Validity of Antineoplastic Procedure Codes in the Danish National Patient Registry

The Case of Colorectal Cancer

Martin Oskar Broe,[a] *Peter Bjødstrup Jensen,*[a] *Thea Otto Mattsson,*[b] *and Anton Pottegård*[a]

Background: Procedure codes in the Danish National Patient Registry are used for administrative purposes and are a potentially valuable resource for epidemiologic research. To our knowledge, the validity of antineoplastic procedure codes has only been evaluated in one study.

Methods: We randomly extracted a sample of 420 patients in the Southern Region of Denmark with a diagnosis of colorectal cancer and an oncology contact during 2016–2018. Using the medical record as gold standard, we computed the positive predictive value (PPV) and sensitivity of antineoplastic procedure codes recorded in the Danish National Patient Registry.

Results: We identified 2,243 codes for antineoplastic treatments in the registry and 2,299 in the medical records. We confirmed that 213 of 214 patients with registered therapies in the Danish National Patient Registry received therapy, corresponding to a PPV of "any registration" of 1.00 (95% confidence interval [CI] = 0.97, 1.00). Considering single registrations, the overall PPV was 0.95 (95% CI = 0.94, 0.95), and the overall sensitivity was 0.90 (95% CI = 0.89, 0.91). Number of recorded treatments and treatments administered were strongly correlated. Considering the most frequent single antineoplastic regimens, PPV ranged from 0.90 (95% CI = 0.87, 0.92) for capecitabine to 0.98 (95% CI = 0.95, 1.00) for cetuximab, whereas sensitivity ranged from 0.81 (95% CI = 0.75, 0.87) for 5-fluorouracil and irinotecan (FOLFIRI) regimen to 0.97 (95% CI = 0.94, 0.99) for bevacizumab. Analysis per hospital showed the highest validity of registrations at the University Hospital.

Conclusion: The validity of antineoplastic procedure codes in the Danish National Patient Registry is generally high and thus usable for epidemiologic research.

Keywords: Antineoplastic agents; Antineoplastic treatment; Danish National Patient Registry; Epidemiology; Sensitivity and specificity; Validity

(*Epidemiology* 2020;31: 599–603)

Submitted October 25, 2019; accepted April 30, 2020.

From the [a]Clinical Pharmacology and Pharmacy, Department of Public Health, University of Southern Denmark, Odense, Denmark; and [b]Department of Oncology, Odense University Hospital, Odense, Denmark.

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

The authors report no conflicts of interest.

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

Data access: Data cannot be shared with third parties as case sensitive material is protected by Danish law.

Correspondence: Anton Pottegård, Clinical Pharmacology and Pharmacy, Department of Public Health, University of Southern Denmark, JB Winsløws Vej 19, 2 DK-5000 Odense C, Denmark. E-mail: apottegaard@health.sdu.dk.

ISSN: 1044-3983/20/3104-0599
DOI: 10.1097/EDE.0000000000001208

The Danish National Patient Registry is a unique nationwide registry, covering all hospital admissions, treatments, and diagnoses at Danish hospitals since 1977.[1] Data can be linked at the individual level by using the Civil Personal Registry,[2] which enables patient identification if needed. The Patient Registry thus constitutes a central resource to Danish epidemiologic research.[1] One such use could be to conduct analysis on the use patterns and effects of use of antineoplastic treatments. Chemotherapy is not recorded in the Danish National Prescription Registry,[3] and it is currently neither included in the newly established Registry for Hospital Medication (Sygehusmedicinregisteret).[4] This leaves researchers with using procedure codes registered in the Patient Registry. However, the validity of antineoplastic procedure codes included in the Patient Registry is not clear. One previous study, Lund et al[5] (2013), concluded that the validity is generally high. However, this study only included 50 colorectal patients with nodal involvement, only assessed the validity of "any/ever" antineoplastic treatment, and provided no details as to the validity of the timing of antineoplastic treatments recorded in the Patient Registry. As such, it is largely unknown to what extent the Patient Registry can facilitate studies on the use of antineoplastic treatments. We therefore aimed to evaluate the validity of antineoplastic procedure codes in the Patient Registry, using colorectal cancer treatments as a case.

METHODS

We conducted this validation study in the Region of Southern Denmark among patients with a diagnosis of colorectal cancer. To obtain the positive predictive value (PPV) and sensitivity of the Patient Registry, we compared the procedure codes in the Patient Registry to the antineoplastic treatments prescription and administration as recorded in the medical records.

Data Sources

Denmark is divided into five regions that are comparable in sociodemographic and health-related characteristics.[6] Each region typically comprises one university hospital and several smaller hospitals. The antineoplastic treatments for the approximately 35,000 prevalent colorectal cancer patients in Denmark[7] are administered by the oncology departments in the regions. The Region of Southern Denmark has four oncology departments, located at Odense University Hospital, Hospital of Southern Jutland, Hospital Little Belt, and South-west Jutland Hospital. In this study, we used medical records obtained from each of the four hospitals, whereas data from the Patient Registry were obtained from the Danish Health Data Authority. We considered the medical record to have the highest validity and therefore our gold standard reference. The Patient Registry has been collecting nationwide administrative data of hospitals admissions since 1977.[1] One of the main purposes is to enable monitoring of diseases and treatments.[1] The Patient Registry uses International Classification of Diseases version 10 (ICD-10) diagnosis codes, and the validity among ICD-10 colorectal cancer diagnoses is found to be very high.[8]

Validation

We estimated the validity of antineoplastic procedure codes in the Patient Registry, on a random sample of colorectal patients who had contact with an oncology department. To identify our sample, we first sampled 800 random patients in the Region of Southern Denmark, with an ICD-10 diagnosis of colorectal cancer and contact with any hospital in the region within 1 May 2016 and 1 May 2018, from the Patient Registry. Because antineoplastic treatments only are administrated by specialists at the oncology departments, we could narrow down the included patients to those with an oncology contact within our study period, which restricted our sample to 431 patients. The initial sample size of 800 was largely arbitrary, mainly due to uncertainty in the proportion ultimately eligible for inclusion. However, it was judged that 431 patients were sufficient to allow reasonably precise estimates of validity, even in subgroups, and it was therefore decided not to expand the material further.

For the 431 patients, we obtained individual medical records directly from the oncology departments in the region. This was possible by using the unique personal identification code (Central Patient Registry number)[2] to link medical records to the data from the Patient Registry. The medical records included all records from nurses and doctors between 1 May 2016 and 1 May 2018. We examined the medical records manually for any antineoplastic treatments in the study period. For all administered therapies, the date and type of antineoplastic treatment was recorded for each individual patient. All data from the medical records were entered into REDCap,[9] which is a secure data capture web application.

We validated each treatment code (type of treatment) separately. For each round of validation, we considered a specific coding instance as valid if it fulfilled one of the three following criteria. First, and most commonly, we accepted a match on the exact code (either specifying a single treatment or a combination treatment). Second, we accepted a match by any combination of codes giving the same result, e.g., validating capecitabine/oxaliplatin (CAPOX); we accepted two single codes specifying capecitabine and oxaliplatin. As a notable exception, we accepted a code as true for oxaliplatin-containing combination treatments even though oxaliplatin was not administrated. This was done based on clinical input, as oxaliplatin is often discontinued, due to side effects, although the remaining treatment is continued. In supplementary analyses, we performed analyses without making this exception. Third, we accepted a match of a code for a single component with a code for a combination including that component; for example, when validating 5-fluorouracil, we accepted a code for the 5-fluorouracil containing regimen 5-fluorouracil/oxaliplatin (FOLFOX).

The handling of combination codes as described above is clinically sensible and ensures consistency in the 2×2 matrices for each validated code. It does, however, mean that some registrations in Patient Registry and the medical records may be involved in the validation of more than one treatment code. Unless otherwise specified, we accepted a one-day deviation between the registration in the Patient Registry and the medical record.

Statistical Analysis

To evaluate the validity of registrations for antineoplastic treatments in the Patient Registry, we applied four different analyses. First, we calculated the positive predictive value (PPV), negative predictive value (NPV), and sensitivity of "any chemotherapy," comparing the registrations in the Patient Registry to the medical record, while accepting a 1-month displacement between registrations. The PPV was defined as the number of confirmed chemotherapy recipients divided by total number of patients recorded as having received chemotherapy according to the Patient Registry. The NPV value was defined as the number of patients confirmed not to have received any chemotherapy divided by total number of patients recorded as not having received chemotherapy according to the Patient Registry. The sensitivity was defined as the number of confirmed chemotherapy recipients divided by the total number of chemotherapy recipients according to the medical records. Second, we assessed how well the number of registrations of antineoplastic treatments per patient correlated with the number of administrated therapies per patient. Third, we investigated the validity of the recording of specific treatment regimens and the differences in validity across individual hospitals (n = 4). Last, we investigated the validity of the timing of registrations by calculating the PPV and sensitivity for any antineoplastic treatment overall and per hospital, allowing a deviation of 5 and 1 days, respectively. We calculated exact 95% confidence intervals for proportions. Analyses were performed using STATA 16 (Stata-Corp, College Station, TX).

TABLE 1. Overview of Patients and Observations of Patients with a Diagnose of Colorectal Cancer in the Region of Southern Denmark During 1 May 2016 to 1 May 2018

n = 420	
Median age and quartiles	71 [64–76]
Male sex	57%
No. of patients receiving treatment according to the medical records	220
No. of observations in Danish National Patient Registry	2,243
No. of observations in the medical records	2,299
Positive predictive value of patients who receives any treatment (95% CI)	1.00 (0.97, 1.00) 213/214
Sensitivity of patients who receives any treatment (95% CI)	0.97 (0.94, 0.99) 213/220
Negative predictive value of patients classified as not having received any treatment (95% CI)	0.97 (0.93, 0.99) 199/206

Approvals

We obtained permission to access the medical journals from the Danish Patient Safety Authority (record no. 3-3013-2494/1). In terms of data protection, the study was registered at the University of Southern Denmark's inventory (record no. 18/12145). Approval from the Ethics Committee was not required.

RESULTS

For our sample of 431 patients with a diagnosis of colorectal cancer and a contact with an oncology department, we obtained all antineoplastic treatment procedure codes from Patient Registry during 1 May 2016 to 1 May 2018. Of the 431 patients, we excluded 11 patients who, while correctly identified as patients with a history of colorectal cancer, had other reasons for their current oncology contact (that is, other cancers). Of the 420 eligible patients, 220 patients had at least one antineoplastic treatment registered in the medical record, comprising in total 2,299 antineoplastic records of antineoplastic treatments. In the Patient Registry, 214 patients had at least one antineoplastic treatment registered, with a total of 2,243 records of antineoplastic treatments (Table 1). The main reasons for an oncology contact not accompanied by any treatment of interest were planned follow-up visits, multidisciplinary team conferences, patients only receiving radiotherapy, and patients whose health condition did not allow antineoplastic treatment. Assessing the validity of having at least one record of antineoplastic treatment in the Patient Registry (Table 1), we found a PPV of 1.00 (0.97–1.00), that is, 213 out of 214 patients in the Patient Registry actually received antineoplastic treatment. The corresponding sensitivity was 0.97 (0.94–0.99), that is, 213 out of 220 patients receiving antineoplastic treatment were captured by the Patient Registry. Conversely, the NPV, that is, the certainty that a patient classified as not receiving any chemotherapy according to the Patient Registry did in fact not receive therapy, was 0.97 (0.93–0.99). In the Figure, we show the percentage of antineoplastic treatments registered in the medical record, which are also found in the Patient Registry, for a given individual. The percentages are generally high, regardless of the number of antineoplastic treatments found in the medical records. With a few exceptions, more than 75% of all individuals had all their treatments recorded in the Patient Registry, and almost

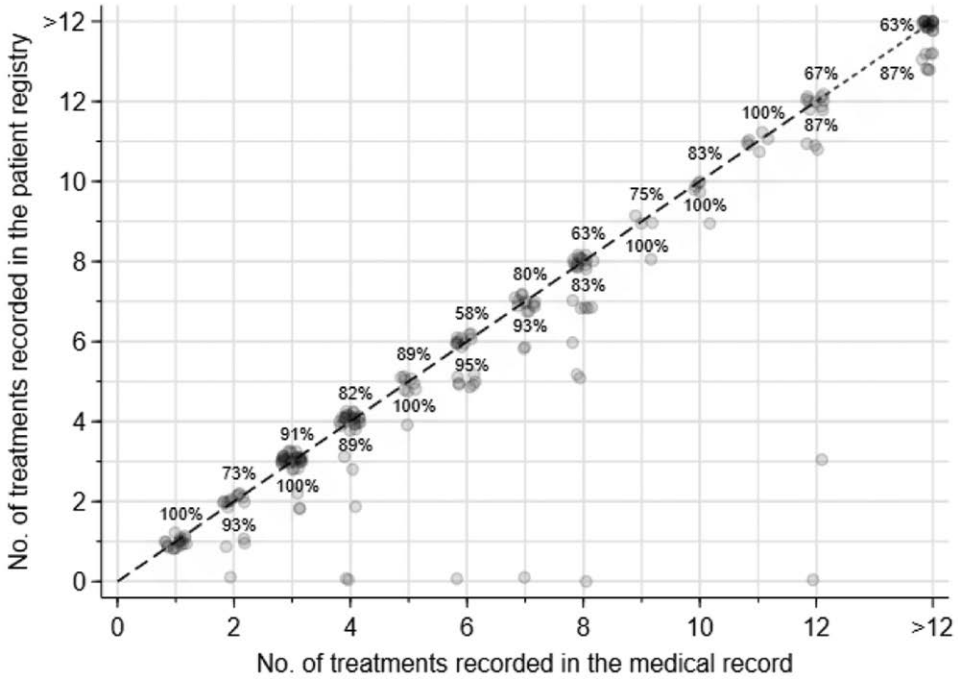

FIGURE. Proportion of correct registered treatments, matched by any treatment, per patient in the Region of Southern of Denmark, accepting 1-day deviation between registrations and using the medical record as the gold standard reference.

everyone had all-but-one treatments recorded (Figure). These estimates did not change if allowing a 30-day deviation between registrations (data not shown). For individual records, the overall PPV was 0.95 (0.94–0.95) and the sensitivity was 0.90 (0.89–0.91) (Table 2). The highest overall PPV and sensitivity were found for the regional university hospital at 0.95 (0.95–0.96) and 0.91 (0.90–0.92), respectively. The overall PPV of the three other (nonuniversity) hospitals within the region ranged from 0.84 (0.76–0.91) to 0.93 (0.91–0.95), and the overall sensitivity ranged from 0.79 (0.71–0.87) to 0.87 (0.83–0.90) (eAppendix; http://links.lww.com/EDE/B671). For the most frequent specific treatments, the overall PPV ranged from 0.90 (95% confidence interval [CI] = 0.87, 0.92) for capecitabine to 0.98 (95% CI = 0.95, 1.00) for cetuximab, whereas the sensitivity ranged from 0.81 (95% CI = 0.75, 0.87) for FOLFIRI regimen (5-fluorouracil/irinotecan) to 0.97 (95% CI = 0.94, 0.99) for bevacizumab (Table 2). Last, we investigated the timing of registrations by allowing a 5-day deviation between a registration in the Patient Registry and the medical record, which did not change the PPV and sensitivity for either the overall values or specific treatments or per hospital (data not shown).

TABLE 2. PPV, Sensitivity, and 95% CIs for All Antineoplastic Procedure Codes and Specific Treatments for All Departments in the Region of Southern Denmark, Accepting 1-Day Deviation Between Registrations and Using the Medical Record as the Gold Standard Reference

	All (n = 420)	
	PPV (95% CI)	**Sensitivity (95% CI)**
All codes	0.95 (0.94, 0.95)	0.90 (0.89, 0.91)
	3,282/3,465	3,282/3,663
All cytostatic	0.95 (0.94, 0.95)	0.90 (0.88, 0.91)
	2,851/3,013	2,851/3,184
All biological	0.95 (0.93, 0.97)	0.90 (0.87, 0.93)
	431/452	431/479
Specific regimes		
CAPOX[a]	0.95 (0.91, 0.98)	0.90 (0.85, 0.94)
	184/194	184/204
FOLFOX[a]	0.95 (0.92, 0.97)	0.90 (0.86, 0.93)
	268/282	268/298
Capecitabine	0.90 (0.87, 0.92)	0.92 (0.90, 0.94)
	563/625	563/610
FOLFIRI	0.95 (0.90, 0.98)	0.81 (0.75, 0.87)
	148/156	148/182
Irinotecan	0.97 (0.95, 0.98)	0.85 (0.81, 0.88)
	356/367	356/420
5-Fluorouracil	0.97 (0.95, 0.98)	0.90 (0.87, 0.92)
	536/552	536/598
Bevacizumab	0.93 (0.89, 0.96)	0.97 (0.94, 0.99)
	233/251	233/241
Cetuximab	0.98 (0.95, 1.00)	0.85 (0.80, 0.90)
	174/177	274/204

[a]Eight percent of CAPOX and 20% of FOLFOX treatments, according to medical records, were without oxaliplatin.

DISCUSSION

In this validation study of antineoplastic procedure codes in colorectal cancer patients in the Danish National Patient Registry, we found a high completeness and validity both for "any use" and for individual administrations. When we specified by the individual antineoplastic regimens or hospitals, we identified some variation; however, the validity generally remained high. Analysis of timing of registrations showed no variation.

The most important strength of this study is the large and randomly extracted sample of patients. Our study also had several limitations. First, we only included one out of five regions in Denmark, although we did include all hospitals within the region. Further, only one author reviewed the medical records, which only included data from 2016 to 2018. Although we found the calendar restriction to be necessary to ensure a high retrieval rate of patient records, it is unknown whether the validity of registrations has changed over time. Also, the selection of patients was by the diagnosis code of colorectal cancer in the Patient Registry and therefore depending on the sensitivity hereof. However, Helqvist et al[8] (2012) found a sensitivity 0.93 (95% CI = 0.91, 0.94) of the ICD-10 colorectal cancer diagnosis. Last, our results are based on patients diagnosed with colorectal cancer, in which treatment is mainly intravenously administered (except capecitabine) in an ambulatory setting. Whether our findings can be extrapolated to other cancer diagnoses and treatments in the Patient Registry, for example, cancers where oral treatment is more common, remains unknown. However, oral chemotherapy is in Denmark prescribed and dispensed via the same systems, and as such it is likely that the results can be generalized to other cancers and other chemotherapy regimens. Nevertheless, it would be valuable for future studies to address the validity of registrations both within other cancers and other aspects of cancer treatment, e.g., radiation or surgery.

To our knowledge, only three previous studies have investigated procedure codes in the Patient Registry. They all showed a high PPV and sensitivity values for the registry.[5,10,11] Nielsson et al[10] (2012) investigated procedure codes for intravenous bisphosphonate administration, Adelborg et al[11] (2016) cardiac procedures, and Lund et al[5] (2013) antineoplastic treatments among colorectal patients. Most of the studies only included patients from University Hospitals, whereas we included all hospitals within the region. Our findings showed a slightly lower validity of specific treatments among nonuniversity hospitals, which emphasizes the need for including all hospitals when doing validation studies.

Like our study, Lund et al[5] (2013) investigated antineoplastic treatments among patients with colorectal cancer. They used both medical records and pharmacy production data to construct a reference standard, where we only used

the medical record. To enhance the likelihood of patients receiving therapy, Lund et al[5] (2013) included patients by a diagnose of colorectal cancer with nodal involvement. In our study, we made no such selection of the diseases stage.

In conclusion, our study show that the validity of antineoplastic procedure codes in the Danish National Patient Registry is generally high and thus usable for epidemiologic research.

ACKNOWLEDGMENTS

We acknowledge the oncological departments in the Region of Southern Denmark for assisting in obtaining medical records.

REFERENCES

1. Schmidt M, Schmidt SAJ, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. *Clin Epidemiol*. 2015;7:449–490.
2. Schmidt M, Pedersen L, Sørensen HT. The Danish Civil Registration System as a tool in epidemiology. *Eur J Epidemiol*. 2014;29:541–549.
3. Pottegård A, Schmidt SAJ, Wallach-Kildemoes H, Sørensen HT, Hallas J, Schmidt M. Data resource profile: the Danish National Prescription Registry. *Int J Epidemiol*. 2017;46:798–798f.
4. The Danish Health Data Authority. [Registry for Hospital Medication] (Sygehusmedicinregisteret). Available at: https://sundhedsdatastyrelsen.dk/da/rammer-og-retningslinjer/sygehusmedicinregisteret. Accessed December 17, 2019.
5. Lund J, Frøslev T, Deleuran T, et al. Validity of the Danish National Registry of Patients for chemotherapy reporting among colorectal cancer patients is high. *Clin Epidemiol*. 2013;5:327–334.
6. Henriksen DP, Rasmussen L, Hansen MR, Hallas J, Pottegård A. Comparison of the five Danish regions regarding demographic characteristics, healthcare utilization, and medication use–a descriptive cross-sectional study. *PLoS One*. 2015;10:e0140197.
7. Danish Colorectal Cancer Group. Annual report - 2017.pdf. Available at: https://dccg.dk/wp-content/uploads/2018/10/DCCG-%C3%85rsrapport-2017.pdf. Accessed February 18, 2019.
8. Helqvist L, Erichsen R, Gammelager H, Johansen MB, Sørensen HT. Quality of ICD-10 colorectal cancer diagnosis codes in the Danish National Registry of Patients. *Eur J Cancer Care (Engl)*. 2012;21:722–727.
9. REDCap. Available at: https://projectredcap.org/. Accessed February 20, 2019.
10. Nielsson MS, Erichsen R, Frøslev T, Taylor A, Acquavella J, Ehrenstein V. Positive predictive values of the coding for bisphosphonate therapy among cancer patients in the Danish National Patient Registry. *Clin Epidemiol*. 2012;4:233–236.
11. Adelborg K, Sundbøll J, Munch T, et al. Positive predictive value of cardiac examination, procedure and surgery codes in the Danish National Patient Registry: a population-based validation study. *BMJ Open*. 2016;6:e012817.