Acta Oncologica, 2015; 54: 1723–1728

informa
healthcare

ORIGINAL ARTICLE

Validity of Norwegian Rectal Cancer Registry data at a major university hospital 1997–2005

SUNNIVA T. SAKKESTAD[1], BJØRN C. OLSEN[1], ANNE KARLICZEK[1,2], OLAV DAHL[3,4] & FRANK PFEFFER[1,2]

[1]*University of Bergen, Department of Clinical Medicine, Bergen, Norway,* [2]*Haukeland University Hospital, Department of Surgery, Bergen, Norway,* [3]*University of Bergen, Department of Clinical Science, Bergen, Norway and* [4]*Haukeland University Hospital, Department of Clinical Oncology, Bergen, Norway*

ABSTRACT

Background. The Norwegian Rectal Cancer Registry (NRCR) has been used extensively to monitor patient treatment and outcomes since its establishment in 1993. Control of data validity is crucial to ensure reliable information, but an audit of the NRCR data validity has not been performed so far. This study aims to validate NRCR data on patients diagnosed in the period 1997–2005, Department of Surgery, Haukeland University Hospital.

Material and methods. The material comprises NRCR data on all 482 patients diagnosed with rectal cancer in the period 1997–2005 at a major Norwegian university hospital. We checked 50 variables for discrepancies by comparing NRCR data with the medical records. All erroneous registrations were recorded.

Results. One hundred patients (21%) had one or more data discrepancies in the registry, and 131 errors (0.5%) were noted in total. Sixteen variables (32%) had no erroneous registrations. Pre-operative CT and type of surgical procedure had the highest proportion of erroneous registrations (2.1%). Recorded errors were grouped into five variable categories: Pre-operative evaluation and adjuvant treatment (40 errors), surgical treatment (44 errors), pathological evaluation (20 errors), complications (7 errors) and oncological outcomes (20 errors). The majority of erroneous registrations (45%) were considered minor in severity, 27% were moderate and 28% were major.

Conclusion. Assessment of the NRCR data from a nine-year period showed a good data validity in this hospital cohort.

Since 1953, data on all new cancer cases in Norway have been routinely entered into the Norwegian Cancer Registry, providing a unique epidemiological insight into the distribution of cancer in a national population [1]. During the last two decades, a number of quality registries have also been established within the framework of the Norwegian Cancer Registry to strengthen the quality of health care for specific cancer diagnoses [2]. The Norwegian Rectal Cancer Registry (NRCR) was established in 1993 as one of the first of these sub-registries with the articulate aim of gathering high quality clinical information to be used in surveilling and ultimately improving rectal cancer treatment [1].

Simultaneously with the establishment of the NRCR, a national effort was made to standardise the surgical treatment of rectal cancer. In this effort, two major improvements were made. First, total mesorectal excision was implemented as the gold standard for operating rectal cancer, second, the surgical treatment was centralised to a restricted number of hospitals and surgeons [3]. These joint efforts led to a significant improvement in both total survival rate and local recurrence rate in the national rectal cancer cohort approximately 20 years after the change in treatment protocol [4,5].

The material in the NRCR is subject to considerable research and surveillance activity, and it has a relatively direct impact on how rectal cancer is treated in Norwegian hospitals. Satisfactory data quality is essential for the registry to be able to continue its important role in rectal cancer treatment by delivering trustworthy data of treatment outcomes. A continuous quality control of registry data is being

---

Correspondence: S. T. Sakkestad, Daniel Hansens gate 3, 5008 Bergen, Norway. Tel: + 47 41640651. E-mail: sunniva.sakkestad@gmail.com

(Received 30 January 2015; accepted 11 March 2015)

ISSN 0284-186X print/ISSN 1651-226X online
DOI: 10.3109/0284186X.2015.1031913

performed within the registry by collecting data from multiple sources such as medical records, pathological laboratories, general practitioners and the National Statistics Office (Statistics Norway), and by sending out reminders to the sources that have not yet provided the most recent clinical data. In addition, the research activity in the registry entails an indirect quality control as data items are checked for errors, incorrect coding and extreme values before any ultimate analysis is carried out [6].

National registries have a pivotal role in assuring cancer treatment quality on a regional, national and international level by comparing outcomes [7]. A systematic approach to the data evaluation is important to get an impression of the data quality as a whole, but there exist relatively few publications that address data quality in specific registries in Norway. Validity, also referred to as accuracy, is one of the key dimensions used to address the data quality in a cancer registry, and it is defined as the proportion of cases with a given characteristic that truly have that attribute [8]. Any discrepancy between registry data and source data will consequently be classified as an error. The proportion of erroneous data elements, or the proportion of cancer cases with an erroneous registration, has been assessed for cancers of the central nervous system, head and neck, prostate and ovary [9–12], respectively. They all found a relatively low proportion of errors ranging from 0.5% to 9.3% which indicates an overall acceptable validity of the registry data. A similar evaluation remains to be done for NRCR data. This study therefore aims to validate NRCR data originating from one major Norwegian hospital in the period 1997–2005.

The study has been approved by the regional ethics committee.

Material and methods

The material comprises NRCR data on all 482 patients diagnosed with rectal cancer during 1997–2005 at a major surgical department in Norway. Reporting policies vary between the hospitals responsible for treating rectal cancer in Norway. In our hospital the clinical data were entered by the operating surgeon using a standardised paper form and sent to the NRCR after initiating the treatment [13]. In addition, the responsible pathologists reported pathological data (including autopsy data) directly from the laboratory in the form of free text or using a standardised template [13]. Patient data were also collected from separate radiation therapy databases, the Causes of Death Register and the Norwegian Patient Register. Data from these multiple sources were finally compiled and entered into the NRCR.

The data comprise 54 variables covering information on pre-operative evaluation, adjuvant treatment, surgery, pathological evaluation of the tumour specimen, complications and long-term outcomes such as local recurrence and distant metastasis. To determine the validity of this NRCR material, all registry data elements were checked for consistency with the medical records and pathology reports. If a particular data element was missing or did not correspond with the source information, the variable was marked as an erroneous registration.

As the variables varied extensively both in design and clinical significance, the severity of the erroneous registrations was evaluated depending on which variable was affected. In Supplementary Table I (available online at http://informahealthcare.com/doi/abs/10.3109/0284186X.2015.1031913), all validated variables are listed according to presumed importance for data validity. Category 1 covers important exposure and outcome variables such as pre-operative radiotherapy, TNM-stage and local recurrence or metastasis. Errors occurring in these were consequently defined as ‘major’. Category 2 comprises mainly of descriptive variables of less prognostic value though still clinically important, including type of surgical procedure and tumour distance from anal verge. Errors occurring in these were therefore defined as ‘moderate’. Finally, Category 3 contains minor descriptive variables that could easily be misinterpreted by the registrar such as calendar dates and numeric parameters, but without potential errors having a substantial impact on the data validity. These errors were defined as ‘minor’. Errors occurring even within the same variable could differ considerably with respect to severity. Discretion was therefore used in the evaluation of data discrepancies. See Supplementary Table I for further details to be found online at http://informahealthcare.com/doi/abs/10.3109/0284186X.2015.1031913.

Four variables out of the original 54 were excluded from the final analysis. Three of these were excluded because their registrations were more or less lacking throughout the whole dataset. Chemotherapy, i.e. whether the patient had received chemotherapy or not, was only implemented as a variable in the NRCR in 2014 and was thus not registered in the study period. Circumferential resection margin (CRM), i.e. the distance between tumour tissue and resection margin, was also not systematically reported in the pathology description in the study period. Date of reoperation, i.e. the calendar date of the reoperation if performed, was excluded due to systematically lacking data. The fourth excluded variable, patient age, had been entered by default in the database according to each patient’s personal identification number and had therefore no possibility for being erroneous. A final 50 variables were thus eligible for analysis.

SPSS version 22 (IBM, Chicago, IL, USA) was used to analyse the data.

Results

This validation showed that 100 patients (21%) had one or more erroneous registrations in the NRCR data, and the number of erroneous registrations in total was 131 constituting 0.5% of all the data elements. Sixteen variables (32%) had no erroneous registrations. The highest proportion of erroneous registrations was 2.1% and occurred in the variables type of surgical procedure and pre-operative computed tomography (CT) of the pelvis. Fifty-nine (45%) errors were considered minor in severity, while 35 (27%) were moderate and 37 (28%) were major. An overview of all validated variables, the proportion of erroneous registrations and their presumed severity is given in Table I.

Errors in pre-operative data

A total of 40 erroneous registrations occurred in variables related to the pre-operative evaluation and adjuvant treatment. Pre-operative CT of the pelvis had the highest frequency of erroneous registrations with 10 errors (2.1%), all of which were false-negative. Of the three erroneous registrations in pre-operative magnetic resonance imaging (MRI), two were false-negative. Four erroneous registrations were noted in pre-operative radiotherapy, of which three were false-positive. Also noted were six erroneous registrations on dosage of pre-operative radiotherapy; five of these represented deviations of 10 Gy or less from the true dosage, whereas one was an overestimation of Gy.

Errors in surgical data

In variables related to the surgical treatment, 44 erroneous registrations were found. Operation status was correct in all patients, as were five other surgical variables. Type of surgical procedure had the highest frequency of erroneous registrations with 10 errors (2.1%). Four of these procedures were truly transanal excisions or transanal endoscopic microsurgeries registered as local excisions in the NRCR data, four were truly Hartmann procedures registered as low anterior resections, and two were truly low anterior resections registered as abdominoperineal resections. The second most frequent erroneous registration in this group was calendar date of operation. Four of these represented time deviations ranging from approximately 1–8 months from the true operation date. In the remaining three patients, calendar date of operation was lacking altogether.

Errors in data concerning pathology, complications and outcomes

There were 20 erroneous registrations in variables related to the pathological evaluation of the tumour specimen. Thirteen (70%) of these were errors in tumour staging data (T-stage and Dukes' classification) of which half represented an erroneous upstaging of the tumour. Seven erroneous registrations occurred in variables related to post-operative complications including peri-operative tumour perforation (0.6%), anastomotic leakage (0.4%) and reoperation (0.4%). They were all false-negative registrations. A total of 20 erroneous registrations occurred in variables related to oncological outcomes. Seven (1.4%) of these were lacking registrations of distant metastasis. No errors were found in local recurrence. Three and four errors were found in the calendar dates for local recurrence and distant metastasis, respectively. The time deviation from the true confirmed date ranged from 10 days to approximately three years. Two of the calendar dates were false-negative, i.e. the local recurrence or distant metastasis had not truly occurred.

Discussion

This validation of NRCR data shows a high level of agreement between NRCR data and the source material. Over 99% of the registered data elements corresponded with the information stated in the medical records, and erroneous registrations ranged from 0% to 2.1% among the variables, thus satisfying the criteria for good internal validity [14]. A total of 16 variables (32%) had no registered errors. Among these were the key variables operation (yes/no), lymph node status (N-stage) and metastasis status (M-stage), which are all critical in describing treatment and outcomes in rectal cancer patients. Approximately three quarters of the erroneous registrations were presumed to have minor or moderate severity. The remaining quarter of erroneous registrations were of a more severe nature, including incorrect T-stages and incorrect or missing information on local recurrence or metastasis. It is evident that such errors can have serious consequences for the data validity.

Erroneous registrations in pre-operative CT of the pelvis, complications and distant metastasis were all false-negative, thus indicating a tendency of under-reporting in this material. Among the erroneous surgical variables, type of surgical procedure was the most frequent with 10 errors. It seems plausible

Table I. Validated variables in the Norwegian Rectal Cancer Registry for 482 patients diagnosed with rectal cancer 1997–2005.

Variable group	Variable	Variable characteristics	No. of erroneous registrations (%)	Severity
Pre-operative evaluation and adjuvant therapy	Year of diagnosis	yyyy	0 (0)	3
	Gender	male/female	3 (0.6)	2
	Curative intent	yes/no	1 (0.2)	2
	Pre-op CT pelvis	yes/no	10 (2.1)	3
	Pre-op MRI pelvis	yes/no	3 (0.6)	3
	Pre-op rectal ultrasound	yes/no	0 (0)	3
	Tumour distance from anal verge I	continuous	2 (0.4)	2
	Tumour distance from anal verge II	categorical	0 (0)	2
	Tumour localisation bowel wall	categorical	4 (0.8)	3
	Tumour fixation bowel wall	categorical	2 (0.4)	3
	Site of pre-op metastasis	categorical	5 (1.0)	1
	Pre-op radiotherapy	yes/no	4 (0.8)	1
	Post-op radiotherapy	yes/no	0 (0)	1
	Dosage pre-op radiotherapy	continuous	6 (1.2)	3
	Dosage post-op radiotherapy	continuous	0 (0)	3
Surgery	Operation	yes/no	0 (0)	1
	Emergency operation	yes/no	0 (0)	2
	Reason for not operating	categorical	0 (0)	3
	Operation, calendar date[1]	dd.mm.yyyy	7 (1.4)	3
	Type of surgical procedure	categorical	10 (2.1)	2 (3)
	Total mesorectal excision	yes/no	4 (0.8)	3
	Distal resection margin	continuous	5 (1.0)	2 (1)
	Tumour ingrowth according to surgeon	yes/no	2 (0.4)	2
	Localisation tumour ingrowth	free text	0 (0)	2
	Local radicality according to surgeon	categorical	1 (0.2)	3
	Number of resected lymph nodes	continuous	2 (0.4)	3
	Resected organ	free text	2 (0.4)	2
	Resected metastasis	categorical	1 (0.2)	2
	Type of anastomosis	categorical	5 (1.0)	3
	Height of anastomosis	continuous	3 (0.6)	3
	Anastomotic suture method	categorical	0 (0)	3
	Rectal stump lavage	yes/no	2 (0.4)	3
	Stoma	yes/no	0 (0)	3
Pathology	T-stage	categorical	8 (1.7)	1
	N-stage	categorical	0 (0)	1
	M-stage	categorical	0 (0)	1
	R-stage	categorical	2 (0.4)	1
	Dukes' classification	categorical	5 (1.0)	1
	Tumour differentiation	categorical	2 (0.4)	2
	Diameter resected tumour	continuous	3 (0.6)	2
	Number of lymph nodes with tumour invasion	continuous	0 (0)	2
Complications	Peri-operative tumour perforation	categorical	3 (0.6)	2
	Anastomotic leakage	yes/no	2 (0.4)	2
	Anastomotic leakage, calendar date[1]	dd.mm.yyyy	0 (0)	3
	Reoperation	yes/no	2 (0.4)	2
Oncological outcomes	Local recurrence and/or metastasis	categorical	7 (1.4)	1
	Site of post-op metastasis	categorical	6 (1.2)	2 (3)
	Local recurrence, calendar date[1]	dd.mm.yyyy	3 (0.6)	1 (3)
	Distant metastasis, calendar date[1]	dd.mm.yyyy	4 (0.8)	1 (3)
	Death, calendar date[1]	dd.mm.yyyy	0 (0)	1

Presumed severity of erroneous registrations are listed to the right, differences in presumed severity within one single variable are listed in parenthesis. 1 = major error, 2 = moderate error, 3 = minor error. [1]Major error if time deviation ≥ 1 month. Minor error if time deviation < 1 month.

that a fair part of these were due to inconsistency or confusion among the registrars about the variable definitions. This is illuminated by the observation that four transanal excisions or transanal endoscopic microsurgeries were registered as local excisions in the NRCR data, presumably because of poorly defined distinctions between these three procedures. Erroneous registrations were otherwise evenly spread among the different variables, implying random rather than systemic errors.

A limitation of our study is the exclusion of three variables due to systematically missing data, among them chemotherapy and CRM. As earlier mentioned, chemotherapy was not formally included in the

NRCR until 2014, thus accounting for the lack of data in our study period. Also, because no standardised pathology reporting form for rectal cancer was used at that time, CRM has not been systematically reported in the study period. This important indicator of surgical quality is therefore severely under-reported in our material.

Another limitation of our study is that the validation is restricted to one hospital only, and it is arguably difficult to generalise the findings to the whole of the NRCR. However, the validated material is of general interest as it includes all patients diagnosed with rectal cancer in a nine-year period at a high-volume hospital. The validation also covers practically all of the variables, and the material is unique in its way of reflecting the long standing registration work of a team of surgeons as opposed to a dedicated person. We believe this detailed insight into the consistency between central registry data and the source material from which they were drawn to be of value to the NRCR and other quality registries.

Quality evaluations of the NRCR data have not yet been published, and this hospital based study is the first to address the validity of NRCR data in a large patient cohort. The material in the Norwegian Cancer Registry as a whole has been evaluated by Larsen et al. in 2009 [6]. This study uses the percentage of morphologically verified diagnoses, that is, the proportion of rectal cancer diagnoses verified by examination of the resected tumour specimen, as a proxy for data validity. The reasoning behind this is that a diagnosis confirmed by a pathologist is more likely to be correct, and a high proportion of such morphologically verified diagnoses could therefore indicate high data validity in the registry in general. Larsen et al. found a morphological verification of 98.4% for rectal, rectosigmoid and anal cancer during the years 2001–2005, which corresponds well with the close to complete degree of data validity found in our study. For all cancer types, Larsen et al. estimated that 93.8% were morphologically verified, a better result than other European cancer registries, but somewhat inferior to that found in the other Nordic countries Finland, Iceland and Sweden [6].

A validation of Swedish Rectal Cancer Registry (SRCR) data has recently been performed [15]. Totally 906 patients were treated with major abdominal surgery between 1995 and 1997 in the study, and 14 selected variables were checked for consistency against the medical records. The authors found that less than 10% of the variables had erroneous registrations. The proportion of erroneous registrations in specific variables was largely comparable between the SRCR and NRCR data. A total of 0.9% of the registrations in pre-operative radiotherapy were erroneous in the SRCR data, as compared to 0.8% erroneous registrations in our material. Erroneous registrations in local recurrence and distant metastasis were 3.6% and 4.7% in the SRCR data, respectively, comparing to 1.4% erroneous registrations for the two variables combined in this study.

Data quality has previously been investigated in other Norwegian quality registries, and the results are in agreement with the present data. In a study of 301 patients with prostate cancer registered during 1957–1986, 0.5% of the registered data elements, corresponding to 6% of the patients, contained errors [11]. In another study investigating the registry data on 729 patients with ovarian cancer registered during 1987–1997, 2% of the patients had a wrongly registered diagnosis when compared with the pathology report [12]. In comparison, 0.5% of the data elements were erroneous in our study, corresponding to 21% of the patients.

The high level of data accuracy found in this evaluation indicates an appropriate number of variables and overall good routines for reporting clinical data to the NRCR. However, under-reporting remains a problem, and our results point towards a tendency of more errors in the end of the study interval in the years 2004 and 2005. This may reflect that reporting in general consumes an increasing amount of time in the healthcare system, and healthcare personnel are indeed under pressure when it comes to reporting data to registries and a number of other authorities alongside their everyday clinical tasks. In this context, errors are bound to happen.

Several measures could be helpful to improve the quality of data. By reducing the time interval between each rectal cancer treatment and the subsequent reporting of clinical data, recall bias is likely to be reduced. Furthermore, it is important to clear out any misconceptions among the reporting clinicians and registrars as to what the different variables represent to make the reporting as uniform and correct as possible. The quantity of variables should also be reviewed regularly and kept to a reasonable number.

Since the establishment of a common Norwegian registry for colon and rectal cancer in 2007 [1], a questionnaire for reporting colorectal cancer patients combining clinical, pathological and epidemiological data has replaced the older templates. As of June 2012, an electronic version of the questionnaire was made available and is now increasingly being put into use at the hospitals treating colorectal cancer in Norway [13]. Starting from January 2015, its use will be mandatory for all hospitals. In the Netherlands, the Dutch Surgical Colorectal Audit, a systematic web-based data collection initiated in 2009, has had remarkable success in simplifying the reporting of surgical colorectal cancer data [16]. Hopefully, the electronic questionnaire that has been employed routinely since 2012 at Haukeland University Hospital will improve the quality of NRCR data by streamlining and quickening the whole reporting process, minimising both under-reporting and recall bias in the reported data.

Conclusion

A proportion of correct data elements exceeding 99% confirms a high degree of validity in this material. It justifies the use of NRCR data in research and quality improvement of rectal cancer treatment.

Acknowledgements

The authors would like to thank the Norwegian Rectal Cancer Registry and the staff at the medical records archive at Haukeland University Hospital.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

[1] Gaard M, Vonen B, Tveit KM, Hoff G, Carlsen E. Eget register for kolorektal kreft. Tidsskr Nor Laegeforen 2007;127:2834–5.

[2] Norwegian Cancer Registry. Om kvalitetsregistre. [cited 2013 Apr 30]. Available from: http://www.kreftregisteret.no/no/Registrene/Kvalitetsregistrene/Om-kvalitetsregistrene/.

[3] Norwegian Directorate of Health. Nasjonalt handlingsprogram med retningslinjer for diagnostikk, behandling og oppfølging av kreft i tykktarm og endetarm. [cited 2014 Dec 15]. Available from: http://helsedirektoratet.no/publikasjoner/nasjonalt-handlingsprogram-med-retningslinjer-for-diagnostikk-behandling-og-oppfolging-av-kreft-i-tykktarm-og-endetarm/Publikasjoner/nasjonalt-handlingsprogram-for-tykk-og-endetarmskreft.pdf .

[4] Wibe A, Carlsen E, Dahl O, Tveit KM, Weedon-Fekjaer H, Hestvik UE, et al. Nationwide quality assurance of rectal cancer treatment. Colorectal Dis 2006;8:224–9.

[5] Guren MG, Körner H, Pfeffer F, Myklebust TÅ, Tandberg Eriksen M, Edna TH, et al. Nationwide improvement of rectal cancer treatment outcomes in Norway, 1993–2010. Acta Oncol (forthcoming).

[6] Larsen IK, Smastuen M, Johannesen TB, Langmark F, Parkin DM, Bray F, et al. Data quality at the Cancer Registry of Norway: An overview of comparability, completeness, validity and timeliness. Eur J Cancer 2009;45:1218–31.

[7] La Vecchia C, Bosetti C, Lucchini F, Bertuccio P, Negri E, Boyle P, et al. Cancer mortality in Europe, 2000–2004, and an overview of trends since 1975. Ann Oncol 2010;21:1323–60.

[8] 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;45:747–55.

[9] Helseth A, Langmark F, Mork SJ. Neoplasms of the central nervous system in Norway. I. Quality control of the registration in the Norwegian Cancer Registry. APMIS 1988;96:1002–8.

[10] Mork J, Thoresen S, Faye-Lund H, Langmark F, Glattre E. Head and neck cancer in Norway. A study of the quality of the Cancer Registry of Norway's data on head and neck cancer for the period 1953–1991. APMIS 1995;103:375–82.

[11] Harvei S, Tretli S, Langmark F. Quality of prostate cancer data in the cancer registry of Norway. Eur J Cancer 1996;32A:104–10.

[12] Tingulstad S, Halvorsen T, Norstein J, Hagen B, Skjeldestad FE. Completeness and accuracy of registration of ovarian cancer in the cancer registry of Norway. Int J Cancer 2002;98:907–11.

[13] Norwegian Gastrointestinal Cancer Group – Colorectal division. Annual report 2013. [cited 2014 Dec 15]. Available from: http://www.kreftregisteret.no/Global/Publikasjoner%20og%20rapporter/Årsrapporter/aarsrapport_2013_tykk-og%20endetarmskreft.pdf.

[14] Gunnarsson U. Quality assurance in surgical oncology. Colorectal cancer as an example. Eur J Surg Oncol 2003;29:89–94.

[15] Jorgren F, Johansson R, Damber L, Lindmark G. Validity of the Swedish Rectal Cancer Registry for patients treated with major abdominal surgery between 1995 and 1997. Acta Oncol 2013;52:1707–14.

[16] Van Leersum NJ, Snijders HS, Henneman D, Kolfschoten NE, Gooiker GA, ten Berge MG, et al. The Dutch surgical colorectal audit. Eur J Surg Oncol 2013;39:1063–70.

Supplementary material available online

Supplementary Table I to be found online at http://informahealthcare.com/doi/abs/10.3109/0284186X.2015.1031913