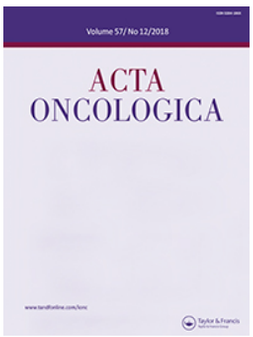

Acta Oncologica

ISSN: 0284-186X (Print) 1651-226X (Online) Journal homepage: http://tandfonline.com/loi/ionc20

Evaluation of the Swedish Colorectal Cancer Registry: an overview of completeness, timeliness, comparability and validity

Peter Moberger, Filip Sköldberg & Helgi Birgisson

To cite this article: Peter Moberger, Filip Sköldberg & Helgi Birgisson (2018) Evaluation of the Swedish Colorectal Cancer Registry: an overview of completeness, timeliness, comparability and validity, Acta Oncologica, 57:12, 1611-1621, DOI: 10.1080/0284186X.2018.1529425

To link to this article: https://doi.org/10.1080/0284186X.2018.1529425

ACTA ONCOLOGICA
2018, VOL. 57, NO. 12, 1611–1621
https://doi.org/10.1080/0284186X.2018.1529425

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Evaluation of the Swedish Colorectal Cancer Registry: an overview of completeness, timeliness, comparability and validity

Peter Moberger, Filip Sköldberg and Helgi Birgisson

[1]Department of Surgical Sciences, Uppsala University, Uppsala, Sweden

ABSTRACT

Background: The Swedish Colorectal Cancer Registry (SCRCR) is a national registry established in 1995 for rectal cancer, and also including colon cancer since 2007. Knowledge of the quality of the registry is vital in order to draw correct conclusions from studies based on the registry. The aim of this study was to assess the completeness, timeliness, comparability and validity of the SCRCR.

Material and methods: Completeness, timeliness and comparability of the registry were estimated. From the SCRCR year 2008, 500 cases were randomly selected to examine the validity of the registry and 486 cases were retrieved. Using hospital patient records as source documents, 130 variables in the SCRCR were reabstracted using the SCRCR registration forms and then compared with the original files.

Result: During the period 2008–2015, the average completeness of the SCRCR was 98.5% for colon cancer and 98.8% for rectal cancer. Timeliness improved between the years 2008 and 2015, with 98% of the patients registered within 12 months for the year 2015. For most of the variables, comparability was estimated to be reproducible and comparable with other registries. Regarding the validity of the registry, when comparing reabstracted data with the original SCRCR data, average agreement was 90%.

Conclusion: The SCRCR can be considered a reliable registry useful for quality assurance and research. Standardization and improvements in journal documentation are needed to improve future evaluation of the source documents.

ARTICLE HISTORY
Received 9 July 2018
Accepted 20 September 2018

Introduction

The Swedish Colorectal Cancer Registry (SCRCR) is a nationwide quality registry that includes data from both the Swedish Rectal Cancer Registry that was launched in 1995 [1] and from the Swedish Colon Cancer Registry that started in 2007 [2].

Today, the SCRCR functions as a single registry using the same forms for registration for both colon and rectal cancer (CRC) although some variables are only applicable to the colon and others to the rectum. The doctor who is responsible for the treatment of the patients also fills in the forms. At each hospital, there is a doctor responsible for the registration process who communicates with the colleagues preforming the registration. The data are registered and monitored at the responsible Regional Cancer Centres located in each of the six healthcare regions in Sweden.

The Umeå Regional Cancer Centre is then responsible for processing and analyzing data from all of Sweden and for publishing yearly reports, including data for individual hospitals that are distributed and discussed at local hospitals and at regional and national meetings. Moreover, data from the SCRCR, together with data from other cancer and disease registries, are used by the National Board of Health and Welfare for comparison of quality and efficiency in the Swedish health care system.

The SCRCR, together with other health care registries, is a valuable source of information for research, quality assurance and benchmarking [3,4]. These open comparisons are widely believed to have led to improvements in many fields of the CRC treatments including diagnostic work-up, surgical treatment and pathologic examination. Data on complications and mortality have been an important source of information for decisions in different regions to centralize treatments. Although it is difficult to prove that the improvements are related to the presence of the SCRCR, it is generally accepted that the population-based improvements in CRC treatment and survival would not have been managed without the SCRCR [2].

The SCRCR has been used extensively for research, with more than 50 publications based on data from the registry. A list of already published works and manuscripts is available on the SCRCR website [5].

It is obvious that it is of major importance to determine the validity of the registry as it is used so extensively for decision-making and research. The validity of the registry has

CONTACT Helgi Birgisson helgi.birgisson@surgsci.uu.se Department of Surgical Sciences, Division of Colorectal Surgery, Uppsala University, Uppsala University Hospital (Entrance 70, 1 tr), SE-751 85, Uppsala, Sweden

Supplemental data for this article can be accessed here.

been partially studied earlier although only for the rectal cancer part of the registry [6–8], with the latest thorough validation performed in 1997.

The aim of this study was to assess the completeness, timeliness, comparability and validity of the SCRCR.

Material and methods

The quality of a registry was assessed regarding four different aspects: completeness, timeliness, comparability and validity.

Completeness means the proportion of cases registered in the SCRCR of all CRCs. In Sweden, all diagnosed cancers are required by law to be registered in the Swedish Cancer Registry. Completeness of the SCRCR is assessed annually by comparison to the *Swedish Cancer Registry*, which has an estimated overall completeness of 96.3% [9]. For the present study, numbers on completeness were retrieved from the annual reports of the SCRCR.

Timeliness is the time from the date of diagnosis of the cancer until it is registered in the SCRCR.

Comparability reflects the adherence in the registry to international standards such as TNM, ICD, pathology reporting standards, Clavien-Dindo and other systems, which makes it possible to reproduce statistics. Comparability was examined by reviewing the definition of the different parameters in the registry [10].

Validity is defined as the proportion of cases in a dataset with a given characteristic (e.g., site and age) that truly have the attribute [11]. For this study, reabstraction was chosen as a validation method.

At the Regional Cancer Centre in Umeå, Sweden, where the SCRCR is managed, 500 cases were randomly selected for reabstraction from a total of 5608 patients that were registered in the CRC for 2008. All 59 hospitals in Sweden treating CRC at that time were involved.

For 48 of the smaller hospitals, which were treating 328 of the patients, letters went out requesting the hospital records of the patients. The investigators visited the 11 larger hospitals, which were treating 172 of the patients.

The steering group of the SCRCR initiated the work in 2010, during the years 2011–2012 the case selection and requests of the hospital records were made, in the years 2013–2014, five surgeons with previous experience in the registration were involved in reabstracting the information from the hospital records, using the registration form and coding instructions that were in use in 2008. These investigators did not have any information about or any access to the original registration.

The registration form comprised 130 variables in five sections, covering preoperative investigations and findings (e.g., date of diagnosis and radiological staging), information related to operative treatment, histopathological examination, postoperative course (including any complications and reoperations) and short-term follow-up (within 30 days). Only variables applicable to each patient are registered, e.g., variables related to surgery are not recorded unless a patient undergoes surgery.

Reabstracted data on paper forms were transferred to an electronic format on the National Information Network for Cancer Care (INCA) platform used for the SCRCR, and the reabstracted dataset was compared with the corresponding (original) dataset from the SCRCR.

Some 'checkbox' variables in the SCRCR containing specifics on a 'parental' variable are encoded only as 'true' or with a missing value. For example, the three variables for pre-therapeutic staging of the primary tumor with CT, MRI and ultrasound, respectively, can only have the value 'true' or a missing value, but are applicable only if the 'parental' variable for pretherapeutic staging of the primary tumor has a value of 1. In these instances, we recoded the 'checkbox' variables to 0 if at least one of them had the value 'true' and the 'parental' variable had the value 1, in order to distinguish from truly missing values. The variables encoding placement of vascular ligatures (Figure 1) are also designed as 'checkboxes' and missing variables were recoded as 0.

Variables concerning histopathological findings were also recoded so that values for 'could not be assessed' (histopathological stage TX, NX, etc.) were regarded as missing values.

The variable for construction of colonic reservoir (a variable no longer in use in the SCRCR) after anterior resection was also recoded to include any reservoir or side-to-end anastomosis in one category.

The results from the comparison of the variables from the two data sets were classified as either matching, non-matching, information missing in both datasets, information missing in the original dataset only, or information missing in the reabstraction dataset only.

Matching variables were regarded as correct. Missing variables in either the original or the reabstracted forms were excluded. If the values were non-matching, it was an incorrect match and defined as an erroneous result. For date of diagnosis, one week of difference was allowed and still considered to be a match.

For the comparisons between the original and datasets, we strived to restrict the analyses to include records were each variable was applicable, e.g., only patients with a record (in the original dataset) of a surgical procedure were included in the assessment of variables concerning postoperative complications. These restriction criteria are shown in Table S1.

The evaluation was done according to the manual for quality registry evaluation available on the website of the Regional Oncology Centre [12] and the outcome was presented in line with the report on the evaluation of the Swedish prostate cancer registry [4].

Long-term follow-up data regarding recurrences and late complications were not included in this study.

Statistical analysis

Data management and statistical analyses were performed using Stata version 15 (StataCorp, College Station, TX, USA). For categorical variables, the percentage of matching results

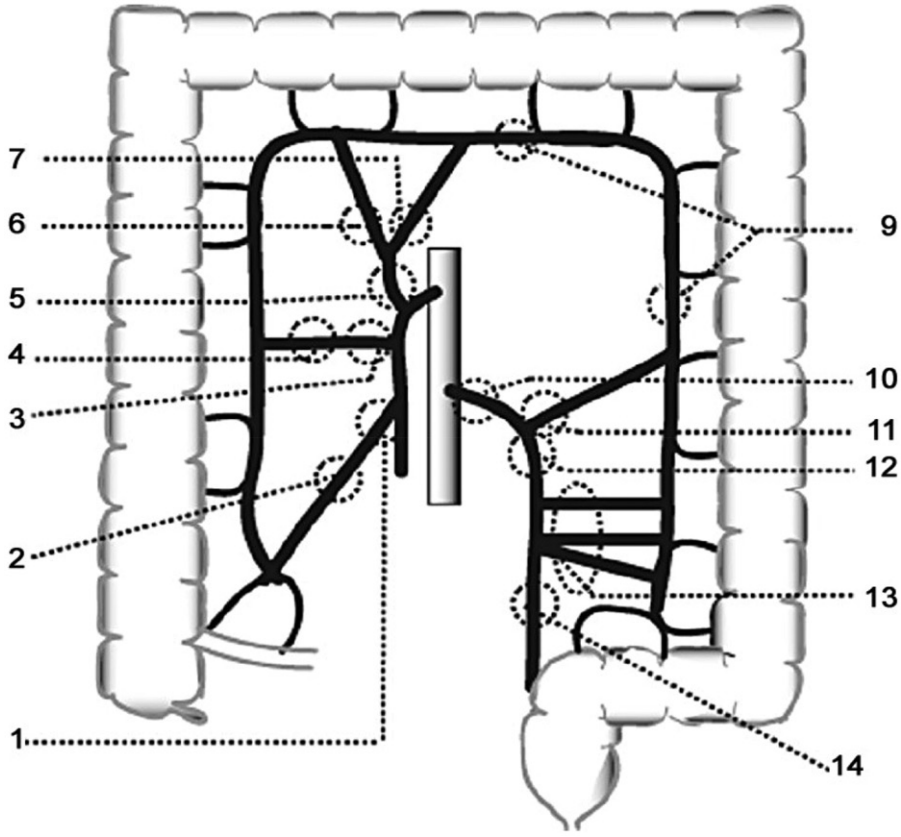

Figure 1. The placements of the proximal vascular ligatures as presented in the registration form of the Swedish Colorectal Cancer Registry.

was calculated and the correlation was analyzed with Cohen's kappa coefficient [13]. For continuous variables, Pearson's correlation coefficient was used to assess the correlation. Similar definitions were used as during the validation of the prostate cancer registry [4] although we in the present analysis adhere to the traditional interpretation of Kappa statistics [14]. Variables with more than 80% agreement or a correlation coefficient greater than 0.8 were classified as almost a perfect agreement, variables with agreement below 80% and correlation of less than 0.6 were regarded as poor agreement. All other variables were regarded as variables with acceptable validity.

Ethics

Evaluation of the quality registries is seen as a necessary part of the quality assurance of the Swedish quality registries and permission from the regional ethical committees are generally not needed [12]. The steering group of the SCRCR initiated and accepted the evaluation of the SCRCR.

Results

Completeness

Between 2008 and 2015, 98.5% of the colon cancer cases and 98.8% of the rectal cancer cases (mean percentages) registered in the Swedish Cancer Registry were also registered in the SCRCR (Table 1).

Table 1. Completeness of the Swedish Colorectal Cancer Registry compared with the Swedish Cancer Registry according to year at diagnosis.

Year	Colon	Rectum
2008	98	100
2009	96	98
2010	98	98
2011	99	99
2012	99.6	99
2013	98.5	98.6
2014	99.3	99.3
2015	99.3	98.5
Mean 2008–2015	98.5	98.8

Timeliness

On average, the time between the dates of diagnosis to entry in the SCRCR in the year 2008 was 221 days and 79% of the cases had been registered within 1 year from diagnosis (Figure 2(a)). For comparison, the timeliness was calculated for the year 2015 and improved to 98% of the cases registered within 1 year from diagnosis (Figure 2(b)). Timeliness for cancer of the colon and cancer of the rectum were almost identical, and the timeliness for the 500 patients selected for reabstraction was identical to the whole of the 2008 year's cohort (data not shown).

Comparability

The variables registered are shown in Table 2. It would be possible to use most of the variables in the registry for comparison with other registries by reviewing the different registration forms.

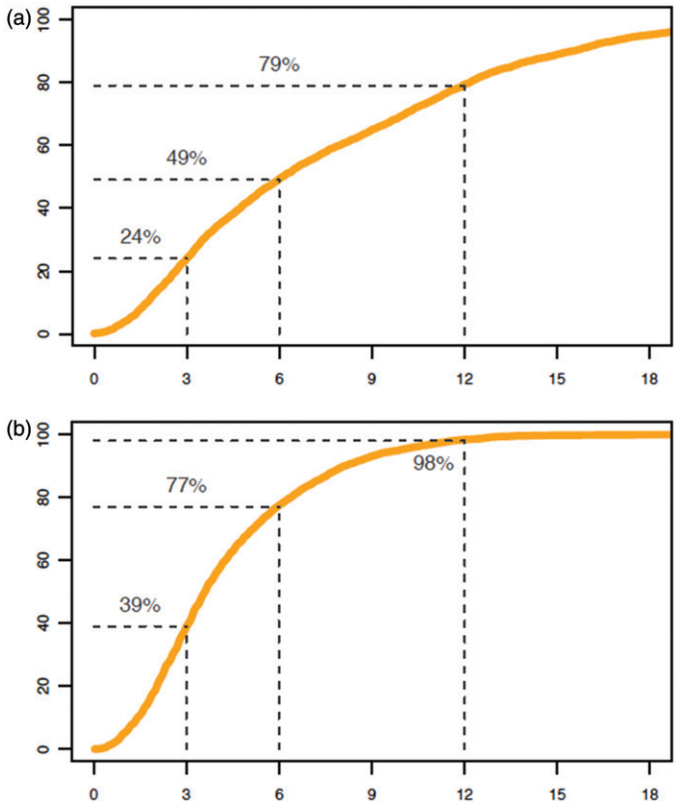

Figure 2. Timeliness calculated as time from diagnosis to the date of entry into the Swedish Colorectal Cancer Registry for (a) the year 2008 and (b) 2015. Y-axis: percentages of patients' entry into the Swedish Colorectal Cancer Registry. X-axis: months.

The variables in the preoperative investigation and findings are mainly diagnostic dates and results of the preoperative imaging documented according to the TNM staging system.

The operative treatment form does register different surgical outcomes, and most of the variables are straightforward and easy to register. For a few variables, such as 'primary inextirpable tumor' and 'curativity as evaluated by surgeon', there could be a risk of variations in how this is estimated by the surgeon.

The histopathologic examination form is based on variables used internationally and based on the TNM system. However, there is some risk of interobserver variability between pathological units in how, for example, vascular invasion is estimated.

Short-term (30 days) follow-up is also based on variables with little risk of having different definitions in other registries. Except for 'patient included in study', different clinics can have different opinions on what should be defined as a study.

Validity

No hospital records were retrieved for 14 patients of the 500 selected for reabstraction, leaving 486 files to be reabstracted (Table 2). The reason for non-retrieval was that copies of the hospital records were not sent for reabstraction from some of the smaller hospitals.

Preoperative investigations and findings

This group had an overall median agreement of 91% and the correlation according to kappa or Pearson's was 0.67 (Table 2).

The only continuous variable in this group, date of diagnosis, showed good agreement according to Pearson's estimation, 0.97, but exact agreement between the original and reabstracted data was only 66% (Table 2).

Operative treatment

The median agreement in this group was 90% and the median correlation was 0.64 (Table 2). To mention some variables with the lowest agreement and correlation, information on mechanical bowel preparation was often missing during reabstraction (Table 3). If a perioperative bowel perforation is recorded, it has to be specified if it is tumor-adjacent or not. In both the original and the reabstraction datasets, eight patients were found to have tumor-adjacent perforation but only three of the patients were exactly the same, with the other patients missing or not estimated to have a tumor-adjacent perforation (Table 3).

Examples of variables with almost perfect agreement are whether preoperative radiation therapy has been given or if the operation was performed laparoscopically (Table 2).

The nine continuous variables had a substantial or almost perfect agreement after correction for extreme outliers (Table 2).

Histopathologic examination

There were 19 variables in this category and the median agreement was 92% and the median correlation was 0.79. The variables with the lowest agreement and correlation were pT4 classification, where the reabstracted data more often found the T4 stage to be T4a when the original stated it was T4b (Tables 2 and 3). The longitudinal resection margin also had low agreement and correlation, where the original was prone to measure it as more than 10 cm when the reabstraction documented it as less than 10 cm (Table 2; Figure 3).

Postoperative course

This group represented 28 variables, of which 19 were categorical. The median agreement was 96% and the median correlation was 0.64.

The variables with the lowest agreement and correlation were surgical complications and other complications, with reabstraction more often stating surgical complication when the original registered no surgical complication (Tables 2 and 3).

Short-term (30 days) follow-up

The median agreement was 88% and the median correlation was 0.66, with the variable 'further follow-up apart from

Table 2. Results of the comparison of the original and reabstracted data for 130 variables in the Swedish Colorectal Cancer Registry.

Variable label	No. of observations expected in the SCRCR	No. of complete records in original data	No. of complete records in reabstracted data	No. of complete records in both original and reabstracted data	% exact agreement (in non-missing records)	Kappa statistic	Pearson's *r*
Preoperative investigations and findings (median value)					(91)	(0.67)	
Hospital code	486	486	474	474	98		
Date of diagnosis	486	486	449	449	66		0.97
Preterapeutic staging, primary tumor	486	486	475	475	80	0.48	
Staging primary tumor, CT scan	356	356	309	309	83	0.64	
Staging primary tumor, MRI	**356**	**356**	**309**	**309**	**94**	**0.87**	
Staging primary tumor, rectal US	356	356	309	309	98	0.62	
Pretherapeutic staging, lung metastasis	486	486	468	468	89	0.60	
Staging lung metastasis, CT scan	411	411	368	368	85	0.67	
Staging lung metastasis, MRI	411	411	368	368	100	0.00	
Staging lung metastasis, chest X-ray	411	411	368	368	89	0.76	
Pretherapeutic staging, liver metastasis	486	486	473	473	91	0.61	
Staging liver metastasis, CT scan	420	420	390	390	91	0.58	
Staging liver metastasis, MRI	420	420	390	390	97	0.71	
Staging liver metastasis, US	420	420	390	390	92	0.62	
Outcome pretherapeutic staging, cT	356	234	126	108	81	0.71	
Outcome pretherapeutic staging, cN	356	255	208	150	85	0.69	
Outcome pretherapeutic staging, cM	**425**	**409**	**373**	**366**	**96**	**0.88**	
Metastasis liver	**420**	**420**	**390**	**390**	**96**	**0.87**	
Metastasis lung	411	411	372	372	96	0.77	
Operative treatment (median value)					(90)	(0.64)	
Pretherapeutic evaluation	486	486	462	462	85	0.69	
Primarily non-resectable tumor	486	486	457	457	90	0.39	
Preoperative radiation therapy	**150**	**147**	**141**	**139**	**96**	**0.91**	
Date of initiation of radiation therapy	71	71	47	47	57		0.98
Radiation dose	**71**	**71**	**63**	**63**	**95**	**0.88**	
Preoperative chemotherapy	**486**	**482**	**446**	**442**	**98**	**0.85**	
Mechanical bowel preparation	*429*	*429*	*179*	*179*	*72*	*0.19*	
Preoperative diversion	429	429	409	409	97	0.78	
Preoperative diversion, temporary intention	*33*	*33*	*23*	*23*	*78*	*0.33*	
Preoperative diversion, stent	**33**	**33**	**24**	**24**	**96**	**0.86**	
Preoperative diversion, stoma	**33**	**33**	**24**	**24**	**96**	**0.86**	
Tumor location	**486**	**486**	**479**	**479**	**100**	**0.99**	
Lower tumor limit by rectoscopy, cm from anal verge	150	148	141	140	61		0.92
Tumor location	**336**	**336**	**327**	**327**	**87**	**0.83**	
Therapeutic intervention	486	485	481	480	97	0.77	
Surgical intervention	**450**	**449**	**434**	**433**	**100**	**1.00**	
Date of therapeutic intervention	**450**	**450**	**431**	**431**	**94**		**0.97**
Ligature placement 1	429	429	419	419	89	0.67	
Ligature placement 2	429	429	419	419	91	0.30	
Ligature placement 3	429	429	419	419	87	0.31	
Ligature placement 4	429	429	419	419	95	0.13	
Ligature placement 5	429	429	419	419	92	0.56	
Ligature placement 6	429	429	419	419	89	0.58	
Ligature placement 7	429	429	419	419	95	0.28	
Ligature placement 8	429	429	419	419	95	0.06	
Ligature placement 9	429	429	419	419	90	0.67	
Ligature placement 10	429	429	419	419	92	0.51	
Ligature placement 12	429	429	419	419	87	0.64	
Ligature placement 13	429	429	419	419	95	0.59	
Ligature placement 14	429	429	419	419	94	0.00	
Resection of other organ	414	414	402	402	89	0.65	
Synchronous resection of liver metastasis	414	414	403	403	99	0.00	
Laparoscopic procedure	**414**	**413**	**400**	**399**	**100**	**0.97**	
Converted (for laparoscopic)	16	16	13	13	100		
Protective stoma	**429**	**427**	**409**	**407**	**98**	**0.91**	
Permanent stoma	**429**	**425**	**405**	**401**	**97**	**0.89**	
Peroperative rectal wash out	115	115	101	101	84	0.68	
Peroperative bowel perforation	414	413	394	393	95	0.45	
Tumor adjacent perforation	*13*	*13*	*7*	*7*	*71*	*0.42*	
Locally radical resection (evaluated by surgeon)	414	414	384	384	89	0.44	
Curative resection (as evaluated by surgeon)	*414*	*414*	*384*	*384*	*79*	*0.49*	
Type of operation	**429**	**429**	**410**	**410**	**97**	**0.88**	
Emergency procedure due to bowel obstruction	75	75	69	69	83	0.58	
Emergency procedure due to bleeding	75	75	69	69	94	0.00	
Emergency procedure due to perforation/abscess	75	75	69	69	90	0.66	
Emergency procedure due to others	75	75	69	69	90	0.31	
Surgical procedure performed	**429**	**429**	**414**	**414**	**90**	**0.88**	
Colonic reservoir	70	62	37	33	82	0.61	
Start of procedure (time)	429	412	313	301	72		0.89

(*continued*)

Table 2. Continued.

Variable label	No. of observations expected in the SCRCR	No. of complete records in original data	No. of complete records in reabstracted data	No. of complete records in both original and reabstracted data	% exact agreement (in non-missing records)	Kappa statistic	Pearson's *r*
End of procedure (time)	429	411	306	293	68		0.90
Duration of procedure, minutes	429	420	306	304	67		0.85
ASA classification	**429**	**417**	**255**	**253**	**89**	**0.82**	
Peroperative bleeding (ml)	**429**	**403**	**331**	**318**	**81**		**0.95**
Bodyweight (kg)	429	373	188	169	77		0.84
Bodyweight not registered	56	52	17	17	100		
Height (cm)	429	336	150	138	88		0.68
Height not documented	93	88	40	38	100		
Histopathologic examination (median value)					(92)	(0.79)	
Adenocarcinoma	486	486	456	456	100	0.00	
T-stage	**429**	**423**	**399**	**396**	**98**	**0.97**	
T1-stage	**28**	**17**	**14**	**11**	**91**	**0.86**	
T3-stage	**246**	**129**	**87**	**77**	**94**	**0.87**	
T4-stage	*88*	*50*	*47*	*33*	*76*	*0.46*	
N-stage	**413**	**401**	**385**	**383**	**97**	**0.95**	
M-stage	**429**	**405**	**334**	**323**	**96**	**0.87**	
Distant metastasis, liver	72	72	56	56	91	0.79	
Distant metastasis, lung	72	72	56	56	89	0.70	
Distant metastasis, other location	**72**	**72**	**56**	**56**	**93**	**0.84**	
No. examined lymph nodes	**414**	**399**	**395**	**384**	**96**		**0.99**
No. positive lymph nodes	**414**	**398**	**396**	**383**	**97**		**0.94**
Clear resection margin	429	419	383	380	89	0.48	
Mucinous tumor	429	354	341	285	91	0.73	
Perineural invasion	429	228	165	147	92	0.77	
Vascular invasion	**429**	**307**	**266**	**241**	**95**	**0.87**	
Differentiation grade	429	428	393	392	86	0.74	
Circumferential resection margin (mm)	414	239	230	195	81		0.41
Longitudinal resection margin (mm)	*414*	*312*	*300*	*253*	*78*		*0.59*
Postoperative course (median value)					(98)	(0.64)	
Postoperative complication	450	448	422	421	84	0.63	
Postoperative infection	450	448	422	421	96	0.55	
Postoperative pneumonia	450	448	422	421	99	0.73	
Postoperative sepsis	450	448	422	421	99	0.00	
Other postoperative infections	450	448	422	421	100		
Cardiovascular complications	450	448	422	421	98	0.72	
Surgical complications	450	448	422	421	89	0.65	
Wound infection	450	448	422	421	95	0.48	
Intraabdominal infection	450	448	422	421	97	0.59	
Wound dehiscence	**450**	**448**	**422**	**421**	**99**	**0.84**	
Bleeding	**450**	**448**	**422**	**421**	**100**	**0.89**	
Anastomotic insufficiency	450	448	422	421	99	0.79	
Stoma complication	450	448	422	421	100	0.00	
Urinary catheter at discharge	450	448	422	421	97	0.58	
Other complications	450	448	422	421	85	0.23	
ICU care	450	446	410	407	93	0.50	
ICU, date of admission	27	27	16	16	75		1.00
ICU, date of discharge	27	26	16	16	75		1.00
Re-operation	**450**	**446**	**414**	**411**	**97**	**0.87**	
Date of first re-operation	**49**	**49**	**42**	**42**	**93**		**1.00**
Date of second re-operation	**49**	**7**	**7**	**4**	**100**	**1.00**	**1.00**
Wound dehiscence as cause for re-operation	49	14	8	8	100		
Bleeding as cause of re-operation	49	7	6	6	100		
Infection as cause of re-operation	49	9	2	2	100		
Anastomotic insufficiency as cause of reoperation	49	14	9	9	100		
Unplanned postoperative hospital admission	450	446	397	393	93	0.64	
Death within 30 days of surgery	450	446	393	390	99	0.73	
Short term (30 days) follow-up (median value)					(88)	(0.66)	
Date of discharge	**450**	**443**	**419**	**416**	**89**		**0.96**
Discharged to	450	440	409	403	94	0.75	
Postoperative multidisciplinary meeting	450	437	395	392	86	0.69	
Adjuvant treatment planned	414	399	375	370	89	0.75	
Palliative treatment	207	187	131	122	84	0.66	
Further treatment and examination with curative intention	486	435	170	156	92	0.41	
Further follow up apart from postoperative return	*486*	*463*	*414*	*407*	*79*	*0.36*	
Patient included in study	486	472	401	391	87	0.62	

Bold font depicts variables with exact agreement above 80% and correlation above 0.8.
Italic font depicts variables with exact agreement below 80% and correlation below 0.6.
Extreme outliers have been excluded from the calculations of Pearson's correlation coefficient.
Where no kappa statistic could be calculated, none is reported.

Table 3. Data validity of the Swedish Colorectal Cancer Registry.

	Reabstraction		
	No	Yes	Missing
Mechanical bowel preparation			
Original			
No	111	8	216
Yes	19	58	17
Missing	5	0	52
Tumor adjacent perforation			
Original			
No	2	1	2
Yes	1	3	5
Missing	6	4	2

	Reabstraction			
	pT4a	pT4b	pT4a/b not mentioned	Missing
p T4 stage				
Original				
pT4a	7	2	4	1
pT4b	6	18	6	5
pT4a/b not mentioned	5	8	15	4
Missing	0	1	1	0

	Reabstraction		
	No	Yes	Missing
Surgical complications			
Original			
No	320	35	24
Yes	10	56	3
Missing	1	0	1
Anastomotic insufficiency			
Original			
No	403	3	27
Yes	3	12	0
Missing	1	0	1

Cross-tables with comparison of original and reabstracted data for selected variables with agreement <80% and correlation below 0.6, and anastomotic insufficiency.

postoperative return' as the only variable with low agreement and correlation (Table 2).

Missing values

Several variables had a high degree of missing values in both the original and the reabstracted datasets. These variables are only documented if an event has occurred. Examples of this are preoperative bowel deviation, bleeding as a cause for reoperation and date of intensive care treatment (Table 2).

Missing values were rarely seen in the original dataset (1.7%). Pretherapeutic TNM and circumferential as well as longitudinal resection margin were missing in more than 10% of cases (Table 2). It was more common that variables were missing only in the reabstracted dataset (9.4%). Several variables had a high number of missing data only in the reabstracted dataset, such as mechanical bowel preparation, body weight and length, operation time and rectal washout (Table 2).

Comparing missing values in patients with rectal cancer ($n = 150$) vs. colon cancer ($n = 336$), missing values were seen for 1.7% of the variables for rectal cancer and 1.5% for colon cancer. In the reabstracted dataset, the percentages of missing variables were 9.1% for rectum and 8.7% for colon.

No difference was seen comparing university vs. regional hospitals. Missing values in the original dataset in the visited hospitals was 2.0%, compared with 1.5% in the non-visited hospitals and missing values only for the reabstracted dataset were seen for 7.7% of the variables in the visited hospitals compared with 9.3% for the non-visited.

Discussion

This study reveals that the SCRCR is a detailed and well-functioning quality registry with high coverage and low levels of missing data. However, the study also identifies several areas for potential improvement.

With more than 98% coverage, the completeness of the SCRCR can be regarded as very good and it is in line with other Swedish quality registries, such as the National Prostate Cancer Registry, which has 98% completeness [4], and the National Registry for Oesophageal and Gastric Cancer, which has a completeness of 95.5% [3]. For international comparisons, the Dutch surgical colorectal audit registered 95% of the cases found with colorectal cancer at the national cancer registry in 2011 [15]. According to the annual report of the national bowel cancer audit in England and Wales, 95.3% of cases registered in the office for national statistics cancer registration dataset were also registered in the audit for the years 2015–2016 [16]. The reason for this good coverage in Sweden can mainly be attributed to the control system of the Regional Cancer Centres, which monitor the completeness of the SCRCR against the Swedish Cancer Registry for the region that they serve. When a patient is discovered in the cancer registry and not found in the SCRCR, an electronic registration form is sent by the INCA system to the responsible registrar at the hospital where the patient was diagnosed. The deadline for completing the registration for each year is the end of February of the upcoming year. The completeness of the cancer quality registries are reported at the hospital level in the yearly reports and published on the website of the regional cancer center. This further encourages the hospitals to finish the registration in time [17].

To estimate the completeness, the Swedish quality registers, including SCRCR, have decided to compare themselves to the *Swedish Cancer Register*, which is regarded as the gold standard. However, it is known from a study comparing records from 1998 in the *Hospital Discharge Register* to the *Swedish Cancer Register*, that an estimated 3.7% of cancer cases are missing in the *Swedish Cancer Register* [9]. The completeness of the *Swedish Cancer Register* regarding colorectal cancer specifically is not known, but it is possible that there is a higher risk of underreporting when the cancer is not histologically verified (e.g., in older individuals, not receiving any treatment, with metastatic disease only diagnosed by radiology). It is also known that Sweden is the only Nordic country still not using death certificates as a source of information on cancer diagnosis – the proportion of non-registered cancer cases in Sweden is therefore higher as compared to the other Nordic countries [18]. It has been estimated that around 4% of all cancer diagnoses that would be

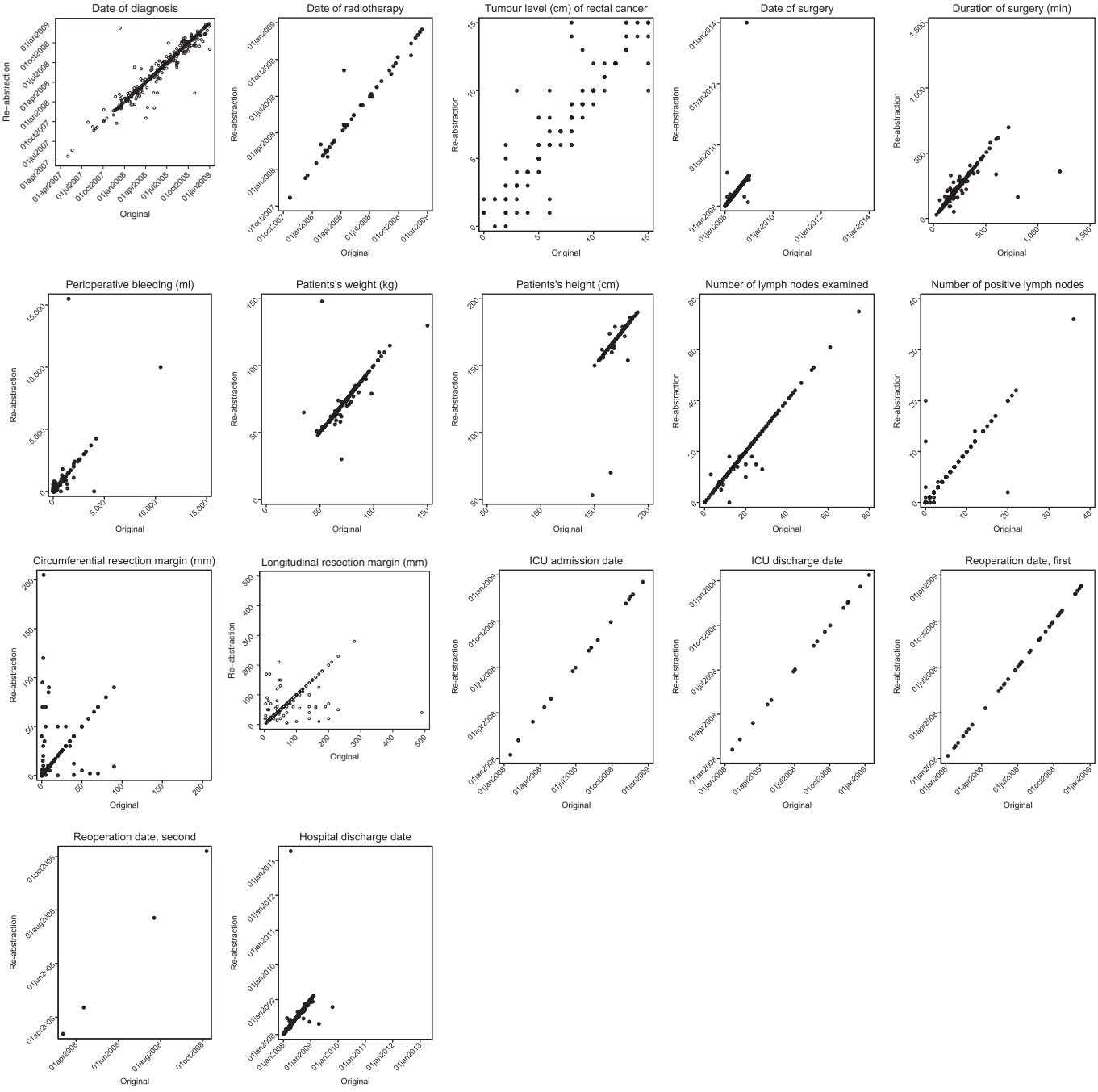

Figure 3. Scatterplots of two continuous variables with poor agreement or correlation found during the reabstractation of the Swedish Colorectal Cancer Registry.

death certificate-initiated could be missing in Sweden [19]. This does not, however, have a great effect on the clinical cancer registries because these are mainly untreated cases. Refer Parkin and Bray for further techniques available for estimating completeness of a Cancer Registry [20].

The timeliness, calculated as the cumulative proportion registered in the SCRCR from the date of diagnosis, reveals that it takes considerable time for the patients to be registered in the SCRCR. The results are good in comparison with the National Prostate Cancer Registry, where cumulative percentage registered was 80% after 1 year [4] and with the National Registry for Oesophageal and Gastric Cancer [3], which showed a cumulative registration rate of 78%. Although improvements were seen for the year 2015, this reveals that the cancer quality registries are currently not suitable for real-time estimation of the cancer treatment. Because the registration forms have different sections, as described in the material and methods, it is not possible to complete the registration until the treatment and follow-up visits have been done. An effort to change the logistics of the registration was commenced in the year 2017 and each section of the current registration form has its own electronic registration form available on the INCA platform. Each form will be possible to fill in after each completed element of

the treatment process. Moreover, a separate form for radiologic estimates and oncologic treatments was added.

The comparability was estimated to be in line with international standards. The majority of the variables are objective, with some of the variables at risk for subjective estimation by the surgeon or the pathologist. Dates, TNM classification and measurements based on the international units can be considered a secure manner of registration that should not vary between different quality registries. There is, however, always a risk that those who do the registration are not aware of the definitions or a change in definition, such as when a new version of the TNM grading system is introduced. It would be a great improvement if different national quality registries were to cooperate on formation of international standards and definitions in line with the TNM staging system.

The source documents for reabstraction and validation are the patient files. The possibility to properly evaluate the registry depends on how well information has been recorded in these files. Missing information in 9.4% of the variables in the reabstracted dataset, compared with only 1,7% of the original dataset, reveals that the documentation of variables in the patient files is incomplete. Since the year 2008, many hospitals have introduced standard forms for multidisciplinary meeting notes, operation and pathology reports, and radiologists answer the reports in TNM. In 2008, the surgeons were responsible for all the registrations. This is gradually changing, and the aim is that the pathologic, oncologic and radiologic clinics will be responsible for reporting data related to their speciality.

When different sections of the registration were analyzed, preoperative investigations and findings generally had good agreement or correlation, except for 'staging of the primary tumor with CT'. It is known that the preoperative T and N staging are often suboptimal [21] but also there is a possibility that the interpretation of the CT reports can be different between doctors if it is not standardized. However, a big effort has been under way in Sweden to improve the preoperative staging of the primary tumor, not only for rectal cancer with MRI but also for colonic cancer with CT scan. This has been done with working group seminars for radiologists and by discussing all cases at preoperative multidisciplinary meetings.

For operative treatment, some variables are difficult to assess from the patient files. The location of proximal ligature often had good agreement but poor correlation. The level of ligature is somewhat difficult to assess, since several ligatures can be registered in one operation (Figure 1). The biggest difficulties for the validators were to assess whether it was arteria mesenterica inferior or arteria hemorrhoidalis superior that was ligated at anterior resection. This was not always clear, and the terminology differed between operators. This further shows the need for standardized operation reports, as has been further discussed by Boström et al. [22].

Perioperative perforation of the bowel close to the tumor was classified as a poor variable. This is of course a major problem, as the variable is important because it is thought to predict local recurrences in rectal cancer and peritoneal carcinomatosis in colon cancer. The question is if the surgeons are reluctant to document a perforation, or if they do not know that it only applies to a perforation of bowel where the tumor is located and not, for example, a perforation of the small bowel during adhesiolysis. This is something surgeons and researchers have to be aware of.

For the postoperative course, some variables were hard to find information about in the files, as is shown in the variable of postoperative multidisciplinary meeting. Surgical complications in general and other complications were classified as poor variables and are good examples of variables with broad definitions and the possibility of great variability between from the registrars. The complications that are supposed to be registered are from Clavien 2 and upwards [23], and this might not be fully understood. A recent study on 1507 patients reported to the SCRCR in 2007–2013 revealed a substantial underreporting of anastomotic leakage after rectal resection, with 52 out of 180 leakages (29%) not reported [24]. The authors stressed the importance of adherence to internation definition of anastomotic leakage as defined by the international study group of rectal cancer [25]. In the present study the exact agreement of anastomotic leak was 94% (including agreement for the absence of anastomotic leak) and of reoperation due to leakage to be 100%, this is in line with earlier study which showed that reoperations and date of reoperation were of good quality [6].

Overall, the accuracy of the registry is satisfying. The variable with the least agreement was 'date of diagnosis'. A one-week difference was considered close enough for a match, but there was still only a 44% agreement. In the instructions for the registry in 2008, it should be stated that the date of diagnosis should be the date that a morphologic diagnosis was established. This might not be well known, and the date might be interpreted arbitrarily. For example, one might choose the date of radiology suspecting the diagnosis, the date of colonoscopy/rectoscopy, the date of first visit at surgical department, etc. It is of great importance that registrars cohere to the same definition of the date of diagnosis as this is a variable much used when times from diagnosis to first treatment is calculated.

From the scatterplots, we can see outliers that are obviously not correct and are caused by human errors, such as a patient's height registered to only 60 cm in the original or the date of surgery in January 2014 in the reabstracted version. These are examples of variables that should be made impossible to register incorrectly through the use of a digitalized control function when the data are registered.

To explain why the results have not been presented earlier, it can be clarified that the present evaluation have met many obstacles during the process, originally the decision was made to do the validation in 2008 and the cohort was selected in 2010. Originally the plan was to have copies of the patient records sent to the research group for validation, however many larger clinics were reluctant to copy so many records. Moreover, the researchers originally engaged in the project did quit and the present study group attended the project in 2012. The decision was made to visit the clinics

with the work finished in 2014. Since then data has been entered into the re-abstracted database, statistics have been made and the text has been written. During this time, the other cancer registers have been validated [3,4] and a manual for validation of the quality registers has been constructed [12].

One lesson is that validation is very time-consuming and the selection of hospitals to visit can be done more pragmatically than was done in the present analysis. This was, for example, done during the evaluation of the prostate cancer registry, in which only 12 hospitals and 5 private health care providers were visited [4].

Although the year of validation was 2008, it is still very relevant to know the quality of the registered data from this year, as there is ongoing research including the diagnostic year of 2008. However, it is highly relevant to do a new validation to be able to see whether there are improvements in the agreements and correlations of the variables, also including oncologic treatment, long term complications and recurrence variables which is crucial when evaluating the true quality of oncologic care.

A lesson learned is that standardized hospital record documentation should be prioritized in health care. This would reduce the amount of missing variables in the hospital records and improve the comparability between hospitals and the quality of the registered data. It would even enable the hospital record system to directly link relevant data to the quality registries, a work in slow progress because of the multitude of different hospital record systems running in Sweden.

Interobserver variability was not estimated between the five investigators that did the reabstraction. This is relevant as there are variables that could be interpreted differentially between investigators although all of them used the same registration manual for their work, something that future validation work should take into account.

Conclusion

The SCRCR can be considered a reliable registry useful for quality assurance and research. Standardization and improvements in journal documentation are needed to improve future evaluation of the source documents.

Acknowledgments

We thank our supervisor Professor Lars Påhlman, who was engaged in this work but passed away in 2015. We thank our surgical colleagues Svante Nordgren and Björn Öjerskog, who participated in the reabstraction of data.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study was supported by the Swedish Cancer Society and grants from the Uppsala University Hospital (ALF), Uppsala, Sweden.

References

[1] Pahlman L, Bohe M, Cedermark B, et al. The Swedish rectal cancer registry. Br J Surg. 2007;94:1285–1292.
[2] Kodeda K, Nathanaelsson L, Jung B, et al. Population-based data from the Swedish Colon Cancer Registry. Br J Surg. 2013;100:1100–1107.
[3] Linder G, Lindblad M, Djerf P, et al. Validation of data quality in the Swedish National Register for Oesophageal and Gastric Cancer. Br J Surg. 2016;103:1326–1335.
[4] Tomic K, Sandin F, Wigertz A, et al. Evaluation of data quality in the National Prostate Cancer Register of Sweden. Eur J Cancer. 2015;51:101–111.
[5] Regionala cancercentrum i samverkan. Publikationer från svenska kolorektacancerregistret (SCRCR) [Publications from the Swedish Colorectal Cancer Registry (SCRC)] [updated 2017 February 17; cited 2017 December 19]. Available from: https://www.cancercentrum.se/globalassets/cancerdiagnoser/tjock–och-andtarm-anal/kvalitetsregister/aktuell_forskning_kolorektal_februari-2017.pdf
[6] Gunnarsson U, Seligsohn E, Jestin P, et al. Registration and validity of surgical complications in colorectal cancer surgery. Br J Surg. 2003;90:454–459.
[7] Jorgren F, Johansson R, Damber L, et al. Risk factors of rectal cancer local recurrence: population-based survey and validation of the Swedish rectal cancer registry. Colorectal Dis. 2010;12:977–986.
[8] Jorgren F, Johansson R, Damber L, et al. Validity of the Swedish Rectal Cancer Registry for patients treated with major abdominal surgery between 1995 and 1997. Acta Oncol. 2013;52:1707–1714.
[9] Barlow L, Westergren K, Holmberg L, et al. The completeness of the Swedish Cancer Register: a sample survey for year 1998. Acta Oncol. 2009;48:27–33.
[10] Regionala cancercentrum i samverkan. INCA variable description – national quality register for colorectal cancer since 2007 [INCA variabelbeskrivning – nationellt kvalitetsregister för kolorektal cancer from 2007] [cited 2017]. Available from: http://www.cancercentrum.se/globalassets/cancerdiagnoser/tjock–och-andtarm-anal/kvalitetsregister/Variabelbeskrivning-tjockoch-andtarms-cancer-2014-03-13.pdf
[11] Bray F, Parkin DM. Evaluation of data quality in the cancer registry: principles and methods. Part I: comparability, validity and timeliness. Eur J Cancer. 2009;45:747–755.
[12] Lambe M, Rejmyr-Davis M. Validering av kvalitetsregister på INCA [Validation of quality registers on the INCA platform] 2014 [cited 2017]. Available from: http://www.cancercentrum.se/globalassets/vara-uppdrag/kunskapsstyrning/kvalitetsregister/manual-for-validering-av-kvalitetsregister-inom-cancer.pdf
[13] Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20:37–46.
[14] Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med. 2005;37:360–363.
[15] Van Leersum NJ, Snijders HS, Henneman D, et al. The Dutch surgical colorectal audit. Eur J Surg Oncol. 2013;39:1063–1070.
[16] National Bowel Cancer Audit. Annual report 2017 (version 2) [cited 2018 June 1]. Available from: www.nboca.org.uk/content/uploads/2017/12/NBOCA-annual-report-2017-v2.pdf
[17] Regionala cancercentrum i samverkan. Annual reports from the Swedish Colorectal Cancer Registry [cited 2018 Jun 1]. Available from: http://www.cancercentrum.se/samverkan/cancerdiagnoser/tjocktarm-andtarm-och-anal/tjock–och-andtarm/kvalitetsregister/rapporter/
[18] Pukkala E, Engholm G, Højsgaard Schmidt LK, et al. Nordic Cancer Registries - an overview of their procedures and data comparability. Acta Oncol. 2018;57:440–455.
[19] Mattsson B, Wallgren A. Completeness of the Swedish Cancer Register. Non-notified cancer cases recorded on death certificates in 1978. Acta Radiol Oncol. 1984;23:305–313.
[20] Parkin DM, Bray F. Evaluation of data quality in the cancer registry: principles and methods Part II. Completeness. Eur J Cancer. 2009;45:756–764.

[21] Sjövall A, Blomqvist L, Egenvall M, et al. Accuracy of preoperative T and N staging in colon cancer–a national population-based study. Colorectal Dis. 2016;18:73–79.
[22] Boström P, Rutegård J, Haapamaki M, et al. Arterial ligation in anterior resection for rectal cancer: a validation study of the Swedish Colorectal Cancer Registry. Acta Oncol. 2014;53:892–897.
[23] Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg. 2004;240:205–213.
[24] Rutegard M, Kverneng Hultberg D, Angenete E, et al. Substantial underreporting of anastomotic leakage after anterior resection for rectal cancer in the Swedish Colorectal Cancer Registry. Acta Oncol. 2017;56:1741–1745.
[25] Rahbari NN, Weitz J, Hohenberger W, et al. Definition and grading of anastomotic leakage following anterior resection of the rectum: a proposal by the International Study Group of Rectal Cancer. Surgery 2010;147:339–351.