European Journal of Cancer 178 (2023) 191–204

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Original Research

Measuring variation in the quality of systemic anti-cancer therapy delivery across hospitals: A national population-based evaluation

Jemma M. Boyle [a,b,*], Jan van der Meulen [a,b], Angela Kuryba [b], Thomas E. Cowling [a,b], Christopher Booth [c], Nicola S. Fearnhead [d], Michael S. Braun [e,f], Kate Walker [a,b,1], Ajay Aggarwal [a,g,1]

[a] *Department of Health Services Research and Policy, London School of Hygiene and Tropical Medicine, London, UK*
[b] *Clinical Effectiveness Unit, Royal College of Surgeons of England, London, UK*
[c] *Department of Oncology, Queen's University, Kingston, Canada*
[d] *Department of Colorectal Surgery, Cambridge University Hospitals, Cambridge, UK*
[e] *Department of Oncology, The Christie NHS Foundation Trust, Manchester, UK*
[f] *School of Medical Sciences, The University of Manchester, Manchester, UK*
[g] *Department of Oncology, Guy's and St. Thomas' NHS Foundation Trust, London, UK*

Received 8 August 2022; received in revised form 10 October 2022; accepted 19 October 2022
Available online 2 November 2022

KEYWORDS
Systemic anti-cancer therapy;
Variation;
Quality of care;
Toxicity;
Colorectal cancer;
Performance indicator;
Quality improvement

Abstract ***Aim:*** To date, there has been little systematic assessment of the quality of care associated with systemic anti-cancer therapy (SACT) delivery across national healthcare systems. We evaluated hospital-level toxicity rates during SACT treatment as a means of identifying variation in care quality.

Methods: All colorectal cancer (CRC) patients receiving SACT within 106 English National Health Service (NHS) hospitals between 2016 and 2019 were included.

Severe acute toxicity rates were derived from hospital administrative data using a validated coding framework. Variation in hospital-level toxicity rates was assessed separately in the adjuvant and metastatic settings. Toxicity rates were adjusted for age, sex, comorbidity, performance status, tumour site, and TNM staging.

Abbreviations: CAPOX, Capecitabine and oxaliplatin; CRC, Colorectal cancer; CTCAE, Common Terminology Criteria for Adverse Events; FOLFOX, 5-FU and oxaliplatin; 5-FU, 5-fluorouracil; HES, Hospital Episode Statistics; ICD-10, International Classification of Diseases, 10th revision; NBOCA, National Bowel Cancer Audit; NHS, National Health Service; ONS, Office for National Statistics; OPCS-4, Office of Population Censuses and Surveys Classification of Surgical Operations and Procedures, 4th revision; RCT, Randomised controlled trial; SACT, Systemic Anti-Cancer Therapy; 2SD, 2 standard deviations; 3SD, 3 standard deviations; UK, United Kingdom.

* *Corresponding author:* Clinical Effectiveness Unit, Royal College of Surgeons of England, London, WC2A 3PE, UK.
E-mail address: jboyle@rcseng.ac.uk (J.M. Boyle).

[1] Joint senior author.

https://doi.org/10.1016/j.ejca.2022.10.017

Results: Eight thousand one hundred and seventy three patients received SACT in the adjuvant setting, and 7,683 patients in the metastatic setting. Adjusted severe acute toxicity rates varied between hospitals from 11% to 49% for the adjuvant cohort, and from 25% to 67% for the metastatic cohort.

Compared to the national mean toxicity rate in the adjuvant cohort, six hospitals were more than two standard deviations (2SD) above, and four hospitals were more than 2SD below. In the metastatic cohort, six hospitals were more than 2SD above, and seven hospitals were more than 2SD below the national mean toxicity rate.

Overall, 12 hospitals (12%) had toxicity rates more than 2SD above the national mean, and 11 (10%) had rates more than 2SD below.

Conclusion: There is substantial variation in hospital-level severe acute toxicity rates in both the adjuvant and metastatic settings, despite risk-adjustment. Ongoing reporting of this performance indicator can be used to focus further investigation of toxicity rates and stimulate quality improvement initiatives to improve care.

1. Introduction

The delivery of systemic anti-cancer therapy (SACT) is a complex care process which includes appropriate patient selection and optimisation, tailoring treatment doses, and the monitoring and management of toxicities [1]. Whilst randomised controlled trials (RCTs) have established the efficacy of SACT, there has been little or no systematic assessment of the quality of SACT delivery within routine care. Much of the available literature on the quality of SACT delivery focuses on access to treatments rather than on outcome measures [2–4].

The only performance indicator currently reported and monitored at hospital level is 30-day mortality after the final SACT treatment, which is more a proxy for the appropriate selection of patients for SACT treatment than a measure of quality of care [5]. Several studies have suggested that the rate of unplanned hospital admissions during SACT could be used as a potential measure of quality [36]. A study in breast cancer patients showed a hospitalisation rate of 43% in patients during SACT with about three quarters of the admissions confirmed as SACT-related events [7].

We have previously validated an indicator of severe acute toxicity (at least Grade 3 according to the Common Terminology Criteria for Adverse Events (CTCAE)) derived from hospital administrative data in colorectal cancer (CRC) [8]. The indicator uses a coding framework to identify specific diagnostic codes during the timeframe of chemotherapy administration which are likely to represent SACT-related severe acute toxicity. As part of this work, we found variations in toxicity rates across different SACT regimens in line with those seen in RCTs. In addition, the rates of toxicity were associated with anticipated risk factors, for example, higher rates in those with comorbidities.

In the current study, we evaluate the use of this measure as a national-level performance indicator to assess hospital variation in severe acute toxicity rates for CRC patients receiving SACT. The indicator will be used to identify potentially outlying hospital performance and benchmark best practice to support quality improvement processes in SACT delivery.

2. Methods

2.1. Data sources

In this national population-based evaluation, we used National Bowel Cancer Audit (NBOCA) data [9], Hospital Episode Statistics (HES) data [10,11], Systemic Anti-Cancer Therapy (SACT) data [12], and Office for National Statistics (ONS) mortality data [13] linked at patient level for CRC patients diagnosed and treated in the English National Health Service (NHS). SACT and HES data were available up until 31 March 2020.

NBOCA is a prospective mandatory database for all newly diagnosed CRC patients in the English NHS. Data items in NBOCA were used to determine sex, age, Eastern Cooperative Oncology Group performance status [14], tumour site, staging (TNM), date of diagnosis, and date of surgery.

HES is a routinely collected administrative dataset of all admissions to English NHS hospitals [11,15]. Diagnoses are coded using the International Classification of Diseases, 10th revision (ICD-10) [16] and procedures are coded using the Office of Population Censuses and Surveys Classification of Surgical Operations and Procedures, 4th revision (OPCS-4) [17]. Data from HES were used to identify admissions for severe acute toxicity (see below), to determine the number of comorbidities according to the RCS Charlson comorbidity score [18] and as one of the sources of SACT information (see below).

HES was also used to determine socioeconomic status. This was derived from the Index of Multiple Deprivation (IMD) which ranks 32,482 geographical areas of England according to their level of deprivation across seven domains [19]. Patients are allocated to an IMD quintile (IMDQ) based on the national ranking of the area corresponding to their postcode.

The SACT dataset captures detailed drug-level information for SACT administered in any inpatient, day case, outpatient, or community setting, including individual administration dates [12]. The SACT dataset was also used to determine the NHS hospital trust (the organisational unit of NHS hospitals in England that can be located on one or more sites) that delivered SACT. Data submission is mandatory for all chemotherapy providers within the English NHS, excluding a small proportion of privately treated patients. One hundred and six hospitals were identified as delivering SACT. These hospitals needed to have treated at least ten patients during the inclusion period (see below) to be included in further analyses.

2.2. Study population

We defined two distinct cohorts of patients aged 18 years and above with a primary diagnosis of CRC (ICD-10: C18, C19 and C20). These included patients identified in the NBOCA database and undergoing treatments at an English NHS hospital during the inclusion period (1 April 2016 to 31 March 2019). These cohorts define the 'denominator' of the performance indicator.

The first cohort included patients with pathological stage III CRC who had received adjuvant SACT according to SACT or HES within the 4-month period after major resection, as per previous work [20]. Patients in the stage III cohort were restricted to those receiving capecitabine and oxaliplatin (CAPOX), 5-fluorouracil (5-FU) and oxaliplatin (FOLFOX), or single agent fluoropyrimidine (capecitabine or 5-FU alone), as per national guidelines [21].

The second cohort included patients who were diagnosed with stage IV CRC and had commenced SACT within the 4-month period after diagnosis, according to SACT or HES. Of the 6,810 patients (89%) with SACT records and therefore drug-level information, approximately 46% of patients received an oxaliplatin-based regimen, 26% an irinotecan-based regimen, 15% single agent fluoropyrimidine, 8% irinotecan with a targeted therapy (e.g. bevacizumab, cetuximab), 5% fluoropyrimidine with a targeted therapy, less than 1% a targeted therapy alone, and less than 1% other agents (e.g. raltitrexed). For stage IV patients, SACT was restricted to treatments given continuously (gaps of no more than 8 weeks between cycles) for a maximum of 12 months.

The pooled reporting of stage III and IV patients was deemed inappropriate given the heterogeneity of the groups in terms of disease burden, underlying fitness, differences in chemotherapy regimens used, and reported differences in overall toxicity rates reported by our prior research [8].

2.3. Definition of the performance indicator

The coding framework defined patients who had experienced severe acute toxicity (the 'numerator' of the performance indicator) according to the presence of predefined ICD-10 diagnostic codes in HES indicative of a SACT-related toxicity (supplementary Table 1) [8]. Severe toxicity was defined as those patients with a selected ICD-10 diagnostic code who required an overnight hospital admission between the administration of the first cycle of SACT and up until 8 weeks after the administration of the last cycle of SACT.

For the small proportion of patients undergoing a surgical procedure during this timeframe, the date of surgery was used as the cut-off for identifying toxicities to ensure that post-operative complications were not captured.

The indicator captures all admissions to any English NHS hospital, regardless of whether or not the hospital provides chemotherapy. Toxicities were attributed to the hospital providing the chemotherapy. Planned and unplanned admissions were included to ensure that we captured all possible SACT-related admissions (direct hospital admission, acute oncology referral, or self-referral).

Table 1
Distribution of types of toxicities following SACT administration for stage III and IV cohorts [52,61].

Toxicity type	Stage III (n = 8173) (%)	Stage IV (n = 7683) (%)
Gastrointestinal	13.2	23.1
Infection	10.4	24.2
Cardiovascular	6.2	14.1
Metabolic & endocrine	5.2	10.4
Constitutional	5.0	10.0
Renal	4.9	9.1
Haematology	4.1	12.0
Pain	3.7	6.5
Neurological	2.6	3.9
Neutropenic sepsis	2.4	7.6
Respiratory	1.2	1.6
Line complications	1.2	3.5
Bleeding	1.1	3.0
Dermatology & rheumatology	0.8	2.3

2.4. Statistical power

We calculated the statistical power to detect an important difference (defined as a 50% increase in the toxicity rate compared to the overall national rate) in the rate of severe acute toxicity in a typical hospital and the overall national rate in England, for each cohort. We calculated the median number of patients per year receiving SACT at each hospital over the 3-year study period and used this to calculate the statistical power for reporting periods of 1-, 3- and 5-years. The definition of an important difference is arbitrary but was chosen because it has been used in previous work and represents a substantial absolute increase in the rate of toxicity [22]. A 5% significance level was used for testing differences between the hospital-level and national rates because it corresponds to the commonly used 95%-control limits of funnel plots (see below).

2.5. Fairness

We determined to what extent we could adjust for the case-mix factors that are likely to affect the risk of severe acute toxicity based on a combination of review of existing literature and expert clinical input. Only factors outside the control of the healthcare provider were adjusted for. We assessed the data completeness of these factors. Missing values for case-mix factors were imputed with multiple imputation [23]. We carried out indirect risk-adjustment, using multivariable logistic regression modelling to obtain expected numbers of severe acute toxicity events per hospital [24]. The association between patient and clinical characteristics and toxicity were assessed separately for each cohort using Wald tests to calculate p values with the significance level set at 0.05.

Calibration of the risk-adjustment model across deciles of predicted risk was assessed using the Hosmer–Lemeshow test, constructing an F-statistic to carry out the test in multiply imputed data [24]. The C-statistic was used to assess model discrimination, combining estimates across imputed datasets using Rubin's rules [23,25].

2.6. Variation between hospitals

Within our cohort of 106 hospitals, we used funnel plots to identify outlying hospitals defined as those with results more than two standard deviations (2SD) (corresponding to 95%-control, or inner, funnel limits), or three standard deviations (3SD) (corresponding to 99.8%-control, or outer, funnel limits) below or above the overall national rate. This is equivalent to carrying out statistical tests comparing a specific hospital's result with the overall national rate using a two-sided 5% or 0.02% significance level, respectively [24,26,27]. Fully-adjusted funnel plots were generated for each cohort.

3. Results

3.1. Study cohorts

Between 1 April 2016 and 31 March 2019, 8,173 patients received adjuvant SACT for stage III CRC. Of these 8,173 patients, 2,074 (25%) had a severe acute toxicity identified according to the indicator. In addition, 7,683 patients received SACT within 4 months of a diagnosis of stage IV CRC. Of these 7,683 patients, 3,625 (47%) had a severe acute toxicity identified. Table 1 summarises the different types of toxicity identified from the coding framework, according to organ system.

3.2. Statistical power

For the stage III cohort, 97 out of 106 English NHS hospitals had treated more than ten patients over the 3-year inclusion period and were included in further analyses. The annual volumes of patients who received adjuvant SACT in each hospital varied considerably with a median value of 24 (range 5–132, interquartile range 15–33). Similarly, 98 hospitals were included for the stage IV cohort, with a median annual volume of 22 (range 5–142, interquartile range 13–32).

The statistical power to detect an increase of 50% compared to the overall national mean rate for different reporting periods (1-, 3-, and 5-year) are presented in supplementary Table 2. These power calculations demonstrate that a 3-year reporting period achieves approximately 70% power in the stage III and 99% power in the stage IV cohort to detect a 50% increase compared to the overall national rate. A 1-year reporting period could have been chosen for the stage IV cohort but, for consistency, the same reporting period was used for both.

3.3. Fairness

For risk-adjustment, the following case-mix factors were identified within the literature and from clinical expertise: age, sex, number of comorbidities, performance status, tumour site, staging, and socioeconomic status [5,28]. All of these case-mix factors are typically accessible from routinely collected datasets. We found that their completeness rate is high in the English NHS (Tables 2a and 2b).

Tables 2a and 2b summarise for the stage III and stage IV cohorts the results of the logistic regression models that capture the associations between the case-mix factors and risk of severe acute toxicity. In both cohorts, severe acute toxicity was increased for female sex, those with more than 2 comorbidities, and advanced T- and N-stage disease. In the stage III cohort, rectal cancer was associated with increased toxicity whereas in the stage IV cohort it was associated with reduced toxicity. In the stage IV cohort, rectal cancer patients were significantly younger and fitter (according

Table 2a
Patient and tumour characteristics and associated severe acute toxicity for patients with stage III disease.

	Stage III (adjuvant) cohort (n = 8173)				
	Number (%)	Severe acute toxicity (%)	Unadjusted odds ratio	Adjusted odds ratio[a] (95% CI)	p value[b]
Patient characteristics					
Age (years)					**0.010**
<60	2369 (29)	566 (24)	1.0	1.0	
60–69	2720 (33)	692 (25)	1.09	1.09 (0.96–1.24)	
70–79	2631 (32)	721 (27)	1.20	1.14 (1.00–1.31)	
≥80	453 (6)	95 (21)	0.85	0.78 (0.61–1.01)	
Sex					**<0.001**
Male	4647 (57)	1072 (23)	1.0	1.0	
Female	3526 (43)	1002 (28)	1.32	1.35 (1.22–1.49)	
Socioeconomic status (IMDQ)					0.969
1 (most deprived)	1159 (14)	293 (25)	1.0	1.0	
2	1460 (18)	374 (26)	1.02	1.02 (0.85–1.22)	
3	1691 (21)	427 (25)	1.00	1.02 (0.85–1.21)	
4	1958 (24)	489 (25)	0.98	1.01 (0.85–1.20)	
5 (least deprived)	1896 (23)	489 (26)	1.03	1.05 (0.89–1.25)	
Missing	*9 (0.1)*	–	–		
RCS Charlson score					**<0.001**
0	4985 (61)	1181 (24)	1.0	1.0	
1	2377 (29)	640 (27)	1.19	1.19 (1.07–1.34)	
≥2	811 (10)	253 (31)	1.46	1.48 (1.25–1.75)	
Performance status					0.546
0	4925 (67)	1217 (25)	1.0	1.0	
1	1946 (27)	526 (27)	1.12	1.06 (0.94–1.20)	
≥2	472 (6)	130 (28)	1.15	1.08 (0.87–1.36)	
Missing	*830 (10)*	–	–	–	
Tumour characteristics					
Site					**0.004**
Colon	6147 (75)	1540 (25)	1.0	1.0	
Rectosigmoid	524 (6)	139 (27)	1.08	1.17 (0.95–1.44)	
Rectum	1502 (18)	395 (26)	1.07	1.25 (1.09–1.44)	
Pathological T-stage					**<0.001**
T1	215 (3)	50 (23)	1.0	1.0	
T2	760 (9)	168 (22)	0.94	0.91 (0.63–1.30)	
T3	4565 (56)	1083 (24)	1.03	1.00 (0.72–1.39)	
T4	2631 (32)	773 (29)	1.37	1.34 (0.96–1.88)	
Missing	*2 (<0.1)*	–	–	–	
Pathological N-stage					**0.002**
N1	5338 (65)	1280 (24)	1.0	1.0	
N2	2835 (35)	794 (28)	1.23	1.18 (1.06–1.31)	

[a] Adjustment for all other variables in the table.
[b] Wald test from multivariable model.

to performance status) than those patients with colon cancer. In the stage IV cohort, poor performance status was associated with increased toxicity but this association was not statistically significant in the stage III cohort.

Following this analysis, we included age, sex, RCS Charlson comorbidity score, performance status, tumour site, T-stage, and N-stage in the logistic regression models used to adjust for case-mix factors when assessing between-hospital variation in rates of severe acute toxicity. Due to the debate around its appropriateness in case-mix adjustment, and the fact it was not associated with increased toxicity, socioeconomic status was included in the model as a sensitivity analysis.

There was no evidence of a lack of calibration for either the stage III cohort (p = 0.711) or the stage IV cohort (p = 0.952), according to the Hosmer–Lemeshow test. The C-statistic of discrimination was 0.58 (95% CI: 0.57 to 0.59) for the stage III cohort and 0.64 (95% CI: 0.62 to 0.66) for the stage IV cohort.

3.4. Variation between hospitals

The unadjusted rates of severe acute toxicity after adjuvant SACT in stage III CRC patients varied considerably between the 97 included hospitals, ranging from 11% to 47% with ten hospitals outside the 95%-funnel limits, including one outside the 99.8%-funnel limits (Fig. 1A). Adjusting for case-mix factors had little effect on the variation in severe acute toxicity rates. Adjusted severe acute toxicity rates ranged from 11% to 49% with the same outlying hospitals (Fig. 1B). This

Table 2b
Patient and tumour characteristics and associated severe acute toxicity for patients with stage IV disease.

	Stage IV (metastatic) cohort (n = 7683)				
	Number (%)	Severe acute toxicity (%)	Unadjusted odds ratio	Adjusted odds ratio[a](95% CI)	p value[b]
Patient characteristics					
Age (years)					**0.026**
<60	2630 (34)	1249 (47)	1.0	1.0	
60–69	2331 (30)	1073 (46)	0.94	0.92 (0.82–1.03)	
70–79	2189 (28)	1071 (49)	1.06	0.96 (0.85–1.09)	
≥80	533 (7)	232 (44)	0.85	0.74 (0.61–0.90)	
Sex					**0.017**
Male	4640 (60)	2111 (46)	1.0	1.0	
Female	3043 (40)	1514 (50)	1.19	1.12 (1.02–1.23)	
Socioeconomic status (IMDQ)					
1 (most deprived)	1206 (16)	615 (51)	1.0	1.0	0.069
2	1399 (18)	683 (49)	0.92	0.93 (0.80–1.09)	
3	1602 (21)	727 (45)	0.80	0.82 (0.71–0.96)	
4	1732 (23)	803 (46)	0.83	0.86 (0.74–0.96)	
5 (least deprived)	1732 (23)	793 (46)	0.81	0.83 (0.72–0.97)	
Missing	*12 (0.2)*	–	–		
RCS Charlson score					
0	4674 (62)	2146 (46)	1.0	1.0	**0.021**
1	2076 (28)	998 (48)	1.09	1.07 (0.97–1.19)	
≥2	759 (10)	395 (52)	1.29	1.24 (1.06–1.45)	
Missing	*174 (2)*	–	–	–	
Performance status					
0	3730 (54)	1616 (43)	1.0	1.0	**<0.001**
1	2255 (33)	1135 (50)	1.32	1.31 (1.18–1.46)	
≥2	902 (13)	500 (55)	1.65	1.61 (1.38–1.87)	
Missing	*796 (10)*	–	–	–	
Tumour characteristics					
Site					**<0.001**
Colon	4929 (64)	2505 (51)	1.0	1.0	
Rectosigmoid	510 (7)	235 (46)	0.83	0.83 (0.69–1.00)	
Rectum	2244 (29)	885 (39)	0.63	0.65 (0.58–0.72)	
Pre-treatment T-stage					**0.004**
T1	14 (0.2)	6 (43)	1.0	1.0	
T2	369 (6)	157 (43)	1.04	0.96 (0.34–2.74)	
T3	3619 (54)	1566 (43)	1.07	0.97 (0.34–2.72)	
T4	2699 (40)	1372 (51)	1.43	1.17 (0.42–3.31)	
Missing	*982 (13)*	–	–	–	
Pre-treatment N-stage					**0.001**
N0	1186 (18)	516 (44)	1.0	1.0	
N1	2935 (44)	1319 (45)	1.07	1.09 (0.96–1.25)	
N2	2610 (39)	1274 (49)	1.26	1.30 (1.12–1.50)	
Missing	*952 (12)*	–	–	–	

[a] Adjustment for all other variables in the table.
[b] Wald test from multivariable model.

corresponded to one hospital being 3SD above, five hospitals being 2SD above, and four hospitals being 2SD below, the national mean toxicity rate. A sensitivity analysis including socioeconomic status in the risk-adjustment did not change the outlying hospitals (results not presented).

The unadjusted rates of severe acute toxicity after SACT for stage IV CRC also varied considerably between the included 98 English NHS hospitals, ranging from 26% to 65% with 12 hospitals outside the 95%-funnel limits (Fig. 2A). Adjusting for case-mix factors had little effect on the variation in severe acute toxicity rates (25%–67%) and outlying hospitals, with 13 hospitals outside the 95%-funnel limits (Figure 2B). This corresponded to six hospitals being 2SD above, and seven hospitals being 2SD below, the national mean toxicity rate. A sensitivity analysis including socioeconomic status in the risk-adjustment generated two new outlying hospitals (results not presented).

Across both cohorts, 22 different hospitals were identified as having rates of severe acute toxicity more than 2SD from the national mean toxicity rate (only 1 hospital had rates more than 2SD for both cohorts). The Pearson correlation coefficient comparing the adjuvant and metastatic rates of toxicity for each hospital was 0.2 (p = 0.090).

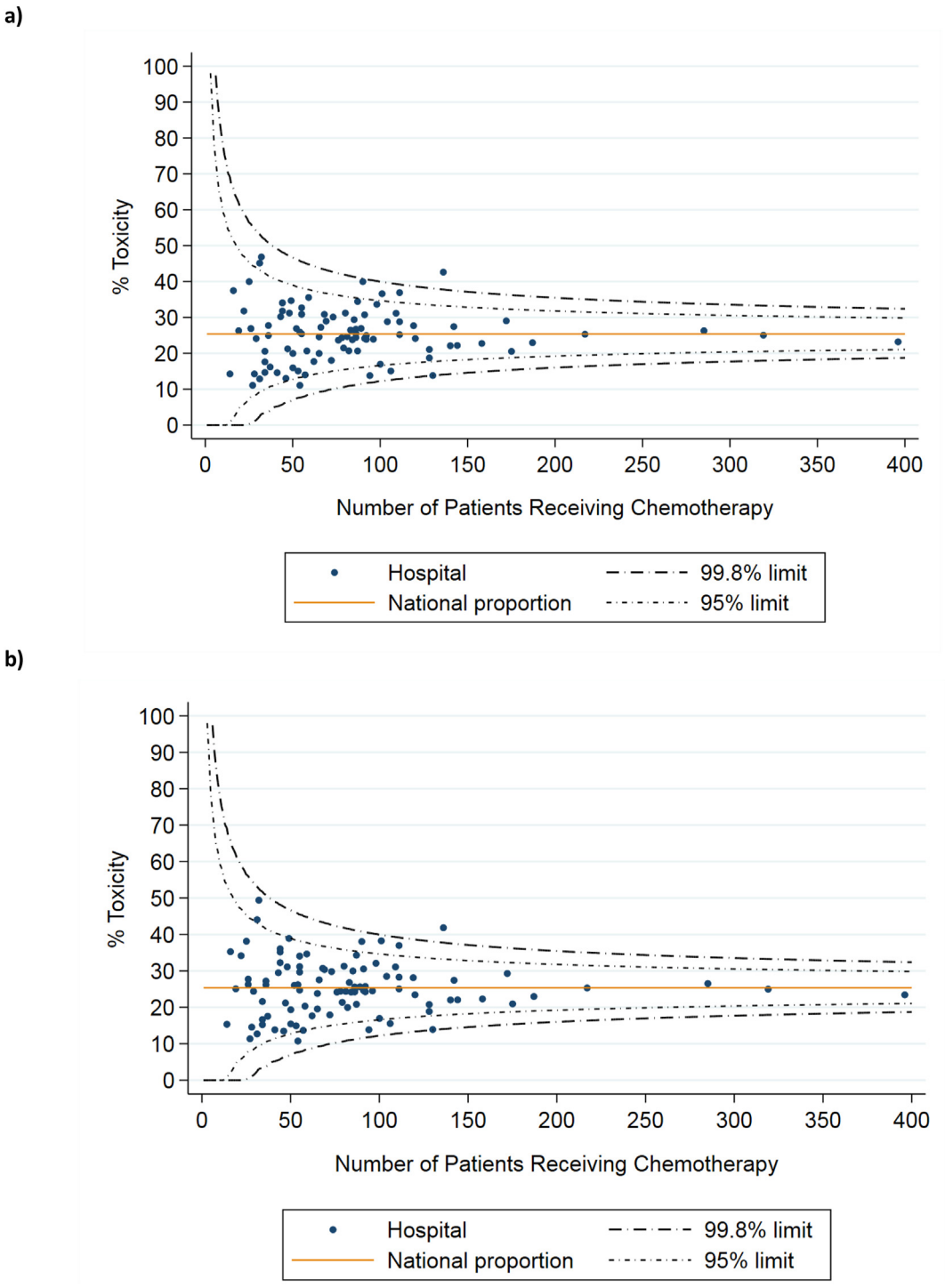

Fig. 1. Funnel plot showing a) unadjusted and b) adjusted rates of severe acute toxicity by English NHS hospital for patients receiving SACT for stage III colorectal cancer.

4. Discussion

This study demonstrates how diagnostic coding in hospital administrative data can be used to derive a hospital-level performance indicator of SACT toxicity across hospitals treating CRC patients. We used a previously validated coding framework based on a predefined set of specific ICD-10 codes and found considerable variation in severe acute toxicity rates between hospitals in both the adjuvant and metastatic setting. We found that hospital rates of severe acute toxicity requiring an overnight hospital admission (equivalent to at least grade 3 CTCAE) varied between 11% and 49% in stage III patients receiving adjuvant chemotherapy,

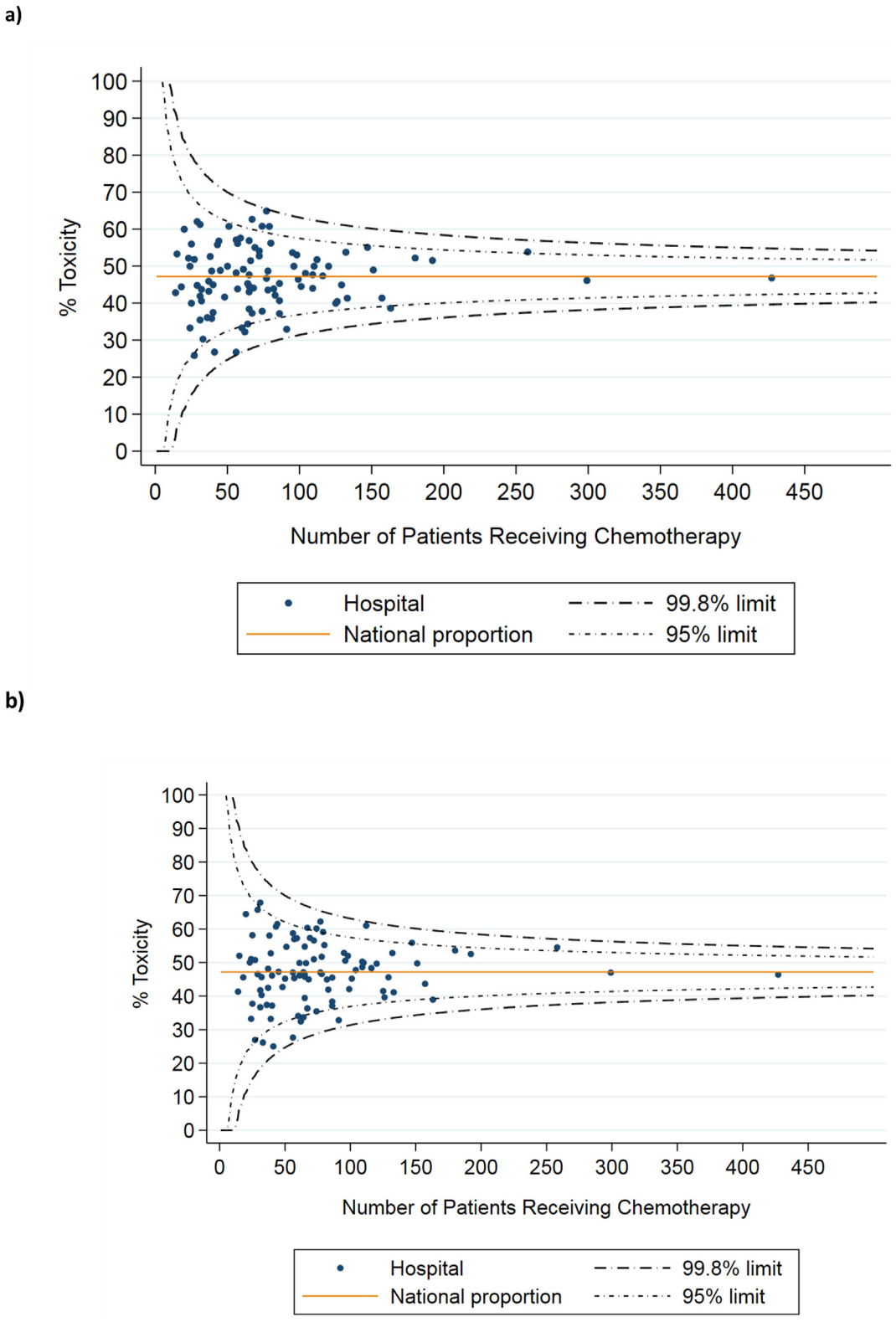

Fig. 2. Funnel plots showing a) unadjusted and b) adjusted rates of severe acute toxicity by English NHS hospital for patients receiving SACT for stage IV colorectal cancer.

and were even higher for stage IV patients with rates varying between 25% and 67%. We identified 22 potentially outlying hospitals, even after adjustment for important case-mix factors.

This hospital performance indicator will be used as part of a publicly reported outlier program in the United Kingdom (UK) from 2022. Within the national outlier process, hospitals are grouped as 'alerts' (greater or less

than 2SD above the national mean) or 'alarms' (greater or less than 3SD above the national mean) [29]. 'Alarm' hospitals are contacted to acknowledge the potential outlier status and start by corroborating data completeness and quality [30]. Once the data is verified, hospitals are expected to formulate a formal response and action plan to understand which factors might be driving unwarranted variation (Fig. 3) [6,27,31]. 'Alert' hospitals are monitored and become potential outliers if they are identified as such in consecutive years. The outlier process is entirely publicly reported. Further work to establish best practice is required and might involve, for example, evaluating survival outcomes in outlying hospitals to evaluate the impact of differential toxicity rates.

These results show that this performance indicator can be used to further explore unwarranted variation in toxicity rates, as well as triggering and guiding initiatives to improve the quality of SACT delivery on a national scale. As the performance indicator is derived from administrative diagnostic coding, the risk of information bias or clinical data manipulation is reduced. The linked datasets included patient and tumour characteristics that allowed the development of a case-mix adjustment model with good calibration and adequate discriminatory power.

The use of ICD-10 codes in the coding framework makes it internationally applicable as it can be applied in different health systems that use ICD-10 codes within their hospital administrative data. In addition, whilst this study has focused on CRC, the coding framework can be applied across different tumour types and regimens, including targeted therapies and immunotherapy.

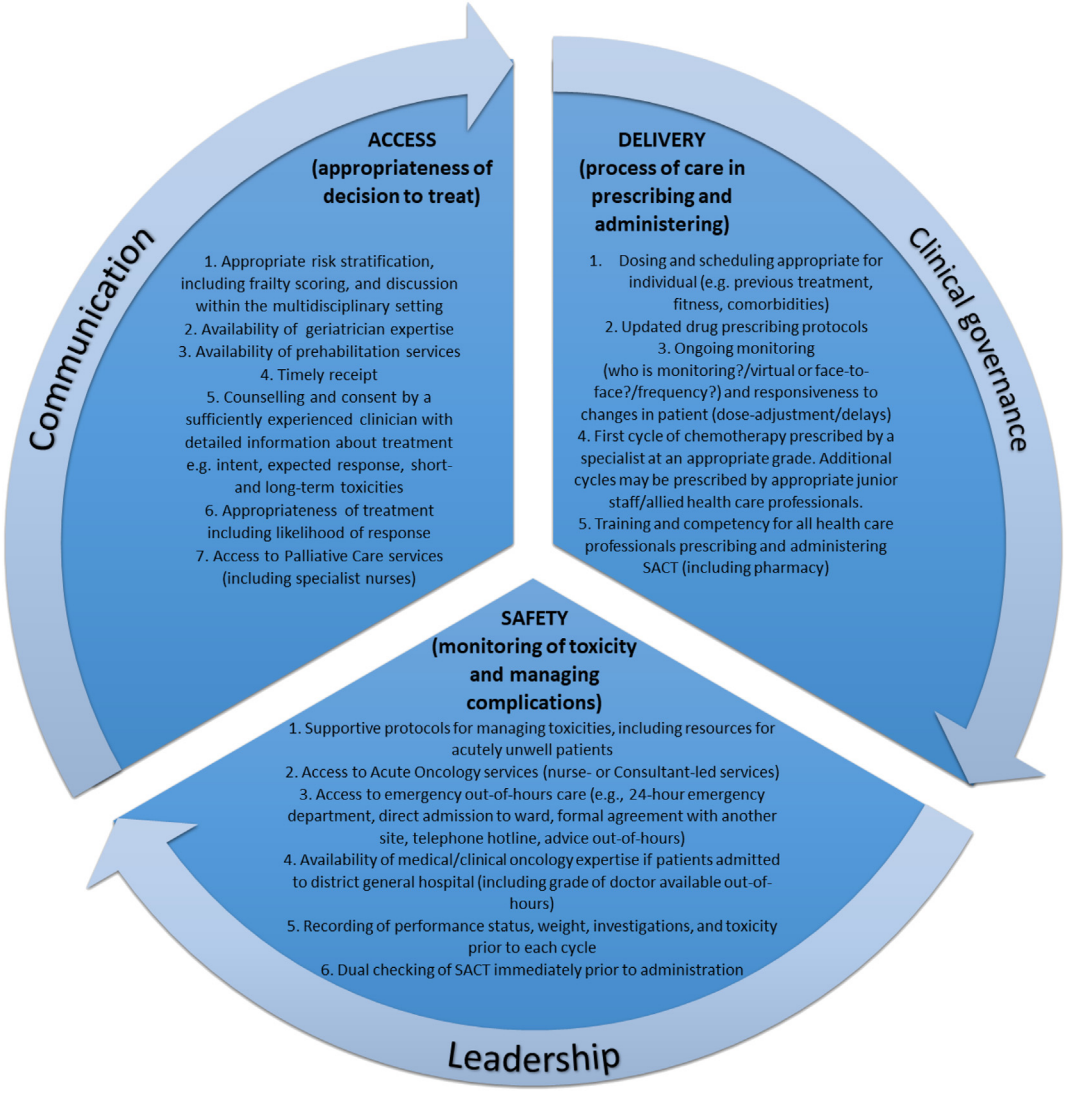

Fig. 3. Quality improvement conceptual framework highlighting potential areas within the SACT care pathway that may represent sources of variation in care.

4.1. Differences in hospital-level severe acute SACT toxicity

The higher rates of toxicity in patients with advanced disease has been previously observed [32]. This is likely due to differences in baseline characteristics, for example, poorer performance status in those with stage IV disease, and the wider range of SACT drugs used within this cohort. Fig. 3 also considers the points along the SACT care pathway that may contribute to the between-hospital variation in toxicity rates.

Outlying rates of severe acute toxicity may reflect differences in the assessment and selection of patients regarding fitness or appropriateness of treatment. For example, patients with a high risk of severe acute toxicity might be identified through comprehensive risk-stratification and discussion within the multidisciplinary team, and there may be differential access to specialist geriatrician and prehabilitation services [33–35]. In addition, patient choice is an important and complex aspect of the decision-making process to consider, with both patient-related factors (e.g. social, cognitive, and psychological issues) and clinician-related factors (e.g. communication and clinician biases) influencing whether SACT is given, as well as what choice of SACT. Patients should be appropriately counselled and consented for SACT treatments by a sufficiently experienced clinician, including the potential short- and long-term side-effects.

Furthermore, unwarranted variation in rates of toxicity may represent under- or over-treatment of patients, as well as inappropriate regimen use and dosing, inadequate or outdated protocols, insufficient monitoring, or failure to recognise and address early signs of toxicity. Despite clear national guidelines in the UK regarding which SACT drugs are approved for which indication [36–41], we recognise that different providers may have variability in prescribing practices beyond this. For example, even current national guidelines do not offer guidance between the use of FOLFOX, CAPOX, or single agent fluoropyrimidine [21].

In addition, the IDEA collaborative results were published during the study period and surveys have highlighted ongoing variation in clinical practice with regards to the choice and duration of combination adjuvant chemotherapy with shifts towards 3 months of treatment for low-risk disease, particularly in the UK [42,43]. For the metastatic cohort, there may be variation in which providers participate in clinical trials. However, adjusting for choice or duration of regimen which are under the control of the provider was deemed inappropriate given that this could represent poor clinical practice.

Finally, the available infrastructure and clinical pathways within hospitals may play a role, for example, variability in access to acute oncology services, emergency services, and the availability of specialist on-site advice out of hours, as well as more generalised disparities in clinical expertise, availability, and training [44]. For example, acute oncology services have been shown to improve outcomes, although substantial variation remains in whether hospitals have access to these services [45]. Acute oncology services may increase planned admissions. Although the vast majority (91%) of admissions for toxicity included in this study were unplanned, this would be an important area to further investigate for potential outliers.

Only one hospital was identified as a low outlier for both cohorts of patients. When comparing both cohorts, there was evidence of a weak association between the rates of toxicity in the adjuvant and metastatic cohorts within each hospital. However, we would not necessarily expect these to align as the two cohorts are very different as evidenced by their mean toxicity rates. First, this may reflect the reduced fitness of patients in the stage IV cohort, and more complex care pathways they undertake, which mean that potential poor treatment, patient selection, and deficient supportive care, are more likely to be exposed. Second, there are differences in care pathways between the two cohorts which might account for differences in toxicity rates within the same provider. For example, the administration and monitoring of adjuvant chemotherapy may be nurse-led in contrast to chemotherapy given for metastatic disease. Third, patient-clinician communication, decision-making, and patient choice are likely to be more complex and nuanced within the metastatic setting which may also influence toxicity rates.

4.2. Strengths and limitations

The strengths of this study include ensuring that the performance indicator meets a pre-defined set of essential criteria ('validity', 'statistical power', and 'fairness' as detailed in the Methods) [46]. The validity of the indicator was demonstrated in an earlier study comparing toxicity across SACT regimens [8]. Although within this study we have reported the overall incidence of severe toxicity after SACT at each hospital, we are also able to detail specific individual toxicities according using our indicator (e.g. neutropenic sepsis, diarrhoea, and line complications – see Table 1). This is hugely important for providing detailed feedback to hospitals to facilitate quality improvement.

In addition, routinely available national clinical cancer data linked to SACT data and hospital administrative data provided over 95% case ascertainment across all English NHS hospitals with good recording of comorbidities, performance status, staging, and detailed SACT information. This also allowed the capture of all hospital admissions, regardless of whether the hospital

provided chemotherapy, and we assigned the toxicity to the hospital delivering the chemotherapy [47].

This, and the good calibration of the risk-adjustment model, meant that toxicity rates between hospitals could be adjusted for important case-mix factors known to influence toxicity, enhancing the fairness of hospital-specific reporting [48]. There is a debate surrounding the complexities of the inclusion of deprivation in risk-adjustment, with varying practice between different reporting programmes [49]. A sensitivity analysis showed minor changes in results for the stage IV cohort which need to be considered. Finally, an overview of potentially actionable areas has been identified as a starting point for targeted local quality improvement (Fig. 3).

The first limitation of this study is the reliance on ICD-10 diagnostic codes in hospital administrative data. However, these diagnostic codes in HES have been shown to be accurate compared to clinical notes, thereby supporting its use for healthcare performance assessment and research [50]. In addition, we used two independent data sources to capture information about the use of SACT. Our previous validation work has shown excellent agreement between the two data sources (SACT and HES data) [20]. Second, there is a possibility that some of the variation between hospitals is due to chance alone. However, we would only expect five hospitals to lie more than two standard deviations from the national mean by chance.

Third, the coding framework is best suited for studying differences between groups of patients (e.g. those treated in different hospitals or receiving different SACT regimens) rather than estimating absolute rates of toxicity which may be over-estimated. However, to limit the likelihood of overestimation we restricted the indicator to only those diagnoses likely to reflect SACT toxicity and to the precise time period of chemotherapy administration (excluding any post-operative period in a small proportion of patients). In addition, previous studies have suggested the vast majority of hospital admissions during SACT treatment are SACT-related [7].

Finally, other studies have demonstrated that mental health status, nutritional status, and laboratory values (e.g. blood tests) were also important predictors for SACT toxicity in older patients [48,51]. However, this information is not routinely included in hospital administrative data and so could not be included as part of the risk-adjustment. As a result, a certain level of 'residual confounding' will need to be accepted, irrespective of which patient groups are being compared.

4.3. Implications

There are several implications of using this performance indicator for a national assessment of the delivery of SACT. First, our study shows that the performance indicator can be used to compare SACT toxicity between hospitals [27]. In the English NHS, similar hospital-level reporting is already routinely available for patients who had surgery or radiotherapy for prostate, bowel, and oesophageal cancer as part of a programme of national clinical audits [9,52,53]. Outcome reporting programmes are also being established internationally [54].

Second, the performance indicator allows ongoing monitoring of severe acute toxicity events within hospitals which will inform continuous local quality improvement processes. Evidence has shown that quality improvement initiatives are more likely to produce positive effects if continuous monitoring and feedback is undertaken [27,28]. As per previous work, the indicator allows specific individual toxicities to be described in detail which can further inform quality improvement processes [8].

Third, an improved understanding of the risks of SACT will inform the counselling of patients and strengthen the process of 'shared decision making' in day-to-day practice, particularly for novel SACT drugs.

In addition to supporting direct patient care, public reporting of severe acute toxicity rates can also provide transparency around best practice through benchmarking to guide patients in making informed choices about the hospital in which they will receive their SACT treatment [55,56]. This avoids the reliance on surrogate markers of care quality (e.g. presence of robotic surgery), and further stimulates quality improvement through competitive mechanisms or regulation by reducing information asymmetry regarding care quality [57–59]. This transparency can also guide investments, with outcomes considered as part of pay-for-performance schemes in order to support greater value in care delivery [60].

Finally, the coding framework developed to identify severe acute toxicity was designed to be broad in order to make it applicable to all types of SACT, including traditional cytotoxic chemotherapy, immunotherapy, and targeted agents [8]. This means that the performance indicator can be used in a wide range of clinical settings and expanded across most cancer types, following appropriate validation.

5. Conclusion

We have evaluated the use of a national performance indicator derived from linked clinical and hospital administrative datasets to assess hospital variation in severe acute toxicity rates for hospitals providing SACT for CRC patients in order to stimulate and support quality improvement. This approach can be applied across different cancer types and in many different countries where similar regional or national clinical and administrative hospital datasets are available.

Funding

This study was undertaken alongside the National Bowel Cancer Audit. The audit is commissioned by the Healthcare Quality Improvement Partnership (HQIP) as part of the National Clinical Audit and Patient Outcomes Programme, and funded by NHS England and the Welsh Government (https://www.hqip.org.uk/national-programmes). AA was supported by a National Institute for Health Research (NIHR) Advanced Fellowship [NIHR300599] and TEC by a fellowship from the Medical Research Council [MR/S020470/1]. Neither HQIP or any of the funders had any involvement in the study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication. The researchers had full independence from HQIP.

Author contributions

Jemma M. Boyle – conceptualization, formal analysis, investigation, data curation, writing – original draft, writing – review and editing, visualisation.

Angela Kuryba – writing – review and editing.

Thomas E. Cowling – formal analysis, writing – review and editing, visualisation.

Jan van der Meulen – writing – review and editing, supervision.

Nicola S. Fearnhead – writing – review and editing.

Kate Walker – conceptualisation, methodology, writing – review and editing, supervision.

Michael S. Braun – conceptualisation, methodology, writing – review and editing, supervision.

Ajay Aggarwal – conceptualisation, methodology, writing – review and editing, supervision.

Chris Booth – writing – review and editing.

Conflict of interest statement

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.

Acknowledgements

This study was based on data collected by the National Bowel Cancer Audit linked to the Systemic Anticancer Therapy database (https://www.chemodataset.nhs.uk/home) made available by the National Cancer Registration and Analysis Service and Hospital Episode Statistics Admitted Patient Care made available by NHS Digital (https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics).

This work uses data provided by patients and collected by the NHS as part of their care and support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejca.2022.10.017.

References

[1] Aggarwal A, Nossiter J, Parry M, et al. Public reporting of outcomes in radiation oncology: the national prostate cancer audit. Lancet Oncol 2021;22(5):e207–15.
[2] Shen S, Krzyzanowska MK. A decade of research on the quality of systemic cancer therapy in routine care: what aspects of quality are we measuring? J Oncol Pract 2015;11(1):55–61.
[3] Mellett C, O'Donovan A, Hayes C. The development of outcome key performance indicators for systemic anti-cancer therapy using a modified Delphi method. Eur J Cancer Care 2020;29(4):e13240.
[4] Donabedian A. The quality of care: how can it be assessed? JAMA 1988;260(12):1743–8.
[5] Wallington M, Saxon EB, Bomb M, et al. 30-day mortality after systemic anticancer treatment for breast and lung cancer in England: a population-based, observational study. Lancet Oncol 2016;17(9):1203–16.
[6] Enright KA, Taback N, Powis ML, et al. Setting quality improvement priorities for women receiving systemic therapy for early-stage breast cancer by using population-level administrative data. J Clin Oncol 2017;35(28):3207–14.
[7] Krzyzanowska MK, Enright K, Moineddin R, et al. Can chemotherapy-related acute care visits be accurately identified in administrative data? J Oncol Pract 2018;14(1):e51–8.
[8] Boyle JM, Cowling TE, Kuryba A, et al. Development and validation of a coding framework to identify severe acute toxicity from systemic anti-cancer therapy using hospital administrative data. Cancer Epidemiol 2022;77:102096.
[9] National bowel cancer audit. Available: https://www.nboca.org.uk/ [accessed 14.01.22].
[10] NHS Digital. Available: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics [accessed 17.01.22].
[11] Herbert A, Wijlaars L, Zylbersztejn A, Cromwell D, Hardelid P. Data resource profile: hospital episode statistics admitted patient care (HES APC). Int J Epidemiol 2017;46(4). 1093–93i.
[12] Bright CJ, Lawton S, Benson S, et al. Data resource profile: the systemic anti-cancer therapy (SACT) dataset. Int J Epidemiol 2019;49(1): 15–15l.
[13] Office for National Statistics. Deaths. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths [accessed 17.01.22].
[14] Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the Eastern cooperative oncology group. Am J Clin Oncol 1982;5(6):649–55.
[15] Ayanian JZ, Zaslavsky AM, Fuchs CS, et al. Use of adjuvant chemotherapy and radiation therapy for colorectal cancer in a population-based cohort. J Clin Oncol 2003;21(7):1293–300.
[16] NHS Digital TRUD. NHS classifications ICD-10. Available: https://isd.digital.nhs.uk/trud3/user/guest/group/0/pack/28 [accessed 22.03.22].
[17] The Health and Social Care Information Centre. Chemotherapy regimens clinical coding standards and guidance OPCS-4 April 2017. 2017. Available: https://classbrowser.nhs.uk/ref_books/ChemRegClinCodingStandGuidApl2017.pdf. [Accessed 22 March 2022].

[18] Armitage JN, van der Meulen JH. Identifying co-morbidity in surgical patients using administrative data with the Royal College of Surgeons Charlson Score. Br J Surg 2010;97(5):772–81.
[19] Camma C, Giunta M, Fiorica F, et al. Preoperative radiotherapy for resectable rectal cancer: a meta-analysis. JAMA 2000;284(8):1008–15.
[20] Boyle JM, Kuryba A, Braun MS, et al. Validity of chemotherapy information derived from routinely collected healthcare data: a national cohort study of colon cancer patients. Cancer Epidemiol 2021;73:101971.
[21] National Institute for Health and Care Excellence. Colorectal cancer. NICE guideline [NG151]. 2020. Available: https://www.nice.org.uk/guidance/ng151 [accessed 17.01.22].
[22] 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(9905):1674–7.
[23] White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Stat Med 2011;30(4):377–99.
[24] Spiegelhalter DJ. Funnel plots for comparing institutional performance. Stat Med 2005;24(8):1185–202.
[25] Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. BMC Med Res Methodol 2009;9(1):57.
[26] Keating NL, Cleveland JLF, Wright AA, et al. Evaluation of reliability and correlations of quality measures in cancer care. JAMA Netw Open 2021;4(3). e212474-e74.
[27] Grote H, Toma K, Crosby L, et al. Outliers from national audits: their analysis and use by the care quality commission in quality assurance and regulation of healthcare services in England. Clin Med 2021;21(5):e511.
[28] Wilson BE, Jacob S, Yap ML, et al. Estimates of global chemotherapy demands and corresponding physician workforce requirements for 2018 and 2040: a population-based study. Lancet Oncol 2019;20(6):769–80.
[29] Outlier Management for National Clinical Audits. Healthcare quality improvement partnership. Available: https://www.hqip.org.uk/wp-content/uploads/2021/11/Appendix-10-HQIP-Outlier-guidance-v4.pdf [accessed 22.03.22].
[30] Shahian DM, Normand SL. What is a performance outlier? BMJ Qual Saf 2015;24(2):95–9.
[31] Kamal AH, Power S, Patierno SR. Addressing issues of cancer disparities, equity, and inclusion through systemized quality improvement. JCO Oncol Pract 2021;17(8):461–2.
[32] Du XL, Osborne C, Goodwin JS. Population-based assessment of hospitalizations for toxicity from chemotherapy in older women with breast cancer. J Clin Oncol 2002;20(24):4636–42.
[33] Silver JK, Baima J. Cancer prehabilitation: an opportunity to decrease treatment-related morbidity, increase cancer treatment options, and improve physical and psychological health outcomes. Am J Phys Med Rehabil 2013;92(8):715–27.
[34] Kalsi T, Babic-Illman G, Ross PJ, et al. The impact of comprehensive geriatric assessment interventions on tolerance to chemotherapy in older people. Br J Cancer 2015;112(9):1435–44.
[35] Wildiers H, Heeren P, Puts M, et al. International Society of Geriatric Oncology consensus on geriatric assessment in older patients with cancer. J Clin Oncol 2014;32(24):2595–603.
[36] National Institute for Health and Care Excellence. Capecitabine and oxaliplatin in the adjuvant treatment of stage III (Dukes' C) colon cancer [TA 100]. 2006. Available: https://www.nice.org.uk/guidance/ta100 [accessed 31.10.18].
[37] National Institute for Health and Care Excellence. Cetuximab and panitumumab for previously untreated metastatic colorectal cancer. [TA439]. 2017. Available: https://www.nice.org.uk/guidance/ta439/chapter/1-Recommendations [accessed 06.10.22].
[38] National Institute for Health and Care Excellence. Trifluridine-tipiracil for previously treated metastatic colorectal cancer. [TA405]. 2016. Available: https://www.nice.org.uk/guidance/ta405 [accessed].
[39] National Institute for Health and Care Excellence. Aflibercept in combination with irinotecan and fluorouracil-based therapy for treating metastatic colorectal cancer that has progressed following prior oxaliplatin-based chemotherapy. [TA307]. 2014. Available: https://www.nice.org.uk/guidance/ta307 [accessed 06.10.22].
[40] National Institute for Health and Care Excellence. Guidance on the use of capecitabine and tegafur with uracil for metastatic colorectal cancer. [TA61]. 2003. Available: https://www.nice.org.uk/guidance/ta61 [accessed 06.10.22].
[41] National Institute for Health and Care Excellence. Cetuximab, bevacizumab and panitumumab for the treatment of metastatic colorectal cancer after first-line chemotherapy: cetuximab (monotherapy or combination chemotherapy), bevacizumab (in combination with non-oxaliplatin chemotherapy) and panitumumab (monotherapy) for the treatment of metastatic colorectal cancer after first-line chemotherapy. [TA242]. 2012. Available: https://www.nice.org.uk/guidance/ta242/chapter/1-Guidance [accessed 06.10.22].
[42] Iveson T, Hanna C, Iveson P, et al. The early impact of the IDEA collaboration results: how the results changed prescribing practice. JNCI Cancer Spectr 2021;5(4).
[43] Hanna C, Boyd K, Jones R. P-335 Self-reported prescribing practices in the setting of adjuvant treatment for colorectal cancer. Ann Oncol 2020;31:S198.
[44] Enright KA, Krzyzanowska MK. Benefits and pitfalls of using administrative data to study hospitalization patterns in patients with cancer treated with chemotherapy. J Oncol Pract 2016;12(2):140–1.
[45] Navani V. How has acute oncology improved care for patients? Curr Oncol 2014;21(3):147–9.
[46] Geary R, Knight H, Carroll F, et al. A step-wise approach to developing indicators to compare the performance of maternity units using hospital administrative data. BJOG An Int J Obstet Gynaecol 2018;125(7):857–65.
[47] Neuss M, Rocque G, Zuckerman D, et al. Establishing a core set of performance measures to improve value in cancer care: ASCO consensus conference recommendation report. J Oncol Pract 2016;13(2):135–40.
[48] Hurria A, Togawa K, Mohile SG, et al. Predicting chemotherapy toxicity in older adults with cancer: a prospective multicenter study. J Clin Oncol 2011;29(25):3457–65.
[49] Tran LD. Social risk adjustment in health care performance measures. JAMA Netw Open 2020;3(6). e208020-e20.
[50] Burns EM, Rigby E, Mamidanna R, et al. Systematic review of discharge coding accuracy. J Public Health 2012;34(1):138–48.
[51] Extermann M, Boler I, Reich RR, et al. Predicting the risk of chemotherapy toxicity in older patients: the chemotherapy risk assessment scale for high-age patients (CRASH) score. Cancer 2012;118(13):3377–86.
[52] Gollins S, Moran B, Adams R, et al. Association of Coloproctology of Great Britain & Ireland (ACPGBI): guidelines for the management of cancer of the colon, rectum and anus (2017) – multidisciplinary management. Colorectal Dis 2017;19(S1):37–66.
[53] Leong K, Hartley J, Karandikar S. Association of Coloproctology of Great Britain & Ireland (ACPGBI): guidelines for the management of cancer of the colon, rectum and anus (2017) – follow up, lifestyle and survivorship. Colorectal Dis 2017;19(S1):67–70.
[54] Rechel B, McKee M, Haas M, et al. Public reporting on quality, waiting times and patient experience in 11 high-income countries. Health Policy 2016;120(4):377–83.
[55] Marshall MN, Shekelle PG, Leatherman S, Brook RH. The public release of performance data: what do we expect to gain? A review of the evidence. JAMA 2000;283(14):1866–74.

[56] Spinks TE, Walters R, Feeley TW, et al. Improving cancer care through public reporting of meaningful quality measures. Health Aff 2011;30(4):664–72.
[57] Aggarwal A, Lewis D, Mason M, et al. Effect of patient choice and hospital competition on service configuration and technology adoption within cancer surgery: a national, population-based study. Lancet Oncol 2017;18(11):1445–53.
[58] Berwick DM, James B, Coye MJ. Connections between quality measurement and improvement. Medical care 2003;41(1 Suppl): I30–8.
[59] Hibbard JH, Stockard J, Tusler M. Does publicizing hospital performance stimulate quality improvement efforts? Health Aff 2003;22(2):84–94.
[60] Scott A, Liu M, Yong J. Financial incentives to encourage value-based health care. Med Care Res Rev 2018;75(1):3–32.
[61] National Cancer Institute. Common terminology criteria for adverse events (CTCAE). Available: https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_5x7.pdf [accessed 22.03.22].