving excellent care: 'The patient needs to feel they are still receiving excellent care.' Although patient opinion regarding the cancer care received by the MDTs are sparsely reported, the association between MDT cancer care and improved outcomes is well recognised forming the preferred model of care for patients with cancer in many countries. [177-179]

- ***Patient recovery and quality of life***

The majority of participants felt their recovery following the operation (time taken for recovery and how 'close to ones normal self' one can feel following the operation) (90.1%), and their post-operative quality of life (including patient experience of discomfort and other side-effects) (87.2%) were important factors in their perception of 'quality of surgery'. In keeping with this and in relation to their personal experiences, some participants complained they 'had no idea of recovery

time or subsequent limitations.' Others affirmed the importance of the patient recovery in their perception of quality of the surgery: 'the way the patient feels obviously affects the way they feel the surgery went.' Patients concern for quality of life with cancer can reportedly often dominate over their concern for survival statistics.[180] Another study assessing patient experience of quality of care in Michigan showed no correlation between quality patient experience and morbidity/mortality rates. [181] However no published study has directly reported patient opinion in this area, and responses to our survey indicated 'surgeon survival outcome statistics' were in fact an important factor contributing to patient confidence in the operating surgeon (supported by 85.4% of participants).

- ***Operative volume and standardisation***

Participants predominately felt that a surgeon's annual operative volume (68.29%), a surgical centre's operative volume (number of cases operated per year) (70.7%), the number of operating surgeons within operative team (68.2%), and standardisation of surgery (surgeons within centres performing the same pre-specified key steps of operation) (53.6%) were important contributing factors to their perception of the quality of surgery. The perception that a surgeons' operative experience gives the patient a sense of confidence and safety was expressed: 'I would feel more confident if the surgeon was very experienced at this type of surgery.' Other particpants' felt it was important for patient's to be aware of operative volume and that they often did not have access to these figures: 'Good to know how many operations are performed yearly, but often we are not able to find out'. The explanation for patient perception of their association between team size and operative quality was that a larger operative team would facilitate higher centre operative volume: 'The amount of surgeons on the team matters as more surgeons allow more operations to be completed.' Participants explained they perceived efficiency of surgery to be associated with standardisation and that this in turn would help to manage the higher operative volume. One participant expressed that standardisation of surgery 'Must be a smoother way of handling a higher operative volume'. Patient's perception within this study concurs with the direction of findings in published

literature regarding operative volume and post-operative outcomes, often considered surrogate markers for 'quality of surgery'. In a systematic review of 32 studies involving 15 surgical interventions there was tentative evidence of a surgeon volume-outcome relationship. This effect was most pronounced with larger correlations identified in colorectal cancer, bariatric surgery, and breast cancer studies.[151]

Patient perception regarding challenges to quality of surgery

Prominent challenges to quality of surgery included: insufficient funding, insufficient communication and clinical contact with certain team members, anxiety over waiting for diagnosis and uncertainty over bed status, and; insufficient investigation and screening for gastro-oesophageal cancer. Likert-scale responses showing proportion of participants agreeing with specific challenges are represented in figure 7 (Chapter 5).

Insufficient funding and resources

The majority of participants felt that a lack of funding (73.2%), insufficient beds for operations to be performed (e.g. Intensive Care bed spaces) (87.8%), and a shortage of imaging in follow-up appointments following surgery (51.3%) were significant challenges to quality of surgery. A general lack of resources and support for staff was cited as a challenge to quality of surgery: 'Lack of support for surgeons and staff. Time and equipment pressures'. A lack of postoperative imaging and long waits for certain imaging modalities had also caused concern: 'I have been told they don't do scans afterwards. I haven't had any'. Another participant felt there was a shortage of surgeons with sufficient experience operating on patients who had previously received neo-adjuvant chemoradiotherpy: 'Unfamiliarity of some surgeons with patients who have received pre-op chemo, and chemo-radiotherapy.' Patient opinion is in keeping with a recent interview study of NHS directors in which under-investment within the NHS has been recognised to reduce the quality of surgical and

imaging services provided according to a qualitative study in which NHS directors were interviewed. [131]

Insufficient awareness, screening and investigation of Upper GI Cancer

Most participants also expressed a concern regarding insufficient screening for Upper GI cancer in the UK (90.3%), slow investigation of Upper GI cancer due to a slow pace of investigations from primary care in UK (66.0%), and a lack of awareness of Upper GI cancer within other specialties (e.g. radiographers not having wedges for neck support during scans) and non specialist centres (65.9%). Concerns regarding the risk of disease progression due to late diagnosis secondary to a lack of screening was expressed: 'Until patients have symptoms no preventative screening appears to be done and by then the disease is, sadly, often well advanced.' The importance of screening in detecting oesophagogastric cancer at an earlier stage and the consequent impact on survival has long been recognised and is reinforced within published literature.[147] Indeed the higher 5-year survival rates from gastric cancer reported in Asia (69%) compared to those in the Western World (10-30%), are thought to be attributable to the widespread availability of screening programmes in Asia.[147] Even once the patient has developed symptoms, slow investigation can often follow, which some worry may adversely impact survival: 'When I first felt ill and couldn't swallow the first doctor put me on a 10 week waiting list for an endoscopy. A few days later I saw another doctor in the surgery who got me in for one the following week. I don't think I would be here now if I had listened to the first doctor.' With concern other specialities have insufficient awareness of gastro-oesophageal cancer some participants described uncomfortable situations: 'I had to have a barium swallow a few days post op and was terrified I was going to choke as the radiographer insisted I lie flat'. Worse survival from various cancers in the United Kingdom, in comparison with other that found in European countries, has previously been attributed to more advanced cancer stage at presentation. [182] In a review of patient-mediated and practitioner-mediated risk factors for delayed presentation or referral of symptomatic cancer, lower socio-economic status was found to be associated with delay in diagnosis of upper gastrointestinal. [182] Similar to concerns

reported by participants in this chapter, ‘misdiagnosis' occurring through treating patients symptomatically was previously identified as a practitioner-mediated factor contributing to delayed diagnosis including. Other factors contributing to delayed diagnosis included: inadequate patient examination; inappropriate investigations, and; failure to follow-up on inconclusive results. [182]

Insufficient communication or clinical contact

Other factors considered challenges to quality of surgery by the majority of stakeholders included: Poor communication of surgical team with the patient's relative (51.2%); a lack of contact with specific members of multi-disciplinary team (e.g. specialist nurse/oncologist) (56.1%); a lack of appropriate clinical cover when operating surgeon away on holiday (56.1%), and; a long wait for counselling for patients following surgery and insufficient mental health support (51.2%). One being unable to contact their specialist oncology nurse one participant describes their frustration: ‘Specialist Nurses tend to work Monday to Friday. No cover at weekends. Had to go to casualty department where staff don't always have the knowledge.’ Insufficient awareness of the emotional impact of the disease in conjunction with a long wait for counselling was distressing for some participants: ‘The emotional impact of cancer doesn't seem to be recognised and counselling wasn't offered. I eventually got it 3 years after the op but would have helped much sooner.’ Similar to findings in this chapter, there is a recognition within the oncology community, that currently patient-centered communication and shared decision making are suboptimal and need to be improved.[183] There is also growing recognition of the increased prevalence of mental health problems including depression and anxiety amongst patients with cancer and the potential negative impact of this on their treatment, recovery quality of life and survival.[184] In keeping with our study findings within this chapter, it has also been suggested that the heightened anxiety observed in post-operative patients may be due to reduced clinical consultations and support following treatment. [184]

Anxiety due to waiting for diagnosis and uncertainty over bed status

The majority of participants felt that patient concern and anxiety on morning of surgery due to uncertainty over ITU bed status (90%), and the time of diagnostic uncertainty (e.g time waiting for diagnosis following endoscopy when waiting for histology results) (66%) were significant challenges to the quality of surgery. Expressing their upset with 'waiting' one participant explains: 'The worst part between that and treatment is the frustration of waiting. Perhaps there would be less need for mental health support if people didn't get so upset because of delays.' The stress of the uncertainty regarding the ITU bed status was also expressed: 'Very stressful not knowing if the operation would go ahead because of lack of ITU bed'. The patient dissatisfaction due to waiting for surgery is in keeping with that reported in a qualitative study of patients undergoing elective surgery in which those who had their operations later in the day were significantly less likely to report a positive experience.[185]

Other challenges to quality of surgery

Operative Variation and patient factors

Variation in patient factors such as their co-morbidities, underlying body mass index (BMI) and fitness were described as an important challenge, in some cases influencing the ease of an operation. The relative ease of operating on a less heavy individual was expressed: 'Fitness of patient. Weight of patient. My surgeon said more than once how much easier it was performing an operation on a thin person like me. Other health issues. Age.' The challenge of increased patient BMI and associated impact on surrogate measures for 'surgical quality' following gastric cancer surgery has been reported in published literature, with one retrospective study finding obesity to be associated with long operation time, increased blood loss, and slower recovery following laparoscopic gastric resection.[186]

Maintaining surgical performance and statistics

The difficulty in surgeons' maintaining their operative performance to the highest level and reflecting this in their patients' survival statistics was considered a key challenge. On being asked regarding their opinion of challenges to quality of surgery one participant answered: 'Always performing to the very highest standard. It is someone's life that is being dealt with.' The difficulty of maintaining adequate operative performance is a recognised phenomenon. In one review of surgical performance and misconduct, it is noted that a surgeon 'may be a perfect gentleman in the theatre suite but have an unacceptably high operative complication rate.'[187]

Effective control of post-operative symptoms

Following an oesophagogastric oncological resection, the adequate control of post-operative symptoms can become a key challenge to maintaining the patient's quality of life. One participant explained a significant challenge to quality of surgery is to:

'Successfully control post op and long term acid reflux'. Often surgical outcomes are conveyed soley in terms of oncological outcomes such as recurrence and survival. The importance of symptom control including pain and dysphagia following gastric cancer surgery were reinforced by a narrative review of 'quality of life' in gastric cancer patients. Other reported important factors to be considered influencing post-operative quality of life included: emotional well-being, social and financial status, and body image.[188]

Patient perception regarding challenges to quality of surgery in Upper GI cancer trials

Key emergent challenges to quality of surgery in Upper GI cancer trials included: Over stringent standardisation; concern over potential for one trial arm to be inferior; training of trial surgeons; negative impact of publishing surgical outcomes/statistics, and; limitations of utilising operative volume in centre selection for trials.

Over stringent standardisation

The concept that standardising surgery may not allow for operative variation required for each individual or hospital was an emergent theme. Participants cautioned that surgeons may be 'enforced to follow a particular method of surgery as opposed to one which is bespoke to individuals needs'. In keeping with this concern another explained: 'Every patient is different. Every hospital is different'. This thematic challenge is in keeping with the philosophy of certain trialists who advocate for pragmatic trials. In one review of trial typology the author argues that for pragmatic trials, attempts to standardise surgery often create difficulties, and ensuring that each step was delivered as planned would be unrealistic.[128]

Concern over potential for one trial arm to be inferior

Participants had concerns one arm of the trial may be inferior with regards to quality of surgery: 'Inclusion in inferior arm of trial'. Similar patient concerns regarding the potential for psychological distress due to randomisation to a trial arm with inferior outcomes has been previously reported within oncology trials.[148]

Training and learning of trial surgeons

The potential consequences to a surgeon not having the adequate level of skill or training was a prominent emergent challenge to quality of surgery in trials. Commenting on their perceived challenges to quality of surgery in trials one participant explained: 'The unknown and the ability and skill of the surgeon to overcome problems.' This concept is also well known within published literature with a surgeons' learning process, particularly for novel trial techniques, thought to confound comparison of outcomes between trial arms within certain surgical randomised control trials. [189]

Negative impact of publishing surgical outcomes/statistics

Many participants were concerned that surgeons may be reluctant to participate in oncology trials due to this potentially adversely affecting their published survival statistics. Accordingly, one participant explained 'Surgeons may not wish to enter trials if it would potentially adversely affect their survival stats.' Indeed surgeons have objected to the publication of their outcomes and survival statistics, with some surgeons snubbing the idea claiming it would disincentivise them from operating on high-risk cases.[190] Some cardiothoracic surgeons felt so offended by publishing named surgeons' outcome that they offered their resignation when this was implemented.[191]

Limitations of utilising operative volume in centre selection for trials

Overall, most participants were either uncertain or disagreed (70.7%) that 'limitations of using operative volume to select surgical centres to partake in trials' (e.g. statistics may not represent complexity of cases) is a challenge to quality of surgery in Upper GI cancer trials. It was interesting participants did not feel this was a challenge, as within published literature the creation of a highly controlled, selected and standardised environment has long been considered a limitation to the external validity and generalisability of the results of surgical RCTs.[133]

Strategies to overcome the challenges to quality of surgery

Strategies to overcome identified challenges to quality of surgery included: Further education/training for non-specialist centres in gastro-oesophageal cancer; screening program introduction; physical, psychological and spiritual preparation and support; credentialing of centres; monitoring of surgery; centralisation of services; improved funding, and; improved availability of trial arm outcome data. Likert-scale responses showing proportion of participants agreeing with specific strategies are represented below in figure 8 (Chapter 5).

Education and training

The majority of participants (94.1%) supported the strategies of further education or training for community health services and GPs in recognising symptoms of Upper GI cancer to aid in early diagnosis. The same majority (94.1%) agreed with raising awareness of Upper GI cancer and potential postoperative complications amongst other health care professionals and non-specialist centres. Many participants expressed concern regarding GPs prescription of indigestion medication rather than further investigations of symptoms of upper gastro-intestinal cancer as a justification for this strategy: 'Definitely. I had no typical symptoms. Too many GP's prescribe indigestion medications, sometimes on repeat prescriptions, but do not do a follow up.' In keeping with participants' proposed strategy of improved education within primary care in our study, one review of diagnosis and treatment of gastrointestinal disease in primary care advised updating knowledge and skills of primary care physicians via continuing medical education is the only way to improve adherence with standards and quality of care for patients.[149] Enhancing the experience, training and competence of the surgeon themselves, also became an emergent strategy to overcome challenges to quality of surgery. One participant explained how the quality of surgery performed may improve with training: 'Clearly, the quality of surgery (or, at least, the outcome of surgery) is enhanced by the experience, understanding, training and competence of the surgeon. If those factors can be enhanced then quality may be improved.' The importance of training, experience and a surgeon's

learning curve and their association with patient outcomes have been extensively discussed within published literature.[150,151]

Improved investigations and screening programmes

Introducing a screening program in the UK for Upper GI cancer to aid early detection of disease (80.5%) and non-invasive breath testing to facilitate early detection of Upper GI cancer (75.4%) were strategies supported by most participants to overcome challenges to quality of surgery. The importance of improved investigation and screening in effort to prevent disease progression was emphasised: 'This is badly needed as the disease can be quite advanced once diagnosed once symptoms present themselves'. Participants explained they had been impressed by the non-invasive breath test they had witnessed: 'I was in a working group on the early development of these and very impressed.' Advances in technology are bringing improved screening and early diagnosis to the forefront within the fiscally stretched UK healthcare system with one app piloted in Manchester offering smokers and ex-smokers CT scans in supermarket car parks quadrupled early diagnosis rates of lung cancer and has since been rolled out over North-Manchester.[152] Non-invasive breath testing for gastro-oesophageal cancer, currently being developed and piloted in primary care, may help to provide enhanced levels of early cancer detection.[153]

Physical, psychological and spiritual preparation and support

Maximising one's pre-operative fitness with patients receiving physiotherapy and cardiovascular training prior to the operation to enhance fitness (63.4%), and improving psychological support with provision of a counselling service to the patient and relative/next of kin before and following operation (70.7%) gained support of the majority of participants. A patients ability to undertake their own personal research into the planned procedure, having the courage to ask questions and being proactive in enhancing their pre-operative fitness were considered important strategies: 'Do your homework and research, ask loads of questions as well as the difficult ones about experience and success rates. Get yourself a fit as you

can before surgery.' The perceived link between pre-operative fitness and improved post-operative outcomes was also reinforced: 'The fitter the patient the more likely it is they have a better outcome and shorter recovery time'. Although there is known relationship between physical fitness and health across almost all clinical contexts, larger prospective studies are required to evaluate the impact of such improvements on disease-specific surgical outcomes.[192] The need to raise awareness of the life-changing impact of the oeosophagogastric surgery pre-operatively and the potential impacts of this on their psychological state was also emphasised: 'More attention to the mental side following the operation, and more specific information is needed in advance about 'life-changes'. The need for counselling to mitigate this was emphasised by one participant expressing their opinion regarding peri-operative counselling: 'Very important, such a strange and devastating effect on the body.' This was reinforced within one review which expanded on the importance of psychological support and counselling to aid in all aspects of patient care from communicating the diagnosis of oesophageal cancer to the patient, to coping with all aspects of peri-operative management including chemo-radiotherapy. Overall the majority disagreed (61%) that spiritual support offered to patients peri-operatively was an important strategy to overcome challenges to quality of surgery. Although many did not feel they needed spiritual support peri-operatively, others felt it should be avaliable at the discretion of each patients' choice: 'Should be down to the individual, I spoke with the hospitals religious team'.

Communication between the surgical team, patients' and their relatives

The majority of participants (63.5%) supported the strategy of providing a personal appointment for the patients' relative with the operating surgeon peri-operatively, allowing them to ask questions or discuss concerns/issues. One participant explains the importance of communication with the surgeon: 'It would be helpful for the patients relative to be fully aware of complications or issues arising from the surgery. It would lessen their stress.' Although little is published regarding communication between the surgical team with patient's relatives, within one qualitative survey study of 48 patients following pancreatic cancer resection the importance of

emotional support from the operating surgeon is emphasised.[193] In addition, long term strategies are proposed to improve communication between the patient and the surgical team post-operatively. [193] Just under half of participants agreed with the strategy of having patient contact with anaesthetic team prior to an operation for patients to discuss options and express preferences (e.g. music playing in anaesthetic/operating room) (48.8%). Although not popular with the majority, one participant explained how helpful they found pre-operative contact with the anaesthetic team: 'I was given an out patient appointment with the anaesthetist pre op and she spent quite a time explaining to me what would happen. She was lovely and I found it so helpful.' Some participants felt strongly regarding the benefit of communication between post-operative and pre-operative patients. They felt that a meeting between the two cohorts would allow those who have experienced the surgery to reassure those still awaiting surgery and to answer their questions: 'Allow post op patients to help reassure pre op patients as I did on our ward'.

Improved funding and resources

The need for improved funding streams to facilitate training and adequate staffing are important to help overcome challenges to quality of surgery according to some stakeholders: 'NHS needs more money for training and operating hospital. Adequate staffing would go a long way'. Some participants felt that improved information technology (IT) with appropriate support would allow staff to be more time efficient: 'Invest in more efficient I.T. and administrative support, so that time is not wasted by clinical staff in resolving administrative matters.' Recognised as a 'make or break issue' for quality of patient care within the NHS, an estimated £900 million per year is required by 2024 within the budget of Health Education England in order to ensure adequate staffing levels are maintained.[194]

Centralisation of services

Participants reinforced the importance of having specialised units with expertise in this condition to overcome challenges to quality of surgery. When asked regarding

strategies to overcome challenges one participant advocated: ‘Specialist Units with specific expertise.’ Other participants felt treatment in "centre of excellence" could improve the quality of surgery. These sentiments were similar to those shared by participants in a questionnaire-based study of 186 specialist cancer care recipients in London and Manchester, who identified ‘highly trained staff’ with ‘specialised skills in cancer’ as important factors related to centralisation of specialist surgical services.[154]

Credentialing of centres

Credentialing Upper GI centres by specifying that they should perform a certain number of cases per year (operative volume) was supported by just less than half (49.2%) of participants to overcome challenges to quality of. Limitations of using case volume alone for credentialing surgeons or centres for participation in trials has been previously recognised and noted as one of the key challenges to quality assurance of surgery in trials.[34] However, given the lack of reliable assessment tools for assessing quality of oesophagogastric cancer resections until recently, operative volume has been the most commonly employed method to credential trial centres and surgeons even in those surgical trials with the strictest SQA measures. [129]

Monitoring of surgery

The majority of participants supported monitoring of surgery with a structured method of assessing Upper GI surgeons’ skill/competencies (73.1%), and/or monitoring of surgery to check standards of operating are being met (78.0%). Although the majority agreed on the importance of operative monitoring one participant cautioned that such monitoring may affect outcomes: ‘Successes rates could be affected if being monitored for good results.’ Patient approval of this strategy to improve monitoring of quality of surgery using a structured method in Upper GI trials concurs with one of the expert FOSQAT consensus (Chapter 8). Although historically trials have been lacking in methods of objectively monitoring

surgical quality, [34] a reliable assessment tool for assessing the process and product of the 2-stage eosophagectomy has now been developed. [24]

Centralisation of oncology surgical services

Participants felt that including centres in which services were centralised within Upper GI trials allows inclusion of 'Areas of Excellence', which in turn can improve quality of surgery. There may be an element of recall bias in this however as all participants are survivors of surgery in UK hospitals which are already centralised for oesophagogastric cancer surgery. A qualitative study revealed positive patient perceptions of centralisation of specialist cancer surgical services with important factors behind this including: highly trained staff; waiting time for cancer surgery; and access to staff members from various disciplines with specialised skills in cancer.[154]

Improved funding

Improving funding and medical leadership were considered important strategies to improve quality of surgery: 'Sufficient funding must be provided and all treatment must be led by medical professionals rather than politicians and administrators'. The importance of improved funding for driving advancements in surgery and within upper GI trials is well recognised. In keeping with this, a recent investment of £20 million in 2014 was allocated between 22 projects, ranging from analysing how robotics can be used to improve surgical outcomes to looking at the long term benefits of gastric band surgery.[195]

Improved availability of trial arm outcome data

Improved availability to patients of statistics and clinical outcomes justifying the need for a clinical trial was proposed. Some felt this would help to recruit future patients to participate in trials: 'Patients are anxious about life and death so trials need to be sold very convincingly. Statistics about comparable outcomes would help

but in the nature of trials may not be very definitive.' However, nine participants stated they were unable to comment on this question; of which two participants cited their reason as having insufficient experience of clinical trials.

Appendix J: Ethics letter of approval for ADDICT study (ICREC: 18IC4247)

Imperial College London

Imperial College Research Ethics Committee
Imperial College London
Room 221
Medical School Building
St Marys Campus
London
W2 1PG
Tel: +44 (0)207 594 1872

researchethicscommittee@imperial.ac.uk

8th January 2019

Dear Professor GB Hanna

Study Title:. Challenges to recruitment of patients and monitoring of surgery in the ADDICT trial and development of consensus mitigating strategies

ICREC reference: 18IC4947

The above study was approved by your Head of Department on 15/03/18and by the Joint Research Compliance Office on 08/01/19

Under the Imperial College Research Ethics Committee process, a study that has been reviewed by the Joint Research Compliance Office and Head of Division/Department (or Principal), where no significant ethical issues have been identified in the protocol or ethics application, can be approved without requiring it to go to full committee.

Documents
The documents reviewed were:

- ICREC Application form
- Addict agreement Professor Young Woo
- Focus Plan (v1 23/02/18)
- Consent Form – Focus Group (v3 23/12/18)
- Consent Form – Survey (v3 13/12/18)
- Consent Form – Workshop (v3 13/12/18)
- Consent Form – Workshop S1 (v3 13/12/18)
- Participant Information Sheet – Focus Group (v4 07/01/19)
- Participant Information Sheet – Stakeholder opinion (v4 07/01/19)
- Participant Information Sheet – Workshop (v4 07/01/19)
- Participant Information Sheet – Workshop S1 (v4 07/01/19)
- Protocol (v1 12/12/18)
- S1 Questionnaire (v2 23/12/18)
- S2 Questionnaire (v2 23/12/18)
- Workshop Plan (v1 23/02/18)

Yours sincerely,

Ruth Nicholson,
Research Governance Manager,
Imperial College London

Appendix K: Detailed qualitative analysis of 'Phase 1 Presentation of results of expert stakeholder interview study to ADDICT' (Chapter 6)

Key emergent themes within perceived challenges to monitoring of surgery in ADDICT included: (i) Limitations of photographic monitoring, (ii) difficulty in training of trial teams to standardise monitoring process, and (iii) poor trial stakeholder engagement with trial education website. The quality of photographs taken was often not felt to be satisfactory and stakeholders reflected on the limitation of photographs for determining the quality of surgery: 'The quality of the images used to monitor quality were felt to be sub-optimal.' As a related sub-theme, stakeholders attributed difficulty in photographic monitoring to a lack of training in trial image capture: 'The method of picture taking is not educated well.' Another participant expressed uncertainty that trial surgeons adhere to standardised monitoring methods. Lack of stakeholder utilisation of the monitoring instructions on the trial website was also mentioned as a barrier to standardised monitoring. Emergent strategies to improve monitoring of surgery in ADDICT included: (i) Standardisation of monitoring process; (ii) Reminding ADDICT trial stakeholders regarding standardisation process and the available online resources; (iii) Utilising video monitoring; (iv) Increased surgeon collaboration and initiation of clinical trials, and; (v) Improving funding available for surgical trials. The importance of educating ADDICT trial investigators on how to monitor surgery including how to take operative photographs in a standardised manner is highlighted. In efforts to standardise the monitoring process in ADDICT several strategies have been implemented including on-line demonstrations of operative photographs: 'Yes, we downloaded a standardised operative picture to the trial website.' Newsletters and emails have also been introduced in efforts to remind ADDICT surgeons regarding the standardisation of monitoring within ADDICT. One stakeholder explained video monitoring facilitates assessment of the quality of the surgery relative to operative photographs. Increased surgeon collaboration is advised in effort to attract more support from the government for surgical trials and surgical quality assurance. One ADDICT stakeholder explained the societal paradigm shift required for this would

require increased effort on behalf of the surgical community: 'So our surgeons need to try harder to get more attention, more interest from a society and that in a sense I think surgeons could collaborate more, to do important clinical trials.' The importance of increased funding for surgical trials was also emphasised in order to improve the supporting infrastructure available to facilitate surgical oncology trials.

Emergent challenges to recruitment in ADDICT included: (i) changing epidemiology of gastric cancer; (ii) trial surgeons overburdened with clinical work; (iii) concern patients may choose different hospitals on trial proposal; (iv) personal issues of some surgeons not wishing to collaborate; (v) lack of funding; (vi) insufficient support for surgical trials/quality improvement; (vii) lack of clinician clinical equipoise for the ADDICT trial, and; (viii) reluctance to expand recruitment of centres due to concerns over operative resections. A reduction in overall gastric cancer incidence and a relative increase in detection of early gastric cancer was attributed as one of the challenges to recruitment, as described by one stakeholder: 'The first challenge is changing incidence of gastric cancer in Korea. The incidence is actually decreasing, a little bit decreasing from 3-4 years ago.' The heavy volume of work Korean surgeons have to do in their daily practice was cited as a reason surgeons may loose interest in the trial. This work burden is further compounded by the challenges of explaining the trial to patients within a busy clinic. The perception some clinicians in smaller hospitals may fear losing their patients to other hospitals if they explain the ADDICT trial to their patients was prominent. The issue of some clinicians not wanting to help others with their research due to the perception that they are from a larger tertiary hospital was also introduced: 'Also probably some have an ego problem, you are also a big hospital, so 'I don't want to help others work maybe.' The limitation of the current Korean fiscal situation with insufficient funding to support investigators with research nurses was also an emergent theme. A concern that support for quality assurance of surgery is not evident within the government agenda was expressed. Some surgeons, due to their experience of patients having advanced disease following surgery, have a lack of clinical equipoise for this trial, believing that D2 surgery must be done for their patients: 'Some surgeons are very reluctant to participate in this trial because he strongly believe

you know that D2 must be done.' Concerns also existed regarding the operative standards of D2 surgery in Europe leading to a reluctance to expand the trial to European centres. Key emergent strategies improve recruitment of patients in ADDICT include: (i) Expanding recruitment to other centres; (ii) Limiting recruitment of trial centres within pre-specified range of trial monitoring teams, and; (iii) Regular seminars and newsletters to maintain interest of trial investigators. Similar to the strategies advocated to improve monitoring of surgery, ADDICT stakeholders advised regularly reminding investigators of the importance of recruiting patients to the ADDICT trial. Reminding trial stakeholders through regular seminars and newsletters was proposed in order to maintain their interest and enhance recruitment. One stakeholder explains how, in endeavour to enhance recruitment, multiple centres throughout Korea were invited to participate in ADDICT. Following from this, efforts to globalise the trial were initiated including inviting centres in China and Japan. The ADDICT stakeholder further explained that recruitment was not extended to Europe and the US due to concerns over surgical quality control. It was explained that there were concerns credentialing threshold requirements for surgeons to have performed at least 50 gastrectomies prior to trial entry would not be met in European centres due to lower operative volumes: 'So technically because, for the surgical quality control, the conditions, the pre-requisite condition of the surgeons is that at least 50 cases of gastrectomy.' The importance of expanding recruitment within a manageable geographical radius of Korea, allowing adequate surgical monitoring within their limited trial funding budget was reinforced.

Challenges to recruitment and monitoring of surgery in the ADDICT trial and mitigating strategies **Please answer the following open and closed questions. For the closed questions please indicate your level of agreement or disagreement with the items suggested, and please write the reasons for your choices where possible.** **Please note all answers to this questionnaire will remain confidential and anonymised**
Demographics
Name:
Date of birth:
Role as expert stakeholder within ADDICT trial (e.g. Oncologist, Surgeon, Trial Methodologist/statistician, Trial Manager):
CHALLENGES TO RECRUITMENT OF PATIENTS TO THE ADDICT TRIAL
Q1. What is your opinion regarding recruitment of patients for the ADDICT trial?
Q2. What do you feel are the challenges to recruitment of patients in the ADDICT trial?

Q3. Please rate the following factors in terms their contribution to reduced recruitment of patients in the ADDICT trial:	**Strongly Disagree**	**Disagree**	**Undecided**	**Agree**	**Strongly Agree**	**Comments**
• Q3.1. A lack of clinical equipoise (surgeon belief that there is genuine clinical uncertainty between the two trial arms) amongst trial surgeons	☐	☐	☐	☐	☐	
• 3.2. Insufficient research funds to support investigators including insufficient funds to hire more research nurses	☐	☐	☐	☐	☐	
• Q3.3. Patients choosing not to partake in the study for personal reasons/beliefs	☐	☐	☐	☐	☐	
• Q3.4. Over-stringent selection of surgical centres	☐	☐	☐	☐	☐	
• Q3.5. Over-stringent standardisation of surgery	☐	☐	☐	☐	☐	
• Q3.6. Reduced interest amongst trial surgeons	☐	☐	☐	☐	☐	
• Q3.7. Burden of routine clinical work on trial surgeons restricting their participation with trial	☐	☐	☐	☐	☐	
• Q3.8. Concern amongst trial surgeons that explaining the trial may cause patients to seek treatment in another centre.	☐	☐	☐	☐	☐	
• Q3.9. Personal factors meaning some trial surgeons may prefer not to help with other colleagues' research	☐	☐	☐	☐	☐	
• Q3.10. Reducing incidence and change in clinical stage of presentation of gastric cancer	☐	☐	☐	☐	☐	

Q3.11. Any other comments?						
Q4. Please indicate your disagreement or agreement as to whether the following factors lead to a reluctance to expand recruitment including more centres internationally:	**Strongly Disagree**	**Disagree**	**Undecided**	**Agree**	**Strongly Agree**	**Comments**
• Q4.1. Difference in surgeons' conception and skill to perform D2 gastrectomy internationally	☐	☐	☐	☐	☐	
• Q4.2. Difficulty in assessing quality of surgery across centres internationally in selecting new centres for participation. (2)	☐	☐	☐	☐	☐	
Q4.3. Any other comments?						
POTENTIAL STRATEGIES TO OVERCOME CHALLENGES TO RECRUITMENT IN THE ADDICT TRIAL						
Q5. What strategies would you propose to try to overcome challenges to recruitment of surgeons in the ADDICT trial?						
Q6. Please state your disagreement or agreement regarding the following potential strategies to overcome challenges to recruitment:	**Strongly Disagree**	**Disagree**	**Undecided**	**Agree**	**Strongly Agree**	**Comments**
• Q6.1. Expand number of centres/surgeons included in trial internationally within selection criteria (participating surgeons having previously performed 50 gastrectomies)	☐	☐	☐	☐	☐	

• Q6.2. Increase the number of participating centres within the geographically accessible area for the surgical quality monitoring teams	☐	☐	☐	☐	☐	
• Q6.3. Trial centres and surgeons increasing collaboration and working towards shared goals	☐	☐	☐	☐	☐	
• Q6.4. Structured, regular and frequent communication between trial centres and trial committee with appropriately senior feedback mechanism and a positive reward system.	☐	☐	☐	☐	☐	
• Q6.5. Flexible standardisation of surgical procedures focusing on those aspects affecting safety or survival	☐	☐	☐	☐	☐	
Q6.6. Any other comments?						
Q7. Please indicate your disagreement or agreement as to whether adjusting centre selection according to the following factors may improve recruitment in the ADDICT trial:	**Strongly Disagree**	**Disagree**	**Undecided**	**Agree**	**Strongly Agree**	**Comments**
• Q7.1. Operation complexity	☐	☐	☐	☐	☐	
• Q7.2. Disease prevalence	☐	☐	☐	☐	☐	
• Q7.3. Hazard ratio (e.g. for post-operative complications or 30-day mortality) between centres	☐	☐	☐	☐	☐	
Q7.4. Any other comments?						

Q8. Please indicate your level of disagreement or agreement as to whether the following strategies would be useful in maintaining the interest of trial investigators to improve recruitment:	**Strongly Disagree**	**Disagree**	**Undecided**	**Agree**	**Strongly Agree**	**Comments**
• Q8.1. Conduct frequent investigator seminars	☐	☐	☐	☐	☐	
• Q8.2. Send regular updates on newsletters	☐	☐	☐	☐	☐	
Q8.3. Any other comments?						

CHALLENGES TO MONITORING OF SURGERY IN THE ADDICT TRIAL

Q9. What do you feel are the challenges to monitoring of surgery in the ADDICT trial?

Q10. Please indicate your level of disagreement or agreement regarding the importance of the following challenges to monitoring of surgery in the ADDICT trial include:	**Strongly Disagree**	**Disagree**	**Undecided**	**Agree**	**Strongly Agree**	**Comments**
Q10.1. Quality of some videos is insufficient to assess quality of surgical procedure	☐	☐	☐	☐	☐	

Q10.2. Quality of some photographic series is insufficient to assess quality of surgical procedure	☐	☐	☐	☐	☐	
Q10.3. Trial teams are unfamiliar with method of taking and submitting standardised videos/photographs	☐	☐	☐	☐	☐	
Q10.4. Feasibility and practical problems in recording trial videos to monitor quality of surgery	☐	☐	☐	☐	☐	
Q10.5. Feasibility and practical problems in taking photographic series to monitor quality of surgery	☐	☐	☐	☐	☐	
Q10.6. Surgeons pride leading to difficulty in providing and receiving feedback	☐	☐	☐	☐	☐	
Q10.7. Regional culture amongst surgeons leading to difficulty in providing and receiving feedback	☐	☐	☐	☐	☐	
Q10.8. Insufficient resources for monitoring of surgery	☐	☐	☐	☐	☐	
Q10.9. Surgeons' reluctance to be monitored	☐	☐	☐	☐	☐	
Q10.10. Surgeons' concerns over adverse outcomes associated with monitoring of surgery	☐	☐	☐	☐	☐	
Q10.11. Surgeons' concerns over potential litigation associated with monitoring of surgery	☐	☐	☐	☐	☐	
Q10.12. Potential negative impact of monitoring of surgery on generalizability of results	☐	☐	☐	☐	☐	

Q10.13. Potential negative impact of monitoring of surgery on surgeon recruitment	☐	☐	☐	☐	☐	
Q10.14. Any other comments?						
POTENTIAL STRATEGIES TO OVERCOME CHALLENGES TO MONITORING OF SURGERY IN THE ADDICT TRIAL						
Q11. What strategies would you suggest to overcome challenges to monitoring of surgery in the ADDICT trial?						
Q12. Please indicate your level of disagreement or agreement regarding the following potential strategies to overcome challenges to monitoring of surgery in ADDICT:	**Strongly Disagree**	**Disagree**	**Undecided**	**Agree**	**Strongly Agree**	**Comments**
Q12.1. Trial teams utilising standardised method of recording operative videos/photographs as per instructions on the ADDICT trial website	☐	☐	☐	☐	☐	
Q12.2. Newsletters and messages to trial stakeholders reminding them of standardised methods for intra-operative monitoring	☐	☐	☐	☐	☐	
Q12.3. Monitoring of surgery (photographs and/or videos) through external peer review within the ADDICT trial to reinforce adherence to protocol	☐	☐	☐	☐	☐	

Q12.4. The process for monitoring of surgery within the ADDICT trial should be standardised	☐	☐	☐	☐	☐	
Q12.5. Monitoring of surgical quality should consist of a graded monitoring system, adjusting stringency of monitoring according to performance and monitoring of initial cases	☐	☐	☐	☐	☐	
Q12.6. Monitoring of surgery should be kept anonymous and confidential	☐	☐	☐	☐	☐	
Q12.7. There should be a structured, regular and appropriately senior feedback mechanism with positive reward system for monitoring of surgery	☐	☐	☐	☐	☐	
Q12.8. Improved system of funding investigator initiated surgical trials with adequate funding for quality assurance of surgical procedures	☐	☐	☐	☐	☐	
Q12.9. Technology and artificial intelligence assistance in monitoring surgical quality	☐	☐	☐	☐	☐	
Q12.10. Any other comments						
Q13. Please indicate your disagreement or agreement regarding your preference for the optimal method for monitoring of surgery within ADDICT:	**Strongly Disagree**	**Disagree**	**Undecided**	**Agree**	**Strongly Agree**	**Comment**
Q13.1. Use of a structured objective assessment tool involving review of operative photographs, review of unedited videos or intra-operative monitoring by a visiting surgical team	☐	☐	☐	☐	☐	

Q13.2. Review of unedited operative videos	☐	☐	☐	☐	☐	
Q13.3. Review of random selection of recorded unedited trial videos	☐	☐	☐	☐	☐	
Q13.4. Review of operative photographs	☐	☐	☐	☐	☐	
Q13.5. Intra-operative monitoring by visiting surgical team	☐	☐	☐	☐	☐	
Q13.6. Intra-operative monitoring using video recording and real-time data transfer to Trial Committee	☐	☐	☐	☐	☐	
Q13.7. Histopathology assessment of quality of resected specimens	☐	☐	☐	☐	☐	
Q13.8. Review of post-operative complications/outcomes and lymph node yield	☐	☐	☐	☐	☐	
Q13.9. Review of operative notes	☐	☐	☐	☐	☐	
Q13.10. Review of case report forms.	☐	☐	☐	☐	☐	
Q13.11. Any questions?						

Appendix M. ADDICT trial survey – Likert Scale of trial stakeholder responses for closed questions

Question	Strongly disagree	Disagree	Undecided	Agree	Strongly Agree
3.0. Stakeholder perception of factors contributing to reduced recruitment of patients in the ADDICT trial:					
3.1. A lack of clinical equipoise (surgeon belief that there is genuine clinical uncertainty between the two trial arms) amongst trial surgeons	1 (12.5%),	1 (12.5%),	3 (37.5%),	3 (37.5%)	0 (0%)
3.2. Insufficient research funds to support investigators including insufficient funds to hire more research nurses	0 (0%)	2 (25.0%),	1 (12.5%),	3 (37.5%)	2 (25.0%)
3.3. Patients choosing not to partake in the study for personal reasons/beliefs	1 (12.50%)	2 (25.0%)	1 (12.50%),	2 (25.0%)	2 (25.0%)
3.4. Over-stringent selection of surgical centres	1 (12.5%)	5 (62.5%)	0 (0%)	2 (25.0%)	0 (0%)
3.5. Overly stringent standardisation of surgery	1 (12.5%)	5 (62.5%)	2(25.0%)	0 (0%)	0 (0%)
3.6. Reduced interest amongst trial surgeons	0 (0%)	0 (0%)	2 (25.0%)	4 (50.0%)	2 (25.0%)
3.7. Burden of routine clinical work on trial surgeons restricting their participation within trial	0 (0%)	2 (25.0%)	0 (0%)	5 (62.5%)	1 (12.5%)
3.8. Concern amongst trial surgeons that explaining the trial may cause patients to seek treatment in another centre	0 (0%)	(4 (50.0%)	0 (0%)	4 (50.0%)	0 (0%)
3.9 Personal factors meaning some trial surgeons may prefer not to help with colleagues' research	0 (0%)	5 (62.5%),	1 (12.5%),	2 (25.0%),	0 (0%))
3.10 Reducing incidence and change in clinical stage of presentation of gastric cancer	0 (0%)	1 (12.50%)	2 (25.0%),	3 (37.5%),	2 (25.0%)
4.0 Factors leading to reluctance to expand recruitment including more centres internationally					
4.1. Difference in surgeons' conception and skill to perform D2 gastrectomy internationally	0 (0%)	1 (12.50%),	0 (0%),	6 (75.0%),	1 (12.5%)
4.2. Difficulty in assessing quality of surgery across centres internationally in selecting new centres for participation.	0 (0%)	1 (12.5%),	0 (0%),	6 (75.0%),	1 (12.5%)
Q6. Potential strategies to overcome challenges to recruitment include:					
6.1. Expand number of centres/surgeons included in trial internationally within selection criteria (participating surgeons having previously performed 50 gastrectomies)	0 (0%)	(12.5%),	1 (12.5%),	4 (50.0%),	2 (25.0%)
6.2. Increase the number of participating centres within the geographically accessible area for the surgical quality monitoring teams	0 (0%)	2 (25.0%),	0 (0%),	5 (62.5%),	1 (12.5%)
Q6.3. Trial centres and surgeons increasing collaboration and working towards shared goals	0 (0%)	1 (12.5%)	0 (0%),	6 (75.0%),	1 (12.5%)
Q6.4. Structured, regular and frequent communication between trial centres and trial committee with appropriately senior feedback mechanism and a positive reward system.	0 (0%)	0 (0%),	0 (0%),	5 (62.5%),	3 (37.5%)
Q6.5. Flexible standardisation of surgical procedures focusing on those aspects affecting safety or survival	0 (0%)	2 (25.0%)	2 (25.0%)	4 (50.0%)	0 (0%)
Q7. Stakeholder perception of adjusting centre selection according to the following factors to attempt to improve recruitment in the ADDICT trial:					
7.1. Operation complexity	1 (12.5%)	1 (12.5%)	1 (12.5%),	5 (62.5%)	0 (0%)
7.2. Disease prevalence	1 (12.5%)	1 (12.5%)	1 (12.5%),	5 (62.5%)	0 (0%)
7.3. Hazard ratio (e.g. for post-operative complications or 30-day mortality) between centres	1 (12.5%)	1 (12.5%)	1 (12.5%)	5 (62.5%)	0 (0%)
Q8. To maintain interest of trial investigators to improve recruitment the following strategies would be useful:					

Q8.1. Conduct frequent investigator seminars	0 (0%)	1 (12.5%)	2 (25.0%),	4 (50%)	1 (12.5%)
Q8.2. Send regular updates on newsletters	0 (0%)	0 (0%)	0 (0%)	6 (75.0%)	2 (25.0%)
Q10. Indicate your level of disagreement or agreement to the following potential challenges to monitoring of surgery in the ADDICT trial?	**Strongly disagree**	**Disagree**	**Undecided**	**Agree**	**Strongly Agree**
10.1. Quality of some videos is insufficient to assess quality of surgical procedure	0 (0%)	1 (12.5%),	1 (12.5%)	5 (62.5%),	1 (12.5%)
10.2. Quality of some photographic series is insufficient to assess quality of surgical procedure	0 (0%)	1 (12.5%),	1 (12.5%)	5 (62.5%),	1 (12.5%)
10.3. Trial teams are unfamiliar with method of taking and submitting standardised videos/photographs	0 (0%)	1 (12.5%),	2 (25%)	4 (50%)	1 (12.5%)
10.4. Feasibility and practical problems in recording trial videos to monitor quality of surgery	0 (0%)	2 (25%)	3 (37.5%)	3 (37.5%),	0 (0%)
10.5. Feasibility and practical problems in taking photographic series to monitor quality of surgery	0 (0%)	2 (25.0%)	2 (25.0%)	4 (50.0%),	0 (0%)
10.6. Surgeons pride leading to difficulty in providing and receiving feedback	0 (0%)	3 (37.5%)	2 (25.0%)	3 (37.5%)	0 (0%)
10.7. Regional culture amongst surgeons leading to difficulty in providing and receiving feedback	0 (0%)	3 (37.5%)	2 (25.0%)	2 (25.5%)	1 (12.5%)
10.8. Insufficient resources for monitoring of surgery	0 (0%)	2 (25.0%)	3 (37.5%)	2 (25.0%)	1 (12.5%)
10.9. Surgeons' reluctance to be monitored	0 (0%)	2 (25.0%)	4 (50.0%)	2 (25.0%)	0 (0%)
10.10. Surgeons' concerns over adverse outcomes associated with monitoring of surgery	1 (12.5%)	1 (12.5%)	3 (37.5%)	3 (37.5%)	0 (0%)
10.11. Surgeons' concerns over potential litigation associated with monitoring of surgery	1 (12.5%)	1 (12.5%)	5 (62.5%)	1 (12.5%)	0 (0%)
10.12. Potential negative impact of monitoring of surgery on generalizability of results	0 (0%)	5 (62.5%)	3 (37.5%)	0 (0%)	0 (0%)
10.13. Potential negative impact of monitoring of surgery on surgeon recruitment	0 (0%)	4 (50.0%)	4 (50.0%)	0 (0%)	0 (0%)
Q12. Stakeholder level of disagreement or agreement with potential strategies to overcome challenges to monitoring of surgery in ADDICT include:					
12.1. Trial teams utilising standardised method of recording operative videos/photographs as per instructions on the ADDICT trial website	0 (0%)	0 (0%),	0(0%),	5 (62.5%),	3 (37.5%)
12.2. Newsletters and messages to trial stakeholders reminding them of standardised methods for intra-operative monitoring	0 (0%)	0 (0%),	0 (0%)	4 (50.0%)	4 (50.0%)
12.3. Monitoring of surgery (photographs and/or videos) through external peer review within the ADDICT trial to reinforce adherence to protocol	0 (0%)	1 (12.5%)	0 (0%)	5 (62.5%)	2 (25.0%)
12.4 The process for monitoring of surgery within the ADDICT trial should be standardised	0 (0%)	0 (0%)	2 (25.0%),	5 (62.5%)	1 (12.5%)
12.5. Monitoring of surgical quality should consist of a graded monitoring system, adjusting stringency of monitoring according to performance and monitoring of initial cases	0 (0%)	1 (12.5%)	1 (12.5%)	5 (62.5%)	1 (12.5%)
12.6. Monitoring of surgery should be kept anonymous and confidential	0 (0%)	1 (12.5%)	1 (12.5%),	3 (37.5%)	3 (37.5%)
12.7. There should be a structured, regular and appropriately senior feedback mechanism with positive reward system for monitoring of surgery	0 (0%)	2 (25.0%)	1 (12.5%)	4 (50.0%)	1 (12.5%)
12.8. Improved system of funding investigator initiated surgical trials with adequate funding for quality assurance of surgical procedures	0 (0%)	0 (0%)	1 (12.5%)	2 (25.0%)	5 (62.5%)
12.9. Technology and artificial intelligence assistance in monitoring surgical quality	0 (0%)	2 (25.0%)	4 (50.0%)	2 (25.0%)	0 (0%)
Q13. Stakeholder level of disagreement or agreement with optimal method for monitoring of ADDICT trial					
13.1. Use of a structured objective assessment tool involving review of operative photographs, review of unedited videos or intra-operative monitoring by a visiting surgical team	0 (0%)	0 (0%)	0 (0%)	8 (100.0%))	0 (0%)
13.2. Review of unedited operative videos	0 (0%)	3 (37.5%)	1 (12.5%)	4 (50.0%)	0 (0%)

13.3. Review of random selection of recorded unedited trial videos	0 (0%)	2 (25.0%)	1 (12.5%)	4 (50.0%)	1 (12.5%)
13.4. Review of operative photographs	0 (0%)	1 (12.5%)	0 (0%)	7 (87.5%))	0(0%)
13.5. Intra-operative monitoring by visiting surgical team	0 (0%)	5 (62.5%)	0 (0%)	3 (37.5%)	0 (0%))
13.6. Intra-operative monitoring using video recording and real-time data transfer to Trial Committee	0 (0%)	5 (62.5%)	1 (12.5%)	2 (25.0%)	0 (0%)
13.7. Histopathology assessment of quality of resected specimens	0 (0%)	2 (25.0%)	0 (0%)	6 (75.0%)	0 (0%)
13.8. Review of post-operative complications/outcomes and lymph node yield	0 (0%)	0 (0%)	0 (0%)	7 (87.5%)	1(12.5%)
13.9. Review of operative notes	0 (0%)	1 (12.5%)	0 (0%)	6 (75.0%)	1 (12.5%)
13.10. Review of case report forms.	0 (0%)	1 (12.5%)	0 (0%)	6 (75.0%)	1 (12.5%)

Appendix N: Detailed qualitative analysis of written survey for ADDICT study (Chapter 6)

6.3.2.1. Challenges to recruitment of patients in the ADDICT trial

The key stakeholder perceived barriers to recruitment of patients in ADDICT included: (i) Reducing incidence and change in clinical stage of presentation of gastric cancer; (ii) Clinical burden on Korean surgeons; (iii) Insufficient funds for surgical trials; (iv) Reduced interest amongst trial surgeons; (v) Difference in surgeons' conception and ability to perform D2 gastrectomy internationally; (vi) Difficulty in assessing quality of surgery across centres internationally in selecting new centres for participation, and; (vii) Patient education and equipoise.

(i) Reducing incidence and change in clinical stage of presentation of gastric cancer

The bulk (62.5%) of trial stakeholders agreed the reducing incidence of gastric cancer was a significant challenge reducing recruitment in ADDICT, particularly within Korea. One ADDICT stakeholder explained this was due to there being a 'relatively small number of patients who are candidates for the trial'.

Epidemiological studies have shown the incidence of gastric cancer in Europe and the United States, along with other *Helicobacter Pylori* related diseases, has declined over recent years with advances in diagnosis and treatment.[196] Attributed to rapid development, improved sanitation, diet and methods of food preservation there is a reducing incidence of gastric cancer in Japan and South Korea, and likely other Asian area where have occurred. The incidence of gastric cancer in Japan has rapidly fallen by approximately 60% between 1965 and 1995 across all age groups.[196]

(ii) Clinical burden on Korean surgeons

The majority of stakeholders were in agreement (80%) that burden of clinical work could reduce participation of surgeons within ADDICT. Some feeling that those in

smaller volume hospitals lack the manpower and support infrastructure to engage in the trial: 'This is particularly true for small volume hospital because they lack of support and human resource from the hospital.' Similar disincentives for clinicians partaking in research were identified within a previously published review of recruitment finding recruiters 'time constraints' were an important factor acting against their recruitment of participants.[161] One RCT of cardiac rehabilitation for patients with bowel cancer revealed salient reasons for under-recruitment included: over-estimation of the number of patient admissions; other reasons were i) not assessing all patients for eligibility, ii) not completing a screening form for eligible patients and iii) patients who signed a screening form being lost to the study before consenting and randomisation.[162] Although these challenges have not specifically been mentioned by ADDICT stakeholders, it is possible that some of them may account for the mechanism of suboptimal recruitment within ADDICT. For example the identified 'burden of clinical work' and 'reduced interest amongst trial surgeons' (questions 3.6 and 3.7), both of which had support from 75% of ADDICT stakeholders, may have led to stakeholders 'not assessing all patients for eligibility'.

(iii) Insufficient funds for surgical trials

The majority of trial stakeholders (62.5%) agreed this factor contributed to reducing recruitment in ADDICT. One ADDICT stakeholder explained this difficulty: 'Reasonable support of research grant to have a research nurse is practically important but it is still difficult in Korea to get research funding for surgery clinical trial for pure investigator initiated trial.' Similar barriers to recruitment have previously been identified in published literature with one narrative review of recruitment finding similarly that 'a lack of resources' could be a significant disincentive to active trial stakeholder recruitment.[161]

(iv) Reduced interest amongst trial surgeons

75% of stakeholders felt that 'reduced interest amongst ADDOICT trial surgeons' was an important challenge to recruitment. One stakeholder explained that Korea

surgeons were difficult to encourage to participate in trial recruitment: 'Maybe. It is general issue for Korean surgeons. Difficult to encourage.' Other stakeholders felt some patients who would meet criteria for the ADDICT trial would receive a D2 gastrectomy under the Japanese guidelines. This fact can reduce the clinical equipoise of ADDICT trial stakeholders making them reluctant to recruit these patients into the trial for fear of going against the Japanese guidelines. One stakeholder summarised his perceived conflict between the ADDICT trial protocol and the JCOG guidleines: 'T1N1 ~ T1. T2 and N0, N1 are currently recommended as D2 LND according to JCOG.'

In one qualitative study, researchers' examined factors influencing how team working affected recruitment in oesophagogastric oncology trials. Several influential factors were identified including: the multidisciplinary team (MDT) meeting; leadership of the trial, and; the recruitment process. Similarly to ADDICT stakeholders in our study, interviewees perceived their clinical colleagues to have strong treatment preferences, which led to scepticism regarding whether the treatments were being described to patients in a balanced manner. [163]

(v) Difference in surgeons' conception and skill to perform D2 gastrectomy internationally

The majority (87.5%) of stakeholders supported this as a contributing factor leading to reluctance to expand trial recruitment. In keeping with this, the perception that D2 gastrectomy has limitations when performed within Western centres is commonly held and documented within published literature.[7] Concerns regarding gastrectomy for gastric cancer were first documented by the famous Canadian surgeon(LH Appleby), who went on to describe the Appleby technique of left upper quadrantectomy for cancer: *"Judged by the standards of colon, breast and rectal surgery for cancer, the operation as commonly practiced for gastric cancer Is wholly Inadequate."* [197] This concern was further highlighted following the British Medical Research Council (MRC) trial and D2 versus D1 gastrectomy trial in the Netherlands, in which both studies showed that mortality and morbidity were significantly higher

in those recieving D2 resection compared to D1, and there was no evidence of a survival advantage in the D2 group.[45,51] This observation was widely attributed to the fact that these trials were introduced too hastily with insufficient credentialing of surgeons, and thus the surgeons who took part had very little experience of performing D2 gastrectomy.[198]

(vi) Difficulty in assessing quality of surgery across centres

Similarly to ADDICT stakeholders concerns regarding surgeons within Western centres ability to perform D2 gastrectomy, the majority (87.5%) of stakeholders felt that 'difficulty in assessing quality of surgery across centres' led to a reluctance to recruit more centres internationally. The inherent challenge of monitoring quality of surgery across multiple centres has been long recognised, and historically 'surgical details collected through self-report by the surgeon' was considered the only feasible option of data collection.[10] Although ADDICT utilises more advanced monitoring technology including review of operative video recordings and photographs following lymphadenectomies at key operative stages, concerns regarding the feasibility of such monitoring if the trial expanded to Western countries was evident. This perceived challenge is likely closely related to the concern of ADDICT stakeholders above regarding the difference in surgeons' understanding and ability to perform D2 gastrectomy within Western centres.

(vii) Patient education and equipoise

Patient education was also mentioned by ADDICT stakeholders, indicating that without potential trial participants having sufficient knowledge of the underlying concepts of the ADDICT trial, they would feel reluctant to participate in a trial. The difficulty of explaining the trial to patients given their lack of knowledge relating to lymphadenectomy, became an emergent challenge: 'Explaining LN dissection to subject is bit challenging, most of the patients have no idea what is LN dissection and what is it for.'

Other challenges to recruitment of patients in ADDICT

The majority of trial stakeholders completing the survey did not feel the following factors were contributing challenges to recruitment of patients in ADDICT: (i) A lack of clinical equipoise (surgeon belief that there is genuine clinical uncertainty between the two trial arms) amongst trial surgeons; (ii) Patients choosing not to partake in the study for personal reasons/beliefs; (iii) Over-stringent selection of surgical centres and/or standardisation of surgery; (iv) Concern amongst trial surgeons that explaining the trial may cause patients to seek treatment in another centre, and; (v) Personal factors meaning some trial surgeons may prefer not to help with colleagues' research. In response to the question of whether 'overly stringent selection of surgical centres' was a contributing challenge to recruitment in ADDICT, one stakeholder rebutted: 'This is a pragmatic trial and low volume hospitals are welcome to participate though there is a surgeon criteria.'

Some ADDICT stakeholders (n=4, 50%) felt that through their explaining the trial, it would push some patients to decide to seek treatment in another centre: ADDICT stakeholders proposed a possible rationale of this process by explaining: 'In some regional areas, people tend to want to get simply adequate surgery'. One of the potential reasons trial stakeholders feel that explaining ADDICT to patients may cause patients to seek treatment elsewhere, is that it may hinder the doctor-patient relationship at a critical moment in which trust is paramount. Within published literature similar concepts are expressed with clinicians acting as recruiters facing concerns over potential threats to the doctor–patient relationship and a loss of professional autonomy.[161] This idea is further supported in another review suggesting 'patients do not wish to spend most of their consultation discussing an opportunity to participate in research which may or may not benefit them.' In private healthcare scenarios it is reported this factor would be even more relevant, with doctors' having little incentive to introduce distractions to that consultation.[199]

The majority of stakeholders (n=5, 62.5%) participating in this survey disagreed with 'personal factors meaning some trial surgeons may prefer not to help with their

colleagues' research' as a contributing challenge to recruitment in ADDICT. However one stakeholder who felt this could be an important factor explained: 'It is difficult to say but could be a real story.' Although stakeholders mainly disagreed that this personal factor may contribute to reduced recruitment in ADDICT, there are reports within published literature that such factors can impact on healthcare services and clinical work. A review of challenges to collaboration in the Norwegian health system found a lack of appropriate collaboration between providers impeded clinical work, and resulted in inadequate rehabilitation services lengthening the institutional stay for older patients.[200]

6.3.2.2. Proposed strategies to overcome challenges to recruitment of surgeons in the ADDICT trial

Recruitment of patients has historically been a challenge in surgical investigator initiated trials, [10,164] and the recruitment period has been extended and centres expanded within ADDICT in order to meet the recruitment target. Description of strategies to improve recruitment of patients in oncology trials are rare in published literature. A summary of the key stakeholder proposed strategies to overcome challenges to recruitment in ADDICT included: (i) Additional human resources at trial centres; (ii) Frequent investigator seminars and newsletters/updates (iii) Structured, regular and frequent communication between trial centres and trial committee; (iv) Trial centres and surgeons working towards shared goals; (v) Expand number of centres/surgeons included in ADDICT internationally; (vi) Flexible standardisation of surgical procedures focusing on those aspects affecting safety or survival, and; (vii) Adjusting centre selection according to operation complexity, disease prevalence and/or hazard ratio between centres.

(i) Additional human resources at trial centres

Multiple stakeholders felt extra human resources would be useful to support investigators and aid recruitment: 'More clinical research assistants to help manage the process of this trial'. The shortage of and need for additional human resources to facilitate surgical trials is recognised internationally, and in the UK the National Cancer Research Network (NCRN) seek to identify centres where allocation of clinical sessions, research nurses and other resources will enhance recruitment in surgical trials. [160] The NRCN states that they recognise UK surgeons are busy and in order optimise recruitment they seek to support them by providing adequate research infrastructure.[160]

(ii) Frequent Investigator seminars and newsletters/updates

Strategies proposed to maintain interest of investigators (Table 20) and encourage standardisation of surgical skills and monitoring of surgery (Table 19) included frequent investigator seminars and newsletters/updates (supported by 62.5% and 100% of ADDICT stakeholders respectively). One stakeholder advocated: 'Seminars should be held frequently for training on standard surgery of the method of operative pictures taken. Video seminar could help the quality of surgery.' Although improved communication between trial coordinators and trial sites has been previously proposed, one review of methods to improve trial recruitment and retention finds evidence to be limited regarding effectiveness of this strategy.[157]

(iii) Structured, regular and frequent communication between trial centres and trial committee

Structured, regular and frequent communication between trial centres and trial committee with appropriately senior feedback mechanism and a positive reward system gained universal support by trial stakeholders (n=8, 100%) (See Table 20). Within the ADDICT survey the exact details of the 'positive reward system' to improve recruitment were not further delineated. Possible incentives were suggested in one review of trial recruitment including: improved care for participants; altruism; career advancement; co-authorship of scientific outputs, and; the opportunity to keep up to date with current research.[161] Other ADDICT stakeholders recommended giving investigators the contact details of the trial coordinating centre at the opening trial meeting (See Table 19), in order to facilitate rapid resolution of problems faced by trial stakeholders: '…giving the trial details at the trial opening meeting and contact point during the trial to solve any challenges they encounter right at the moment.'

(iv) Trial centres and surgeons increasing collaboration and working towards shared goals

'Trial centres and surgeons increasing collaboration and working towards shared goals' gained support of 87.5% of stakeholders. A similar ethos was advocated in one

qualitative study to improve recruitment in the oesophagogastric oncology trials, in which shared study leadership was felt to positively influence healthcare professionals' willingness to participate.[163] Other ADDICT stakeholders promoted focusing on enhancing surgeons' enthusiasm for trials (See Table 19), through establishment of an oncology group similar to that existing in Japan (Japanese Clinical Oncology Group - JCOG): 'Organizing clinical trial group which has stronger enthusiasm and commitment like JCOG group in Japan.'

(v) Expand number of centres/surgeons included in ADDICT internationally

Expanding the number of centres/surgeons included in ADDICT internationally within selection criteria (participating surgeons having previously performed 50 gastrectomies) was supported by the majority of trial stakeholders (75%). Similar strategies have previously been utilised within trials struggling to meet required recruitment targets. Within the Support and Treatment After joint Replacement (STAR) trial, assessing effectiveness of a new care pathway for patients with chronic pain after a knee replacement, implementing additional recruitment of sites was based on site feasibility assessments.[201] However there was no credentialing process for additional sites recruited within the STAR trial, whereas ADDICT stakeholders stipulate that participating sites must contain surgeons who meet benchmarking criteria of having performed a minimum of 50 gastrectomies. Increasing the number of participating centres within the geographically accessible area for the surgical quality monitoring teams gained support of the majority (n=6, 75%) of stakeholders. This strategy shares more similarities to that proposed by the pilot qualitative study within the STAR trial, in that feasibility of monitoring is considered rather than credentialing of surgeons.[201] One concerned stakeholder highlighted the importance of defining authorship early on in trial prior to opening the trial to new centres: 'Authorship should be defined before spreading institutes.'

(vi) Adjusting centre selection according to operation complexity, disease prevalence and/or hazard ratio between centres

Adjusting centre selection according to operation complexity, disease prevalence and/or hazard ratio between centres (e.g. for post-operative complications or 30-day mortality) was supported by the majority of stakeholders (62.5%). Hazard ratios have been utilised previously to justify expanding trial recruitment when hazard rations have not differed significantly between trial centres, and within interim trial analysis.[202] Adjusting centre selection and recruitment according to disease prevalence and operation complexity are scarcely reported in published literature, however they were mentioned by several expert trial stakeholders within the interview study (Chapter 4, Section 4.3).

Other strategies to improve recruitment of patients in ADDICT

Dividing stakeholders, only 50% supported 'Flexible standardisation of surgical procedures focusing on those aspects affecting safety or survival' as a strategy to improve recruitment. Being an explanatory trial and already containing multiple methods to standardise trial surgical procedures, ADDICT stakeholders may not be keen to introduce more flexibility into this process. Within published literature however such flexibility is supported, as evidenced by one review of typology and monitoring of surgical interventions within 160 RCTs. The author of this review of trial typology recommends adjusting standardisation according to trials design. Great detail to define and standardises interventions is advocated for explanatory trials. On the contrary in pragmatic trials, often in multicentre studies with large numbers of surgeons, they advise such attempts would likely create difficulties, and ensuring that each step was delivered as planned would be unrealistic.[128]

6.3.2.3. Challenges to monitoring of surgery in the ADDICT trial

Monitoring of surgical quality within the ADDICT trial has consisted of intraoperative photographs distinguishing between D1+ and D2 procedures and a short 1-2 minute video clip on completion of lymph node dissection.

Key challenges to monitoring of surgery in ADDICT include: (i) insufficient quality of video and/or photographic series; (ii) Trial teams being unfamiliar with methods of taking/submitting standardised videos/photographic series, and; (iii) insufficient monitoring members.

(i) Insufficient quality of video and/or photographic series

The majority (70%) of stakeholders felt that insufficient quality of either the photographic or video monitoring (See Table 20) posed a challenge to the monitoring of surgical quality. The difficulty in clearly distinguishing between the resections in the two trial arms was emphasised: 'Exact D1+ and D2 should be done for the success of this trial. Even though photos were taken after D1 or D2, it is hard to discriminate.' In part this was attributed to the quality of the photographs taken: 'Monitoring of surgery is doing by the evaluation of OP picture which is uploaded to ADDICT trial Website. However the quality of picture is not satisfying.' Similar difficulties were encountered in a study assessing reliability of 2-stage oesophagectomy assessment tools within the ROMIO trial, in which 31 videos and 53 photographic series were rated by 3 surgeons. This research showed that a high proportion of video and photographic data was absent or insufficient in order to rate performance. The operative time for esophagectomy and the potential intrusiveness of audio-visual recordings, given the restricted surgical access and limited operative field in open surgery, were challenges encountered in capturing the image data.[24]

(ii) Trial teams unfamiliar with methods of taking/submitting standardised videos/photographic series

The majority of stakeholders (62.5%) felt that the trial teams were unfamiliar with methods of taking and submitting standardised videos/photographs. However, less trial stakeholders were convinced this was secondary to feasibility issues with either the photographic (50%) or video (37.5%) monitoring method. The difficulty in feasibility of monitoring of surgery in randomised controlled trials has been highlighted in a previously published narrative review of recruitment.[10]

(iii) Insufficient monitoring members

Other stakeholders felt there are an insufficient number of monitoring members in the trial committee to adequately monitor surgical quality. One stakeholder explained a seminal challenge to monitoring of surgery in ADDICT was 'A lack of monitoring members.' The considerable resources required in order to adequately monitor and record operative interventions has been previously recognised.[122]

Other challenges to monitoring od surgery in trials

This majority (n=6, 62.5%) were either undecided or disagreed that 'surgeons' pride' or 'regional culture amongst surgeons' posed a challenge to monitoring of surgical quality. Stakeholder responses offer an explanation for this in indicating there is anonymity within the monitoring process and no specific feedback system in place: 'Quality is just judged blindly and no feedback.' Although not specific to providing or receiving feedback, cultural forces within the surgical profession and their interaction with external bodies has been known to effect surgeon engagement/participation with educational interventions. This was clearly demonstrated in a qualitative study of the enforced regionalisation of hepatobiliary services (HPB) in Canada, in which the consequent tensions led to a lack of engagement by HPB surgeons.[203] In a similar manner, it is possible that cultural factors may have played a role in reducing surgeon adherence to the monitoring process.

Most trial stakeholders (n=5, 62.5%) were either undecided or disagreed with 'insufficient resources' as a challenge to monitoring of surgical quality. With a paradigm favouring a proactive response, one stakeholder poignantly explains that in fact effort on behalf of those surgeons already involved is what is required: 'Self assertion and effort in participating surgeons could be only way to improve surgical quality.' The majority (50%) were undecided with regard to 'surgeons' reluctance to be monitored' as a contributing challenge. One stakeholder expresses his disagreement with this point due to the fact that most surgeons have been providing monitoring evidence already: 'Anyway participating surgeons send photos and videos and expose their surgery.' The majority of respondents were either undecided or in disagreement with the potential challenges of 'adverse outcomes' (n=5, 62.5%) or 'potential litigation' (n=7, 87.5%) associated with monitoring of surgical quality. One stakeholder reflects on his experience citing he has not encountered adverse outcomes as a result of monitoring: 'We have never had this kind of issue. We don't collect whole video but just a final result of surgery'

With regard to the potential deleterious impact of monitoring on generalizability of results or on recruitment of surgeons participating in ADDICT, stakeholders were either undecided (37.5% and 50%) or disagreed (62.5% and 50%) respectively. One stakeholder explained that this was not an 'expert only' trial, thus implying generalizability should not be affected by the monitoring process: 'This trial is not expert only trial but more open trial.' Within published literature few have reflected on the challenges posed by endeavouring to monitor surgical interventions (See Chapter 2). In one narrative review the author reflects on the complexity of surgical interventions consisting of multiple interacting components. [204] Given the complexity and the inherent variability in surgical interventions the author questions when is variation in form substantial enough to be worth assessing, and how standardised should interventions be? [204] In another narrative review the author argues that if trials are performed in a highly controlled environment some surgeons argue that the results of these trials are not generalizable to or helpful for the patients they encounter in their practices.[205]

In recognition of the unique importance and challenge of monitoring surgery in trials, monitoring adherence to protocol is specified in the CONSORT NPT guidelines.[12] Despite this monitoring of surgery is rarely done in practice as demonstrated by a large systematic review of submitted trial protocols to the EORTC.[14] This was re-affirmed by our findings within our systematic review and meta-analysis of oesophagogastric trials (see chapter 3), in which only 11 (39%) of the 28 RCTs over the years 2000-2018 utilised methods to monitor surgical quality. Perhaps the historical challenges to monitoring of surgery are most evident by their scarce inclusion within RCTs and their respective protocols.

6.3.2.4. Strategies to overcome challenges to monitoring in the ADDICT trial

The ADDICT trial already utilises extensive monitoring of surgical quality control with short videos following surgical dissection and photographs of key operative stages. This was described succinctly by one stakeholder: 'ADDICT is a unique trial to collect every surgical photo and or video documentation. It could itself improve general quality of the surgery.'

Key proposed strategies to overcome challenges to monitoring od surgery in ADDICT included: (i) regular communication between trial investigators and trial committee; (ii) standardised monitoring of surgery through external peer; (iii) utilising an anonymous, graded monitoring system; (iv) Improved trial funding; (v) utilising a structured and appropriately senior feedback mechanism, and; (vi) utilising optimal methods to monitor surgery in the ADDICT trial.

(i) Regular communication between trial investigators and trial committee

Trial stakeholders universally agreed (n=8, 100%) that newsletters and messages are an appropriate method to remind trial investigators of ADDICT surgical monitoring approaches. Stakeholders felt that standardising the monitoring process was important. In order to achieve this, stakeholders proposed frequent seminars including training on standardisation with operative videos, and/or use of a manual explaining the monitoring process: 'Seminars should be held frequently for training on standard surgery and the method of operative picture taken. Video seminar could help the quality of surgery.'

(ii) Standardised monitoring of surgery through external peer

ADDICT stakeholders universally agreed (n=8, 100%) with the use of standardised monitoring methodology of recording operative videos/photographs as per the ADDICT trial website. The majority of stakeholders agreed that surgical monitoring should both be standardised as a process (n=6, 75%) and be performed through external peer review (n=7, 87.5%). Although merits of specific forms of monitoring of surgical interventions have rarely been debated in published literature, broad discussions regarding the optimal approached for such measures have been discussed. The author of one review on monitoring within trials, similarly suggests that in order to succeed the monitoring of surgical procedures must involve a rigorous and formal system of assessment, and be championed by surgical practice leaders. [205] In the first published study to assess the reliability of a structured objective assessment tool to rate the quality of the Ivor-Lewis oesophgagectomy within an RCT, both photographic and video oesophagectomy tools were found to be reliable with high G-coefficients and levels of internal consistency.[24]

(iii) Utilising an anonymous, graded monitoring system

The majority of stakeholders (n=6, 75%) agreed operative monitoring should consist of an anonymous, graded monitoring system with adjustment of stringency of monitoring according to performance during initial cases. Tailoring the quality assurance (QA) monitoring process according to individual cases, trial units and

performance has been previously utilised within radiotherapy oncology trials. In a review of radiotherapy quality assurance in oncology trials the National Cancer Institute Work Group on Radiotherapy Quality Assurance recommends adjusting the intensity of QA to the clinical trial objectives (include general credentialing of trial units, trial-specific credentialing, and individual case review).[31]

(iv) Improved trial funding

Improving funding for surgical trials, in particular ensuring adequate finance for SQA within the trial, also reached majority agreement (n=7, 87.5%). One stakeholder expressed they felt improving trial funding would help with surgical quality through improving the quality of care: 'it is potentially important for care quality in general'. With increased funding more resources could possibly be allocated to re-imburse surgeons to monitor operative interventions and/or employ more research nurses. One systematic review of feasibility of surgical randomised controlled trials advised improved funding of surgical trials is important to validate the efficacy of surgical innovations, and to investigate the efficacy of surgical procedures.[206] Although the costs of running surgical RCTs are high in the short-term, they allow appropriate funding of treatments with proven efficacy in the long-term which may help to improve the allocation of resources and to lower the costs of healthcare.[206]

(v) Utilising a structured and appropriately senior feedback mechanism

The use of a structured, regular and appropriately senior feedback mechanism with a positive reward system for monitoring of surgery was a popular strategy with majority agreement (n=7, 62.5%). Although few have reported on feedback mechanisms between the trial monitoring teams and trial surgeons within trials, one qualitative study assessed provision of feedback to consultant anaesthetists relating to patient-reported quality of recovery indicators in a large London teaching hospital.[207] It found the most popular feedback mechanism was to make feedback specifically relevant to the recipient supported professional learning within a supportive and open collaborative environment. [207]

(vi) Utilising optimal methods to monitor surgery in the ADDICT trial

The majority of stakeholders supported the following strategies as optimal methods of monitoring quality of surgery in ADDICT: Review of random selection of recorded unedited trial videos (n=5, 62.5%); use of a structured objective assessment tool in monitoring (n=8, 100%); review of operative photographs (n=7, 87.5%), histopathology assessment of quality of resected specimens (n=6, 75%); review of post-operative complications/outcomes and lymph node yield (n=8, 100%); review of operative notes (n=7, 87.5%), and review of case report forms as (n=7, 87.5%). Post-operative complications are already a 'secondary end point and included in this trial' possibly accounting for its support.

Review of unedited operative videos divided ADDICT stakeholder opinion with only 50% (n=4) agreeing that it would be an optimal method of monitoring. The majority (n=5, 62.5%) disagreed that intra-operative monitoring by a visiting surgical team would be an optimal monitoring strategy. Regarding intra-operative monitoring by a visiting surgical team, stakeholders expressed concerns regarding its complexity, practical difficulties and that litigation may be an issue. Time, resources and practicality were similarly cited as reasons for forgoing the review of unedited videos within the monitoring process.

Video and photographic monitoring were utilised within the randomised trial of open versus minimally invasive oesophagectomy (ROMIO) trial. In development of quality assessment tools for 2-stage oesophagectomy within the ROMIO study quality of data from intra-operative videos and photographs was often absent or of insufficient quality.[24] Similar to findings in this qualitative study of ADDICT trial stakeholders, authors planning the ROMIO trial recommended clear instructions for data capturing, adequate monitoring resources and strong engagement from participating surgeons are required to improve the monitoring process.[24] Although this study demonstrated development of a reliable structured objective assessment tool in monitoring 2-stage oesophagectomy,[24] no such tool has as yet been

developed for D1+ or D2 gastrectomy as would be required for monitoring in the ADDICT trial.

Other strategies to overcome challenges to monitoring of surgery in ADDICT

Stakeholders were generally undecided (n=4, 50%) regarding use of 'technology and artificial intelligence (AI) assistance in monitoring surgical quality' as a strategy and one stakeholder remarked that this is a 'challenging and provocative idea.' Contrary to the findings in this study, one review of artificial intelligence (AI) in surgery proposed intraoperative monitoring of such different types of data using AI could lead to real-time prediction and avoidance of adverse events. It also advised integration of pre-, intra-, and post-operative data may help to monitor patients' recovery and predict complications.[208]

Appendix O: Delphi Survey for the ADDICT trial

Delphi Round 1: Expert Consensus on mitigating strategies to overcome challenges to monitoring and recruitment in the ADDICT trial

Name: …… Date of Birth: …….. Role within ADDICT Trial: ……

Please indicate your level of agreement or disagreement with the following mitigating strategies to overcome challenges to monitoring of surgery and recruitment of patients in the ADDICT trial. Please also indicate the reasons for your opinion in the comments section. We are aiming for an agreement level of 70% between experts, which may require more than one Delphi rounds.

PROPOSED STRATEGIES TO OVERCOME CHALLENGES TO RECRUITMENT IN THE ADDICT TRIAL	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	Comments
Q1. Potential strategies to overcome challenges to recruitment in the ADDICT trial include:						
Expand number of centres/surgeons included in trial internationally within selection criteria (participating surgeons having previously performed 50 gastrectomies)	☐	☐	☐	☐	☐	
Increase the number of participating centres within the geographically accessible area for the surgical quality monitoring teams	☐	☐	☐	☐	☐	
Trial centres and surgeons increasing collaboration and working towards shared goals	☐	☐	☐	☐	☐	
Flexible standardisation of surgical procedures focusing on those aspects affecting safety or survival	☐	☐	☐	☐	☐	
	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	Comments
Providing contact details to trial stakeholders at the trial opening meeting and contact points during the trial to enable trial centres to overcome any challenges they encounter at the moment they arise.	☐	☐	☐	☐	☐	
Making more clinical research assistants available to trial centres to help manage the process of the trial	☐	☐	☐	☐	☐	
Organizing a clinical trial group in Japan which has strong enthusiasm and commitment like the Korean Cancer Oncology Group (KCOG)	☐	☐	☐	☐	☐	
Q2. Please rate whether you agree or disagree that adjusting centre selection according to the following factors may improve recruitment in the ADDICT trial:						
• Operation Complexity	☐	☐	☐	☐	☐	
• Disease Prevalence	☐	☐	☐	☐	☐	
• Hazard ratio (e.g. for post-operative complications or 30-day mortality) between centres	☐	☐	☐	☐	☐	

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	Comments
Q3. Please indicate your level of agreement or disagreement that adjusting standardisation according to the following factors may improve recruitment in the ADDICT trial:						
• **Disease process studied in the trial**	☐	☐	☐	☐	☐	
• **Trial design (explanatory versus pragmatic)**	☐	☐	☐	☐	☐	
Q4. To maintain interest of trial investigators to improve recruitment the following strategies would be useful :						
• **Conduct frequent investigator seminars**	☐	☐	☐	☐	☐	
• **Send regular updates on newsletters**	☐	☐	☐	☐	☐	
PROPOSED STRATEGIES TO OVERCOME CHALLENGES TO MONITORING OF SURGERY IN THE ADDICT TRIAL	**Strongly Disagree**	**Disagree**	**Undecided**	**Agree**	**Strongly Agree**	**Comments**
Q5. Potential strategies to overcome challenges to monitoring of surgery in ADDICT include:						
Trial teams utilising standardised method of recording operative videos/photographs as per instructions on the ADDICT trial website	☐	☐	☐	☐	☐	
Educating trial surgical teams in photographic and video monitoring of surgery and giving them a manual for this process	☐	☐	☐	☐	☐	
Frequent trial stakeholder seminars should be conducted regarding standardised surgery for the ADDICT trial	☐	☐	☐	☐	☐	
Operative videos should be reviewed in trial stakeholder seminars to monitor the quality of surgery in the ADDICT trial	☐	☐	☐	☐	☐	
Newsletters and messages to trial stakeholders reminding them of standardised methods for intra-operative monitoring	☐	☐	☐	☐	☐	
Monitoring of surgery (photographs and/or videos) through external peer review within the ADDICT trial to reinforce adherence to protocol	☐	☐	☐	☐	☐	
The process for monitoring of surgery within the ADDICT trial should itself be standardised	☐	☐	☐	☐	☐	
Monitoring of surgical quality should consist of a graded monitoring system, adjusting stringency of monitoring according to performance and monitoring of initial cases	☐	☐	☐	☐	☐	
Monitoring of surgery should be kept anonymous and confidential	☐	☐	☐	☐	☐	

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	Comments
There should be a structured, regular and appropriately senior feedback mechanism with positive reward system for monitoring of surgery	☐	☐	☐	☐	☐	
Improved system of funding investigator initiated surgical trials with adequate funding for quality assurance of surgical procedures	☐	☐	☐	☐	☐	
Technology and artificial intelligence assistance in monitoring surgical quality	☐	☐	☐	☐	☐	
Q6. The optimal method for monitoring of surgery in the ADDICT trial includes:						
Use of a structured objective assessment tool involving review of operative photographs, review of unedited videos or intra-operative monitoring by a visiting surgical team	☐	☐	☐	☐	☐	
Review of unedited operative videos	☐	☐	☐	☐	☐	
Review of random selection of recorded unedited trial videos	☐	☐	☐	☐	☐	
Review of operative photographs	☐	☐	☐	☐	☐	
Intra-operative monitoring by visiting surgical team	☐	☐	☐	☐	☐	
	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	
Intra-operative monitoring using video recording and real-time data transfer to Trial Committee	☐	☐	☐	☐	☐	
Histopathology assessment of quality of resected specimens	☐	☐	☐	☐	☐	
Review of post-operative complications/outcomes and lymph node yield	☐	☐	☐	☐	☐	
Review of operative notes	☐	☐	☐	☐	☐	
Review of case report forms	☐	☐	☐	☐	☐	

Appendix P: On-going work in the ADDICT Trial

Further work on this project extends to 'Phase 4' (please see initial project phases Chapter 6), in which we are sending the list (Table 21) of ADDICT stakeholder consensus strategies to overcome challenges to recruitment and monitoring of surgery to the trial committee. At present we are awaiting the ADDICT trial committee review and input regarding these consensus strategies, to consider which would be suitable for implementation within a cohort of patients within the ADDICT trial as previously planned.

Phase 5. Assess utility of mitigating strategies: Survey Questionnaire (S2) and preliminary data analysis

In order to allow assessment of usability and feasibility of the implemented strategies, after a period (approximately 2-4 months depending on nature of strategies implemented) following implementation, a further written survey will be conducted with a purposive, representative sample of 10 trial stakeholders to assess their opinion on the utility and feasibility of the implemented strategies.
Additionally in order to assess the efficacy of the implemented strategies, participant data from the cohort within which mitigating strategies were implemented (please see phase 4) will be requested for analysis from trial results database. Outcomes within the strategy implementation cohort will be compared against a similar trial cohort who did not receive the implemented strategies. For this analysis relevant data will be requested from the trial committee including: lymph node yield, post-operative morbidity, operation duration, in hospital mortality, R0 resection rate, trial surgeon adherence to the implemented mitigating strategies and recruitment of trial surgeons.

Analysis of quantitative variables

SPSS (IBM) software will be used to analyse perioperative outcomes between the cohorts with and without strategy implementation. Statistical analysis methodology may vary according to consensus mitigating strategies implemented within study trial cohort. Outcome variables compared between groups will include: lymph node yield, post-operative morbidity, operation duration, in hospital mortality, R0 resection rate, trial surgeon adherence to the implemented mitigating strategies and recruitment of trial surgeons. Non-parametric tests (e.g. Mann Whitney-U) will be employed to compare non-normally distributed variables between groups with and without implemented mitigating strategies, and parametric tests (MANOVA or t-test) for normally distributed variables.

Appendix Q: ROMIO Quality Assurance Video/Photographic Assessment tools

ROMIO Quality Assurance Video Rating Scale

ROMIO Centre:..... ROMIO Case Reference:..... Surgeon:……..

Please tick the appropriate descriptions of the safety, efficiency, and quality of the end product for each task, according to the key below.

i) TECHNICAL SAFETY

Safe	No adverse events or near misses occurred.
Near miss	Potential harms were narrowly avoided.
Unsafe	Adverse event(s) that resulted in reversible harm occurred.
Dangerous	Adverse event(s) that resulted in permanent harm occurred.

ii) OPERATIVE EFFICIENCY

Optimal	Purposeful and progressive movements throughout.
Adequate	Some unnecessary movements, but generally progressive.
Inefficient	Repeated, unproductive, movements.
Poor	Wrong movements that compromised patient safety.

iii) QUALITY OF THE END PRODUCT

Complete	Anatomical structure is clearly demonstrated following complete dissection of all associated lymphatic (LN) tissue.
Incomplete	Incomplete LN clearance of the anatomical structure (quantify if possible please)

TASK 1: DIAPHRAGMATIC HIATUS

i) Safety	**Safe**	**Near miss**	**Unsafe**	**Dangerous**	**Comments**
	☐	☐	☐	☐	.....

ii) Efficiency	**Optimal**	**Adequate**	**Inefficient**	**Poor**	**Comments**
	☐	☐	☐	☐	.....

iii) Quality of end product	**Complete**	**Incomplete**	**Not performed**	**Comments**
Right crus	☐	☐	☐	.....
Left crus	☐	☐	☐	.....
Aorta	☐	☐	☐	.....
Pericardium	☐	☐	☐	.....
Right lung	☐	☐	☐	.....
Left lung	☐	☐	☐	.....

TASK 2: ABDOMINAL LYMPHADENECTOMY

i) Safety	**Safe**	**Near miss**	**Unsafe**	**Dangerous**	**Comments**
	☐	☐	☐	☐	.....

ii) Efficiency	**Optimal**	**Adequate**	**Inefficient**	**Poor**	**Comments**
	☐	☐	☐	☐	.....

iii) Quality of end product	**Complete**	**Incomplete**	**Not performed**	**Quantify if incomplete**

Portal vein	☐	☐	☐	.....
Proper hepatic artery	☐	☐	☐	.....
Common hepatic artery	☐	☐	☐	.....
Coeliac artery	☐	☐	☐	.....
Left gastric artery (stump)	☐	☐	☐	.....
Left gastric vein (stump)	☐	☐	☐	.....
Proximal splenic artery	☐	☐	☐	.....
Distal splenic artery	☐	☐	☐	.....
Splenic vein	☐	☐	☐	.....
Splenic hilum (if appropriate)	☐	☐	☐	.....

Continued…

TASK 3: THORACIC LYMPHADENECTOMY

i) Safety	**Safe**	**Near miss**	**Unsafe**	**Dangerous**	**Comments**
	☐	☐	☐	☐	.....

ii) Efficiency	**Optimal**	**Adequate**	**Inefficient**	**Poor**	**Comments**
	☐	☐	☐	☐	.....

iii) Quality of end product	**Complete**	**Incomplete**	**Not performed**	**Quantify if incomplete**
Carina	☐	☐	☐	.....
Right main bronchus	☐	☐	☐	.....
Left main bronchus	☐	☐	☐	.....
Right pulmonary veins	☐	☐	☐	.....
Left pulmonary veins	☐	☐	☐	.....
Pericardium	☐	☐	☐	.....
Aorta	☐	☐	☐	.....

TASK 4: RECONSTRUCTION

i) Safety	**Safe**	**Near miss**	**Unsafe**	**Dangerous**	**Comments**
	☐	☐	☐	☐	.....

ii) Efficiency	**Optimal**	**Adequate**	**Inefficient**	**Poor**	**Comments**
	☐	☐	☐	☐	.....

iii) Quality of end product	**Yes**	**No**	**Borderline**	**Comments**
Viable color of gastric tube	☐	☐	☐	.....
Lesser curve cleared of LN tissue	☐	☐	☐	.....
Tension free anastomosis	☐	☐	☐	.....
Appropriate approximation of sutures	☐	☐	☐	.....

ANY OTHER COMMENTS

ROMIO ASSESSOR (Print):…… SIGN:…… DATE:……

ROMIO Quality Assurance Photographic Rating Scale

ROMIO Centre:..... ROMIO Case Reference:..... Surgeon:.....

Instructions: Please tick the appropriate description for the quality of the end product for each task.

Key: **Complete** Anatomical structure is clearly demonstrated following complete dissection of all associated lymphatic (LN) tissue.
Incomplete Incomplete LN clearance of the anatomical structure (Quantify if incomplete).

Task 1 Diaphragmatic hiatus	**Complete**	**Incomplete**	**Not performed**	**Quantify if incomplete**
Right crus	☐	☐	☐	.....
Left crus	☐	☐	☐	.....
Aorta	☐	☐	☐	.....
Pericardium	☐	☐	☐	.....
Right lung	☐	☐	☐	.....
Left lung	☐	☐	☐	.....

Task 2 Abdominal lymphadenectomy	**Complete**	**Incomplete**	**Not performed**	**Quantify if incomplete**
Portal vein	☐	☐	☐	.....
Proper hepatic artery	☐	☐	☐	.....
Common hepatic artery	☐	☐	☐	.....
Coeliac artery	☐	☐	☐	.....
Left gastric artery (stump)	☐	☐	☐	.....
Left gastric vein (stump)	☐	☐	☐	.....
Proximal splenic artery	☐	☐	☐	.....
Distal splenic artery	☐	☐	☐	.....
Splenic vein	☐	☐	☐	.....
Splenic hilum (if appropriate)	☐	☐	☐	.....

Task 3 Thoracic lymphadenectomy	**Complete**	**Incomplete**	**Not performed**	**Quantify if incomplete**
Carina	☐	☐	☐	.....
Right main bronchus	☐	☐	☐	.....
Left main bronchus	☐	☐	☐	.....
Right pulmonary veins	☐	☐	☐	.....
Left pulmonary veins	☐	☐	☐	.....
Pericardium	☐	☐	☐	.....
Aorta	☐	☐	☐	.....

Task 4 Reconstruction	**Yes**	**No**	**Borderline**	**Quantify if incomplete**
Viable color of gastric tube	☐	☐	☐	.....
Lesser curve cleared of LN tissue	☐	☐	☐	.....
Tension free anastomosis	☐	☐	☐	.....

Appropriate approximation of sutures	☐	☐	☐	……

ROMIO ASSESSOR (Print):…… SIGN:…… DATE:……

Appendix R: Challenges to recruitment of patients and monitoring of surgery in Neo-AEGIS

Challenges to recruitment of patients and monitoring of surgery in Neo-AEGIS

Introduction

The Neo-AEGIS trial is a multicentre randomised controlled trial (RCT) comparing pre-operative and post-operative chemotherapy (MAGIC regimen) versus pre-operative chemoradiotherapy (CROSS regimen) for advanced adenocarcinoma of oesophagus and junction.[209] Each included patient within Neo-AEGIS receives a radical en-bloc resection and regional lymphadenectomy for their locally advanced adenocarcinoma (cT2-3, N0-3, M0) of the oesophagus and junction (AEG). Use of SQA measures within gastro-oesophageal oncology RCTs have been shown to influence outcome, with one recent systematic review showing credentialing of surgeons through assessment of operative reports and performance monitoring, reduced variation in lymph-node harvest, a proposed marker of the extent and quality of surgical resection. Furthermore, standardisation of surgical techniques and credentialing of surgeons was shown to be associated with reduced adjusted in-hospital mortality.[1] Although post-operative complications are recorded on case report forms for up to 90 days following surgery in Neo-AEGIS, no specific strategy for monitoring quality of surgery has yet been implemented. Often Phase III trials in the UK and Europe involving adenocarcinoma of the oesophagus face challenges in recruiting sufficient numbers of patients in order to reach sufficient statistical power due to lower relative incidence of this cancer in the UK/Europe and other factors inherent to surgical RCTs.

As part of research into SQA in Oncology Trials (Chapter 4), key stakeholders including surgeons, trial methodologists, trial managers and oncologists were interviewed to gain insight into key challenges to SQA. Following this interview process a Delphi process with a purposive sample of oncology trial stakeholders was

conducted (Chapter 4.4) gaining expert consensus on challenges to SQA in oncology trials and potential mitigating solutions.

We aim to identify the challenges to recruitment of participants and monitoring of surgical quality and potential mitigating strategies, through focus group discussions, a tailored detailed survey questionnaire and Delphi consensus involving key stakeholders within the Neo-AEGIS trial. Additionally, we will assess Neo-AEGIS stakeholder opinion of relevant previously developed expert consensus mitigating strategies from the interview study (Chapter 4, Section 4.4) within the Neo-AEGIS Delphi process. We then will aim to implement consensus strategies within a cohort of the Neo-AEGIS trial and evaluate their usability and effect on clinical outcomes.

Objectives

vi. Investigate challenges to monitoring of surgery and recruitment of patients within an active oesophagogastric RCT.

vii. Gain trial stakeholder consensus on strategic solutions to overcome challenges to monitoring of surgery and recruitment of patients within an active oesophagogastric RCT.

METHODS

A similar methodology to that previously utilised in chapter 6 for ADDICT was implemented within Neo-AEGIS. The preliminary focus group (Phase 1) is detailed below followed by the proceeding Phases of the study and ongoing/future work.

Phase 1. Preliminary Focus Group discussion (FG1)

A purposive sample of 10 key stakeholders were identified within the Neo-AEGIS trial consisting of surgeons, oncologists, trial methodologists and trial managers. Stakeholders were selected by the chief investigator of Neo-AEGIS on the basis of stakeholder experience and role within Neo-AEGIS. Initially approached using an email from the chief investigator, stakeholders were asked if they would like to participate in a focus group regarding challenges faced by stakeholders within the trial, and areas that may require further attention within the next stage of the study – the survey questionnaire.

RESULTS

Focus Group Results

10 Neo-AEGIS stakeholders including 2 surgeons, 1 oncologist, and 7 trial managers undertook the focus group in two separate face-face meetings. Median age: 40. M:F ratio 3:7. The first focus group with two surgeons lasted a duration of 32 mins 35 seconds, and the second with the oncologist and 7 trial managers lasted 19 mins and 50 seconds

Emergent themes were categorised into 4 main sub-categories of: (i) Challenges to monitoring of surgery in Neo-AEGIS; (ii) challenges to recruitment of patients; (iii) strategies to overcome challenges to monitoring of surgery, and; (iv) strategies to overcome challenges to recruitment of patients in Neo-AEGIS. (See Table 26 for summary)

Table 26: Focus Group Neo-AEGIS on challenges to monitoring of surgery and recruitment of patients within Neo-AEGIS

CATRGORY	FOCUS GROUP NEO-AEGIS – EMREGENT THEMES
Challenges to recruitment	
	Geographical location – cost and inconvenience of long-distance travel for treatment
	Stringent inclusion criteria for patients in Neo-AEGIS
	Patients choosing not to participate in the study for personal reasons/beliefs
	Patients choosing to seek treatment locally
	Clinical equipoise lacking amongst investigators
	Lack of patient equipoise
	Surgeons within some trial centres not being fully engaged with the trial
	Organisation of the clinic acting as first patient contact point
	The specialty of the chair/lead for the multidisciplinary oncology team meetings
	Lack of infrastructure to support clinical trials within some trial centres
Challenges to monitoring	
	Unclear documentation
	No strict quality assurance measures for surgery
	Potential wide variety of surgical approaches
	Lack of consensus on standardisation of oesophageal cancer resection
	Potential for outlier centres in regard to surgical quality within Neo-AEGIS
	GDPR regulations
	Establishing optimum method to monitor surgical quality
	Logistics difficulties of monitoring quality of surgery
	Cost of monitoring quality of surgery
Strategies to overcome challenges to recruitment	
	Centralisation of treatment provision
	Explanation of treatment options
	Regular patient communication and clear explanations
	Investigating how centres with low recruitment levels could improve
	Approval of FLOT regimen
	Trial leaders (Chief or Principle investigators) visiting other centres
	Conduct frequent investigator seminars, phone calls and steering group meetings
	Send regular updates on newsletters
Strategies to overcome challenges to monitoring	
	Standardised monitoring and reporting of complications
	Clear documentation
	Monitoring of surgery with photographic images or short videos following operative resection
	Structured objective assessments within monitoring process
	Internal and external peer review monitoring of quality of surgery
	Academic investigation to illicit optimal method to assess quality of surgery
	Recording a comprehensive dataset of intra-operative and post-operative events/markers
	Define and gain consensus agreement on operative/procedural standards prior to monitoring quality of surgery

CHALLENGES TO RECRUITMENT OF PATIENTS IN NEO-AEGIS

Geographical location

The impact of geographical location and long travel distances was felt to influence patients on deciding whether or not to participate with large commuting distances required in some circumstances. The trial team does not have funding to pay commuting costs and such payments may be prohibitive for patients.

'We get people from all over Ireland. Some people find it difficult travelling up and down for treatment. So they opt to have treatment somewhere near where they live and then have the surgery at James' hospital.'

'So with the travel that can be quite a big barrier.'

Stringent inclusion criteria for patients in Neo-AEGIS

The fact that a proportion of patients do not meet the stringent inclusion criteria for the trial was mentioned by stakeholders (n=2).

'So they have to go through a series of, ehm, they have to be a certain cancer, and they have to pass a certain few tests, there respiratory tests have to be a certain function and there cardiac function has to be also.'

Patients choosing not to participate in the study for personal reasons/beliefs

Trial stakeholders felt that a patient's personal beliefs played a significant role in adversely affecting recruitment in some circumstances. Other patients have strong beliefs that certain modalities unavailable within the trial, which they feel may be curative. Some stakeholders have recognised that certain patients prefer to complete all of their treatment before surgery, allowing them to focus on their

recovery following surgery. The chemotherapy following surgery in one of the trial is therefore prohibitive to their participation of for these candidates.

'Often patients are quite suspicious actually of clinical trials, because they're seen as experimental.'

'Some patients, in fact most patients know about immunotherapy, and can they have it up front, and it will cure them type of thing. I know a lot of patients who are like that..'

Patients choosing to seek treatment locally

The cultural phenomena that people have loyalty to their local hospitals and a desire to received family support was also explained.

'…local family support, plus there's always a loyalty to, we're very colloquial, so there's always a loyalty to the local hospital.'

Clinical equipoise lacking amongst investigators

In some countries investigators may be reluctant to recruit patients to Neo-AEGIS due to their wish to give FLOT to patients randomised to the chemotherapy arm of the trial, which has not yet gained ethics approval. Referring to one centre which prefers chem-radiation followed by observation rather than the Neo-AEGIS trial regime, this stakeholder explains how this lack of equipoise within this centre means it is unable to participate in the trial.

'I think in the UK, the equipoise is an issue at the moment, because investigators are not putting patient's in the trial because the investigator wants to give FLOT to the

patient, if they are randomised to the A arm. That wouldn't be an issue in Ireland because it is an option over here'

'I mean one centre, because they are very attracted to chemo-radiation and a watch and wait policy, chemo-radiation in all and a watch and wait in some. So they didn't feel this needed to be tested or it didn't appeal to them.'

Lack of patient equipoise

Some patients have the perception that chemo-radiotherapy must be better as it contains two modalities.

'They feel they're getting chemo and radiotherapy so it must, must be better, or they just want to get all of their treatment beforehand, and have the surgery and be done.'

Surgeons within some trial centres not being fully engaged with the trial

With low recruitment levels in some centres, Neo-AEGIS stakeholders felt that surgeons in some centres did not share their commitment to the trial.

'..something specific in Cork to do with surgeons, not sure if they were fully engaged with the trial..'

Organisation of the clinic acting as first patient contact point

For newly diagnosed patients, stakeholders felt that the first point of contact (e.g. whether the clinic is a medical oncology clinic as opposed to a surgical clinic or a multidisciplinary clinic) was crucial in determining potential recruitment into Neo-AEGIS.

'…you know whether it is a medical oncology clinic as apposed to a multi-disciplinary clinic, or a surgical clinic is probably critical..'

The specialty of the chair/lead for the multidisciplinary oncology team meetings

It was perceived by stakeholders that the speciality of the MDT chairing the MDT was also an important factor, potentially introducing biases

'…..the chairing of the MDT may be important in terms of biases and so on.'

Lack of infrastructure to support clinical trials within some trial centres

Some trial centres were felt to have insufficient infrastructure to support a clinical trial and that this would adversely influence recruitment.

'One centre has no good set-up for clinical trials, the smallest centre.'

CHALLENGES TO MONITORING OF SURGERY IN NEO-AEGIS

Unclear documentation

Certain intraoperative factors could be more clearly documented facilitating the role of trial monitoring staff according to one stakeholder.

'For me it would just be very precise documentation whether patient had scarring from opening, or previous adhesions.'

No strict quality assurance measures for surgery

The fact that at present there are no clear, stringent quality assurance measures for surgery within Neo-AEGIS was highlighted by one stakeholder.

‘Well, my opinon first, eh, there’s no monitoring, there’s no standardisation of surgery within Neo-AEGIS, but we’re very keen to be a pilot how that might be done for a cohort of patients.’

Potential wide variety of surgical approaches

Stakeholders cautioned that within other trial centres a broad variety of different surgical approaches could be being executed leading to concerns over quality of surgery.

‘…within other centres, within the broader dimension of the trial. We have no real knowledge or understanding of the breadth of possible different approaches that are being taken.’

Lack of consensus on standardisation of surgical approach for oesophageal cancer resection

One stakeholder explains that internationally there is vast heterogeneity in operative approach and techniques in performing oesophagectomy. The current situation in oesophagogastric surgery in which evidence are lacking and standards absent was also portrayed.

‘Oesophageal cancer is done different ways by different people, from completely minimally invasive to hybrid operations, to transhiatal to transthoracic, ….. reflects the international lack of consistency and standardisation.’

‘We don’t have standards now, to say, whether it will help to improve the situation. We don’t have evidence to go with.’

Potential for outlier centres in regard to surgical quality within Neo-AEGIS trial

Referring to centres participating in Neo-AEGIS the one stakeholder explains his concern that there may be outlying centres

'So my guess is that for quite a lot of what's done there is surgical quality by those kind of surrogate criteria in the trial, but there may be outliers also, and that would be a concern.'

GDPR regulations

The recent implementation of GDPR governing the use of patient data within trials as a challenge to recorded monitoring (video/photographic)

'So we are, so the main difficulty will be GDPR, we'd have to get every patient to agree to that, obviously ethics and all of that,'

Establishing optimum method to monitor surgical quality

Expressing the importance of establishing what you measure to reflect surgical quality one stakeholder then mentions recurrence

'Almost certainly there's going to be quality improvement there, but how you measure it and what you measure it against is the challenge..'

Logistics difficulties of monitoring quality of surgery

Within a clinical trial one stakeholder anticipated difficulties with multiple data management processes, including data storage and editing of recorded material.

‘There may be issues storing the data, editing, editing the material, storing it.’

Cost of monitoring quality of surgery

Stakeholders also predicted the complex processes of data management involved in monitoring quality of surgery would be costly and thus thorough prior investigation is required into potential implications of such monitoring on oncological outcommes.

‘…cost, cost, there’s cost involved, so. I think it probably lends itself to academic investigation…. without being absolutely certain of its implications on oncological quality and outcomes’

STRATEGIES TO OVERCOME CHALLENGES TO RECRUITMENT IN NEO-AEGIS

Centralisation of treatment provision

Provision of radiotherapy at main hospital treatment centre would allow patients to access this locally rather than travelling to others centres

‘The radiotherapy is going to be an option here, in the next few weeks. They won’t have to travel to Rathka for their radiotherapy any longer because they’ll be able to get their radiotherapy at the St Luke’s centre in St James’.’

Explanation of treatment options

A thorough explanation of treatment options within the trial including the fact these treatments have been used successfully for a number of years was noted to help reassure patients a facilitate recruitment

'When you explain that these are treatments that are used for quite a number of years, it kind of helps.'

Regular patient communication and clear explanations

Trial managers explained how they felt that their regular communication with patients via telephone and through inviting them to clinics multiple times to answer their questions allayed their anxiety facilitating recruitment. Clear explanations for patients ensuring they are aware they have the choice to opt out of the trial at any stage was also felt to help.

'The good thing is, we do bring them back, a few times, and go through everything with them. They are given our details, and we phone them an awful lot as well.'

'…that makes them more comfortable is, you know, they can opt out at any time. And if we feel from our side that its not correct for them, then we can opt them out.'

Investigating how centres with low recruitment levels could improve

With more open investigator one stakeholder explained how they felt it would be useful to investigate and ask centres with low recruitment for the reasons behind this.

'…investigating how other centres could open up or investigating why other centres have failed, and I don't know what the reasons for that are.'

Approval of FLOT regimen

Stakeholders were hopeful that once regulatory authorities and ethics had approved FLOT for use in some countries this would facilitate recruitment.

‘But were expecting approval in the next 2 or 3 weeks now, and unless of course the UK regulatory authorites have further queries, but the Irish regulatory authorities have approved the amendment…’

‘And if that’s the case the UK are looking for funding to run through to 2020, so we need that to happen as well for the trial to meet its accrual target.’

Trial leaders (Chief or Principle investigators) visiting other centres

The importance of the trial leader visiting other centres to help harnas and propel the support for the trial was also emphasised.

‘I think there’s probably an awful lot of work involved in being a PI or Co-PI of a major international trial. …..But also I think just some people are very good at going around and meeting, and going to every centre and drumming up support.’

‘…yes, its about leaders in each centre, buying into it and getting there colleagues on board, its like you said.’

Conduct frequent investigator seminars, phone calls and steering group meetings

Stakeholders recognised the importance of regular phone seminars and meetings in stimulating investigator interest in the trial and thereby improving recruitment. It was also recognised that visiting trial centres in person for this same purpose is also important.

‘We phone every month, so there were 4 or 5 people from the UK, one from France, one from Denmark and the cancer administration, so they are very useful. They are useful’

'…heading off the Korea and Japan and all these places to connect and engage with people who he is collaborating with for clinical trials. So that personal touch is probably, well it is important..'

Send regular updates on newsletters

Regularly updating stakeholders on trial progress was felt to be important.

'…drumming up support. So its down to phone calls and newsletters and so on.'

STRATEGIES TO OVERCOME CHALLENGES TO MONITORING IN NEO-AEGIS

Standardised monitoring and reporting of complications

One stakeholder explained the improvement in morbidity monitoring with new definitions of complications and monitoring them at 30, 60 and 90 days. Advising of the importance of monitoring complications within the process of monitoring surgical quality within trials one stakeholder explains that Neo-AEGIS is the first RCT using consensus complication reporting guidelines.

'…so I suppose the biggest change in that is the ESA data has released the definitions. So that has improved the monitoring of post-operative complications for all upper GI patients.'

'..this is the first randomised trial that is using the oesophageal cancer consensus complication guidelines as per Don Low's group of which we're part. So there is certainly consistency of reporting of complications which is good.'

Clear documentation

One stakeholder explains that clear documentation of blood loss, lymph node yield and previous surgery and associated adhesions is essential for monitoring in trial patients.

'…that increases their nodal yield and just the resection margins and clear documentation of blood loss, which is something that is documented quite well…. previous adhesions from previous surgeries'

Monitoring of surgery with photographic images or video

When asked their opinion on use of photographic images of operative bed following resection one stakeholder replied positively indicating either this or a short video would be the inevitable direction of operative monitoring. He also indicated that one method for open surgery may be for the operating surgeon to wear a head camera.

'Could even be something that you'd have, you know those movie cameras on your head, or whatever (gesturing to a head cam on his head), you know what I mean. That's I think, you know, inevitably how its going to go..'

Structured objective assessments within monitoring process

When asked regarding the use of structured objective assessments, using set criteria in order to assess quality of surgery using recorded images/short unedited video of operative bed within the monitoring process, stakeholders felt this would be useful following agreement on operative standards first.

‘I think that’s good but I think you know Ravi’s saying. He’s saying the gold standard has to be agreed by, first.’

Internal and external peer review monitoring of quality of surgery

Two stakeholders reported their pride in the operative of their patients following cancer resections and on this basis explained they would accept internal and external peer review of their results.

‘..we would be happy to have internal and external audit on that basis.‘

Academic investigation to illicit optimal method to assess quality of surgery

In order to further clarify the effect of monitoring surgical outcomes on oncological outcomes, one investigator indicates his support for the process and advocates further academic investigation of the area.

‘I think it probably lends itself to academic investigation as you’re doing and as you’re proposing’

Recording a comprehensive dataset of intra-operative and post-operative events/markers

Stakeholders assert the comprehensive nature of the post-operative complications data set established for Neo-AEGIS, which they feel should reflect surgical outocmes.

‘As I said, there has been a comprehensive data set that has been gathered to reflect surgical outcomes, for the first time in this trial.’

Define and gain consensus agreement on operative/procedural standards prior to monitoring quality of surgery

Stakeholders felt it is important prior to monitoring of surgery, Neo-AEGIS trial surgeons would need to agree upon trial operative standards. He explains that without this there would be a significant disputes.

'So I think something would have to be agreed upon as to what the gold standard would be, and I think if everyone agreed with that and if they participated then I think its fine ok, lets do this.'

'Or else there would be a lot of argument, because some people do a lot more than others without it, there being a proven advantage to it.'

Conclusion

From the focus group of Neo-AEGIS a 19 emergent challenges to recruitment and monitoring were identified and 16 strategies to overcome these challenges. These were utilised to develop an on-line survey with open and closed questions which has been circulated to Neo-AEGIS trial stakeholders (oncologists, surgeons, trial managers and trial methodologists) in order to more thoroughly explore their opinion regarding challenges to recruitment and monitoring of surgery in Neo-AEGIS and strategies to overcome these challenges.

Neo-Aegis: On-going and planned research methodology

Phase 2. Survey questionnaire (S1)

A larger group of stakeholders (approximately 10-15) will be targeted with a semi-structured survey questionnaire, developed based upon the findings of the preliminary focus group, which will include closed and open-ended questions relating to recruitment and monitoring quality of surgery. Relevant challenges and potential strategies from the Imperial Research Group project exploring international expert opinion on 'Mitigating strategies to overcome challenges to quality assurance of surgery in oesophagogastric oncology trials', (IRAS ID: 220221) will be incorporated into this survey.

Phase 3. Workshop

Following from the survey questionnaire, the workshop will be conducted with a purposive sample of approximately 10 key stakeholders from the trial identified by the chief investigator. The purpose of this is to reach agreement amongst trial stakeholders on strategies, to improve monitoring of quality of surgery and recruitment, for implementation in the Neo-AEGIS trial. This process will involve a Delphi on-line survey, which may require more than 2 rounds, in which stakeholder opinions are sought and documented in order to reach a target agreement level of 70%. Relevant consensus strategies from the Imperial Research Group project exploring international expert opinion on 'Mitigating strategies to overcome challenges to quality assurance of surgery in oesophagogastric oncology trials', (IRAS ID: 220221) will be incorporated into this Delphi process in Neo-AEGIS.

Phase 4. Recommendation of mitigating strategies for implementation within Neo-AEGIS trial

Following the Delphi process we expect to have a list of proposed strategies to address the challenges to monitoring of surgery and recruitment in the Neo-AEGIS trial which are deemed feasible for implementation. Once agreement is reached on selected mitigating strategies, they will be recommended for implementation within

a cohort of approximately 40 participants* within the Neo-AEGIS trial, following local ethics and trial committee approval.

Phase 5. Assess utility of mitigating strategies: Survey Questionnaire (S2) and preliminary data analysis

In order to allow assessment of usability and feasibility of the implemented strategies, after a period (approximately 2-4 months depending on nature of strategies implemented) following implementation, a further written survey will be conducted with a purposive, representative sample of 10 trial stakeholders to assess their opinion on the utility and feasibility of the implemented strategies.

Additionally in order to assess the efficacy of the implemented strategies, participant data from the cohort within which mitigating strategies were implemented (please see phase 4) will be requested for analysis from trial results database including: lymph node yield; post-operative morbidity; operation duration; in hospital mortality; R0 resection rate; trial surgeon adherence to the implemented mitigating strategies, and; recruitment of trial surgeons.

GRANT

Successfully awarded the Great British Sasakawa Foundation (GBSF) travel grant, the GBSF funded travel and accommodation for investigator (JB) and supervisor (PB) for a 2 week research trip to Japan for the Phase 3 (expert interview study).