MAYO CLINIC

MAYO CLINIC PROCEEDINGS:
INNOVATIONS, QUALITY & OUTCOMES

ORIGINAL ARTICLE

Accuracy of a Cancer Registry Versus Clinical Care Team Chart Abstraction in Identifying Cancer Recurrence

Elsa A. Sutton, MD; Benjamin C. Kamdem Talom, BA; Daniel K. Ebner, MD, MPH; Taylor M. Weiskittel, MS; William G. Breen, MD; Roman O. Kowalchuk, MD; Heather J. Gunn, PhD; Courtney N. Day, MS; Eric J. Moore, MD; Sara J. Holton, CTR, BS Health/Health Care Administration/Management; Kathryn M. Van Abel, MD; Chadi N. Abdel-Halim, MD; David M. Routman, MD; and Mark R. Waddle, MD

Abstract

Objective: To evaluate the completeness and reliability of recurrence data from an institutional cancer registry for patients with head and neck cancer.

Patients and Methods: Recurrence information was collected by radiation oncology and otolaryngology researchers. This was compared with the institutional cancer registry for continuous patients treated with radiation therapy for head and neck cancer at a tertiary cancer center. The sensitivity and specificity of institutional cancer registry data was calculated using manual review as the gold standard. False negative recurrences were compared to true positive recurrences to assess for differences in patient characteristics.

Results: A total of 1338 patients who were treated from January 1, 2010, through December 31, 2017, were included in a cancer registry and underwent review. Of them, 375 (30%) had confirmed cancer recurrences, 45 (3%) had concern for recurrence without radiologic or pathologic confirmation, and 31 (2%) had persistent disease. Most confirmed recurrences were distant (37%) or distant plus locoregional (29%), whereas few were local (11%), regional (9%), or locoregional (14%) alone. The cancer registry accuracy was 89.4%, sensitivity 61%, and specificity 99%. Time to recurrence was associated with registry accuracy. True positives had recurrences at a median of 414 days vs 1007 days for false negatives.

Conclusion: Currently, institutional cancer registry recurrence data lacks the required accuracy for implementation into studies without manual confirmation. Longer follow-up of cancer status will likely improve sensitivity. No identified differences in patients accounted for differences in sensitivity. New, ideally automated, data abstraction tools are needed to improve detection of cancer recurrences and minimize manual chart review.

 ■ Mayo Clin Proc Inn Qual Out 2024;8(3):225-231

Cancer databases serve an essential role in the evaluation of cancer treatment trends and outcomes. Large databases, such as the National Cancer Database and Surveillance, Epidemiology, and End Results are the backbone for many clinical research studies for cancer in the United States. A 2022 PubMed search for National Cancer Database returned over 1600 results.[1] Although useful for many purposes, these databases are not designed to track cancer recurrence which is a critical outcome of cancer treatment. Similarly, institutions and departments caring for these patients do not routinely collect this information, often citing limited resources.[2] For example, a 2014 review of over 700,000 patients with various types of cancer found, on average, that hospitals had incomplete recurrence information recorded for 56.7%-66.7% of patients.[3]

To address this limitation, researchers often must conduct manual chart review which is an expensive and time-consuming

From the Department of Radiation Oncology, Mayo Clinic, Rochester, MN (E.A.S., B.C.K.T., D.K.E., W.G.B., R.O.K., S.J.H., D.M.R., M.R.W.); Department of Molecular Pharmacology & Experimental Therapeutics, Mayo Clinic, Rochester, MN (T.M.W.); Department of Quantitative Health Sciences, Mayo

Affiliations continued at the end of this article.

process. Alternative means of extracting recurrence data that have been attempted, such as using billing codes or Medicare claims as indicators of recurrence, have been found to be inconsistent and unreliable.[4] Thus, there exists a need for a reliable and accurate method for recording recurrence status of patients treated for cancer that does not rely solely on manual data extraction by clinical departments or care teams.

Although accurate prospective collection is likely the ideal solution, this does not account for historical recurrence data, patients not being followed, or patients followed in different departments. One possible solution to collect retrospective data on cancer recurrence is to leverage existing institutional or multicenter cancer registries that employ trained data extraction specialists. At our institution, these specialists conform with the American College of Surgeons guidelines to record cancer recurrence status if identified during their review, but do not seek out this information as it is not required. Currently, no data exists on the accuracy and completeness of recurrence status collected using this method. Thus, an evaluation is warranted to gain insight on whether this method may be adequate, or if a new method must be developed. The goal of this project is to evaluate an institutional cancer registry for the completeness and accuracy of recurrence data collected per the American College of Surgeons for a group of patients with head and neck cancer (HNC) treated with radiation therapy (RT) vs trained chart review by clinical care teams in departments of radiation oncology and otolaryngology (ENT).

METHODS

Cohort Selection and Data Collection

After institutional review board approval, cohorts of patients were identified who were treated for HNC. Patient level and recurrence data was collected, also as outlined below. We identified head and neck (H&N) cancer patients through the Mayo Clinic Cancer Registry (Cohort 1), and then cross-referenced this registry with H&N cancer patients within the radiation oncology database (Cohort 2), which contains all patients treated with RT. Next, to identify interobserver differences in trained medical team abstraction, we identified an additional cohort of patients who were also present in a third H&N otolaryngology—head and neck surgery database (Cohort 3).

Mayo Clinic Cancer Registry—Cohort 1

Since 2004, the Mayo Clinic Cancer Registry has systematically identified and registered patients diagnosed with cancer utilizing Health Level 7 listener software linked to institutional pathology reports which searches ~3500 terms and abbreviations that could refer to a cancer diagnosis. These cases are then reviewed by cancer registry personnel to determine if they are new and if so, they are added to the database. Alternatively, if patient identifiers on the incoming case match a case already abstracted, the patient's chart is further reviewed to determine if the pathology being presented is a new primary or recurrence, as per the North American Association of Central Cancer Registries (NAACCR) national coding guidelines (multiple primary/histology [MPH rules] or solid tumor rules). Specifically, the national coding guidelines define disease recurrence as recurrence at any point after a disease-free interval post curative treatment. Patients were defined as never disease-free if there was no disease-free interval after treatment. Once the first recurrence was documented, no further recurrences were captured.

Additionally, treatment-specific patient lists from radiation oncology and medical oncology databases are routinely reviewed for patients in the existing registry, and any new treatments received are reviewed and cancer status is updated accordingly for recurrences.

Radiation Oncology—Cohort 2

Continuous patients from 2010 to 2018 with a diagnosis of HNC were collected from the cancer registry recorded cases. Patients treated with RT were cross-referenced from an existing institutional radiation oncology database. Manual chart review was then conducted by a trained college-level graduate research intern under close supervision of 2 radiation oncology faculty and 2 radiation oncology residents. Imaging, pathology, and clinical notes were reviewed, and recurrence was recorded using a structured REDCap database. Recurrence date, type of recurrence (local, regional,

Mayo Clin Proc Inn Qual Out ■ June 2024;8(3):225-231 ■ https://doi.org/10.1016/j.mayocpiqo.2024.03.005
www.mcpiqojournal.org

or distant), and distant site of recurrence (spine, non-spine bone, lung/thorax, abdomen/adrenal/liver, pelvis, or central nervous system) were recorded. The location of follow-up information and notes (most or all at Mayo Clinic versus elsewhere via scanned media documents or Epic Care Everywhere) was noted, as was the method of confirmation of recurrence (radiology, pathology, physical examination, or biomarkers). All sources of information, including more obscure sources such as Epic CareEverywhere, scanned records from outside institutions, and documented phone calls, were used to determine cancer status.

Otolaryngology—Head and Neck Surgery—Cohort 3

Patients treated for primary or recurrent oropharyngeal squamous cell carcinoma at Mayo Clinic are actively abstracted into the Departmental REDCap Registry under IRB: 22-000684 by a combination of senior faculty, trainees, and a dedicated abstractor. The earliest year of diagnosis documented in the database is 1990. Abstraction is updated weekly and regularly statistically reviewed for inconsistencies. The database currently houses 2902 discrete patients. Patients were included regardless of staging, treatment intent, or treatment modality. Patients were excluded if they did not have Minnesota research authorization. Variable categories include the following: demographic characteristics, comorbidities, presentation and physical examination, surveillance visit clinical data, risk factors, biopsy data, diagnostic imaging, surveillance imaging, primary management, pathologic data, staging, recurrence, secondary management, oncologic outcome, and functional outcome. Record review for data abstraction includes clinical, radiographic, and pathologic data from Epic and CareEverywhere.

The data used in this study included patients with human papilloma virus(+) oropharyngeal squamous cell cancer who received intent-to-cure surgery with adjuvant therapy in the period from January 1, 2016 to December 31, 2021. Patients with history of HNC or the presence of synchronous primary cancer or distant metastatic disease at time of diagnosis were excluded. After the identification of patients from the Departmental REDcap Registry, patients that matched the radiation oncology database were reviewed manually by the combination of an otolaryngology faculty and resident. The diagnosis of recurrence was obtained from the clinical notes, imaging reports, and pathology reports. The type of recurrence was registered as local, regional, locoregional, or distant. If no biopsies were performed, the diagnosis of recurrence was accepted based on a combination of clinical assessment and radiologic evaluation.

Gold Standard

Once recurrence status was appropriately collected for each cohort from the 3 data sources above, patients included in all 3 groups were identified and compared for discrepancies in recurrence status. When a discrepancy was identified between radiation oncology and otolaryngology, or a recurrence was listed by the cancer registry that was not recorded by one or both of the clinical teams, it was reviewed and discussed by a combined group of physicians from both groups. A final decision was made after chart review and discussion, and a list of gold standard recurrence status was formalized. Of note, when all 3 data sources agreed on recurrence status, no further review or clarification was performed.

Statistical Analysis

Two analyses were completed based on the availability of information. Cohort selection with inclusion and exclusion criteria are shown in the consort diagram in Figure 1. First, the complete set of patients from radiation oncology were compared with the cancer registry (n=1338, Cohort 1). In this analysis, the gold standard was considered the manual chart review by radiation oncology. Patients with persistent disease were excluded because of differences in definitions of persistent disease between the registry and radiation oncology group. Patients with a concern for recurrence but without enough evidence to verify were excluded. Finally, patients with incomplete data were excluded, resulting in the final group for Cohort 2 (n=1128). Patients lost to follow-up were censored at the time of their last appointment or documentation. Accuracy, sensitivity, specificity, positive predictive value, and negative predictive value were calculated and reported. Information on age, sex, time

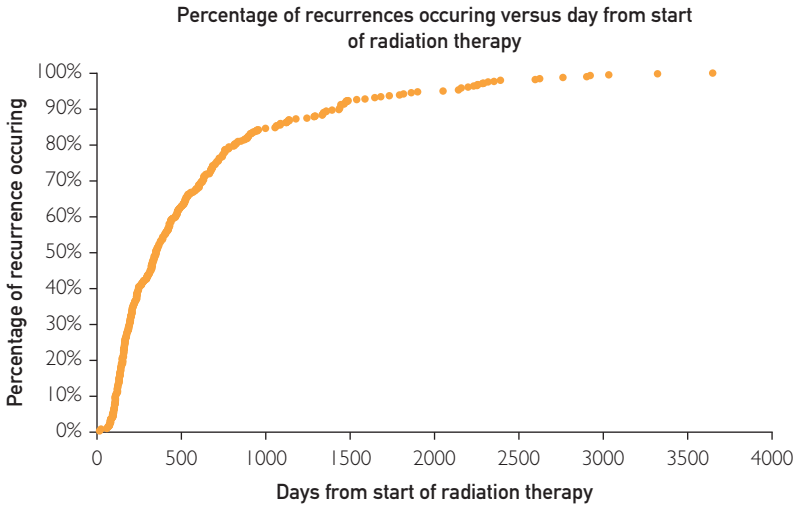

FIGURE 1. Percentages of recurrences recorded over time; time listed in days from the start of radiation therapy.

from radiation to recurrence, year of treatment, and available follow-up data were compared between those captured by the radiation oncology registry versus those missed by the cancer registry to identify possible systematic errors in cancer registry data collection.

A second analysis was completed for patients with data available from all 3 databases, called Cohort 3. Similar to the previous analysis, patients with persistent disease or concern for recurrence without data to verify were excluded from their respective analyses. Using the gold standard as outlined above, accuracy, sensitivity, specificity, positive predictive value, and negative predictive value were calculated for the cancer registry, ENT, and radiation oncology departments.

All data evaluation and statistical analyses were performed using SAS version 9.4 (SAS Institute Inc). $P<.05$ was considered statistically significant.

RESULTS

There were 1338 patients identified in the cancer registry who were treated for HNC with RT at Mayo Clinic from 2010-2018. A total of 139 were excluded for having persistent disease and 45 were excluded because of having

TABLE 1. Demographic Characteristic and Clinical Features of Recurrence

Features	Cohort 1 (n=1338)	Cohort 2 (n=1128)	Cohort 3 (n=262)
Age at diagnosis (y), median (IQR)	61.0 (53.0-68.0)	61.0 (53.0-68.0)	59.0 (52.0-63.0)
Sex			
Female	290 (21.7%)	238 (21.1%)	31 (11.8%)
Male	1048 (78.3%)	890 (78.9%)	231 (88.2%)
Days from RT end to recurrence, median (IQR)[a]	306.0 (123.0-664.0)	387.0 (146.0-732.0)	518.0 (146.5-1067.0)
Pathologic confirmation[a]			
Nonpathology	126 (30.0%)	45 (15.4%)	8 (18.6%)
Pathology+	294 (70.0%)	248 (84.6%)	35 (81.4%)

[a]Among patients with a recurrence

Mayo Clin Proc Inn Qual Out ■ June 2024;8(3):225-231 ■ https://doi.org/10.1016/j.mayocpiqo.2024.03.005
www.mcpiqojournal.org

TABLE 2. Comparison of Sensitivity, Specificity, Positive Predictive Value, and Negative Predictive Value Between the 3 Database Groups as Compared With the Gold Standard

	Radiation Oncology (n=262)	Otolaryngology (n=262)	Cancer Registry (n=1128)
Sensitivity	87.5	93.2	50.0
Specificity	97.7	97.7	100.0
Positive predictive value	87.5	89.1	100.0
Negative predictive value	97.7	98.6	91.1

concern for recurrence with incomplete records to further classify. Twenty-six additional patients were excluded because of incomplete tumor registry records. Of the remaining 1128 patients, 183 (16.2%) were found to have recurrent disease in the cancer registry while 293 (26.0%) were found to have recurrent disease by clinically trained radiation oncology chart abstracters. Using the radiation oncology review as a gold standard for this comparison, the cancer registry was found to have 89.4% accuracy, 60.8% sensitivity (95% CI, 0.55-0.66), 99.4% specificity (95% CI, 0.99-1), 97.3% positive predictive value (PPV) (95% CI, 0.95-1), and 87.8% negative predictive value (NPV) (95% CI, 0.86-0.90).

Demographic characteristics and clinical features of patients with recurrence can be found in Table 1. On average, patients falsely labeled as not recurrent on the cancer registry (false negatives) had cancer recurrence later than patients who were correctly labeled as recurrent (true positives) (mean 1007 vs 414 days from start of radiation to recurrence, $P<.001$). Additionally, the median year of treatment for patients with a false negative was later at 2014, versus 2013 who were correctly identified ($P=.024$). No other statistically significant differences in the groups were identified.

A subset of 262 patients with human papilloma virus+ oropharynx tumors was available with data in all 3 databases including the ENT database (Cohort 3). Patients with evidence of persistent disease or concern for recurrence were excluded from the respective analyses. In Cohort 3, the cancer registry documented recurrence in 21 patients (8.2%), ENT documented recurrence in 44 patients (16.8%), and radiation oncology documented recurrence in 40 patients (15.5%). After a final discussion involving radiation oncology and ENT reviewers, 44 patients were considered to truly have recurrence. Using this consensus as the gold standard, the cancer registry was the least accurate in documentation at 91.8%, followed by radiation oncology at 96.1%, and finally ENT at 96.9%. Radiation oncology review had a sensitivity of 87.5% (95% CI, 77.3-97.8), specificity of 97.7% (95% CI, 95.7-99.7), PPV of 87.5% (95% CI, 77.3-97.8), and NPV of 97.7% (95% CI, 95.7-99.7). The ENT review had a sensitivity of 93.2% (95% CI, 85.7-1), specificity of 97.7% (95% CI, 95.7-99.7), PPV of 89.1% (95% CI, 80.1-98.1), and NPV of 98.6% (95% CI, 97.1-1). A comparison of the sensitivity, specificity, PPV, and NPV can be found in Table 2.

The number and percentage of recurrences recorded over time were collected from the time of RT completion, as displayed in Figure 1. Eighty percent of recurrences were identified by 2 years, 90% by 3.7 years, 95% by 5.5 years, and 99% by 7.5 years from completion of RT. An exploratory analysis was completed estimating the number of additional cancer registry abstractors that would be required to complete annual chart review for recurrence for all patients during their first 2, 5, or 8 years after treatment. Using 2022 numbers of completed cases per registry staff, we assumed a given cancer registry abstractor could complete 120 new cases per week or review 438 cases for recurrence a week. With 45,000 new cases per year, we found that 8 registry staff are required to review new cases. For 2-year, 5-year, and 8-year follow-up, we calculated an additional 5, 11, and 18 required full-time staff to complete chart review for recurrence, representing a doubling in staff for the longest follow-up scenario.

DISCUSSION

In this retrospective review, we identified that the institutional cancer registry's current method of reporting recurrence status has a sensitivity of is ∼50%-60% and a specificity approaching 100% for patients with a diagnosis of HNC. When compared with expert medical team chart review, use of the cancer registry results in a lower sensitivity but a similarly high specificity. Thus, this represents a gap in the complete reporting of recurrences, calling into question data based solely on cancer registry reporting.

Trained cancer registry abstractors are highly skilled at reviewing charts and determining recurrence when informed to do so by the current screening methods. However, after review of the methodology, it is clear that if a recurrence occurs through imaging alone, or if subsequent treatment is not captured in the current treatment databases used, cases will be missed. When missed cases were compared with appropriately identified cases, systematic bias was indeed present. Patients with missed cancer recurrence on average recurred nearly 600 days later than those who were identified appropriately with recurrence. This may be due to a lack of biopsy confirmation of these cases, or a delay in population into treatment databases. With 45,000 new cases per year, we found that 7.8 full-time registry staff would be required to review new cases with adequate accuracy. With up to 8 years of follow-up data, this need would increase to a total of 24.9 full-time employees. Alternatively, however, novel artificial intelligence methods may be able to improve cancer registries without requiring hiring of additional staff or subject matter experts.

Recurrence data inaccuracy is indeed common to institutional databases and not unique to the current one studied. For instance, the Danish Gynecological Cancer Database, a large national multidisciplinary clinical cancer database, was found to only have 71% agreement between recurrence reported in the database and found on manual review of pathology reports.[5] Authors attributed this inaccuracy to low incidence of recurrence reporting since it was an optional field. In the current study, the likelihood of a recurrent case not being recorded increased with increasing time from original radiation treatment, suggesting possibly an inadequate method for detecting recurrences as time from treatment increases.

In our study, an exploratory analysis found a considerable workload increase would be required to manually review each patient annually. This increase in staff would be ∼8 to 25 FTE depending on the number of years patients are followed. For a program already under staffing constraints and financial pressures, this level of increase in staff would not be reasonable.

As previously mentioned, several other methods of recurrence data extraction have been explored with mixed results. Particularly intriguing are computer-based algorithms, as these have the potential to drastically reduce the labor involved in extracting recurrence data and allowing targeted chart review. A number of small studies have used algorithms based on claims data to identify recurrences.[6,7] Although several of these algorithms have been found to have sensitivity over 80% and specificity over 97%, these have been done in single institutions, represent specific use cases with small sample sizes, and are not suitable for widespread use.

A larger multi-institutional study applied an algorithm to health care utilization data of over 2000 women with a diagnosis of breast cancer across the state of Ontario. The algorithm, which utilized procedure data, diagnosis codes, and reported use of systemic therapy and RT, detected recurrence with 85% sensitivity and 94% specificity.[8] However, the algorithm was only applied to women whose original cancer diagnosis was in the previous 4 years, which does not address the difficulties with accurate recurrence detection long after diagnosis. A Danish study of 500 rectal cancer patients had similar success utilizing an algorithm based on patient and pathology registries to identify recurrence with 88% sensitivity and 96% specificity when compared with manual review.[9] Again, however, there were considerable limitations with follow-up time, with only 5 years of follow-up data included. Thus, there remains a need for automated methods by which recurrences may be more broadly and longitudinally detected in a variety of sites and across different institutions and patient populations.

Mayo Clin Proc Inn Qual Out ■ June 2024;8(3):225-231 ■ https://doi.org/10.1016/j.mayocpiqo.2024.03.005
www.mcpiqojournal.org

CONCLUSION

Our study has several limitations that should be noted. First, the study was conducted at a single institution. Additionally, it focused only on HNC, which is a disease group that typically has a relatively high rate of early recurrence when compared with other common disease sites, such as prostate or breast cancer. Accordingly, the sensitivity of detecting later recurring cancers such as breast and prostate cancer may be even lower than the 50%-60% noted on this study of patients with HNC. These factors may affect generalizability of our results to other disease sites and other institutions. The accuracy of the cancer database was compared against chart evaluation by clinicians, whose interpretation of chart data is inherently subject to bias and error. Additionally, the increased workload analysis uses rough estimates which may not be completely accurate.

Notwithstanding the above limitations, this study provides valuable insight into the benefits and shortcomings of utilization of cancer databases for recurrence data. Although specificity is high, the database's poor sensitivity and overall decreased accuracy compared with clinician chart review reports the continued need for manual review of recurrence data before implementation into research studies. Accuracy may be improved by manual review, improving methods of detection of recurrence, or application of new algorithms to assist in real-time recurrence detection.

POTENTIAL COMPETING INTERESTS

The spouse of author Roman Kowalchuk, MD, was previously employed by GE Healthcare.

ACKNOWLEDGMENTS

Elsa A. Sutton and Benjamin C. Kamdem Talom contributed equally to this work as cofirst authors. David M. Routman and Mark R. Waddle contributed equally to this work as colast authors.

Abbreviations and Acronyms: **ENT,** Otolaryngology; **HNC,** head and neck cancer; **NPV,** negative predictive value; **PPV,** positive predictive value; **RT,** radiation therapy

Affiliations (Continued from the first page of this article.): Clinic, Rochester, MN (H.J.G., C.N.D.); and Department of Otolaryngology—Head and Neck Surgery, Mayo Clinic, Rochester, MN (E.J.M., K.M.V.A., C.N.A-H.).

Data Previously Presented: Abstract was previously presented at the ASTRO Annual Meeting in San Antonio, TX, October 23-26, 2022.

Correspondence: Address to Mark Waddle, MD, Department of Radiation Oncology, Mayo Clinic, 200 First Street SW, Rochester, MN 55905 (waddle.mark@mayo.edu; Twitter: @WaddleMD).

REFERENCES

1. NCDB—Search Results—PubMed. PubMed, https://pubmed.ncbi.nlm.nih.gov/?term=ncdb. Accessed December 10, 2022.
2. National Cancer Database Tools. American College of Surgeons. https://www.facs.org/media/buvfsm0p/22_ca_ncdbtools_8-5x11_v01.pdf. Accessed June 16, 2022.
3. In H, Bilimoria KY, Stewart AK,W, et al. Cancer recurrence: an important but missing variable in national cancer registries. *Ann Surg Oncol*. 2014;21(5):1520-1529. https://doi.org/10.1245/s10434-014-3516-x.
4. Hassett MJ, Ritzwoller DP, Taback N, et al. Validating billing/encounter codes as indicators of lung, colorectal, breast, and prostate cancer recurrence using 2 large contemporary cohorts. *Med Care*. 2014;52(10):e65-e73. https://doi.org/10.1097/MLR.0b013e318277eb6f.
5. Juhl CS, Hansen ES, Høgdall CK, Ørtoft G. Valid and complete data on endometrial cancer in the Danish Gynaecological Cancer Database. *Dan Med J*. 2014;61(6):A4864.
6. Warren JL, Yabroff KR. Challenges and opportunities in measuring cancer recurrence in the United States. *J Natl Cancer Inst*. 2015;107(8):djv134. https://doi.org/10.1093/jnci/djv134.
7. Chubak J, Yu O, Pocobelli G, et al. Administrative data algorithms to identify second breast cancer events following early-stage invasive breast cancer. *J Natl Cancer Inst*. 2012;104(12):931-940. https://doi.org/10.1093/jnci/djs233.
8. Holloway CMB, Shabestari O, Eberg M, et al. Identifying breast cancer recurrence in administrative data: algorithm development and validation. *Curr Oncol*. 2022;29(8):5338-5367. https://doi.org/10.3390/curroncol29080424.
9. Colov EP, Fransgaard T, Klein M, Gögenur I. Validation of a register-based algorithm for recurrence in rectal cancer. *Dan Med J*. 2018;65(10):A5507.