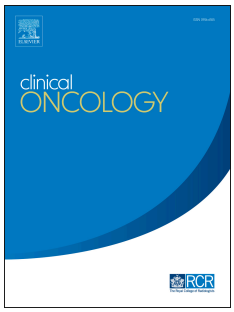
Journal Pre-proof



Quality assurance in multi-modality oesophago-gastric cancer clinical trials: past, present and future perspectives

Jonathan Helbrow, Mary Booth, Bhamini Vadhwana, Richard Adams, Kieran Foley, Christopher J. Peters, Russell Petty, Sarah Gwynne

PII: S0936-6555(25)00226-2

DOI: https://doi.org/10.1016/j.clon.2025.103971

Reference: YCLON 103971

To appear in: *Clinical Oncology*

Received Date: 8 October 2025

Accepted Date: 26 October 2025

Please cite this article as: Helbrow J, Booth M, Vadhwana B, Adams R, Foley K, Peters CJ, Petty R, Gwynne S, Quality assurance in multi-modality oesophago-gastric cancer clinical trials: past, present and future perspectives, *Clinical Oncology*, https://doi.org/10.1016/j.clon.2025.103971.

Title

Quality assurance in multi-modality oesophago-gastric cancer clinical trials: past, present and future perspectives

Authors

Jonathan Helbrow[1,2], Mary Booth[3], Bhamini Vadhwana[4], Richard Adams[5], Kieran Foley[6], Christopher J Peters[4], Russell Petty[7] Sarah Gwynne[2,8,9]

Institutions

1. Department of Oncology, Gloucestershire Hospitals NHS Foundation Trust, Cheltenham, UK
2. National Radiotherapy Trials Quality Assurance (RTTQA) Group, National Institute of Health and Care Research, UK
3. University of Leeds, Leeds, UK
4. Academic Department of Surgery and Cancer, Imperial College London, UK
5. Centre for Trials research, Cardiff University and Velindre Cancer Centre, Cardiff, UK
6. Division of Cancer & Genetics, School of Medicine, Cardiff University, Cardiff, UK
7. Division Of Cancer Research, School of Medicine, University of Dundee, Dundee, UK
8. South West Wales Cancer Centre, Swansea, UK
9. Faculty of Medicine, Life Science and Health, Swansea University, Swansea, UK

Corresponding Author

Jonathan Helbrow
Email: Jonathan.helbrow@nhs.net
ORCID ID: 0000-0003-3865-3966
Department of Oncology
Cheltenham General Hospital
Sandford Road
Cheltenham
GL73 7AR

Author Contributions

JH is the guarantor of the integrity of the entire study.

All authors contributed to study concepts and design, literature search, manuscript preparation and manuscript editing.

Clinical studies, experimental studies / data analysis and statistical analysis are not applicable to this submission.

Acknowledgements

JH and SG are part of the National Radiotherapy Trials Quality Assurance (RTTQA) Group which is funded by the National Institute for Health and Care Research (NIHR). KF receives research funding

from the National Institute of Health and Care Research (NIHR204931), Velindre Cancer Centre and Health and Care Research Wales (HCRW).

Abstract

Clinical trials must ensure the quality of both standard and interventional treatments to rigorously evaluate potential benefits, avoid adverse outcomes, and maintain integrity of results. Quality assurance (QA) endeavours to achieve this and is fundamental to all clinical trial elements, though variation exists between specialties. For radiotherapy (RT) in the UK, the NIHR-funded national Radiotherapy Trials Quality Assurance (RTTQA) group has centralised trial RTQA processes across the RT pathway enabling a robust, consistent, efficient and multidisciplinary approach, replacing piecemeal, trial-by-trial application for QA funding. Meanwhile, the surgical community are moving towards standardised QA processes but are yet to achieve this universally. For SACT, though the importance of QA is recognised, under-reporting persists, and the increasing number and diversity of agents used poses challenges. QA in pathology and radiology is also growing as the complexity of clinical trials increases. Internationally, the EORTC have developed QA processes across domains, but uncertainty and challenge in QA implementation remain. Additionally, while the benefits of trial QA are now recognised, the potential negative effects of QA need to be recognised. Using illustrative examples from contemporary oesophago-gastric cancer studies, we further explore the current status of clinical trial QA cross these specialties.

Key words

Quality assurance, radiotherapy, surgery, systemic therapy, clinical trial, oesophago-gastric cancer

Abbreviations

BICR	Blinded independent central reviews
CTCAE	Common Terminology Criteria for Adverse Events
EORTC	European Organisation of Research and Treatment of Cancer
GHG	Global Harmonization Group
LN	Lymph node
NIHR	National institute for Health and Care Research
OG	Oesophago-gastric
PFS	Progression-free survival
QA	Quality assurance
RCT	Randomised controlled trial
RT	Radiotherapy
RTQA	Radiotherapy quality assurance

RTTQA Radiotherapy Trials Quality Assurance group

SACT Systemic anti-cancer therapy

SOC Standard of care

SQA Surgical quality assurance

Introduction

All interventions, including standard of care, in a randomised clinical trial (RCT) should be delivered to a high standard, to enable rigorous evaluation of potential benefits, avoid adverse outcomes, and ensure the validity of the trial results (1, 2). Clinical trials teams have identified that quality assurance (QA) programmes* are an essential component of high-quality trial delivery (3) and the work to develop the training and delivery of such programmes has become a standard consideration. Furthermore, QA has been shown to result in greater uniformity in clinical practice and an overall improvement in quality of care for centres participating in clinical research, even in the setting of a "negative" trial (4, 5).

Treatment of cancer is increasingly complex, often requiring a multi-modal approach of radiotherapy (RT), surgery, and systemic anti-cancer therapy (SACT) as well as progressively greater input from other specialties such as pathology and radiology. Trial QA varies substantially across these domains, however (3). Here we seek to describe the current landscape of QA in oncology clinical trials, drawing from recent studies in oesophago-gastric (OG) cancer, and referring to challenges and future directions of this landscape.

Radiotherapy

RT quality assurance (RTQA), which seeks to promote protocol adherence and high-quality RT delivery, forms an integral part of modern RT trials, with evidence demonstrating the positive relationship between protocol compliance and outcome across a range of tumour sites (6), and a breadth of strategies encompassing each stage of the RT pathway showing an improvement in standards (5, 7) (for a more detailed description of the RT pathway please see Gwynne *et al* (8)). In the UK, the NIHR-funded national Radiotherapy Trials Quality Assurance (RTTQA) group has, since 2010, centralised trial RTQA processes enabling a robust, consistent, cost-effective and multidisciplinary approach (7, 9). There has been a gradual move from historical, retrospective RTQA, where amending identified variations was not possible, to a prospective or timely-retrospective (typically within two weeks of treatment starting) approach, enabling on-treatment corrections and thus improving standards. Such a shift may have led to more time-, labour- and cost-intensive processes.

Classifying trial-related RTQA by complexity [9] ensures proportionality of the interventions. The pre-accrual and on-trial components of RTQA are described in Table 1.

The SCOPE trials have been the backbone of UK OG RT trials and have included a RTQA programme starting with SCOPE1 in 2008 and each subsequent trial's QA programme informing the next (10). This QA approach increases the reliability of the trial series' results, promotes learning, education and training of clinicians and the community as a whole, and enhances the confidence to implement new techniques in a supported, stepwise and uniform manner (11). In 2001, the US SWOG 9008/Intergroup 0116 RCT comparing chemoradiotherapy after surgery and surgery alone, found that 30% did not complete the CRT because of toxicity and more than 40% of the RT plans had significant errors, leading to some scepticism surrounding the true benefit of RT in the context of optimal surgery (2) and has led to an ongoing debate about the role of RT in stomach cancer ever since.

The need for internationally standardised, collaborative approaches for both RTQA implementation and reporting is well recognised (6, 7, 9, 12). The RTQA Global Harmonisation Group (GHG), of which the UK RTTQA group is a founding member, has sought to do this through publishing consensus guidance on in nomenclature (13), protocol variations definitions (14), and contouring of organs-at-risk (15). A recent UK survey of processes for contouring reviews across 24 trials has shown that, in many areas, approach is consistent and in line with GHG guidance, but differences were seen in reviewer training, assessment of clinical impact of variation, and mitigating inter-reviewer variation, reflecting areas where there is a lack of consensus/guidelines. Competency requirements have since been introduced for new reviewers and criteria to reflect clinical impact are being developed. Next steps include sharing examples of best practice, standardising education and training for new reviewers, developing strategies to mitigate for inter-reviewer variation (10, 16).

Both the UK RTTQA group and GHG continue to adapt to new radiation technologies, such as molecular RT, the use of proton beam and MRI guided RT (7). The role of semi-automated assessment (17) and artificial intelligence (18) may in the future assist with QA workload and reduce inter-reviewer variation (9).

Surgery

In comparison, surgical QA (SQA) initiatives in RCTs are more limited, due to the challenges in heterogeneity of surgical techniques, decision-making, and post-operative management (19). The impact of surgical expertise, which influences clinical outcomes and compromises the validity of the findings, along with learning curves for new techniques (which may be poorly understood), are often not considered. The European Organisation of Research and Treatment of Cancer (EORTC) identified poor SQA measures in RCTs from 1980 to 2013 (20) and a recent review demonstrated no surgical QA for 96% of interventional trials (21). In particular, the variability in lymph node (LN) dissection in OG trials has cast doubt on the reported outcomes. The Dutch D1 vs D2 trial demonstrated 84% of D2 dissection gastrectomies were suboptimal (1), while the US SWOG 9008/Intergroup 0116 RCT identified 54% had an inadequate LN resection rate (2). The Medical Research Council has emphasised the need for optimal surgical performance to reach clinically translatable outcomes (22). Although studies provide videotape and booklets with step-by-step instructions on the operative approach, there is inadequate monitoring during surgery which needs to be addressed. The IDEAL framework (23) for robotic surgery recognises the need for a consensus on the markers of adequate quality of the surgery to allow definitive RCTs to be performed.

Various methods of SQA have been employed. The Jadad scale (24) for QA in RCTs was deemed simplistic and lacking accurate surgical quality indicators (25). Newer platforms such as the Consolidated Standards of Reporting Trials of Non-Pharmacological Treatment (CONSORT-NPT) (26) and Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) (27) have endorsed detailed reporting of interventions for replicability in practice.

However, few studies reported video assessments or review of operative notes as a quality control measure (2, 28, 29). OG RCTs have shown reduced variation in LN harvest and adjusted in-hospital mortality, with surgeon credentialing by operative note assessment and performance monitoring (30). The Neo-AEGIS trial selected high volume surgeons to ensure SQA (31). Most recently, the ROMIO trial presented a strong example of implementing a QA tool to monitor technical performance and quality of oesophagectomies comparing laparoscopic and open approaches. The Objective Structured Assessment of Technical Skills (OSATS) or Hierarchical Task Analysis for Oesophagectomy (HTA-O) were used with intraoperative photography and/or videography (32). The ROMIO trial ensured timely

feedback to the trial surgeons and group sessions to show examples of good technique and areas for improvement (32) .

Overall, reviews show a lack of transparency in QA reporting with 18% reporting entry criteria, 29% providing standardisation of procedures, and 28% undergoing monitoring (33). For RCTs, which are often multi-centre, there is a balance between standardising practice and understanding what is realistically achievable via a pragmatic design.

Recent suggestions of a standardised approach to QA in pragmatic surgical oncology trials aim to ensure replicability, reliability and data integrity. Butterworth *et al* described the first international expert consensus on QA in OG surgical oncology trials, identifying three critical components as shown in Figure 1 (34).

Quality-controlled registries have been recommended to ensure data transparency and monitoring study delivery (35). Surgical QA tools have been developed for gastrectomies and Ivor-Lewis oesophagectomies (36, 37). A neutral party should perform monitoring to reduce the risk of bias and ensure standards are met. The ADDICT trial, currently open to recruitment, is achieving surgical standardisation by video operation manuals and blind review of photo documentation of the surgical field post-resection (38). Operation manuals detail authorised operative steps and those not allowed; a widely accepted method to document SQA in surgical quality (39). Future surgical oncology trials can utilise this framework to address challenges in good SQA (34). There is no consensus on the minimum caseload required for proficiency, and therefore a statement should be included in SQA protocols and assessment of real-time videos should be considered prior to recruiting trial surgeons. Park *et al* demonstrated improved clinical outcomes including operative time and lymph node retrieval in the second 100 robotic cases compared to the first 100 (40).

The recent increase in adoption of robotic platforms has raised new SQA challenges, with valid concerns about the effect this may have on maintaining QA in surgical clinical trials. Varying designs may influence the ease and speed of performing the surgery but does not affect the operative steps and underlying principles of oncological resection. For SQA, operative steps must be defined, and the surgeon must be credentialed in performing robotic

surgery. Factors specific to the system, such as speed and haptic feedback need to be considered in the analysis.

Developments have seen the EORTC establish the SURCARE collaborative platform in 2016 with the European Society of Surgical Oncology and the Japanese Clinical Oncology Group encouraging surgeon credentialing, technical standardisation and central committee review to ensure QA for surgical oncology trials (3, 41). Careful consideration of SQA by achieving standardised protocols will inform more robust trial outcomes. This is critical as outcomes of high quality RCTs inform changes in the delivery of surgical care.

Systemic Anti-Cancer Therapy

SACT QA is important in clinical trials to assess the validity and reproducibility of trial results, and their applicability to individual scenarios in real world clinical settings. Many elements of QA are commonly incorporated into SACT trial protocols. Such elements include: proportion of patients completing specific SACT regimen; proportion of patients experiencing toxicities according to Common Terminology Criteria for Adverse Events (CTCAE); pharmacy accountability; and SACT prescription and administration records. SACT records may include details of preparation and administration of intravenous medication, as well as compliance with oral SACT supportive medications. However, it has been observed that these QA measures are inconsistently collected and reported. Furthermore, where specific QA measures have been recorded according to a trial protocol, these are not always comprehensively presented alongside the trial results. The recently published MATTERHORN phase 3 RCT trial of perioperative FLOT chemotherapy +/- Durvalumab immunotherapy for gastric and junctional cancer provides a good example of where details of SACT compliance, toxicity and treatment delays have been detailed in the main publication and supplementary materials, guidance on dose adjustments and treatment preparation and administration offered in the treatment protocol, as well as the important interplay between pathology QA in disease response assessment (blinded central review) (42).

Publications of SACT RCTs are generally reported according to CONSORT (Consolidated Standards of Reporting Trials) guidelines, a set of regularly updated consensus-based reporting standards (43, 44). In addition, the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) Statement refers to a set of guidelines of items that should be included in trial protocols (27). CONSORT and SPIRIT guidance both detail checklists of the minimum information, which should be included in trial reporting, although these checklists focus on methodological aspects of trial design and participant flow rather than SACT delivery QA. SACT delivery safety indicators are key measures of QA not fully captured by CONSORT and SPIRIT guidance. Like surgery, the Jadad scale (24) has not found traction and is not routinely used in SACT clinical trials, likely as its simplicity prevents comprehensive QA assessment. The EORTC uses QA questionnaires as part of EORTC membership assessment and on-site visits prior to commencement and as part of ongoing monitoring of EORTC clinical trials (3), which covers aspects of prescription, rounding procedures, preparation and administration. The selection of sites is guided by the local provisions to deliver the medical treatment as well as the track record of sites in delivering clinical research (45).

Different types of SACT may necessitate different QA measures. Chemotherapy, small molecule inhibitors, and immunotherapy have different methods of delivery and toxicity profiles. Using dose reductions as a QA measure may be useful in chemotherapy studies but meaningless in studies of flat-dosed immunotherapies; autoimmune toxicities are an important safety measure in immunotherapy studies, but of lesser relevance in investigation of small-molecule drugs. Additionally, differences between curative and palliative treatment approaches may require different QA measures.

Whilst not routinely performed in SACT clinical trials, there may be a role for peer review in the QA of SACT trials where there are degrees of subjectivity, for example if (and to what degree) to dose-reduce SACT. Later phase trials typically provide some recommendations, but direct investigators to apply local practice and clinical judgement. Although such peer-review processes are not routine practice in SACT trials, translation of established peer review practice in RT and surgical trials could be helpfully applied to some aspects of SACT

QA in clinical trials, with implementation of peer-review for "grey-area" scenarios where physician interpretation affects decision-making.

A QA framework could improve the consistency and interpretation of such trials, but this should not be a one-size-fits-all checklist and nuances of different SACT therapies and treatment settings should be taken into consideration. One approach might be to create modular recommendations for each SACT subtype (e.g. cytotoxic chemotherapy, immunotherapy, small molecules etc.) Such a framework should be evidence-based, validated, and updated as appropriate, created using robust methodology and stakeholder engagement. This would then provide a firm foundation for increasing awareness, improving consistency and setting the standard for SACT clinical trial QA.

Other areas

Multi-centre trials also create challenges for two other critical disciplines – pathology and radiology, as detailed in the five pillars of the QA programme at the EORTC (3). Röcken *et al.* undertook a literature review of QA of pathology in clinical trials in 2016 and found no references to participation in an external QA programme for testing of several biomarkers that influence choice of systemic therapy and explores the reasons why such credentialing may be necessary (46). In the UK, the UK NEQAS provides External Quality Assessment/Proficiency Testing for all major aspects of clinical laboratory testing addresses this (47). In addition, the Royal College of Pathologists sets out guidance for reporting, but a UK survey found that preparation and histopathological assessment of OG specimens varied significantly across institutions, with only five out of 32 units who responded 'meeting' or 'exceeding' those guidelines. There was wide variation in how centres defined positive (R1) margins, and how margins and LNs were assessed (48). This led to the HERO group setting out guidelines for OG cancer specimen preparation and assessment, to provide maximum benefit for patient care and standardize reporting to allow benchmarking and improvement of surgical quality (49). Central pathological evaluation, where biomarkers may guide treatment, or where pathological response is a primary or secondary endpoint, provides another QA strategy in this domain.

For radiology, quality of both staging (required for eligibility) and tumour response assessment are essential. For the former, the Union for International Cancer Control (UICC) TNM classification of malignant tumours, and the latter, the RECIST (Response Evaluation Criteria in Solid Tumours) criteria (50), provide globally recognised standards used universally in both routine practice and the clinical trial setting, markedly improving consistency and reproducibility internationally in these domains. However, regarding response assessment, RECIST has limitations, particularly in OG cancer, where the disease is difficult to accurately measure on CT alone. Novel agents also present new challenges because of atypical treatment response patterns, for example those seen with immunotherapy (i.e. pseudoprogression) (51) and an updated irRECIST criteria has therefore been proposed (52).

As with pathology, many trials, including the UK PICCOS trial (exploring the role of pressurised intraperitoneal chemotherapy in stomach, colon, and ovarian cancers) (53), include central radiology review. This strategy has been adopted to ensure the validity of the primary endpoint of peritoneal progression free survival (PFS) in PICCOS, given the difficulties of measuring peritoneal disease response on CT scans (54). Blinded independent central reviews (BICR), where local assessments are reviewed by a central committee, provide important QA, but can add complexity to trial design and increase costs. Radiology QA requirements should be defined for each trial. One example might be that BICRs are mandated for trials with PFS as an endpoint, but local audit may be appropriate for others.

As with RT, the radiology equipment is also subject to QA checks. SCOPE2, a definitive chemoradiotherapy trial for oesophageal cancer utilised PET CT to aid treatment decisions, provided a PET imaging manual and appointed a radiology lead and core laboratory for QA to ensure the standards of PET imaging across sites (55).

Reproducibility in routine clinical practice

Robust QA during a clinical trial ensures the integrity of the trial results, but when these treatments are implemented in routine clinical practice, there is potentially a risk that the absence of such rigorous QA will not result in the same outcomes. In reality, there is already a certain level of QA in place at sites across the country, through national agencies and initiatives (e.g. IRMER for ionising radiation

(https://www.gov.uk/government/publications/ionising-radiation-medical-exposure-regulations-2017-guidance), centralisation of OG surgery (https://data.parliament.uk/DepositedPapers/Files/DEP2010-1407/DEP2010-1407.pdf), National Oesophago-gastric Cancer Audit (https://www.nogca.org.uk/), local morbidity and mortality meetings, Blueteq for SACT prescribing (https://blueteq.com/commissioner-high-cost-drugs-new/). In addition, a positive association between a centre's post-surgical and survival outcomes and its research participation has been demonstrated in colorectal cancer (4). In OG cancer, participation in the SCOPE trials, supported by the QA programme, has been the driver for UK centres implementing new technologies, increasing quality and access to world-leading RT techniques (5). The experience of contouring in UK clinical trials has informed the Royal College of Radiologists guidance on peer review, which has been adopted across UK centres (56). A wider discussion of how results of RCTs relate to real world populations is beyond the scope of this paper.

Impact on study set up and costs

The pressures sites experience in setting up clinical trials is high in the current climate. The RTQA group, as part of work done in the UK to increase efficiency of trial set-up has developed processes to address this (57). Recommendations include time limits for feedback and introduction of streamlining, where pre-accrual QA requirements are modified according to past participation in a similar trial (57).

The costs of QA require acknowledgement. The centralisation of the RTTQA group in 2010, with funding from the NIHR, replaced the need for piecemeal, trial-by-trial application for funding for QA component (7) and centres are now able to claim clinical trial RT QA as a NHS service support cost (57). Currently there is no such centralised funding for surgical or non-commercial trial SACT QA. There is also a need for the cost effectiveness of QA to be addressed. Weber *et al.* explored the effect of differing levels of RTQA on outcome in a simulated head and neck study, specifically looking at quality-adjusted life years (58). They found that despite a 14-fold increase in the cost of the RTQA programme from basic to complex, the higher level of RTQA input was cost effective, as it resulted in better patient outcomes in terms of tumour control and overall survival. Despite the uncertainties of

model, it does act as proof of concept. For all QA programmes, improving efficiency through adjusting the intensity of the QA according to the complexity of the study is important (9).

Conclusion

Trial outcomes are known to be impacted by the quality of the treatments delivered within it, but ensuring quality across centres in a multi-centre trial is challenging. QA programmes for all trial components are essential to ensure that variations from protocol do not lead to poor outcome and invalidate trial results. Learning from QA process across disciplines and across international organisations can help and inform best practice both in clinical trials and routine practice, and further enhancement of QA programmes is likely to lead to improved patient care in both settings.

*Quality assurance (QA): the systematic and independent examination of all trial-related activities and documents. These audits determine whether the evaluated activities were appropriately conducted and that the data were generated, recorded, analysed, and accurately reported according to protocol, standard operating procedures , and good clinical practices (59).

References

1. Bunt AM, Hermans J, Boon MC, van de Velde CJ, Sasako M, Fleuren GJ, et al. Evaluation of the extent of lymphadenectomy in a randomized trial of Western- versus Japanese-type surgery in gastric cancer. J Clin Oncol. 1994;12(2):417-22.
2. Macdonald JS, Smalley SR, Benedetti J, Hundahl SA, Estes NC, Stemmermann GN, et al. Chemoradiotherapy after surgery compared with surgery alone for adenocarcinoma of the stomach or gastroesophageal junction. N Engl J Med. 2001;345(10):725-30.
3. Members of EORTC QAC; Working Groups of Scientific Experts. Multidisciplinary quality assurance and control in oncological trials: Perspectives from European Organisation for Research and Treatment of Cancer (EORTC). Eur J Cancer. 2017;86:91-100.
4. Downing A, Morris EJ, Corrigan N, Sebag-Montefiore D, Finan PJ, Thomas JD, et al. High hospital research participation and improved colorectal cancer survival outcomes: a population-based study. Gut. 2017;66(1):89-96.
5. Gwynne S, Higgins E, Poon King A, Radhakrishna G, Wills L, Mukherjee S, et al. Driving developments in UK oesophageal radiotherapy through the SCOPE trials. Radiat Oncol. 2019;14(1):26.
6. Jomy J, Sharma R, Lu R, Chen D, Ataalla P, Kaushal S, et al. Clinical impact of radiotherapy quality assurance results in contemporary cancer trials: a systematic review and meta-analysis. Radiotherapy and Oncology. 2025;207:110875.
7. Miles E, Wadsley J, Diez P, Patel R, Gwynne S. The National Radiotherapy Trials Quality Assurance Group - Driving up Quality in Clinical Research and Clinical Care. Clin Oncol (R Coll Radiol). 2024;36(5):273-7.

8. Gwynne S, Staffurth J. Principles of cancer treatment by radiotherapy. Surgery (Oxford). 2012;30(4):191-3.
9. Brooks C, Miles E, Hoskin PJ. Radiotherapy trial quality assurance processes: a systematic review. Lancet Oncol. 2024;25(3):e104-e13.
10. Helbrow J, Lewis G, Hurt C, Radhakrishna G, Nicholas O, Hawkins MA, et al. Radiotherapy Quality Assurance in the SCOPE2 Trial: What Lessons can be Learned for the Next UK Trial in Oesophageal Cancer? Clin Oncol (R Coll Radiol). 2024;38:103735.
11. Bridges S, Thomas B, Radhakrishna G, Hawkins M, Holborow A, Hurt C, et al. SCOPE 2 – Still Answering the Unanswered Questions in Oesophageal Radiotherapy? SCOPE 2: a Randomised Phase II/III Trial to Study Radiotherapy Dose Escalation in Patients with Oesophageal Cancer Treated with Definitive Chemoradiation with an Embedded Phase II Trial for Patients with a Poor Early Response using Positron Emission Tomography/Computed Tomography. Clinical Oncology. 2022;34(7):e269-e80.
12. Cox S, Cleves A, Clementel E, Miles E, Staffurth J, Gwynne S. Impact of deviations in target volume delineation - Time for a new RTQA approach? Radiother Oncol. 2019;137:1-8.
13. Melidis C, Bosch WR, Izewska J, Fidarova E, Zubizarreta E, Ulin K, et al. Global harmonization of quality assurance naming conventions in radiation therapy clinical trials. Int J Radiat Oncol Biol Phys. 2014;90(5):1242-9.
14. Global Harmonization Group. Protocol variations definitions: ; 2019 [Available from: https://rtqaharmonization.org/resources/.
15. Mir R, Kelly SM, Xiao Y, Moore A, Clark CH, Clementel E, et al. Organ at risk delineation for radiation therapy clinical trials: Global Harmonization Group consensus guidelines. Radiotherapy and Oncology. 2020;150:30-9.
16. Satram V, Helbrow J, Miles E, Diez P, Gwynne S. Contouring quality assurance for UK clinical trials – challenges and opportunities. Radiother Oncol. 2025;206:S2181–S2182.
17. Gwynne S, Spezi E, Wills L, Nixon L, Hurt C, Joseph G, et al. Toward semi-automated assessment of target volume delineation in radiotherapy trials: the SCOPE 1 pretrial test case. Int J Radiat Oncol Biol Phys. 2012;84(4):1037-42.
18. Simon L, Robert C, Meyer P. Artificial intelligence for quality assurance in radiotherapy. Cancer Radiother. 2021;25(6-7):623-6.
19. Lilford R, Braunholtz D, Harris J, Gill T. Trials in surgery. Br J Surg. 2004;91(1):6-16.
20. Tanis E, Caballero C, Collette L, Verleye L, den Dulk M, Lacombe D, 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. Eur J Surg Oncol. 2016;42(8):1115-22.
21. Robinson NB, Fremes S, Hameed I, Rahouma M, Weidenmann V, Demetres M, et al. Characteristics of Randomized Clinical Trials in Surgery From 2008 to 2020: A Systematic Review. JAMA Netw Open. 2021;4(6):e2114494.
22. Cuschieri A, Fayers P, Fielding J, Craven J, Bancewicz J, Joypaul V, et al. Postoperative morbidity and mortality after D1 and D2 resections for gastric cancer: preliminary results of the MRC randomised controlled surgical trial. The Surgical Cooperative Group. Lancet. 1996;347(9007):995-9.
23. Marcus HJ, Ramirez PT, Khan DZ, Layard Horsfall H, Hanrahan JG, Williams SC, et al. The IDEAL framework for surgical robotics: development, comparative evaluation and long-term monitoring. Nature Medicine. 2024;30(1):61-75.
24. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials. 1996;17(1):1-12.
25. Olivo SA, Macedo LG, Gadotti IC, Fuentes J, Stanton T, Magee DJ. Scales to assess the quality of randomized controlled trials: a systematic review. Phys Ther. 2008;88(2):156-75.

26. Boutron I, Altman DG, Moher D, Schulz KF, Ravaud P. CONSORT Statement for Randomized Trials of Nonpharmacologic Treatments: A 2017 Update and a CONSORT Extension for Nonpharmacologic Trial Abstracts. Ann Intern Med. 2017;167(1):40-7.
27. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med. 2013;158(3):200-7.
28. Balch CM, Durant JR, Bartolucci AA. The impact of surgical quality control in multi-institutional group trials involving adjuvant cancer treatments. Ann Surg. 1983;198(2):164-7.
29. Degiuli M, Sasako M, Ponti A, Vendrame A, Tomatis M, Mazza C, et al. Randomized clinical trial comparing survival after D1 or D2 gastrectomy for gastric cancer. Br J Surg. 2014;101(2):23-31.
30. Markar SR, Wiggins T, Ni M, Steyerberg EW, Van Lanschot JJ, Sasako M, et al. Assessment of the quality of surgery within randomised controlled trials for the treatment of gastro-oesophageal cancer: a systematic review. Lancet Oncol. 2015;16(1):e23-31.
31. Reynolds JV, Preston SR, O'Neill B, Baeksgaard L, Griffin SM, Mariette C, et al. ICORG 10-14: NEOadjuvant trial in Adenocarcinoma of the oEsophagus and oesophagoGastric junction International Study (Neo-AEGIS). BMC Cancer. 2017;17(1):401.
32. Blencowe NS, Skilton A, Gaunt D, Brierley R, Hollowood A, Dwerryhouse S, 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.
33. Blencowe NS, Boddy AP, Harris A, Hanna T, Whiting P, Cook JA, et al. Systematic review of intervention design and delivery in pragmatic and explanatory surgical randomized clinical trials. Br J Surg. 2015;102(9):1037-47.
34. Butterworth JW, Boshier PR, Mavroveli S, Reynolds JV, Kim YW, Hanna GB. A proposed framework of strategies to overcome challenges to surgical quality assurance in oncology trials (SQA-Onc.). Eur J Surg Oncol. 2025;51(5):109593.
35. Öfner D. Quality assurance in surgery—a moral obligation. European Surgery. 2024;56(3):110-5.
36. Harris A, Butterworth J, 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. 2022;275(1):121-30.
37. Harris A, Butterworth JB, Boshier PR, Mavroveli S, Vadhwana B, Peters CJ, et al. Development of a reliable surgical quality assurance tool for gastrectomy in oncological trials. Gastric Cancer. 2024;27(4):876-83.
38. Kim J, Eom BW, Nam BH, Yoon HM, Kim YW. 423. 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):S162.
39. Blencowe NS, Mills N, Cook JA, 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.
40. Park JY, Kim YW, Ryu KW, Eom BW, Yoon HM, Reim D. Emerging Role of Robot-assisted Gastrectomy: Analysis of Consecutive 200 Cases. J Gastric Cancer. 2013;13(4):255-62.
41. Evrard S, Audisio R, 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):911-2.
42. Janjigian YY, Al-Batran SE, Wainberg ZA, Muro K, Molena D, Van Cutsem E, et al. Perioperative Durvalumab in Gastric and Gastroesophageal Junction Cancer. N Engl J Med. 2025;393(3):217-30.
43. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. J Pharmacol Pharmacother. 2010;1(2):100-7.
44. Butcher NJ, Monsour A, Mew EJ, Chan AW, Moher D, Mayo-Wilson E, et al. Guidelines for Reporting Outcomes in Trial Reports: The CONSORT-Outcomes 2022 Extension. Jama. 2022;328(22):2252-64.

45. de Balincourt C, Lacombe D, 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. 2017;86:91-100.
46. Röcken C. Quality assurance in clinical trials--the role of pathology. Virchows Arch. 2016;468(1):83-92.
47. UK NEQAS. https://ukneqas.org.uk/programmes/?keyword=pathology&centre=0&location=0&speciality=0#find-a-programme-block. : DS Creative; [Available from: https://ukneqas.org.uk/programmes/?keyword=pathology&centre=0&location=0&speciality=0#find-a-programme-block. .
48. Pucher PH, Green M, Bateman AC, Underwood TJ, Maynard N, Allum WH, et al. Variation in histopathological assessment and association with surgical quality indicators following oesophagectomy. Br J Surg. 2021;108(1):74-9.
49. Pucher PH, Allum WH, Bateman AC, Green M, Maynard N, Novelli M, et al. Consensus recommendations for the standardized histopathological evaluation and reporting after radical oesophago-gastrectomy (HERO consensus). Dis Esophagus. 2021;34(8).
50. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228-47.
51. Somarouthu B, Lee SI, Urban T, Sadow CA, Harris GJ, Kambadakone A. Immune-related tumour response assessment criteria: a comprehensive review. Br J Radiol. 2018;91(1084):20170457.
52. Bohnsack O, Hoos A, Ludajic K. Adaptation of the Immune Related Response Criteria: Irrecist. Annals of Oncology. 2014;25:iv369.
53. Jones S, Hudson E, Adams R, Murphy J, Peters C, Gwynne S, et al. #1018 Launching the PICCOS trial: PIPAC in cancers of the colon, ovary and stomach2023. A329.2-A p.
54. Fleur Jones SE, Hudson E, Adams R, Murphy J, Peters C, Gwynne S, et al. #1018 Launching the PICCOS trial: PIPAC in cancers of the colon, ovary and stomach. International Journal of Gynecologic Cancer. 2023;33(Suppl 3):A329-A.
55. Mukherjee S, Hurt CN, Adams R, Bateman A, Bradley KM, Bridges S, et al. Efficacy of early PET-CT directed switch to carboplatin and paclitaxel based definitive chemoradiotherapy in patients with oesophageal cancer who have a poor early response to induction cisplatin and capecitabine in the UK: a multi-centre randomised controlled phase II trial. EClinicalMedicine. 2023;61:102059.
56. Radiologists; RCo. Radiotherapy dose fractionation: Fourth Edition 2024 [Available from: https://www.rcr.ac.uk/our-services/all-our-publications/clinical-oncology-publications/radiotherapy-dose-fractionation-fourth-edition/.
57. Hanna CR, Lynskey DM, Wadsley J, Appleyard SE, Anwar S, Miles E, et al. Radiotherapy Trial Set-up in the UK: Identifying Inefficiencies and Potential Solutions. Clin Oncol (R Coll Radiol). 2020;32(4):266-75.
58. Weber DC, Hurkmans CW, Melidis C, Budach W, Langendijk JH, Peters LJ, et al. Outcome impact and cost-effectiveness of quality assurance for radiotherapy planned for the EORTC 22071-24071 prospective study for head and neck cancer. Radiother Oncol. 2014;111(3):393-9.
59. European Medicines Agency. ICH Guideline for good clinical practice E6 (R2). 2016.
60. Clark CH, Aird EG, Bolton S, Miles EA, Nisbet A, Snaith JA, et al. Radiotherapy dosimetry audit: three decades of improving standards and accuracy in UK clinical practice and trials. British Journal of Radiology. 2015;88(1055).
61. Gwynne S, Gilson D, Dickson J, McAleer S, Radhakrishna G. Evaluating Target Volume Delineation in the Era of Precision Radiotherapy: FRCR, Revalidation and Beyond. Clinical Oncology. 2017;29(7):436-8.

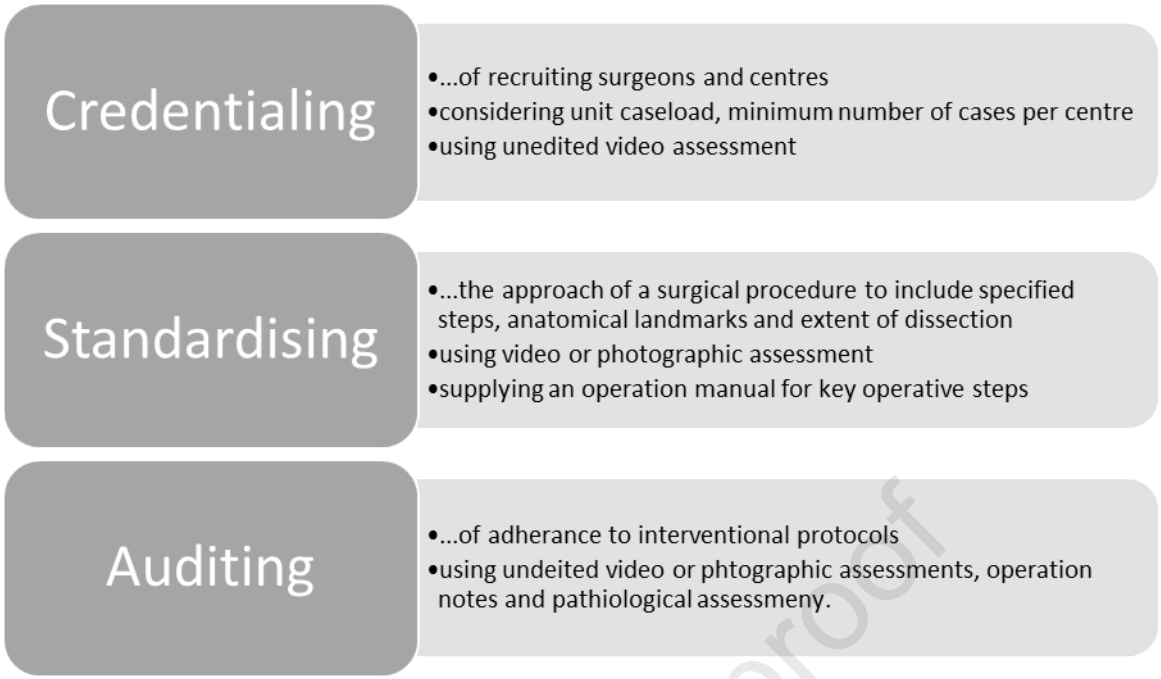
455 62. Segedin B, Petric P. Uncertainties in target volume delineation in radiotherapy - are they
456 relevant and what can we do about them? Radiol Oncol. 2016;50(3):254-62.
457 63. Gwynne S, Spezi E, Sebag-Montefiore D, Mukherjee S, Miles E, Conibear J, et al. Improving
458 radiotherapy quality assurance in clinical trials: assessment of target volume delineation of the pre-
459 accrual benchmark case. Br J Radiol. 2013;86(1024):20120398.
460

461 Figure legends

462 Table 1. Pre-trial on On-trial RTQA requirements (7, 13, 60-63)

463 Figure 1. Critical components identified for surgical QA in OG trials(34).

Pre-accrual RTQA	On-trial RTQA
Facility Questionnaire • *Details of a centre's equipment, procedures and available personnel is collected* **Dosimetry audit** • *Dose measurements are taken under reference conditions by an independent clinical scientist* **RPGD** • *A Radiotherapy Planning Guidance Document is provided, giving detailed instructions for each part of the RT pathway, including contouring and planning* **Benchmarking** • *Centres are required to complete a contouring and planning case prior to recruitment. Submissions are compared with a 'gold standard'. This ensures processes are of satisfactory standard, provides opportunity for learning, and highlights protocol ambiguities*	**Contouring and planning Individual case reviews (ICRs)** • *Contouring and planning ICRs are undertaken for recruited patients* • *This can be prospective (prior to RT start) or retrospective (after RT start), and permits ongoing maintenance of standards* • *The requirements for each trial varies depending on trial and procedure complexity* • *ICRs are usually undertaken by member(s) of the Trial Management Group or RTQA team* • *Streamlining of these processes is sometimes applied if a centre has participated in a preceding, related clinical trial*



Highlights

- Clinical trial quality assurance (QA) is essential to ensure validity of results
- Radiotherapy QA is well established, surgical QA is growing but not yet universal
- QA in systemic therapy is challenging and currently under-reported
- The importance of pathology & radiology QA also needs recognition
- Impact on trial set up, delivery and clinical practice requires consideration

Authorship Contributions

1 guarantor of integrity of the entire study	Jonathan Helbrow Sarah Gwynne
2 study concepts and design	All authors
3 literature research	All authors
4 clinical studies	NA
5 experimental studies / data analysis	NA
6 statistical analysis	NA
7 manuscript preparation	All authors
8 manuscript editing	All authors

Declaration of interests

☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

JH is a clinician part-employed by the NIHR RTTQA group. SG is the lead clinician for the NIHR RTTQA group. We recognise this may act as a competing interest through employment. The remaining authors declare no competing interests.

If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.