

Quality indicators in surgical oncology: systematic review of measures used to compare quality across hospitals

Megan McLeod^{1,2,*} , Kari Leung³, C. S. Pramesh⁴, Peter Kingham⁵, Miriam Mutebi⁶, Julie Torode⁷, Andre Ilbawi⁸, Jade Chakowa⁹, Richard Sullivan¹⁰ and Ajay Aggarwal¹¹ 

¹Department of Health Policy, London School of Economics and Political Science, London, UK

²Department of Otolaryngology—Head & Neck Surgery, Vanderbilt University Medical Center, Nashville, Tennessee, USA

³Department of Oncology, Guy's & St Thomas' NHS Trust, London, UK

⁴Tata Memorial Centre, Homi Bhabha National Institute, Mumbai, India

⁵Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, New York, USA

⁶Department of Surgery, Aga Khan University, Nairobi, Kenya

⁷Institute of Cancer Policy, Centre for Cancer, Society & Public Health, King's College London, London, UK

⁸Department of Universal Health Coverage, World Health Organization, Geneva, Switzerland

⁹City Cancer Challenge, Geneva, Switzerland

¹⁰Institute of Cancer Policy, Global Oncology Group, Centre for Cancer, Society & Public Health, King's College London, London, UK

¹¹Department of Health Services Research and Policy, London School of Hygiene and Tropical Medicine, London, UK

*Correspondence to: Megan McLeod, Department of Otolaryngology—Head & Neck Surgery, Vanderbilt University Medical Center, 1211 Medical Center Drive, Nashville, Tennessee, 37232, USA (e-mail: Megan.mcleod@vumc.org)

Abstract

Background: Measurement and reporting of quality indicators at the hospital level has been shown to improve outcomes and support patient choice. Although there are many studies validating individual quality indicators, there has been no systematic approach to understanding what quality indicators exist for surgical oncology and no standardization for their use. The aim of this study was to review quality indicators used to assess variation in quality in surgical oncology care across hospitals or regions. It also sought to describe the aims of these studies and what, if any, feedback was offered to the analysed groups.

Methods: A literature search was performed to identify studies published between 1 January 2000 and 23 October 2023 that applied surgical quality indicators to detect variation in cancer care at the hospital or regional level.

Results: A total of 89 studies assessed 91 unique quality indicators that fell into the following Donabedian domains: process indicators (58; 64%); outcome indicators (26; 29%); structure indicators (6; 7%); and structure and outcome indicators (1; 1%). Purposes of evaluating variation included: identifying outliers (43; 48%); comparing centres with a benchmark (14; 16%); and supplying evidence of practice variation (29; 33%). Only 23 studies (26%) reported providing the results of their analyses back to those supplying data.

Conclusion: Comparisons of quality in surgical oncology within and among hospitals and regions have been undertaken in high-income countries. Quality indicators tended to be process measures and reporting focused on identifying outlying hospitals. Few studies offered feedback to data suppliers.

Introduction

The Institute of Medicine (IoM) posits six dimensions of healthcare quality; namely, that it should be safe, timely, effective, efficient, equitable, and patient-centred (STEEEP)¹. It is estimated that over 9 million cancer cases worldwide had an indication for surgery in 2018 and that number is projected to increase to more than 13 million by 2040². In cancer care, low-quality treatment leads to poor survival rates, higher rates of morbidity, and higher costs both to the individual and society^{3,4}. Cancer care typically requires STEEEP coordination of several modalities of treatment (that is surgery, chemotherapy, and radiation therapy) to treat the cancer, prevent recurrence, and avoid toxicities. The growing burden of cancer, the complexity of treatment pathways, and the need for multidisciplinary coordination means that developing

indicators that adequately measure all aspects of high-quality oncological care delivery is essential.

Donabedian framed measures of quality according to whether they assess the maintenance of the organizational and physical structure necessary for the provision of quality care, the execution of processes that are known to contribute to high-quality care, or the achievement of outcomes consistent with provision of quality care⁵. These categories are typically referred to as structure, process, and outcome measures respectively. Some models have gone further, leading to 12 domains of quality to appraise cancer care⁶.

With the exponential growth in data collected via electronic medical records, computerized billing, and online disease-specific registries has come an opportunity to measure, report, and compare quality on both a larger and more granular scale

Received: October 17, 2023. Revised: November 16, 2023. Accepted: December 17, 2023

© The Author(s) 2024. Published by Oxford University Press on behalf of BJS Foundation Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

than ever before. In the USA, the American College of Surgeons maintains the National Surgical Quality Improvement Program (NSQIP)⁷. The Netherlands and the UK conduct quality audits of specific diseases, especially cancers^{8,9}.

While most academic studies in the field of quality measurement and improvement have focused on developing and validating quality indicators, analyses have begun to focus on comparing quality across organizations or regions of care provision¹⁰. However, no study has yet explored which quality indicators are being used to assess quality in surgical oncology, which measures meaningfully detect variation between organizations or regions, and what, if anything, comes from identification of this variation¹¹.

This aim of this study was to conduct a review of the quality indicators being routinely used to assess variation in quality in surgical oncology care across hospitals or regions with a view to inform standardization of performance measurement in surgical oncology. Second, it sought to identify the sources of data used in these analyses and the entities responsible for producing and maintaining the databases. Lastly, it sought to describe the aims identified in studies seeking to establish hospital or regional variation in quality indicators and what, if any, feedback the authors of these studies offer to the hospitals or regions they analysed to support quality improvement.

Methods

This review was exempt from Research Ethics Committee review per institutional Research Ethics Policy and Procedures and was approved at the departmental level by the Department of Health Policy.

Search strategy

This study followed the PRISMA and AMSTAR 2 ('a measurement tool to assess systematic reviews, version 2') guidelines^{12,13}. A literature search was performed to identify studies published between 1 January 2000 and 31 October 2023 that utilized quality indicators to assess surgical oncology care. The search strategy was created using the population/intervention/comparator/outcomes/study design (PICOS) framework outlined in the Cochrane Handbook (Table 1)¹⁴. The databases searched were Embase via the Ovid interface and MEDLINE via the PubMed interface. The full search strategy is available in Appendix S1.

Study selection

Inclusion and exclusion criteria were developed according to the published literature on quality in cancer care in consultation with a clinical oncologist and a surgeon who are experts in the

Table 1 Outline of population/intervention/comparator/outcomes/study design ('PICOS')

Population

Hospitals or regions providing cancer care to adult patients (at least 18 years old)

Intervention

Comparing quality of surgical oncology care

Comparator

Other hospitals, regions, or regional/national/international averages

Outcomes

Performance on quality indicators

Study design

Cohort studies

measurement of care quality and experienced in conducting systematic reviews. These criteria are outlined in Table 2.

Original investigations of adult patients (at least 18 years old) that applied surgical quality indicators to detect variation in cancer care at the centre (hospital) or regional level, including those that used quality indicators to benchmark individual centres against a national average, were included.

Studies that only developed or validated quality indicators (and did not compare centres or regions), that only associated achievement of quality indicators with patient or surgical factors (for example co-morbidities, surgical approaches, and preoperative or postoperative interventions), or that solely focused on surgical pathology quality indicators were excluded.

The abstracts and titles of studies retrieved from the literature search were screened for inclusion in full-text review by M.Mc. and K.L. Full-text records were then evaluated against inclusion criteria. Senior reviewers (A.A. and R.S.) were consulted to resolve uncertainties.

Data extraction

Data were extracted by M.Mc. and K.L. using a predefined and piloted data collection form regarding study characteristics (country, number of hospitals and patients included, tumour type(s), tumour stage, source(s) of data, entity that produced or maintained the database, aims of the study, and method of data visualization), quality indicators and their definitions, characteristics of each quality indicator, and the intended use of the study's results. Quality indicator definitions were recorded as reported by the investigators of each study. Uncertainties were resolved by consensus meetings with senior reviewers (A.A. and R.S.). Screening and data extraction were conducted using Covidence software.

Categorization of results

Quality indicators were categorized according to aspect of the care pathway (preoperative, intraoperative, or postoperative), Donabedian domain (structure, process, or outcome), and Quality Indicators in Cancer Care (QICC) domain^{5,6}. This was intended to aid in mapping quality indicators to which provider(s) and resources were required for measurement and/or achievement of the quality indicators. QICC domains were defined according to the description of Chiew *et al.*⁶ with one exception. Two domains—safety errors and adverse events (for example surgical morbidity and mortality) and disease-specific outcomes (for example local recurrence rate)—were felt to be too similar in practice to consistently categorize measures. Therefore, these domains were combined.

Quality indicators identified in the literature were also compared with cancer quality indicators endorsed by the National Quality Forum (NQF), the Centers for Medicare and Medicaid Services (CMS), and Healthcare Improvement Scotland (HIS)^{15–17}.

The intended use of a study's results was categorized as either exploratory (for example evaluating the ability of a metric to detect inter-hospital variation) or presenting results of an established quality improvement or reporting programme (for example the results of the Dutch Surgical Colorectal Audit). Whether a study's authors explicitly commented on providing the results of their study to the studied groups was also noted.

Risk of bias

A risk-of-bias assessment was not performed for this review as it was intended to provide an overview of the quality indicator

Table 2 Inclusion and exclusion criteria

| | Inclusion | Exclusion |
|-----------------------|--|---|
| Date | 1 January 2000–23 October 2023 | Before 2000 |
| Exposure of interest | Quality assessment of surgical oncology care | Not oncological care Not surgical care (including surgical pathology) Development or validation of quality indicators |
| Geographical location | Hospital or regional | National or international |
| Language | English language only | Any language other than English |
| Participants | Adult patients (at least 18 years old) | Paediatric patients |
| Peer review | Peer-reviewed studies | Non-peer reviewed studies Systematic reviews Grey literature |
| Reported outcomes | Variation in quality indicators | Comparison of surgical approaches or case-mix or individual surgeon variation Single surgeon databases |
| Setting | Clinical/administrative/quality data sets | Any study design other than cohort study |
| Study design | Cohort studies | Editorials, letters, news, comments, case reports, commentaries, and narrative reviews |
| Type of publication | Original studies and abstracts | Subgroup and post-hoc analyses Systematic reviews/meta-analyses |

landscape in surgical oncology rather than a conclusion by combining studies.

Statistical analysis

Counts and percentages were calculated for categorical variables. Medians and interquartile ranges were calculated for continuous variables. Continuous variables were also analysed graphically using histograms. All analyses were performed using Stata17.

Results

Of the 3617 articles retrieved, 120 full-text articles were reviewed and 89 met the inclusion criteria. [Figure 1](#) summarizes the literature search and screening process, including the reasons for exclusion during full-text review.

Study characteristics

Characteristics of the included studies are summarized in [Table 3](#). A total of 55 studies (62%) took place in the USA, 15 (17%) in the Netherlands, six (7%) in the UK, and 13 (15%) in other countries.

Most studies reported variation in quality indicators between centres or hospitals (82; 92%), five (6%) reported variation in quality indicators between geographical regions, one (1%) compared a single hospital with the national average, and one (1%) compared provider networks. The median number of hospitals evaluated was 148 (interquartile range 28–809). The median number of patients included was 12 839 (interquartile range 3035–40 892). Most studies (48; 54%) evaluated fewer than 10 000 patients (see [Fig. S1](#)). Only 14 studies (16%) evaluated more than 100 000 patients^{18–30}.

Studies evaluated 16 tumour types in total. Surgery for colorectal tumours was most frequently evaluated (26; 29%)^{19–22,30–52}, followed by breast tumours (14; 16%)^{53–66} and lung tumours (14; 16%)^{23–26,50–52,67–78}. A total of 46 studies (58%) evaluated patients in all cancer stages, 18 studies (20%) included patients with stage I–III cancer, 12 studies (15%) included only patients with stage I–II cancer, and two studies (2%) only included patients with stage III or IV cancer (ovarian)²⁹.

The data used to compare centres or regions primarily came from national (64; 72%) or regional (16; 18%) clinical data

registries ([Table 4](#)). These registries included the US National Cancer Database (22), various Dutch cancer audits (8), and the US NSQIP (6). A total of ten studies (11%) utilized hospital administrative data, nine studies (10%) used medical records, eight studies (9%) used hospital/insurer billing data, and one study (1%) used survey data⁷⁵.

Data collection and maintenance were most often facilitated through collaboration of a specialty society and a not-for-profit organization (24; 27%). Governments seldom took on this role (national governments, nine (10%); and state/regional governments, five (6%)). Specialty societies included the American College of Surgeons, the Dutch Society of Lung Surgeons, and the German Society of General and Visceral Surgery.

Characteristics of quality indicators

In total, the 89 included studies assessed 91 unique quality indicators. [Table 5](#) quantifies the categorization of these quality indicators by Donabedian and cancer care pathway domains. More than half of the quality indicators were process indicators (58; 64%). Process indicators included measures such as percentage of patients achieving negative margins, waiting times for surgery, and adequate number of follow-up visits. Outcome quality indicators were also well represented (26; 29%). Outcome quality indicators included failure-to-rescue rates (that is patient deaths after complications), 30- and 90-day mortality rates, and 5-year overall survival. Few quality indicators were classed as structure indicators (6; 7%) and one (1%) fell into both the structure and outcome domains⁴⁸. Structure indicators included the availability of psychological counselling, mean total episode payment, and regular monitoring of morbidity and mortality. The structure and outcome quality indicator was a combined measure of hospital volume and outcomes⁴⁸. A complete list of unique quality indicators by Donabedian domain is available in [Table S1](#). A total of 30 of the 91 (33%) quality indicators identified in this review were endorsed by the NQF, the CMS, or HIS ([Table S1](#)).

When grouped by which aspect of the care pathway they evaluated, the quality indicators were nearly evenly distributed among preoperative (25; 27%), intraoperative (28; 31%), and

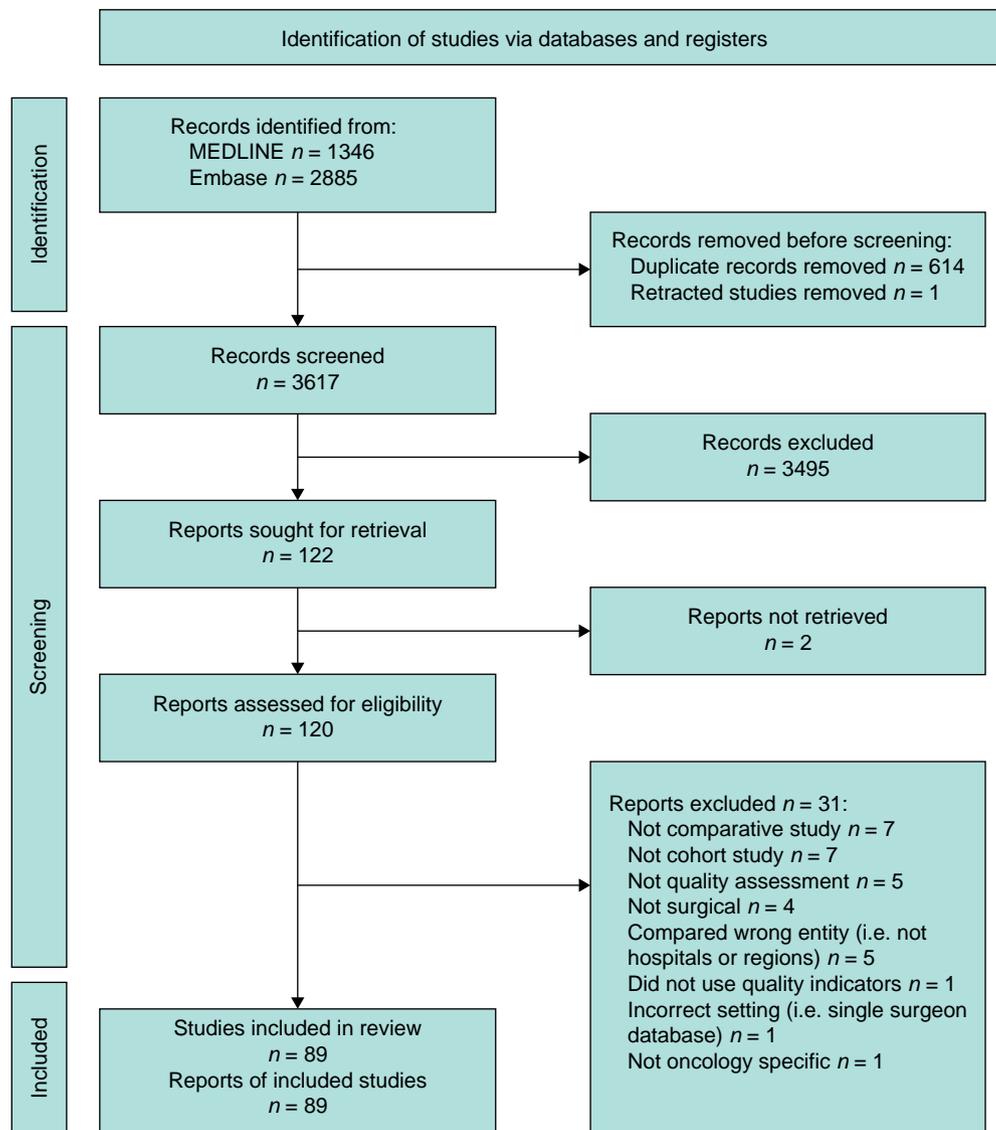


Fig. 1 PRISMA flow chart

This figure illustrates the process of identifying, screening, and including or excluding studies based upon the defined inclusion and exclusion criteria.

postoperative (30; 33%). Preoperative quality indicators were measures such as completion of preoperative functional assessment, clinical staging, and length of waiting times for surgery. Intraoperative measures included sentinel lymph node biopsy rate, mesorectal excision rate (colon cancer), and risk-adjusted margin positivity rate. Postoperative quality indicators included prolonged length of stay, in-hospital mortality, and readmission rate. A total of six (7%) quality indicators fell into more than one category^{28,43,79,80} (for example patient-reported experience of all care received for a surgical episode) and two (2%) did not fit into any of these groupings as they related to the structure of the healthcare delivery system and/or innovation (for example regular monitoring of morbidity and mortality)^{75,80}. A list of quality indicators by clinical flow domain is presented in [Table S2](#).

Quality indicators are listed according to QICC domain in [Table S3](#), but are not quantified because quality indicators often were not meant to evaluate just one QICC domain and so this grouping was subjective. There were also several quality indicators that spanned more than one QICC domain. These are

listed at the bottom of [Table S3](#). Quality indicators are also listed by tumour type in [Table S4](#).

Purpose of surgical performance reporting

A total of 43 studies (48%) sought to identify hospital outliers (typically defined as being 2 or 3 standard deviations from the mean), both positive and negative, 29 studies (33%) aimed to provide evidence of variation in practice across hospitals, 14 studies (16%) were interested in which hospitals met a predetermined benchmark for guideline adherence, a certain outcome, or the like, and three studies had other purposes (comparing a single hospital with the national average (1; 1%)³⁶, evaluating the effect of adjusting for case-mix variation on hospital ranking (1; 1%)⁶³, and proving quality indicator validity by demonstrating the ability to detect variation (1; 1%)³¹) ([Table 6](#)).

Risk/case-mix adjustment

Most of the studies (72; 83%) adjusted for case-mix variation in some manner. A few studies (15, 17%) reported not adjusting for

Table 3 Characteristics of included studies

| | |
|---|----------------------|
| Country | |
| USA | 55 (62) |
| Netherlands | 15 (17) |
| UK | 6 (7) |
| Canada | 3 (3) |
| Belgium | 2 (2) |
| Other | 7 (8) |
| Level of comparison | |
| Between hospital variation | 82 (92) |
| Between geographical region variation | 5 (6) |
| Single hospital to national average | 1 (1) |
| Between provider networks | 1 (1) |
| Tumour types evaluated | |
| Colorectal | 26 (29) |
| Lung | 14 (16) |
| Breast | 14 (16) |
| Multiple | 7 (8) |
| Pancreas | 6 (7) |
| Oesophageal | 3 (3) |
| Head and neck | 3 (3) |
| Prostate | 3 (3) |
| Gynaecological | 3 (3) |
| Gastric | 2 (2) |
| Upper GI tract | 2 (2) |
| Renal | 2 (2) |
| Skin | 2 (2) |
| Other | 2 (2) |
| Stage of patients evaluated | |
| I-IV | 46 (58) |
| I-III | 18 (23) |
| I-II | 12 (15) |
| III-IV | 2 (2) |
| II-IV | 1 (1) |
| Number of hospitals evaluated, median (i.q.r.) | 148 (28–809) |
| Number of regions evaluated, median (i.q.r.) | 5 (5–5) |
| Number of patients evaluated, median (i.q.r.) | 12 839 (3035–40 892) |

Values are n (%) unless otherwise indicated. GI, gastrointestinal; i.q.r., interquartile range.

patient-level factors^{19,20,34–38,43,46,52,53,57,65,73,75,76,81} and two studies (2%) did not mention whether they made adjustments^{71,82}.

While a formal sub-analysis of the variables included in the risk adjustment was not undertaken, most risk-adjusted studies accounted for patient characteristics such as age, sex, race, physical status (for example ASA classification), and co-morbidities (for example Charlson co-morbidity index). Some also adjusted for cancer-related variables such as tumour histological subtype, oncological stage, and neoadjuvant therapy status. A few also included operative factors such as emergent/urgent surgery classification, operating time, and operative approach (typically for outcome measures such as 30-day readmission rate or morbidity rate).

Intended use and feedback to studied hospitals/regions

Most studies (74; 83%) were intended for academic exploratory use. Aims of exploratory studies included highlighting variation in performance across hospitals on a specific metric, investigating associations between hospital factors and performance on quality indicator(s), and comparing the performance of similar metrics (for example 30- versus 90-day mortality rates). The remainder (15; 17%) were descriptive reports of established quality improvement or reporting programmes^{24,33,43,44,46,52–54,56,57,60,64–66,71}. These programmes included clinical audits (for example the Dutch Surgical Colorectal Audit) and regional quality improvement programmes (for example the Michigan Surgical Quality

Table 4 Data and database sources

| | |
|---|---------|
| Data source(s) | |
| National clinical data registry | 64 (72) |
| Regional clinical data registry | 16 (18) |
| Hospital administrative data | 10 (11) |
| Medical records | 9 (10) |
| Hospital/insurer billing data | 8 (9) |
| Patient survey | 1 (1) |
| Entity that produced/maintained database | |
| Specialty society | 15 (19) |
| National government | 9 (10) |
| Not-for-profit organization | 6 (7) |
| Quality collaborative (self-organized group of hospitals) | 6 (7) |
| State or regional government | 5 (6) |
| Insurance organization (public or private) | 3 (3) |
| Other single entity | 3 (3) |
| Joint effort | |
| Specialty society and not-for-profit organization | 24 (27) |
| Quality collaborative and insurance organization | 6 (7) |
| National government and insurance organization | 3 (3) |
| National government and state or regional government | 2 (2) |
| National government, state or regional government, and insurance organization | 2 (2) |
| Other combination | 5 (6) |

Values are n (%).

Collaborative). See Supplemental Results and Table S5 for further description of data visualization approaches in these studies.

The authors of 22 studies (25%) reported feeding the results back to the groups (that is hospitals or regions) they studied (see Table S5 for details)^{14,21,22,33,37,42–44,46,47,49,50,52,53,57,64,71,73,78,81,83,84}, whereas 67 studies (75%) made no mention of feeding the results back to the hospitals or regions studied. Those authors who did report providing feedback fed the results into quality improvement efforts at the regional (9; 39%) or national (6; 27%) level, sent the results to all studied groups (4; 18%), or both sent the results to all studied groups and fed the results into quality improvement efforts (3; 13%). Notably, only one of the 22 studies advocated for public reporting of their outcomes³⁹. The feedback of all other studies was explicitly for the use of healthcare systems.

Discussion

This analysis sought to describe quality indicators used to compare the quality of surgical oncology care across hospitals or regions. Most quality indicators were process measures and were evenly distributed among preoperative, intraoperative, and postoperative care. The data used to make these comparisons primarily came from national or regional data registries as well as collaborations between specialty societies and not-for-profit organizations. The purpose of evaluating inter-hospital or inter-regional variation fell into one of three main categories: identifying outliers; comparing centres with a predetermined benchmark; or supplying evidence of practice variation. Interestingly, only 22 studies (25%) reported notifying those providing the data of the variation they detected or otherwise using the results to inform quality improvement efforts. Only one study advocated for public reporting of their findings³⁹. The majority of these quality indicators were derived from high-income settings.

This review can be utilized to encourage health systems and regions that have already begun employing quality indicators to consider additional previously validated metrics or adapt those that have previously been used. The study team has completed

Table 5 Characteristics of quality indicators

| | |
|---|---------|
| Donabedian domain(s) | |
| Process | 58 (64) |
| Outcome | 26 (29) |
| Structure | 6 (7) |
| Structure and outcome | 1 (1) |
| Total, n | 91 |
| Clinical flow | |
| Preoperative | 25 (27) |
| Intraoperative | 28 (31) |
| Postoperative | 30 (33) |
| Preoperative, intraoperative, and postoperative | 1 (1) |
| Preoperative and postoperative | 2 (2) |
| Preoperative and intraoperative | 1 (1) |
| Intraoperative and postoperative | 2 (2) |
| Total, n | 89* |

Values are n (%) unless otherwise indicated. *Two innovation indicators not classified.

an international Delphi study to establish a minimum set of surgical oncology indicators in all resource settings, including low- and middle-income countries (LMICs)⁸⁵. The present review also provides a framework for healthcare systems interested in starting such initiatives by providing a rigorous overview of the procedures for reporting and feedback, including data sources and governance of reporting programmes. The present review uniquely moves beyond a set of 'ideal indicators' that many consensus initiatives have sought to establish to a set of indicators that have been implemented or tested for validity and fairness. The review also exposes the gaps in current metrics, with a paucity of outcome measures to appraise surgical oncology care used, the need for integrated data architectures, and the routine collection of patient-reported outcome measures (PROMs) all providing opportunities for improving care.

Only 29% of quality indicators in the surgical oncology literature were outcome indicators. This is likely because processes are proxies that are easier to measure than outcomes, which are subject to issues such as loss to follow-up and case-mix variation. Many process indicators were highly specific to tumour-type specialty guidelines. However, some indicators were tumour agnostic (for example whether an adequate lymphadenectomy had been performed). Likewise, achievement of negative margins was measured in several ways, including risk-adjusted margin positivity (RAMP) rate, rate of resections with a negative margin (that is RO resections), and circumferential margin positivity rate. Another theme that emerged was proxy measurement of overtreatment. Two examples of these measures were the rate of radical mastectomy for pre-invasive breast cancer and the proportion of prostate biopsies in men with limited life expectancy. Several studies compared 30- and 90-day mortality/readmission/reoperation rates, suggesting ongoing disagreements regarding the validity of some outcome measures.

Composite measures identified in this review included 'textbook outcome' (a binary indicator achieved when a patient had the ideal outcome as defined by a group of experts), MTL22 ('mortality, transfer, or prolonged length of stay in first 22 postoperative days'), and a combined measure of volume and outcome^{33,36,48,77}. The purported benefit of composite measures is their superior sensitivity as they capture related outcomes.

It is noteworthy that no studies compared PROMs. This might be a factor of limited collection of PROMs, difficulty linking this information to patient records, and/or belief that these measures are less reliable or valid despite evidence demonstrating

Table 6 Study purpose and feedback to studied hospitals/regions

| | |
|--|---------|
| Purpose of surgical performance reporting | |
| Identify outliers | 43 (48) |
| Provide evidence of variation | 29 (33) |
| Compare with benchmark | 14 (16) |
| Other | 3 (3) |
| Feedback to studied hospitals/regions? | |
| Yes | 22 (25) |
| No | 67 (75) |
| Type of feedback | |
| Fed into regional quality improvement | 9 (39) |
| Fed into national quality improvement | 6 (27) |
| Sent to studied hospitals/regions | 4 (18) |
| Sent and fed into quality improvement | 3 (13) |

Values are n (%).

concordance between PROMs and clinician-reported outcome measures⁸⁶. Additionally, measures of access and innovation/improvement were absent. However, collection of these quality indicators is feasible as demonstrated by national audits such as the National Prostate Cancer Audit in the UK⁸⁷.

Clinically meaningful quality indicators need to be valid, feasible to collect, have clinician 'buy-in', and show variation over space and time¹⁰. Each of these components holds challenges. Validating a quality indicator includes ensuring differences in an indicator are reflective of true differences in quality⁸⁸. Validity was typically demonstrated a priori for these quality indicators by an academic study and/or specialty society consensus. Even still, reporting comparative quality indicator performance in a thoughtful manner that is fair (that is adjusted for case-mix variation and volume) and understandable by the public is challenging⁸⁹.

Feasibility of data collection can also pose a challenge. This is attributable in part to barriers to routine performance measurement such as lack of a shared data structure, consistent means of patient identification, and assurance of data completeness, particularly in decentralized healthcare systems such as in the USA⁹⁰. Existing administrative data sets, while less burdensome, may be missing clinically important information such as co-morbidities that are necessary for case-mix adjustment. Purpose-made data sets for quality indicator comparison may be more complete, but with the trade-off of additional burden on clinicians or staff members for data collection and entry, which can impact on completeness.

National clinical data registries were by far the most common source of data (64; 72%), while collaborations between specialty societies and not-for-profit organizations (for example the American Cancer Society) (24; 27%) and specialty societies alone (15; 17%) were the most common entities that produced/maintained databases.

Acknowledging the expectation that many studies would be exploratory, perhaps the most notable result of this analysis is that only 22 (25%) of the included studies reported feeding the results back to those who provided the data or using them to inform quality improvement efforts. Of these, only one advocated for public reporting of their results³⁹. The other 67 studies were performed purely for academic purposes.

If the ultimate aim of measuring and comparing quality is to improve quality, it is worth considering how results of these analyses could be used to that end. Berwick *et al.*⁹¹ proposed two pathways for quality improvement based on the intrinsic motivation of providers (change pathway) or through market competition whereby patients select a provider of their choice.

Feeding the results back to those who provided the data for these studies and using the results to inform quality improvement efforts both rely on hospitals' intrinsic motivation to improve their care delivery^{91–93}. While some healthcare organizations may be responsive to private communication of comparative results, evidence shows that public reporting of hospital performance spurs significantly more quality improvement activities^{94,95}. Several systematic reviews with meta-analyses have been published that show the opposite effect^{95–97}. Even at an individual clinician level, a 2018 study by Vallance *et al.*⁹⁸ found that surgeon-specific outcome reporting in colorectal cancer surgery did not lead to gaming or adverse clinical practices, but did lead to a significant reduction in 90-day mortality.

Additionally, there are now multiple examples of successful public reporting initiatives⁸⁷. The UK Bowel Cancer Audit, for instance, has shown a reduction in 90-day and 2-year mortality, shrunk the variation between providers in use of adjuvant chemotherapy and 30-day readmissions, and increased the proportion using laparoscopic surgery between 2015 and 2019⁹⁹. The evidence demonstrates that public reporting is feasible and has the potential to improve the quality of patient care. However, this requires buy-in from stakeholders, including clinicians and health system administrators, and investment in integrated data systems.

If comparative quality measures are available to those who choose where care is received (that is public reporting), market competition may take its course. Because lower performing hospitals stand to lose revenue, they will be incentivized to improve their quality to attract more patients. New York state, for example, has been publicly reporting short-term cardiac surgery outcomes for over 20 years and many hospitals, concerned about losing market share, have undertaken quality improvement initiatives as a result⁹⁵.

Evidence suggests patients are prepared to travel to alternative, more distant elective secondary care services, but tend to be more sensitive to hospital-level characteristics (for example availability of robotic surgery and overall hospital quality) than to cancer-specific outcome measures^{100,101}. This is likely driven in part by the strength of informal knowledge networks, which tend to focus on the brand and reputation of providers, as well as the relevance of current cancer-specific outcomes when patients consider their own care choices. Further patient engagement will be critical to support rational use of metrics¹⁰².

All studies included in this article were conducted in high-income countries (HICs). This is notable because the context within which a health system operates effects not only how and what care is provided but also dictates what quality indicators are (or can be) collected. For example, though each HIC has a unique health system structure, the infrastructure available to support data collection and analysis is more uniformly robust than that in their LMIC counterparts.

The largest burden of cancer in the next few years will be in emerging economies and LMICs². This review shows a paucity of published measures in this setting. Anecdotally, there is heterogeneity in quality indicators collected in these settings, underscoring a need to develop basic matrices that can be used in diverse settings to ensure the quality of surgical oncology services. This study can be used to support LMICs in selecting quality measures to deploy in their own healthcare systems by providing a menu of validated surgical oncology indicators.

It is notable that 15 studies did not adjust for case-mix variation in their analyses (and 2 studies did not mention whether they did). Risk adjustment is important to ensure the validity of results and

to prevent cream skimming/gaming. Without risk adjustment, a hospital that cares for sicker patients (that is more co-morbidities and higher stage at diagnosis) may appear to have a poorer performance than one that treats fewer complex patients, even if they have the same true level of quality (that is given the same patients they would perform the same). This can frustrate the hospitals/regions being compared and cause them to distrust the results and lose confidence in the process. Alternatively, the hospitals/regions might engage in cream skimming or gaming whereby they selectively provide care to fewer complex patients to inflate their performance^{98,103}.

There are limitations to the analyses presented in this review that should be acknowledged. First, only published inter-hospital and inter-regional quality comparisons could be reviewed. There are likely quality comparisons undertaken for internal audits or other purposes that are not published and so were not included in this review. Embase and MEDLINE were used to ensure that a wide range of published papers was included, but it is acknowledged that some regional country-specific initiatives may not have been included if they were not published or translated into English. Additionally, the categorization of the indicators into QICC domains was not always simple. While uncertainties were resolved via discussion with two subject matter experts, there was an unavoidable element of judgement required in these decisions. For this reason, two outcome-related QICC categories were combined and this review did not quantify the number of indicators in each QICC category, but rather listed the indicators in the category that were judged most appropriate. Lastly, only studies that had full-text availability in English were included.

There is a modest corpus of comparisons of quality indicators in surgical oncology almost entirely drawn from HICs such as the USA, the Netherlands, and the UK. These comparisons can be used to improve quality in the healthcare system by providing feedback to those providing the data via sharing the results or undertaking evidence-informed quality improvement efforts and/or by providing useable information about detected differences in quality to aid decisions about where to receive care and to inspire public trust. However, substantial gaps exist, especially in developing quality indicators for surgical oncology that are relevant across different health systems.

Once this core set of quality indicators is identified, organizations that collect quality-related data might consider generating performance reports to be shared with leaders of hospitals and/or regions whose data they collect. These same organizations could also facilitate knowledge sharing among participating hospitals/regions to improve quality across the board. Defining a core indicator set may also open the door to broader comparisons across data sets and health systems. Tiering the indicator set based upon the robustness of the data and health system infrastructure could aid prioritization of quality improvement efforts in developing health systems.

Funding

The authors have no funding to declare.

Author contributions

Megan McLeod (Conceptualization, Data curation, Formal analysis, Methodology, Project administration, Writing—original draft, Writing—review & editing), Kari Leung (Data curation, Investigation, Project administration, Validation, Writing—review & editing), C. S. Pramesh (Validation, Writing—review & editing), Peter Kingham (Validation,

Writing—review & editing), Miriam Mutebi (Validation, Writing—review & editing), Julie Torode (Conceptualization, Writing—review & editing), Andre Ilbawi (Validation, Writing—review & editing), Jade Chakowa (Conceptualization, Writing—review & editing), Richard Sullivan (Conceptualization, Investigation, Methodology, Resources, Supervision, Validation, Writing—review & editing), and Ajay Aggarwal (Conceptualization, Investigation, Methodology, Resources, Supervision, Validation, Visualization, Writing—review & editing)

Disclosure

The authors declare no conflict of interest.

Supplementary material

Supplementary material is available at BJS Open online.

Data availability

Data available upon request.

References

- Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: The National Academies Press, 2001
- Perera SK, Jacob S, Wilson BE, Ferlay J, Bray F, Sullivan R et al. Global demand for cancer surgery and an estimate of the optimal surgical and anaesthesia workforce between 2018 and 2040: a population-based modelling study. *Lancet Oncol* 2021;**22**:182–189
- Showalter TN, Camacho F, Cantrell LA, Anderson RT. Determinants of quality care and mortality for patients with locally advanced cervical cancer in Virginia. *Medicine (Baltimore)* 2016;**95**:e2913
- Baicker K, Chandra A. Medicare spending, the physician workforce, and beneficiaries' quality of care. *Health Aff* 2004;**23**:W4-184–W4-197
- Donabedian A. Evaluating the quality of medical care. *Milbank Mem Fund Q* 1966;**44**:166–206
- Chiew KL, Sundaresan P, Jalaludin B, Vinod SK. A narrative synthesis of the quality of cancer care and development of an integrated conceptual framework. *Eur J Cancer Care (Engl)* 2018;**27**:e12881
- Ko CY, Hall BL, Hart AJ, Cohen ME, Hoyt DB. The American College of Surgeons National Surgical Quality Improvement Program: achieving better and safer surgery. *Jt Comm J Qual Patient Saf* 2015;**41**:199–204
- Van Leersum NJ, Snijders HS, Henneman D, Kolfshoten NE, Gooiker GA, ten Berge MG et al. The Dutch Surgical Colorectal Audit. *Eur J Surg Oncol* 2013;**39**:1063–1070
- Cornish JA, Tekkis PP, Tan E, Tilney HS, Thompson MR Smith JJ. The National Bowel Cancer Audit project: the impact of organisational structure on outcome in operative bowel cancer within the UK. *Surg Oncol* 2011;**20**:e72–e77
- McGory ML, Shekelle PG, Ko CY. Development of quality indicators for patients undergoing colorectal cancer surgery. *J Natl Cancer Inst* 2006;**98**:1623–1633
- Harden SV, Chiew K-L, Millar J, Vinod SK. Quality indicators for radiation oncology. *J Med Imaging Radiat Oncol* 2022;**66**:249–257
- Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;**372**:n71
- Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017;**358**:j4008
- Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ et al. *Cochrane Handbook for Systematic Reviews of Interventions, Version 6.3 (Updated February 2022)*. 2022. www.training.cochrane.org/handbook (accessed June 2023)
- National Quality Forum. *Measure Search*. <https://www.qualityforum.org/QPS/QPSTool> (accessed 27 February 2023)
- Centers for Medicare and Medicaid Services. *Measures Inventory Tool*. <https://cmit.cms.gov/cmit/#/MeasureInventory> (accessed 27 February 2023)
- Health Improvement Scotland. *Cancer Quality Performance Indicators (QPIs)*. 31 August 2022. https://www.healthcareimprovementscotland.org/our_work/cancer_care_improvement/cancer_qpis/quality_performance_indicators.aspx (accessed 27 February 2023)
- Adam MA, Turner MC, Sun Z, Kim J, Ezekian B, Migaly J et al. The appropriateness of 30-day mortality as a quality metric in colorectal cancer surgery. *Am J Surg* 2018;**215**:66–70
- Antunez AG, Kanters AE, Regenbogen SE. Evaluation of access to hospitals most ready to achieve national accreditation for rectal cancer treatment. *JAMA Surg* 2019;**154**:516–523
- Abdelsattar ZM, Krell RW, Campbell DA Jr, Hendren S, Wong SL. Differences in hospital performance for noncancer vs cancer colorectal surgery. *J Am Coll Surg* 2014;**219**:450–459
- Morris EJ, Taylor EF, Thomas JD, Quirke P, Finan PJ, Coleman MP et al. Thirty-day postoperative mortality after colorectal cancer surgery in England. *Gut* 2011;**60**:806–813
- Almoudaris AM, Burns EM, Bottle A, Aylin P, Darzi A, Vincent C et al. Single measures of performance do not reflect overall institutional quality in colorectal cancer surgery. *Gut* 2013;**62**:423–429
- Khorfan R, Cooke DT, Meguid RA, Backhus L, Varghese TK, Farjah F et al. Institutional factors associated with adherence to quality measures for stage I and II non-small cell lung cancer. *J Thorac Cardiovasc Surg* 2021;**162**:649–660.e8
- Lin CC, Smeltzer MP, Jemal A, Osarogiagbon RU. Risk-adjusted margin positivity rate as a surgical quality metric for non-small cell lung cancer. *Ann Thorac Surg* 2017;**104**:1161–1170
- Moore CB, Cox ML, Mulvihill MS, Klapper J, D'Amico TA, Hartwig MG. Challenging 30-day mortality as a site-specific quality metric in non-small cell lung cancer. *J Thorac Cardiovasc Surg* 2019;**158**:570–578.e3
- Ivanovic J, Maziak DE, Gilbert S, Shamji FM, Sundaresan RS, Ramsay T et al. Assessing the status of thoracic surgical research and quality improvement programs: a survey of the members of the Canadian Association of Thoracic Surgeons. *J Surg Educ* 2011;**68**:258–265
- Liu JB, Berian JR, Liu Y, Ko CY, Weber SM. Trends in perioperative outcomes of hospitals performing major cancer surgery. *J Surg Oncol* 2018;**118**:694–703
- Greenberg CC, Lipsitz SR, Neville B, In H, Hevelone N, Porter SA et al. Receipt of appropriate surgical care for Medicare beneficiaries with cancer. *Arch Surg* 2011;**146**:1128–1134
- Lawson KA, Saarela O, Abouassaly R, Kim SP, Breau RH, Finelli A. The impact of quality variations on patients undergoing

- surgery for renal cell carcinoma: a National Cancer Database study. *Eur Urol* 2017;**72**:379–386
30. Manisundaram N, DiBrito SR, Hu CY, Kim Y, Wick E, Palis B et al. Reporting of circumferential resection margin in rectal cancer surgery. *JAMA Surg* 2023;**158**:1195–1202
 31. Hardt J, Doyon F, Klinger C, Buhr H, Post S. MTL, a composite measure for objectively profiling hospitals on surgical performance in colorectal cancer surgery. *Int J Colorectal Dis* 2019;**34**:889–898
 32. Kanters A, Mullard AJ, Arambula J, Fasbinder L, Krapohl G, Wong SL et al. Colorectal cancer: quality of surgical care in Michigan. *Am J Surg* 2017;**213**:548–552
 33. Lucas DJ, Ejaz A, Bischof DA, Schneider EB, Pawlik TM. Variation in readmission by hospital after colorectal cancer surgery. *JAMA Surg* 2014;**149**:1272–1277
 34. Merkow RP, Bilimoria KY, Cohen ME, Richards K, Ko CY, Hall BL. Variability in reoperation rates at 182 hospitals: a potential target for quality improvement. *J Am Coll Surg* 2009;**209**:557–564
 35. Merkow RP, Hall BL, Cohen ME, Wang X, Adams JL, Chow WB et al. Validity and feasibility of the American College of Surgeons colectomy composite outcome quality measure. *Ann Surg* 2013;**257**:483–489
 36. Kanters AE, Cleary RK, Obi SH, Asgeirsson T, Evilsizer SK, Fasbinder LG et al. Uptake of total mesorectal excision and total mesorectal excision grading for rectal cancer: a statewide study. *Dis Colon Rectum* 2020;**63**:53–59
 37. McSorley S, Lowndes C, Sharma P, Macdonald A. Unplanned reoperation within 30 days of surgery for colorectal cancer in NHS Lanarkshire. *Colorectal Dis* 2013;**15**:689–694
 38. Morris E, Quirke P, Thomas JD, Fairley L, Cottier B, Forman D. Unacceptable variation in abdominoperineal excision rates for rectal cancer: time to intervene? *Gut* 2008;**57**:1690–1697
 39. Giesen L, Olthof P, Elferink M, Verhoef C, Dekker J. Surgery for rectal cancer: differences in resection rates among hospitals in the Netherlands. *Eur J Surg Oncol* 2021;**47**:2384–2389
 40. Massarweh NN, Hu CY, You YN, Bednarski BK, Rodriguez-Bigas MA, Skibber JM et al. Risk-adjusted pathologic margin positivity rate as a quality indicator in rectal cancer surgery. *J Clin Oncol* 2014;**32**:2967–2974
 41. Abdelsattar ZM, Wong SL, Birkmeyer NJ, Cleary RK, Times ML, Figg RE et al. Multi-institutional assessment of sphincter preservation for rectal cancer. *Ann Surg Oncol* 2014;**21**:4075–4080
 42. Chung KP, Chang YJ, Lai MS, Kuo RN-C, Cheng SH, Chen L-T et al. Is quality of colorectal cancer care good enough? Core measures development and its application for comparing hospitals in Taiwan. *BMC Health Serv Res* 2010;**10**:1–11
 43. Henneman D, Snijders H, Fiocco M, van Leersum NJ, Kolfshoten NE, Wiggers T et al. Hospital variation in failure to rescue after colorectal cancer surgery: results of the Dutch Surgical Colorectal Audit. *Ann Surg Oncol* 2013;**20**:2117–2123
 44. Schootman M, Lian M, Pruitt S, Hendren S, Mutch M, Deshpande AD et al. Hospital and geographic variability in two colorectal cancer surgery outcomes: complications and mortality after complications. *Ann Surg Oncol* 2014;**21**:2659–2666
 45. Ratnayake I, Park J, Biswanger N, Feely A, Musto G, Decker K. Colorectal cancer surgery quality in Manitoba: a population-based descriptive analysis. *Curr Oncol* 2021;**28**:2239–2247
 46. Patel SH, Hu CY, Massarweh NN, You NY, McCabe R, Dietz D et al. Circumferential resection margin as a hospital quality assessment tool for rectal cancer surgery. *J Am Coll Surg* 2020;**230**:1008–1018.e5
 47. Snijders H, Henneman D, Van Leersum N, ten Berge M, Fiocco M, Karsten TM et al. Anastomotic leakage as an outcome measure for quality of colorectal cancer surgery. *BMJ Qual Saf* 2013;**22**:759–767
 48. Kolfshoten NE, Marang-van de Mheen PJ, Wouters MW, Eddes E-H, Tollenaar RAEM, Stijnen T et al. A combined measure of procedural volume and outcome to assess hospital quality of colorectal cancer surgery, a secondary analysis of clinical audit data. *PLoS One* 2014;**9**:e88737
 49. Breslin TM, Caughran J, Pettinga J, Wesen C, Mehlinger A, Yin H et al. Improving breast cancer care through a regional quality collaborative. *Surgery* 2011;**150**:635–642
 50. Bertoy L, Harbaugh CM, Millis MA, Zhuo L, Gutsche N, Beck G et al. Positive margin rates for colorectal cancer vary significantly by hospital in Michigan: can we achieve a 0% positive margin rate? *Surg Open Sci* 2023;**16**:37–43
 51. Xu W, Wells CI, McGuinness M, Varghese C, Keane C, Liu C et al. Characterising nationwide reasons for unplanned hospital readmission after colorectal cancer surgery. *Colorectal Dis* 2023;**25**:861–871
 52. Kalata S, Lam GT, Alnajjar RM, Clark MJ, He C, Welsh RJ et al. Adequate lung cancer surgery lymphadenectomy within a statewide quality collaborative: quality improvement in action. *J Thorac Cardiovasc Surg* 2023;**166**:241–250.e3
 53. Gort M, Broekhuis M, Otter R, Klazinga NS. Improvement of best practice in early breast cancer: actionable surgeon and hospital factors. *Breast Cancer Res Treat* 2007;**102**:219–226
 54. Boughey JC, Hieken TJ, Jakub JW, Degnim AC, Grant CS, Farley DR et al. Impact of analysis of frozen-section margin on reoperation rates in women undergoing lumpectomy for breast cancer: evaluation of the National Surgical Quality Improvement Program data. *Surgery* 2014;**156**:190–197
 55. Laronga C, Gray JE, Siegel EM, Lee J-H, Fulp WJ, Fletcher M et al. Florida initiative for quality cancer care: improvements in breast cancer quality indicators during a 3-year interval. *J Am Coll Surg* 2014;**219**:638–645.e1
 56. Ponti A, Mano MP, Tomatis M, Baiocchi D, Barca A, Berti R et al. Audit system on quality of breast cancer diagnosis and treatment (QT): results of quality indicators on screen-detected lesions in Italy, 2011–2012. *Epidemiol Prev* 2015;**39**:40–47
 57. van der Heiden-van der Loo M, de Munck L, Visser O, Westenend PJ, van Dalen T, Menke MB et al. Variation between hospitals in surgical margins after first breast-conserving surgery in the Netherlands. *Breast Cancer Res Treat* 2012;**131**:691–698
 58. McCahill LE, Single RM, Bowles EJA, Feigelson HS, James TA, Barney T et al. Variability in reexcision following breast conservation surgery. *JAMA* 2012;**307**:467–475
 59. van der Heiden-van der Loo M, Siesling S, Wouters M, van Dalen T, Rutgers E, Peeters P. The value of ipsilateral breast tumor recurrence as a quality indicator: hospital variation in the Netherlands. *Ann Surg Oncol* 2015;**22**:522–528
 60. Jeevan R, Cromwell D, Trivella M, Lawrence G, Kearins O, Pereira J et al. Reoperation rates after breast conserving surgery for breast cancer among women in England: retrospective study of hospital episode statistics. *BMJ* 2012;**345**:e4505
 61. Talsma A, Reedijk A, Damhuis R, Westenend P, Vles W. Re-resection rates after breast-conserving surgery as a performance indicator: introduction of a case-mix model to allow comparison between Dutch hospitals. *Eur J Surg Oncol* 2011;**37**:357–363

62. Vos EL, Lingsma HF, Jager A, Schreuder K, Spronk P, Vrancken Peeters MTFD et al. Effect of case-mix and random variation on breast cancer care quality indicators and their rankability. *Value Health* 2020;**23**:1191–1199
63. Gray JE, Laronga C, Siegel EM, Lee J-H, Fulp WJ, Fletcher M et al. Degree of variability in performance on breast cancer quality indicators: findings from the Florida initiative for quality cancer care. *J Oncol Pract* 2011;**7**:247–251
64. Ratnayake I, Hebbard P, Feely A, Biswanger N, Decker K. Assessment of breast cancer surgery in Manitoba: a descriptive study. *Curr Oncol* 2021;**28**:581–592
65. Ishizaki T, Imanaka Y, Hirose M, Kuwabara K, Ogawa T, Harada Y. A first look at variations in use of breast conserving surgery at five teaching hospitals in Japan. *Int J Qual Health Care* 2002;**14**:411–418
66. Wakeam E, Byrne JP, Darling GE, Varghese TK Jr. Surgical treatment for early small cell lung cancer: variability in practice and impact on survival. *Ann Thorac Surg* 2017;**104**:1872–1880
67. Hu Y, McMurry TL, Wells KM, Isbell JM, Stukenborg GJ, Kozower BD. Postoperative mortality is an inadequate quality indicator for lung cancer resection. *Ann Thorac Surg* 2014;**97**:973–979
68. Geraci T, Baratta V, Young J, Milman S, Dunican A-M, Jones RN et al. Lobectomy for lung cancer at Veterans Administration Medical Center versus academic medical center. *Ann Thorac Surg* 2017;**103**:1715–1722
69. Farjah F, Backhus L, Cheng A, Englum B, Kim S, Saha-Chaudhuri P et al. Failure to rescue and pulmonary resection for lung cancer. *J Thorac Cardiovasc Surg* 2015;**149**:1365–1373.e3
70. Ten Berge M, Beck N, Heineman DJ, Damhuis R, Steup WH, van Huijstee PJ et al. Dutch Lung Surgery Audit: a national audit comprising lung and thoracic surgery patients. *Ann Thorac Surg* 2018;**106**:390–397
71. Thornblade LW, Wood DE, Mulligan MS, Farivar AS, Hubka M, Costas KE et al. Variability in invasive mediastinal staging for lung cancer: a multicenter regional study. *J Thorac Cardiovasc Surg* 2018;**155**:2658–2671.e1
72. Vrijens F, De Gendt C, Verleye L, Robays J, Schillemans V, Camberlin C et al. Quality of care and variability in lung cancer management across Belgian hospitals: a population-based study using routinely available data. *Int J Qual Health Care* 2018;**30**:306–312
73. Farjah F, Varghese TK, Costas K, Krishnadasan B, Farivar AS, Hubka M et al. Lung resection outcomes and costs in Washington state: a case for regional quality improvement. *Ann Thorac Surg* 2014;**98**:175–182
74. Udelsman BV, Madariaga ML, Chang DC, Kozower BD, Gaissert HA. Concordance of clinical and pathologic nodal staging in resectable lung cancer. *Ann Thorac Surg* 2021;**111**:1125–1132
75. van Roessel S, Mackay TM, van Dieren S, van der Schelling GP, Nieuwenhuijs VB, Bosscha K et al. Textbook outcome: nationwide analysis of a novel quality measure in pancreatic surgery. *Ann Surg* 2020;**271**:155–162
76. Bernard A, Cottenet J, Pages PB, Quantin C. Is there variation between hospitals within each region in postoperative mortality for lung cancer surgery in France? A nationwide study from 2013 to 2020. *Front Med (Lausanne)* 2023;**10**:1110977
77. Algera MD, Slangen BFM, van Driel WJ, Wouters MWJM, Kruitwagen RPFM. Textbook outcome as a composite outcome measure to compare hospital performances regarding cytoreductive surgery for ovarian cancer: a nationwide population-based study. *Gynecol Oncol* 2023;**174**:89–97
78. Algera MD, Baldewpersad Tewarie NMS, Driel WJV, van Ham MAPC, Slangen BFM, Kruitwagen RPFM et al. Case-mix adjustment to compare hospital performances regarding complications after cytoreductive surgery for ovarian cancer: a nationwide population-based study. *Int J Gynecol Cancer* 2023;**33**:534–542
79. Van den Bruel A, Francart J, Dubois C, Adam M, Vlayen J, De Schutter H et al. Regional variation in thyroid cancer incidence in Belgium is associated with variation in thyroid imaging and thyroid disease management. *J Clin Endocrinol Metab* 2013;**98**:4063–4071
80. Abdollah F, Ye Z, Miller DC, Linsell SM, Montie JE, Peabody JO et al. Understanding the use of prostate biopsy among men with limited life expectancy in a statewide quality improvement collaborative. *Eur Urol* 2016;**70**:854–861
81. Burks FN, Hu JC, Telang D, Liu A, Hawken S, Montgomery Z et al. Repeat prostate biopsy practice patterns in a statewide quality improvement collaborative. *J Urol* 2017;**198**:322–328
82. Spencer BA, Miller DC, Litwin MS, Ritchey JD, Stewart AK, Dunn RL et al. Variations in quality of care for men with early-stage prostate cancer. *J Clin Oncol* 2008;**26**:3735–3742
83. Merkow RP, Bilimoria KY, Chow WB, Merkow JS, Weyant MJ, Ko CY et al. Variation in lymph node examination after esophagectomy for cancer in the United States. *Arch Surg* 2012;**147**:505–511
84. Liu JB, Schuitevoerder D, Vining CC, Berger Y, Turaga KK, Eng OS. Benchmarking perioperative outcomes of cytoreductive surgery for cancer: implications for quality measurement. *Ann Surg Oncol* 2020;**27**:5039–5046
85. McLeod M, Torode J, Leung K, Bhoo-Pathy N, Booth C, Chakowa J et al. Quality indicators for evaluating cancer care in low-income and middle-income country settings: a multinational modified Delphi study. *TLO* 2024
86. Thong AE, Ying Poon B, Lee JK, Vertosick E, Sjoberg DD, Vickers AJ et al. Concordance between patient-reported and physician-reported sexual function after radical prostatectomy. *Urol Oncol* 2018;**36**:80.e1–80.e6
87. Aggarwal A, Nossiter J, Parry M, Sujenthiran A, Zietman A, Clarke N et al. Public reporting of outcomes in radiation oncology: the National Prostate Cancer Audit. *Lancet Oncol* 2021;**22**:e207–e215
88. Geary R, Knight H, Carroll F, Gurol-Urganci I, Morris E, Cromwell DA et al. A step-wise approach to developing indicators to compare the performance of maternity units using hospital administrative data. *BJOG* 2018;**125**:857–865
89. Walker K, Neuburger J, Groene O, Cromwell DA, van der Meulen J. Public reporting of surgeon outcomes: low numbers of procedures lead to false complacency. *Lancet* 2013;**382**:1674–1677
90. Nolte E. *International Benchmarking of Healthcare Quality: A Review of the Literature*. Santa Monica, CA: RAND Corporation, 2010
91. Berwick DM, James B Coye MJ. Connections between quality measurement and improvement. *Med Care* 2003;**41**:I30–I38
92. Foy R, Skrypak M, Alderson S, Ivers NM, McInerney B, Stoddart J et al. Revitalising audit and feedback to improve patient care. *BMJ* 2020;**368**:m213
93. Brehaut JC, Colquhoun HL, Eva KW, Carroll K, Sales A, Michie S et al. Practice feedback interventions: 15 suggestions for optimizing effectiveness. *Ann Intern Med* 2016;**164**:435–441
94. Hibbard JH, Stockard J, Tusler M. Does publicizing hospital performance stimulate quality improvement efforts? *Health Aff (Millwood)* 2003;**22**:84–94

95. Fung CH, Lim YW, Mattke S, Damberg C, Shekelle PG. Systematic review: the evidence that publishing patient care performance data improves quality of care. *Ann Intern Med* 2008;**148**:111–123
96. Campanella P, Vukovic V, Parente P, Sulejmani A, Ricciardi W, Specchia ML. The impact of public reporting on clinical outcomes: a systematic review and meta-analysis. *BMC Health Serv Res* 2016;**16**:296
97. Metcalfe D, Rios Diaz AJ, Olufajo OA, Massa MS, Ketelaar NABM, Flottorp SA et al. Impact of public release of performance data on the behaviour of healthcare consumers and providers. *Cochrane Database Syst Rev* 2018;**(9)**:CD004538
98. Vallance AE, Fearnhead NS, Kuryba A, Hill J, Maxwell-Armstrong C, Braun M et al. Effect of public reporting of surgeons' outcomes on patient selection, "gaming," and mortality in colorectal cancer surgery in England: population based cohort study. *BMJ* 2018;**361**:k1581
99. Lilford R, Mohammed MA, Spiegelhalter D, Thomson R. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. *Lancet* 2004;**363**:1147–1154
100. Aggarwal A, Lewis D, Mason M, Sullivan R, van der Meulen J. Patient mobility for elective secondary health care services in response to patient choice policies: a systematic review. *Med Care Res Rev* 2017;**74**:379–403
101. Aggarwal A, Han L, Boyle J, Lewis D, Kuyruba A, Braun M et al. Association of quality and technology with patient mobility for colorectal cancer surgery. *JAMA Surg* 2022;**158**:e225461
102. Aggarwal A, Bernays S, Payne H, van der Meulen J, Davis C. Hospital choice in cancer care: a qualitative study. *Clin Oncol* 2018;**30**:e67–e73
103. Friesner DL, Rosenman R. Do hospitals practice cream skimming? *Health Serv Manage Res* 2009;**22**:39–49