ORIGINAL ARTICLE

WILEY

Positive predictive values of hematological procedure codes in the Danish National Patient Registry—A population-based validation study

Tine Bichel Lauritsen[1] 🄳 | Jan Maxwell Nørgaard[1] |
Mathilde Egelund Christensen[2,3] | Susanne Oksbjerg Dalton[4,5,6] |
Lene Sofie Granfeldt Østgård[7,8]

[1]Department of Hematology, Aarhus University Hospital, Aarhus, Denmark

[2]Department of Hematology, Rigshospitalet, Copenhagen, Denmark

[3]The Research Unit for General Practice and the Section of General Practice, Department of Public Health, University of Copenhagen, Copenhagen, Denmark

[4]Danish Cancer Society Research Center, Danish Cancer Society, Copenhagen, Denmark

[5]Department of Clinical Oncology and Palliative Care, Zealand University Hospital, Roskilde, Denmark

[6]Institute of Clinical Medicine, Copenhagen University, Copenhagen, Denmark

[7]Department of Hematology, Odense University Hospital, Odense, Denmark

[8]Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark

Correspondence
Tine Bichel Lauritsen, Department of Hematology, Aarhus University Hospital, Palle Juul-Jensens Boulevard 99, 8200 Aarhus N, Denmark.
Email: tinelaur@rm.dk

Abstract

Background: The Danish National Patient Registry holds data on hematological procedure codes including date and type of treatment from all hematological departments in Denmark. The validity of the hematological procedure codes remains to be clarified before they are used in epidemiological research.

Patients and Methods: Using the Danish Myelodysplastic Syndromes Database, we identified 897 patients diagnosed with myelodysplastic syndromes or chronic myelomonocytic leukemia treated at five Danish Hospitals between 1 January 2012 and 30 April 2019. From the Danish National Patient Registry, we ascertained information about hematological procedure codes and date of procedure registered on each patient and generated random samples. Using medical record review as the reference standard, we validated procedure codes in the Danish National Patient Registry and calculated positive predictive values (PPVs) with 95% confidence intervals (CIs) for each procedure code.

Results: A total of 523 medical records (99% of the total sample) were available for review. PPVs for specific procedure codes ranged from 71% to 100%. The overall PPV was 91% (95% CI: 88%–92%), reflecting PPVs of 95% (95% CI: 92%–97%) for low-dose-chemotherapy, 90% (95% CI: 81%–96%) for high-dose chemotherapy, 99% (95% CI: 93%–100%) for allogeneic stem cell transplantation, 75% (95% CI:

62%–85%) for immuno-modulating agents, 80% (95% CI: 74%–85%) for growth factors, and 99% (95% CI: 99%–100%) for bone marrow examination. The accuracy of coding was consistent across geographic regions and year of registration/coding.

Conclusions: Hematological procedure codes reported to the Danish National Patient Registry had high PPVs and are suitable for epidemiological research.

KEYWORDS
hematology, procedure codes, Registry, validity

Keypoints

- We conducted a nationwide validation study of hematological procedure codes recorded in the Danish National Patient Registry during 2012–2019.
- Using medical record review as a reference standard, we calculated positive predictive values (PPVs) of hematological procedure codes in the Danish National Patient Registry demonstrating PPVs ranging between 71% and 100%.
- Overall PPVs of first-time hematological procedure codes were highest for low-dose-chemotherapy (95%), high-dose chemotherapy (90%), and allogeneic stem cell transplantation (99%), but lower for immuno-modulating agents (75%) and hematopoietic growth-factors (80%).
- The Danish National Patient Registry constitutes an important data source for future pharmaco-epidemiological studies of patients with hematological diseases.

Plain language summary

The Danish National Patient Registry (DNPR) is a nationwide hospital registry that holds information on in-hospital treatment procedure codes, for example, hematological chemotherapy codes since 1999. We examined the data quality of hematological procedure codes recorded in the DNPR in a cohort of patients with myelodysplastic syndromes or chronic myelomonocytic leukemia treated at five Danish Hospitals between January 1, 2012 and April 30, 2019. From the DNPR, we ascertained information about hematological procedure codes and date of procedure registered on each patient. Considering information in the medical records as the golden standard, we reviewed medical records to ensure that procedure codes recorded in the DNPR were truly administered. We reviewed 523 medical records. Overall, data quality was high. Of 845 procedure codes recorded in the DNPR, 765 (91%) were administered according to the medical records. Procedure codes on high-dose chemotherapy, bone marrow transplantation, and low-dose chemotherapy had the highest data quality while procedure codes on immune-modulating agents and hematopoietic growth-factors were less accurate. In conclusion, data quality of hematological procedure codes reported to the DNPR was high and the codes are suitable for epidemiological research.

1 | INTRODUCTION

Danish population-based health care registries are frequently used in epidemiological cancer studies on disease etiology, cancer prognosis, and safety and efficacy of various treatment modalities.[1,2] Within hematology, several detailed clinical quality databases exist.[3–7] The main aim of these databases is to monitor the quality of care of Danish patients with hematological cancers. The databases hold detailed and valid clinical information on date and type of diagnosis, laboratory values, cytogenetic and mutational status and disease stage.[7,8] Treatment information is also available,

however, the degree of treatment details and the quality of data differs greatly among the hematological databases.[3,5–8] As a supplement to these clinical quality databases, data on type and time of in-hospital and outpatient treatment modalities can be retrieved from The Danish National Patient Registry (DNPR).[9] This registry includes prospectively collected data from all Danish hospitals on in-hospital treatments e.g. chemotherapy and surgical procedures since 1999.[9] The evidence on the validity of procedure codes, especially for anti-neoplastic treatment, in the DNPR is however sparse[10,11] and is particularly lacking for treatment modalities used for hematologic cancer patients.

Because misclassification may impact study findings, the quality of registry-based research largely depends on the data validity. We therefore examined the positive predictive values (PPVs) of hematological procedure codes across Danish hospitals in the DNPR using medical record review as the reference standard.

2 | MATERIALS AND METHODS

2.1 | Setting

In Denmark a universal and tax-supported health-care system ensures free and equal access for all residents to general practitioners and hospitals, including treatment of cancer.[1] Denmark is divided into five geographic and health administrative regions (North Denmark Region, Central Denmark Region, Region of Southern Denmark, Region Zealand, and Capital Region of Denmark) which are considered comparable regarding sociodemographic characteristics, medication use, and healthcare utilization.[1] Patients with hematological diseases are treated at one of 10 hematological departments in Denmark and no hematological patients are treated at private hospitals or clinics. Registration of the hematological procedure codes is performed by the treating physician or nurse in collaboration with a secretary and it is mandatory by law for each hospital department to submit their data electronically to the DNPR monthly.[9]

2.2 | Study population and study variables

We performed a nationwide population-based validation study. Our study-population comprised all patients diagnosed with either myelodysplastic syndromes (MDS) or chronic myelomonocytic leukemia (CMML) and registered in the Danish Myelodysplastic Syndromes Database. This database is a nationwide clinical cancer database that captures nearly 100% of patients diagnosed with MDS or CMML since 2010.[4] We restricted our study-population to cover five hospitals in three Danish regions and due to study feasibility we applied different study periods for the different hospitals. Our final study-population thus included patients treated at Aarhus University Hospital and Holstebro Regional Hospital (Central Denmark Region) between January 1, 2012 and April 30, 2019; patients treated at Odense University Hospital (Southern Denmark Region) between November 15, 2015 and April 30, 2019; patients treated at Rigshospitalet (Capital Region of Denmark) between November 5, 2016 and April 30, 2019, and patients treated at Herlev Hospital (Capital Region of Denmark) between May 20, 2016 and April 30, 2019.

Using a 10-digidt identification number assigned to all Danish citizens upon birth or immigration, by the Danish Civil Registration System,[2] we linked our study-population with data from the DNPR. This was done to retrieve information on all first-time procedure codes, including date of procedure on hematopoietic growth-factors,

TABLE 1 Hematological procedure codes in the Danish National Patient Registry validated in the present study

Procedure	Codes
Low-dose chemotherapy	
Low-dose cytarabine	BWHA158
Azacitidine	BWHA256
High-dose-chemotherapy	
Remission induction chemotherapy[a]	BWHA3
High-dose cytarabine	BWHA301
Anthracycline+cytarabine[b]	BWHA303
Allogeneic stem cell transplantation	
Myeloablativ HSCT	BOQF11/BOQF12
Non-myeloablativ HSCT	BOQF21/BOQF22
Immuno-modulating agents	
Lenalidomide	BWHB82
Ciclosporine	BOHJ20
Growth-factors	
Erythropoiesis-stimulating agent	BOHE10
Granulocyte-colony-stimulating-factor	BOHE20
Bone marrow examination	KTNE25A

Abbreviation: HSCT, hematopoietic stem cell transplantation.
[a]Covers any type of remission induction chemotherapy, forexample, "3 + 10," "High-dose cytarabine," "MITO-FLAG," etc.
[b]Covers treatment with an anthracycline in combination with cyterabine such as "3 + 10," "3 + 8," "2 + 5," etc.

low-dose chemotherapy, high-dose chemotherapy, allogeneic stem cell transplantation (allo-HSCT), immunomodulating agents, and bone marrow biopsies registered in the DNPR following a first-time diagnosis of MDS or CMML (specific codes are available in Table 1). For validation, we identified a random sample of 100 patients for each code, or the highest obtainable number if fewer patients were available (Figure 1). For the low-dose chemotherapy agent azacitidine, which usually is administered daily for five consecutive days in 21 days' cycles, we also validated procedure codes for injection number 6, 11, 16, and 21, as these represent day 1 in treatment cycle 2, 3, 4, and 5. This was done to evaluate if the validity of the coding changed during the treatment course.

To calculate the negative predictive value (NPV) of "no treatment," we sampled 100 patients who had no registration with any of the included procedure codes in our study except from codes encoding bone marrow biopsy, since all patients were diagnosed using this procedure.

2.3 | Medical record review

We considered treatment information in the medical records as the reference standard and validated procedure codes in the DNPR against the medical records to examine PPVs of each procedure code. The registration of a procedure code was classified as 1) correct, if
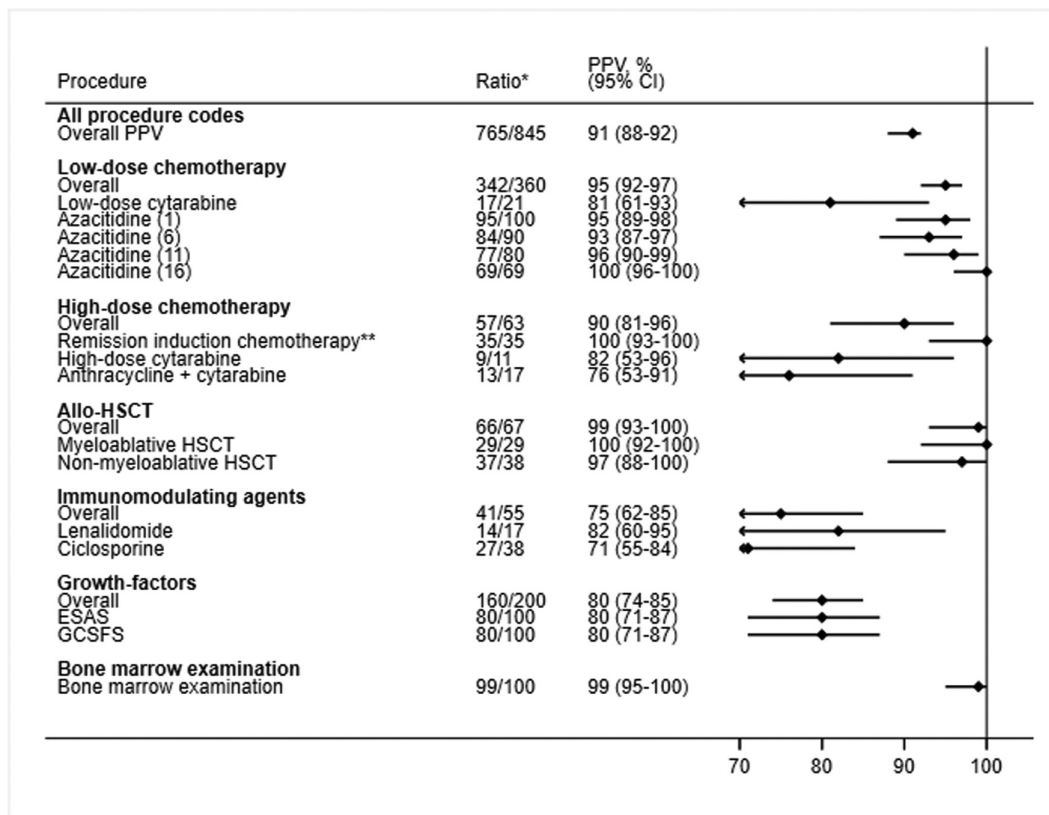
FIGURE 1 Positive predictive values of first-time hematological procedure codes recorded in the Danish National Patient Registry. Correct coding was defined as a procedure that was administered for the first time according to the medical records and within a one-day window between the date of registration in the Danish National Patient Registry and the medical records. CI, confidence interval; ESAS, erythropoiesis-stimulating agents; GCSFS, granulocyte-colony-stimulating-factors; HSCT, allogeneic hematopoietic stem cell transplantation; PPV, positive predictive value. *Number of correct codes divided by the total number of medical reviews; ** Covers any type of remission induction chemotherapy, for example, "3 + 10," "High-dose cytarabine," "MITO-FLAG," etc.

there was a maximum of a one-day window between the registration in the DNPR and the medical records and if the procedure was administered for the first time according to the medical records, 2) partly correct, if the procedure was administered within a one-day window between the registration in the DNPR and the medical records, but not for the first time according to the medical records, or as 3) incorrect, if absent in the medical records. To examine the NPV we similarly reviewed the 100 medical patient records that according to the DNPR had none of the pre-specified procedure codes except from the code encoding bone marrow biopsy.

The adjudication was performed by two physicians with hematologic experience (MA and TBL) during April 2021 through September 2021.

2.4 | Statistical analysis

The PPVs and NPV were calculated with 95% confidence intervals (CIs) using the Jeffreys method.[12] The PPV was defined as the number of confirmed first-time treatment recipients in the medical records divided by the total number of patients recorded as having received

first-time treatment according to the DNPR. The NPV was defined as the number of patients confirmed not to have received any treatment in the medical records divided by the total number of patients recorded as not having received treatment according to the DNPR. PPVs were also calculated separately by calendar year (2012–2016 vs. 2017–2019) and by geographic Region (Central Region Denmark, Region of Southern Denmark and Capital Region of Denmark). The latter was done to investigate whether coding practice differed among the different hospitals as the Regions use different electronic medical record systems, which could result in different coding practice

Finally, we also estimated PPVs including both correct and partly correct answers as correct.

We collected and managed study data using REDCap (Research Electronic Data Capture) hosted at Aarhus University Hospital, Central Region Denmark. REDCap is a secure, web-based software platform designed to support data capture for research studies.[13] Stata version 16 was used for the sampling process and the statistical analyses.

The study was approved by the Danish Data Protection Agency (record number: 1-16-02-321-18) and the Danish Patient Safety Authority (record number 3-3013-2960/1). In accordance with Danish law, no approval from the Ethics Committee was required.
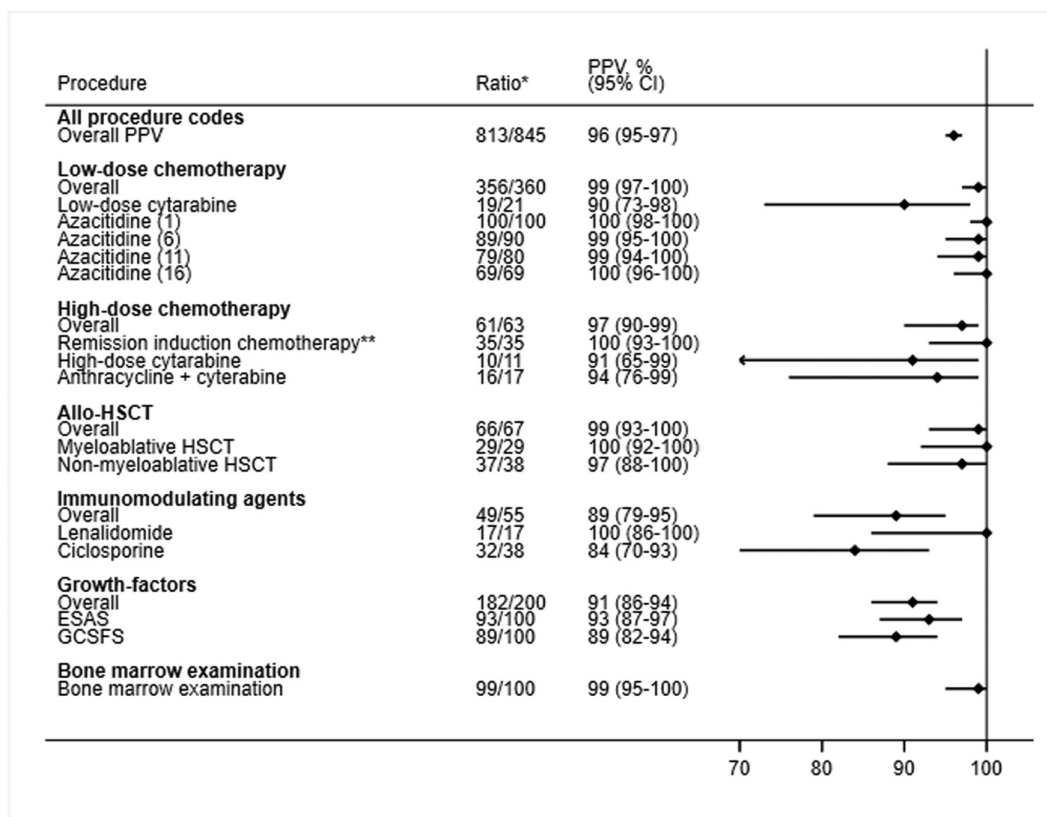
FIGURE 2 Positive predictive values of first-time hematological procedure codes recorded in the Danish National Patient Registry. Correct coding was defined as a procedure that was administered within a one-day window between the date of registration in the Danish National Patient Registry and the medical records. CI, confidence interval; ESAS, erythropoiesis-stimulating agents; GCSFS, granulocyte-colony-stimulating-factors; HSCT, allogeneic hematopoietic stem cell transplantation; PPV, positive predictive value.*Number of correct codes divided by the total number of medical reviews; ** Covers any type of remission induction chemotherapy, for example, "3 + 10," "High-dose cytarabine," "MITO-FLAG," etc.

3 | RESULTS

We retrieved hematological procedure codes from the DNPR on 897 patients from the Danish Myelodysplastic Syndromes Database who had either MDS or CMML. We sampled 527 patients, and of these, medical records were available for 523 (99%) patients. In total, 765 out of 845 first-time procedure codes in the DNPR were correctly coded corresponding to an overall PPV of first-time procedure codes of 91% (95% CI: 88%–92%). Overall PPVs and PPVs for each procedure code are provided in Figure 1. PPVs ranged from 71% to 100%. In general, PPVs were high for bone marrow examination (99% [95% CI: 95%–100%]), low-dose-chemotherapy (95% (95% CI: 92%–97%)), high-dose chemotherapy (90% (95% CI: 81%–96%)) and allo-HSCT (99% (95% CI: 93%–100%)). PPVs for growth-factors (80% (95% CI: 74%–85%)) and immune-modulating agents (75% (95% CI: 62%–85%)) were lower. Noticeably, correct coding of the administration of azacitidine remained high over treatment courses 2, 3, 4, and 5 with PPVs exceeding 93%.

When considering both correct and partly correct answers as correct (accepting that the treatment was not administered for the first time according to the medical record), the overall PPV for all treatments and procedures increased to 96% (95% CI: 95%–97%) (Figure 2). Additionally, when the date of first-time procedure code according to the DNPR did not correspond to the first time treatment was administered in the medical records, we noticed that the true first treatment-date according to the medical records in the majority of cases was within 30 days prior to the first-time procedure code date in the DNPR (Table 2). Stratified analysis by geographical region and calendar year agreed with our main findings (Tables 3 and 4). The negative predictive value of not receiving any treatment was 99% (95% CI: 95%–100%).

4 | DISCUSSION

In this study, we found high agreement between the majority of procedure codes in the DNPR and the actual administered treatment according to the medical patient records.

To the best of our knowledge, no previous studies on the validity of hematological procedure codes in the DNPR exist. Two prior studies examined the validity of antineoplastic codes in the DNPR among colorectal cancer patients.[10,11] Lund et al. reported on the validity of

TABLE 2 Proportion of first-time procedure codes in the Danish National Patient Registry where the procedure was administered within ± 1 day, −2/−30 days, and more than 30 days ago according to the first true treatment date in the medical record

Time between the first registration in the DNPR and the true first treatment date in the medical record	Proportion (number)
Azacitidine1	
Within ±1 day	95% (95)
Within −2 to −30 days	4% (4)
More than 30 days ago	1% (1)
Azacitidine6	
Within ±1 day	94% (84)
Within −2 to −30 days	6% (5)
More than 30 days ago	–
Azacitidine11	
Within ±1 day	97% (77)
Within −2 to −30 days	3% (2)
More than −30 days ago	–
Azacitidine16	
Within ±1 day	100 (69)
Within −2 to −30 days	–
More than −30 days ago	–
Low-dose cytarabine	
Within ±1 day	89% (17)
Within −2 to −30 days	–
More than −30 days ago	11% (2)
Remission induction chemotherapy	
Within ±1 day	100% (35)
Within −2 to −30 days	–
More than −30 days ago	–
High-dose cytarabine	
Within ±1 day	90% (9)
Within −2 to −30 days	–
More than −30 days ago	10% (1)
Anthracycline + cytarabine	
Within ±1 day	81% (13)
Within −2 to −30 days	–
More than −30 days ago	19% (3)
Erythropoiesis stimulating agents	
Within ±1 day	86% (80)
Within −2 to −30 days	12% (4)
More than −30 days ago	2% (2)
Granulocyte colony stimulating factors	
Within ±1 day	90% (80)
Within −2 to −30 days	6% (5)
More than −30 days ago	4% (4)
Lenalidomide	
Within ±1 day	84% (27)

TABLE 2 (Continued)

Time between the first registration in the DNPR and the true first treatment date in the medical record	Proportion (number)
Within −2 to −30 days	10% (9)
More than −30 days ago	6% (2)
Ciclosporine	
Within ± 1 day	84% (27)
Within −2 to −30 days	10% (9)
More than −30 days ago	6% (2)

antineoplastic procedure codes in the DNPR among 50 colorectal cancer patients treated at two University Hospitals in Denmark and found that the overall measures of validity (sensitivity, specificity, PPV, and NPV) for identifying the receipt of any chemotherapy were high with measures ranging between 87% and 100%. Broe et al. investigated PPVs and sensitivity of antineoplastic procedure codes recorded in the DNPR among 431 colorectal cancer patients treated in the Southern Region of Denmark between May 1, 2016 and May 1, 2018.[10] They found an overall PPV for single registrations of 95% (95% CI: 94%–95%) and an overall sensitivity of 90% (95% CI: 89%–91%). Furthermore, they found a NPV of 97% (95% CI: 93%–99%). In line with these studies, we found similar PPVs for low-dose chemotherapy (95% (95% CI: 92%–97%)) and more advanced treatments like high-dose chemotherapy (90% (95% CI: 81%–96%)) and allo-HSCT (99% (95% CI: 93%–100%)). These are expensive treatments with potentially life-threatening side-effects, which may explain the high accuracy of these procedure/treatment codes. For growth-factors and immune-modulating agents, we found a lower overall PPV which may reflect that these treatments often are self-administered, and therefore, not as accurately coded as chemotherapy procedure codes.

In studies examining the safety and effectiveness of different treatment modalities, high PPVs of treatment procedure codes are of particular importance to ensure that only individuals who truly received the treatment of interest are included in the study.[14] In studies with a comparison group, a high NPV is also important to ensure that unexposed patients did not receive any treatment, as contamination of the reference group would bias the associations toward unity. In this study the NPV of not receiving any treatment was 99% (95% CI: 95%–100%). We examined PPVs stratified by calendar year (2012–2016 vs. 2017–2019) and found similar PPVs for most of the procedure codes suggesting that data can be used to study changes in, that is, azacitidine treatment or allo-HSCT over time.

The present study has obvious strengths. We conducted a comprehensive medical record review validating procedure codes from 3 out of 5 regions in Denmark during a time period of 7 years. Given the homogeneity of the Danish Health-care system,[1] it is reasonable to assume that our results also reflect coding practice in the two Danish regions we did not include in our study.

Still, our study may have some limitations. First, the medical record review was performed by only two reviewers, which might

TABLE 3 Positive predictive values of first-time hematological procedure codes in the Danish National Patient Registry, by calendar year

Procedure	2012–2016		2017–2019	
	Ratio[a]	PPV (95% CI)	Ratio[a]	PPV (95% CI)
All procedure codes				
Overall PPV	303/335	90 (0.87–0.93)	462/510	91 (0.88–0.93)
Low-dose chemotherapy				
Overall	148/155	95 (0.91–0.98)	194/205	95 (0.91–0.97)
Low-dose cytarabine	5/5	100 (0.62–1.00)	12/16	75 (0.51–0.91)
Azacitidine (1)	45/46	100 (0.90–1.00)	50/54	93 (0.83–0.97)
Azacitidine (6)	34/38	89 (0.77–0.96)	50/52	96 (0.88–0.99)
Azacitidine (11)	34/36	94 (0.83–0.99)	43/44	98 (0.90–1.00)
Azacitidine (16)	30/30	100 (0.92–1.00)	39/39	100 (0.94–1.00)
High-dose-chemotherapy				
Overall	5/7	71 (0.35–0.94)	52/56	93 (0.84–0.98)
Remission induction chemotherapy[b]	c	c	33/33	100 (0.93–1.00)
High-dose cytarabine	c	c	9/11	82 (0.53–0.96)
Anthracycline + cytarabine	c	c	10/12	83 (0.56–0.96)
Allogenic stem cell transplantation				
Overall	19/19	100 (0.88–1.00)	47/48	98 (0.91–1.00)
Myeloablative HSCT	c	c	26/26	100 (0.91–1.00)
Nonmyeloablative HSCT	c	c	21/22	95 (0.81–1.00)
Immuno-modulating agents				
Overall	12/14	86 (0.62–0.97)	29/41	71 (0.56–0.83)
Lenalidomide	6/6	100 (0.67–1.00)	8/11	73 (0.43–0.92)
Ciclosporine	6/8	75 (0.41–0.94)	21/30	70 (0.52–0.84)
Growth-factors				
Overall	78/98	80 (0.71–0.87)	82/102	80 (0.72–0.87)
Erythropoiesis stimulating agents	44/56	79 (0.67–0.88)	36/44	82 (0.69–0.91)
Granulocyte-colony-stimulating-factors	34/42	81 (0.67–0.91)	46/58	79 (0.68–0.88)
Bone marrow examination				
Bone marrow biopsy	41/42	98 (0.89–1.00)	58/58	100 (0.96–1.00)

Abbreviations: CI, confidence interval; HSCT, allogeneic hematopoietic stem cell transplantation; PPV, positive predictive value.
[a]Number of correct codes divided by the total number of medical reviews.
[b]Covers any type of remission induction chemotherapy, e.g. "3 + 10", "High-dose cytarabine", "MITO-FLAG" etc.
[c]Data not shown due to small numbers.

TABLE 4 Positive predictive values of first-time hematological procedure codes in the Danish National Patient Registry, stratified by regions

Procedure	Central region Denmark		Capital region of Denmark		Region of Southern Denmark	
	Ratio[a]	PPV (95% CI)	Ratio[a]	PPV (95% CI)	Ratio[a]	PPV (95% CI)
All procedure codes						
Overall	289/314	92 (0.89–0.95)	237/270	88 (0.83–0.91)	239/261	92 (0.88–0.94)
Low-dose chemotherapy						
Overall	124/127	98 (0.94–0.99)	89/93	96 (0.90–0.99)	129/140	92 (0.87–0.96)
Low-dose cytarabine	b	b	13/15	87 (0.64–0.97)	c	c
Azacitidine	34/34	100 (0.93–1.00)	24/25	96 (0.83–1.00)	37/41	90 (0.78–0.97)
Azacitidine	29/30	97 (0.85–1.00)	23/23	100 (0.90–1.00)	32/37	86 (0.73–0.95)
Azacitidine	30/30	100 (0.92–1.00)	16/17	94 (0.76–0.99)	31/33	94 (0.82–0.99)
Azacitidine	27/27	100 (0.91–1.00)	13/13	100 (0.83–1.00)	29/29	100 (0.92–1.00)
High-dose-chemotherapy						

TABLE 4 (Continued)

Procedure	Central region Denmark		Capital region of Denmark		Region of Southern Denmark	
	Ratio[a]	PPV (95% CI)	Ratio[a]	PPV (95% CI)	Ratio[a]	PPV (95% CI)
Overall	7/11	64 (0.35–0.86)	29/31	94 (0.81–0.99)	21/21	100 (0.89–1.00)
Remission induction chemotherapy[d]	[c]	[c]	14/14	100 (0.84–1.00)	21/21	100 (0.89–1.00)
High-dose cytarabine	[b]	[b]	8/9	89 (0.59–0.99)	[c]	[c]
Anthracycline+cytarabine	[b]	[b]	7/8	88 (0.55–0.99)	[c]	[c]
Allogenic stem cell transplantation						
Overall	28/29	97 (0.85–1.00)	38/38	100 (0.94–1.00)		
Myeloablative HSCT	11/11	100 (0.8–1.00)	18/18	100 (0.87–1.00)	[c]	[c]
Non-myeloablative HSCT	17/18	94 (0.77–0.99)	20/20	100 (0.88–1.00)	[c]	[c]
Immuno-modulating agents						
Overall	10/12	83(0.56–0.93)	23/33	70 (0.53–0.83)	8/10	80 (0.50–0.96)
Lenalidomid	[b]	[b]	5/7	71 (0.35–0.94)	[b]	[b]
Ciclosporine	[b]	[b]	18/26	69 (0.50–0.84)	[b]	[b]
Growth-factors						
Overall	120/135	89 (0.83–0.93)	58/75	77 (0.67–0.86)	81/90	90 (0.83–0.95)
Erythropoiesis-stimulating agents	43/51	84 (0.73–0.92)	17/25	68 (0.49–0.84)	20/24	83 (0.65–0.94)
Granulocyte-colony-stimulating-factors	34/41	83 (0.69–0.92)	17/25	68 (0.49–0.84)	29/34	85 (0.71–0.94)
Bone marrow examination	43/43	100 (0.94–1.00)	24/25	96 (0.83–1.00)	32/32	100 (0.93–1.00)

Abbreviations: CI, confidence interval; HSCT, allogeneic hematopoietic stem cell transplantation; PPV, positive predictive value.
[a]Number of correct codes divided by the total number of medical reviews.
[b]Data not shown due to small numbers.
[c]No procedure codes with this code.
[d]Covers any type of remission induction chemotherapy, *e.g* "3 + 10", "High-dose cytarabine", "MITO-FLAG" etc.

have biased the evaluation of the codes. We did, however, pre-specify the gold standards for the correct answers for each procedure code to minimize the impact of this issue. Second, in our study, we chose NPV and PPV as the measures of validity. Of importance, predictive values are correlated with the prevalence of a treatment. We could not calculate sensitivity or specificity as the data were sampled from the codes pertinent to the different treatment modalities of interest. However, as discussed previously, we have no reasons to believe that the sensitivity or specificity of the procedures investigated in our study, would be lower than in prior reports of cancer patients (>90%).[10,11] Third, we did not include a sample for each of the procedure codes for calculation of NPV, but our overall NPV of no treatment was high. Fourth, few of the examined procedure codes had a PPV of 100% indicating that the amount of patients receiving a certain treatment may be overestimated which could introduce misclassification bias. Using two or more procedure codes to ascertain patients receiving a specific treatment may increase the accuracy. Finally, our estimates from subgroup analyses should be interpreted with caution due to small numbers.

5 | CONCLUSION

The majority of the hematological procedure codes registered in the DNPR are suitable for identifying patients who receive a specific treatment including date of treatment. Thus, the DNRP constitutes an important data source for future epidemiological studies examining treatment of hematological cancers.

ORCID

Tine Bichel Lauritsen 🔟 https://orcid.org/0000-0002-1730-8533

REFERENCES

1. Schmidt M, Schmidt SAJ, Adelborg K, et al. The Danish health care system and epidemiological research: from health care contacts to database records. *Clin Epidemiol*. 2019;11:563-591.
2. Schmidt M, Pedersen L, Sørensen HT. The Danish civil registration system as a tool in epidemiology. *Eur J Epidemiol*. 2014;29(8): 541-549.
3. Arboe B, Josefsson P, Jørgensen J, et al. Danish national lymphoma registry. *Clin Epidemiol*. 2016;8:577-581.
4. Østgård LS, Nørgaard JM, Raaschou-Jensen KK, et al. The Danish National Acute Leukemia Registry. *Clin Epidemiol*. 2016;8:553-560.

5. da Cunha-Bang C, Geisler CH, Enggaard L, et al. The Danish National Chronic Lymphocytic Leukemia Registry. *Clin Epidemiol*. 2016;8:561-565.

6. Gimsing P, Holmström MO, Klausen TW, et al. The Danish National Multiple Myeloma Registry. *Clin Epidemiol*. 2016;8:583-587.

7. Lauritsen TB, Nørgaard JM, Grønbæk K, et al. The Danish myelodysplastic syndromes database: patient characteristics and validity of data records. *Clin Epidemiol*. 2021;13:439-451.

8. Ostgård LS, Nørgaard JM, Severinsen MT, et al. Data quality in the Danish National Acute Leukemia Registry: a hematological data resource. *Clin Epidemiol*. 2013;5:335-344.

9. Schmidt M, Schmidt SA, 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.

10. Broe MO, Jensen PB, Mattsson TO, Pottegård A. Validity of antineoplastic procedure codes in the Danish National Patient Registry: the case of colorectal cancer. *Epidemiology*. 2020;31(4):599-603.

11. Lund JL, 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.

12. Brown LD, Cai TT, DasGupta A. Interval estimation for a binomial proportion. *Stat Sci*. 2001;16(2):101-133.

13. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. *J Biomed Inform*. 2019;95:103208.

14. Chubak J, Pocobelli G, Weiss NS. Tradeoffs between accuracy measures for electronic health care data algorithms. *J Clin Epidemiol*. 2012;65(3):343-9.e2.