Clinical Epidemiology

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

High Validity of the Danish National Patient Registry for Systemic Anticancer Treatment Registration from 2009 to 2019

Charles Vesteghem[1–3]
Rasmus Froberg Brøndum[1–3]
Ursula G Falkmer[1,3,4]
Anton Pottegård[5]
Laurids Østergaard Poulsen[1,3,4]
Martin Bøgsted[1–3]

[1]Department of Clinical Medicine, Aalborg University, Aalborg, Denmark; [2]Department of Hematology, Aalborg University Hospital, Aalborg, Denmark; [3]Clinical Cancer Research Centre, Aalborg University Hospital, Aalborg, Denmark; [4]Department of Oncology, Aalborg University Hospital, Aalborg, Denmark; [5]Department of Public Health, University of Southern Denmark, Odense, Denmark

Background: The Danish National Patient Registry is a major resource for Danish epidemiology. Only a few studies have been conducted to check the validity of the reporting of systemic anticancer treatments. In this study, we assessed this validity for a range of cancer types over a long period of time.

Patients and Methods: We extracted systemic anticancer treatment procedures from the Danish National Patient Registry for patients with solid malignant tumors treated at the Department of Oncology at Aalborg University Hospital between 2009 and 2019 (12,014 patients with 215,293 drug records). These data were compared to records obtained from the antineoplastic prescription database used at the department. We estimated the sensitivity, positive predictive value (PPV), and F1-score defined as the harmonic mean of the sensitivity and the PPV.

Results: There was an overall high concordance between the two datasets with a sensitivity and a PPV >92%. Treatments for brain, ovarian and endometrial cancers displayed lower concordance (81–89%). The validity was stable over the study period, with a slight drop during 2016–2017. Most drugs had a high validity with F1-scores above 90%. Fluorouracil, gemcitabine, pemetrexed, pembrolizumab, and nivolumab had F1-scores above 97%. Drugs that were introduced in the study period, such as lapatinib, palbociclib, erlotinib, pertuzumab, and panitumumab, yielded lower F1-scores due to the absence of specific registry codes early after introduction.

Conclusion: The Danish National Patient Registry can be used to reliably obtain information about systemic anticancer treatments, keeping in mind limitations for recently introduced drugs and for some types of cancer.

Keywords: antineoplastic agents, registries, Danish National Patient Registry, epidemiology, sensitivity and specificity, validity

Background

Nordic countries have extensive nationwide healthcare registries.[1] These registries are notably used for epidemiological studies.[2] One of the main data sources used to conduct these studies is the Danish National Patient Registry (DNPR) which has been shown to have a high validity for cancer diagnoses.[3] While most of these studies use the diagnoses recorded in the DNPR to analyze patients' trajectories,[4,5] other types of data are available, such as treatment procedure codes. It is of special interest in oncology to study for example the real-world efficacy of systemic anticancer treatments.[6] However, one of the main concerns of studies using the

Correspondence: Charles Vesteghem
Department of Clinical Medicine, Aalborg University, Søndre Skovvej 15, Aalborg, 9000, Denmark
Tel +45 97 66 38 72
Fax +45 97 66 63 23
Email charles.vesteghem@rn.dk

Received: 13 August 2021
Accepted: 28 October 2021
Published: 24 November 2021

DNPR data is the validity of the registration. Some work has already been published to address this concern for these treatments,[7,8] reporting high validity in terms of positive predictive value and sensitivity, but these studies were focused on colorectal cancers and included less than 500 patients. Thus, it remains unknown whether this high validity could be extrapolated to other solid malignant tumor types.

The aim of this study was to investigate the validity, using the same metrics, of systemic anticancer treatment procedure registration over a wide range of solid malignancies and over a long period of time.

Patients and Methods

A retrospective cohort study was conducted on patients with solid malignant tumors treated in the North Denmark Region.

Data Sources

The DNPR is encoded using the Danish Health Care Classification System (SKS)[9] and was used to obtain primary diagnoses and procedure information for both in- and outpatients containing the patient identifier, the admission and discharge dates, and the diagnosis or procedure code. For category-level diagnoses, the SKS encoding is identical to the ICD-10 classification.[10]

The second main data source was the database from the ARIA OIS for Medical Oncology v13.7 prescription software[11] (MedOnc) used at the Department of Oncology, Aalborg University Hospital. The corresponding data include the patient identifier, the start of treatment date, the duration, the drug name, and the dose given for each prescription and are only available for patients treated in the Region North Denmark. The MedOnc dataset was used as the gold standard to evaluate the validity of the DNPR dataset.

Data Extraction

Our focus is on anti-neoplastic agents as defined by the Anatomical Therapeutic Chemical (ATC) classification,[12] ie, drugs with an ATC code starting with "L01". These drugs are referred to here as L01 drugs. The corresponding data were extracted from the DNPR using SKS codes looking at the procedures: "Special medical treatments and treatment principles" (codes starting with "BWH") and "Treatment with antibodies and immunomodulatory therapy" (codes starting with "BOHJ"). These procedures were mapped to ATC codes. Procedures corresponding to drug combinations, ie, multiple ATC codes, in the DNPR data were split into individual drug entries. Drugs administered over consecutive days were grouped into one drug entry with a duration equal to the number of consecutive days. These drug entries are referred to here as drug cycles (see Figure 1).

For MedOnc, the drug names were mapped to ATC codes. The MedOnc prescriptions with no dose given, corresponding to non-administered treatments, were removed from the dataset. The drug entries were grouped in drug cycles, where applicable, in a similar manner to the DNPR dataset.

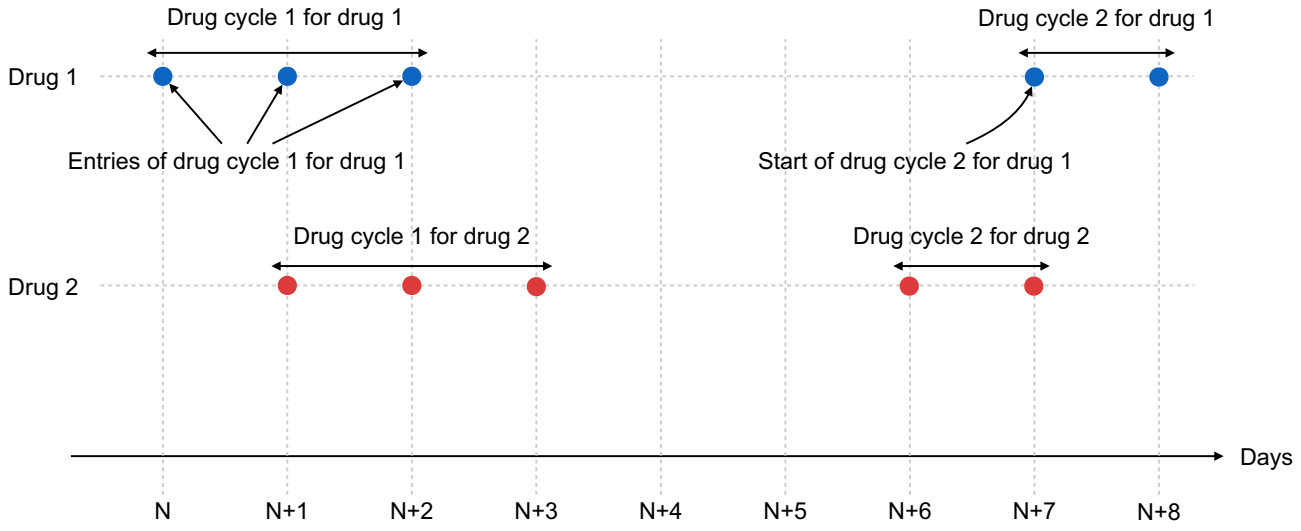

Figure 1 Grouping of drug entries into drug cycles.

Inclusion Criteria

The patients included in this study were identified using the cancer diagnosis codes (ICD-10 codes starting with C) found in the DNPR data as primary diagnosis. The diagnoses were grouped into common cancer types (see Supplementary Table 1). Only patients with a listed cancer type and at least one L01 drug cycle record in either the DNPR or MedOnc were included (see Supplementary Figure 1).

For the DNPR, we considered only L01 drug cycles from procedures performed at the Department of Oncology, Aalborg University Hospital between 2009 and 2019 (11 years). These data cover all systemic anticancer treatments given in the North Denmark Region. For MedOnc, we similarly only considered L01 drug cycles given over the same period.

In Denmark, each citizen is assigned an ID number from the Danish Civil Registration System.[13] The data sets were pseudonymized and linked at the patient level using an encoded version of this number.

Analysis

The comparisons of the two datasets were performed both for patients and for L01 drug cycles. For the patients, matching was performed using the patient identifier and the analyses were stratified by diagnosis. For L01 drug cycles, the ATC code and the start of treatment date were additionally considered for matching and the analyses were stratified by diagnosis, year, and drug.

Following an approach similar to Broe et al[8] the concordance of the datasets was measured using the positive predictive value (PPV) and the sensitivity. The MedOnc data were the gold standard, and the DNPR dataset was the predictive dataset. PPV was defined as the ratio of drug cycles in the intersection between both datasets and in the DNPR dataset, and the sensitivity was defined as the ratio of drug cycles in the intersection between both datasets and in the MedOnc dataset. Additionally, the F_1 score, defined as the harmonic mean of the PPV and sensitivity, was also used as an overall metric for concordance. As a sensitivity analysis, we considered a margin of 1 day for matching on the start date, as used by Broe et al.[8]

The data management and statistical analyses were performed using SAS Enterprise Guide 8.3 (SAS Institute Inc., Cary, NC, USA) and Python 3.8 in Jupyter notebooks,[14] respectively.

Table 1 Study Population Characteristics

Category	Variable	Count	Ratio
Overall	Patients	12,155	100%
Sex	Male	5113	42%
	Female	7042	58%
Age at diagnosis	18–44	878	7%
	45–59	3617	30%
	60–74	6089	50%
	75+	1571	13%
Cancer Diagnosis	Brain	462	4%
	Lung	2621	22%
	Breast	2968	24%
	Gastro-esophageal	620	5%
	Pancreatic	573	5%
	Colorectal	2306	19%
	Ovarian	557	5%
	Endometrial	226	3%
	Prostatic	514	4%
	Urinary	284	2%
	Other	1024	7%

Ethical Approval and Study Registration

According to Danish legislation, ethical approval and patient consent for purely registry-based projects is not required, only registration at the data responsible host institution is needed. The study protocol was registered in the North Denmark Region's research project inventory under the number 2019–41 and thereby complies with relevant data protection and privacy regulations.

Results

Study Population

This study included patients with a broad range of solid malignant tumors, the largest groups being lung, breast, and colorectal cancers, representing two-thirds of the cohort (see Table 1). Female patients accounted for the majority of the patients (58%). Ninety-three percent of the patients were >45 years old at diagnosis.

Matching Patients and Drug Cycles

Almost all patients are present in the intersection between MedOnc and the DNPR, which translates into a large concordance between the two datasets at the patient level, with a PPV and a sensitivity of 98.8% and 98.4%, respectively (see Table 2). However, the matching of brain tumor patients led to a lower sensitivity of 90%.

Table 2 PPV, Sensitivity, and F1-Score for Patients and L01 Drug Cycles per Diagnosis

Cancer Diagnosis	Type	MedOnc	DNPR	Intersection	PPV	Sensitivity	F1-Score
Overall	Patients	12,014	11,965	11,824	98.8%	98.4%	98.6%
	Drug cycles	215,293	216,074	198,888	92.0%	92.4%	92.2%
	With a 1-day margin	215,293	216,074	200,301	92.7%	93.0%	92.9%
Brain	Patients	440	419	397	94.7%	90.2%	92.4%
	Drug cycles	6671	6804	5546	81.5%	83.1%	82.3%
	With a 1-day margin	6671	6804	5610	82.5%	84.1%	83.3%
Lung	Patients	2613	2571	2563	99.7%	98.1%	98.9%
	Drug cycles	34,628	33,240	31,774	95.6%	91.8%	93.6%
	With a 1-day margin	34,628	33,240	32,066	96.5%	92.6%	94.5%
Breast	Patients	2937	2953	2922	99.0%	99.5%	99.2%
	Drug cycles	62,637	63,644	57,198	89.9%	91.3%	90.6%
	With a 1-day margin	62,637	63,644	57,582	90.5%	91.9%	91.2%
Gastro-oesophageal	Patients	619	611	610	99.8%	98.5%	99.2%
	Drug cycles	10,811	10,558	9863	93.4%	91.2%	92.3%
	With a 1-day margin	10,811	10,558	9998	94.7%	92.5%	93.6%
Pancreatic	Patients	573	565	565	100.0%	98.6%	99.3%
	Drug cycles	11,255	11,086	10,693	96.5%	95.0%	95.7%
	With a 1-day margin	11,255	11,086	10,726	96.8%	95.3%	96.0%
Colorectal	Patients	2255	2296	2245	97.8%	99.6%	98.7%
	Drug cycles	55,580	56,928	53,434	93.9%	96.1%	95.0%
	With a 1-day margin	55,580	56,928	53,688	94.3%	96.6%	95.4%
Ovarian	Patients	554	544	541	99.4%	97.7%	98.5%
	Drug cycles	11,901	12,252	10,586	86.4%	89.0%	87.7%
	With a 1-day margin	11,901	12,252	10,681	87.2%	89.7%	88.4%
Endometrial	Patients	225	222	221	99.5%	998.2%	98.8%
	Drug cycles	2800	3003	2511	83.6%	89.7%	86.5%
	With a 1-day margin	2800	3003	2527	84.1%	90.2%	87.1%
Prostatic	Patients	509	512	507	99.0%	99.6%	99.3%
	Drug cycles	3766	3861	3589	93.0%	95.3%	94.1%
	With a 1-day margin	3766	3861	3600	93.2%	95.6%	94.4%
Urinary	Patients	283	275	274	99.6%	96.8%	98.2%
	Drug cycles	3562	3541	3331	94.1%	93.5%	93.8%
	With a 1-day margin	3562	3541	3344	94.4%	93.9%	94.2%
Other	Patients	1006	997	979	98.2%	97.3%	97.8%
	Drug cycles	11,682	11,157	10,363	92.9%	88.7%	90.7%
	With a 1-day margin	11,682	11,157	10,479	93.9%	89.7%	91.8%

Notes: Patients are matched on encrypted CPR number, drug cycles on start date and ATC code. The 1-day margin is on the start date for drug cycles allowing additional matching if the start dates of unmatched drug cycles in MedOnc and the DNPR are 1 day or less from each other.

Matching the drug cycles using the patient identifier, the ATC code, and the start of treatment date generated a PPV and a sensitivity above 92%. Treatments within all diagnoses except brain, ovarian, and endometrial cancers have a sensitivity and a PPV above 89%, with treatments for pancreatic cancer above 95% (see Figure 2). Adding a 1-day margin for the start date improves the performance with a gain of 0.7% for PPV, 0.6% for sensitivity and 0.7% for F1-score.

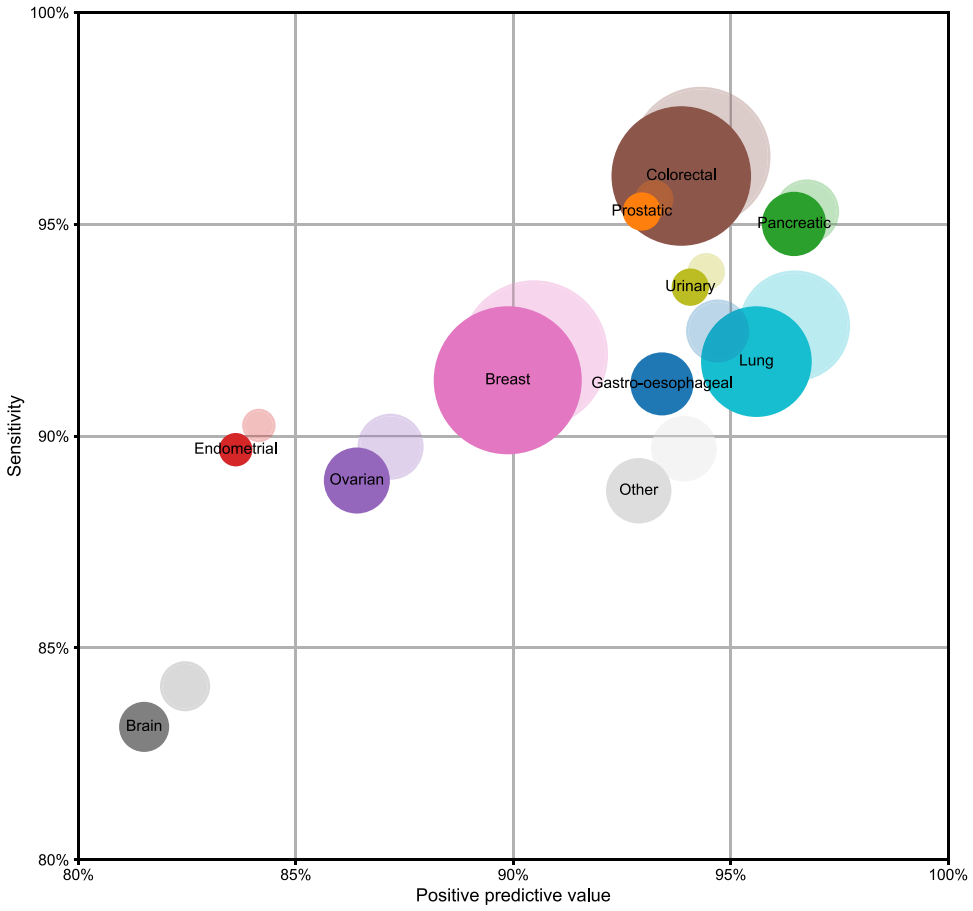

Figure 2 Positive predictive value vs sensitivity for the matching of drug cycles per cancer diagnosis. The area of the circle is proportional to the number of corresponding drug cycles. The lighter circles in the background correspond to the performances with a 1-day margin.

Evolution Over Time

The validity of the registered drug cycles is mostly stable over the 2009–2019 period (11 years) (see Figure 3). Nevertheless, a drop in PPV can be seen for 2016 and 2017. The sensitivity was also negatively impacted in 2012 and 2016. The effect of the 1-day margin, shown as lighter surfaces above both lines in Figure 3, seems to be stable over the period.

Validity per Drug

Looking at the most frequently administered drugs there is a more detailed picture, with most drugs having F1-scores above 90% (see Table 3). Some drugs (fluorouracil, gemcitabine, pemetrexed, pembrolizumab, and nivolumab) have high validity with F1-scores above 97%, while others (temozolomide, pertuzumab, palbociclib, erlotinib and lapatinib) have F1-scores below 80%. The low validity is typically due to a low sensitivity with values below 70%, ie, many entries in MedOnc cannot be matched with corresponding data in the DNPR (see Figure 4). As shown in Table 3, there is a strong correlation between drugs and diagnoses, for example temozolomide and cyclophosphamide are almost exclusively used for brain and breast cancer, respectively.

Discussion

Main Results

The DNPR data can be used as a good proxy for L01 drug cycles when matching the ATC code and start of treatment date. The reporting of drug cycles appears to be reliable across diagnoses, especially for colorectal and pancreatic cancers, but historically not for brain cancers, even though improvements have occurred. Looking at specific drugs, only a few have limited validity among frequently used drugs, including temozolomide.

Using the Start of Treatment Date Only

The duration of the cycle was not considered because the DNPR does not contain this information. However, in the context of a specific treatment for a specific cancer type,

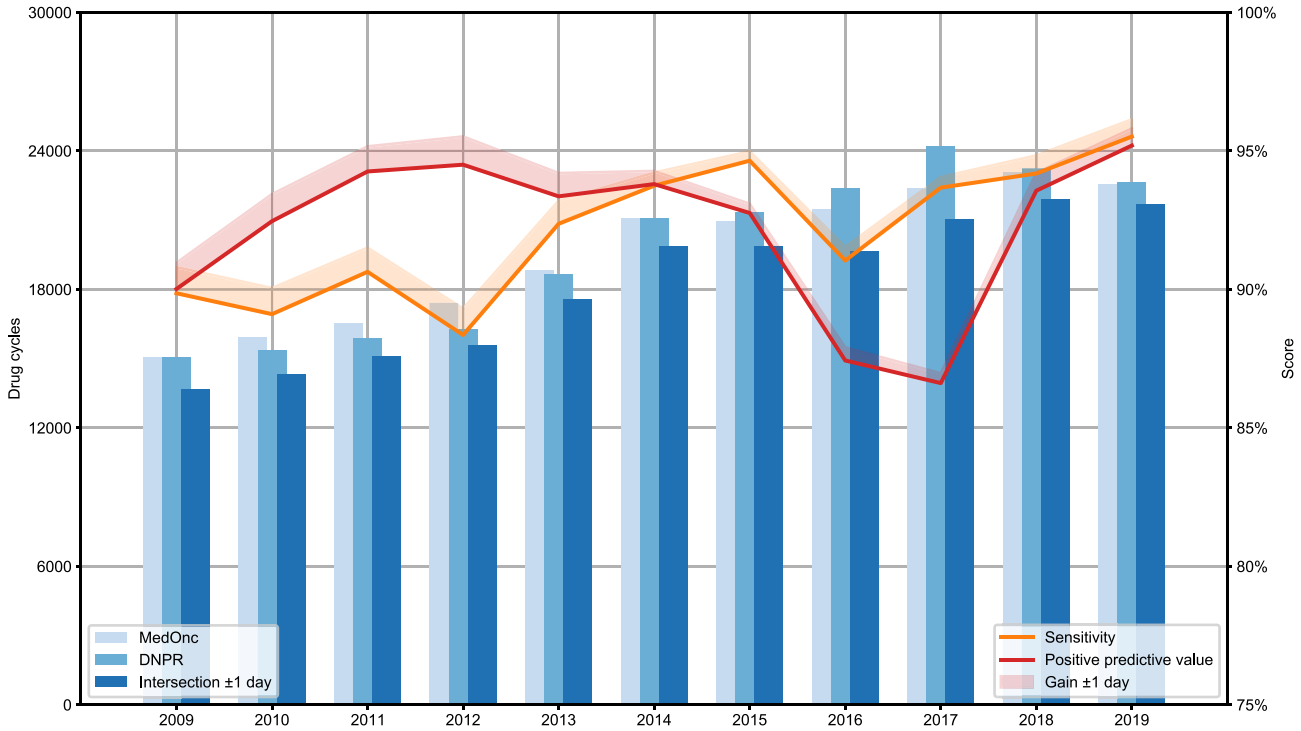

Figure 3 Evolution over time of the validity of the DNPR registrations for L01 drug cycles for systemic anticancer treatments. The lighter surface above each line represents the gain in performance by adding a 1-day margin.

the durations of cycles would be known, especially for adjuvant and neoadjuvant treatments and, to a lesser extent, for palliative treatments. Thus, the whole history of patients could be reconstructed, as a cycle is typically not stopped in the middle but instead cancelled or postponed altogether if the patient is not fit for it.

Temozolomide and Brain Cancer

Temozolomide cycles from the DNPR have a good PPV but a low sensitivity, ie, a significant proportion of these cycles do not seem to have been registered in the DNPR up to 2014 (see Figure 4). This is due to historically poor reporting in the DNPR by administrative personnel. This could be explained by the complexity of the treatment regimen used for glioblastoma[15] and thus point toward reporting issues at the diagnosis level. This poor reporting mechanically impacts the concordance at the patient level, as seen in Table 2.

Recent Drugs

Similar to temozolomide, other drugs, such as pertuzumab, palbociclib, erlotinib, lapatinib, and panitumumab, also display a good PPV with a low sensitivity but for a different reason. Indeed, these are recently introduced drugs for which specific national registry codes were not available when first used, leading to a suboptimal registration at the drug level. For example, pertuzumab was first used in 2012 according to the MedOnc dataset but was only registered in the DNPR with a specific code in 2015.

Cyclophosphamide and Epirubicin

Cyclophosphamide and epirubicin display a low PPV but a high sensitivity. This is due to an error in the registration in 2016 and 2017. These two drugs are administered to breast cancer patients in an adjuvant regimen composed of three cycles of these two drugs followed by three cycles of docetaxel. They were nevertheless registered in the DNPR as given for all six cycles until the registration error was discovered. This can also explain the drop in PPV seen for these years, since they are frequently used drugs to treat breast cancers which is the largest sub-cohort of the study and thus have a significant impact on the overall performance. Outside of these years, the performances are nevertheless good with sensitivities and PPVs above 90%.

Limitations and Strengths

Limitations

MedOnc was used as a reference, but some manual curation was nevertheless needed. We considered MedOnc to

Table 3 Matching Performances for Drug Cycles per Drug Type for Drugs with More Than 500 Cycles in MedOnc

ATC Code	Drug Name	Top 2 Diagnoses	MedOnc	DNPR	Matching	PPV	Sensitivity	F1-Score
L01BC02	Fluorouracil	Colorectal (19,508), Pancreatic (1552)	23,272	22,935	22,644	98.7%	97.3%	98.0%
L01CD01	Paclitaxel	Breast (12,263), Ovarian (2834)	19,001	17,793	17,208	96.7%	90.6%	93.5%
L01CA04	Vinorelbine	Lung (11,098), Breast (5751)	17,779	16,541	16,233	98.1%	91.3%	94.6%
L01XA02	Carboplatin	Lung (8794), Ovarian (3973)	16,099	15,956	15,446	96.8%	95.9%	96.4%
L01XX19	Irinotecan	Colorectal (10,924), Brain (1421)	14,981	15,051	14,500	96.3%	96.8%	96.6%
L01XC03	Trastuzumab	Breast (14,103), Gastro-esophageal (736)	15,094	14,491	14,011	96.7%	92.8%	94.7%
L01XC07	Bevacizumab	Colorectal (7140), Ovarian (2995)	13,924	13,703	13,209	96.4%	94.9%	95.6%
L01XA03	Oxaliplatin	Colorectal (8136), Gastro-esophageal (2826)	12,840	13,063	12,388	94.8%	96.5%	95.6%
L01BC06	Capecitabine	Colorectal (5163), Breast (2785)	11,173	12,394	10,460	84.4%	93.6%	88.8%
L01CD02	Docetaxel	Breast (5378), Prostatic (2703)	9964	9937	9581	96.4%	96.2%	96.3%
L01BC05	Gemcitabine	Pancreatic (5753), Urinary (1768)	9624	9475	9330	98.5%	96.9%	97.7%
L01DB03	Epirubicin	Breast (6925), Gastro-esophageal (1269)	8361	9951	8088	81.3%	96.7%	88.3%
L01AA01	Cyclophosphamide	Breast (6278), Lung (72)	6435	8053	6283	78.0%	97.6%	86.7%
L01XC06	Cetuximab	Colo-rectal (3058), Other (1415)	4754	4710	4483	95.2%	94.3%	94.7%
L01CB01	Etoposide	Lung (3894), Other (273)	4556	4467	4351	97.4%	95.5%	96.4%
L01XA01	Cisplatin	Other (1734), Lung (942)	3785	3652	3497	95.8%	92.4%	94.0%
L01BA04	Pemetrexed	Lung (1936), Other (266)	2244	2253	2216	98.4%	98.8%	98.6%
L01XC18	Pembrolizumab	Lung (1142), Urinary (265)	1688	1640	1623	99.0%	96.1%	97.5%
L01XC17	Nivolumab	Lung (1115), Other (420)	1616	1616	1600	99.0%	99.0%	99.0%
L01AX03	Temozolomide	Brain (2145), Other (78)	2318	1620	1495	92.3%	64.5%	75.9%
L01XX41	Eribulin	Breast (1475), Ovarian (6)	1481	1457	1408	96.6%	95.1%	95.8%
L01DB01	Doxorubicin	Ovarian (774), Endometrial (199)	1260	1246	1202	96.5%	95.4%	95.9%
L01XC13	Pertuzumab	Breast (1554), Other (6)	1560	1073	1035	96.5%	66.3%	78.6%
L01XE33	Palbociclib	Breast (1544), Endometrial (46)	1590	1109	998	90.0%	62.8%	74.0%
L01XX17	Topotecan	Lung (758), Ovarian (184)	1072	1056	979	92.7%	91.3%	92.0%
L01XC08	Panitumumab	Colo-rectal (1220), Ovarian (74)	1350	979	951	97.1%	70.4%	81.7%
L01XE03	Erlotinib	Lung (1223), Breast (87)	1334	1043	889	85.2%	66.6%	74.8%
L01XC14	Trastuzumab emtansine	Breast (746), Colo-rectal (3)	749	736	716	97.3%	95.6%	96.4%
L01CD04	Cabazitaxel	Prostatic (525), Other (3)	528	541	517	95.6%	97.9%	96.7%
L01XE07	Lapatinib	Breast (639), Colo-rectal (9)	648	391	336	85.9%	51.9%	64.7%

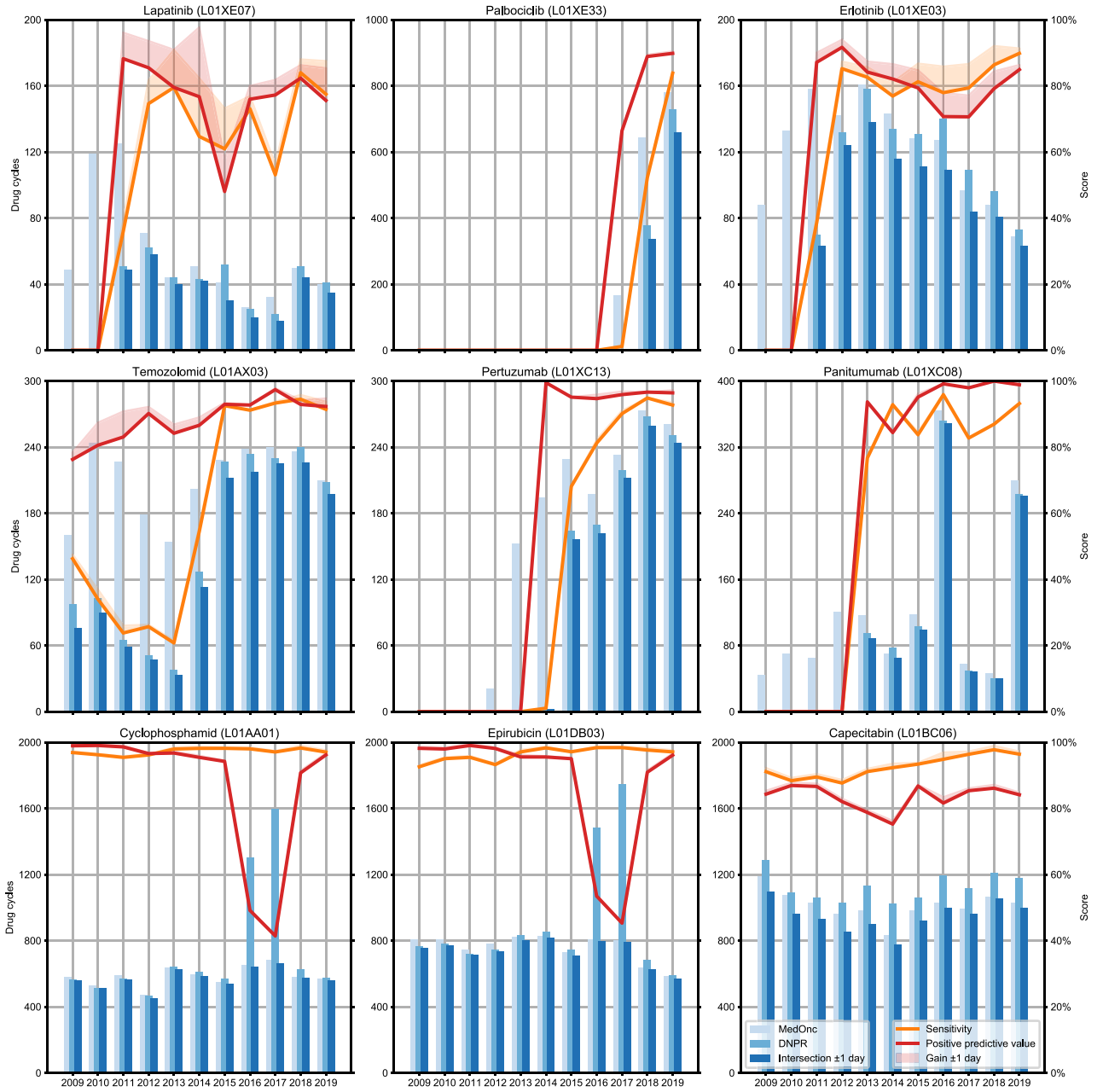

Figure 4 Evolution over time of the validity of the DNPR registrations for bottom 9 performing L01 drugs. Only drugs with more than 500 cycles were considered. The lighter surface above each line represents the gain in performance by adding a 1-day margin.

be a reliable source because it is used in clinical practice to plan, prescribe, and administer treatment; therefore, data entry is expected to be done by doctors and nurses with much more care than in the DNPR, which is an administrative tool filled in by secretaries. However, the DNPR is used for reimbursement of procedures which is a strong incentive to avoid underreporting in this system. The validity of MedOnc compared to patient journals remains unknown but is expected to be similar.

Also, the results shown here might be specific to the North Denmark Region since there might be some spatial and temporal differences across Denmark and Scandinavia in terms of clinical tools and reporting practices. Indeed, Broe et al have reported slight discrepancies between university hospitals and other hospitals,[8] but this study only included data from one university hospital.

We report issues in the DNPR data. However, these issues only affect a limited number of drugs and seem to

have been resolved in recent years. The fact that they are consistent with previously reported results suggests the generalizability of these results.

Strengths

The main strength of this study is its large time span and broad range of cancer diagnoses with low variability in the results, which should guarantee a high level of consistency in the data reported in the DNPR.

Comparison to Other Studies

Only a few articles[7,8] analyzing registration practices are available, and they focus exclusively on colorectal cancers with much smaller cohorts. Broe et al's work[8] is the more directly comparable with ours. For individual drug cycles to colorectal cancer patients, we report a PPV of 94% and a sensitivity of 97% compared to a PPV of 95% and a sensitivity of 90% in Broe et al's study, illustrating the reliability of the MedOnc dataset. Lund et al's study,[7] similarly to our work, reports high validity of the DNPR for fluorouracil, oxaliplatin, and bevacizumab.

Conclusions

This study confirms the validity of the registration of DNPR drug cycles for a large variety of cancer types and antineoplastic drugs, with some limitations for brain cancer and recently introduced drugs. Identified reporting issues, notably for temozolomide, cyclophosphamide, and epirubicin, seem to have been resolved in the latter years of the study period. Therefore, these data can be used for retrospective studies on antineoplastic agent usage across the country.

Acknowledgments

We would like to thank System Administrator of MedOnc, Annette Juul Madsen and Special Consultant Thomas Mulvad Larsen for their help in obtaining and understanding the data needed for this study.

Disclosure

This work was supported by grants from Department of Oncology, Aalborg University Hospital, The Regional Research Fund of North Denmark Region, and from "Det Obelske Familie Fond", no. 50.62 to Ursula G Falkmer. The authors report no other conflicts of interest in this work.

References

1. Furu K, Wettermark B, Andersen M, Martikainen JE, Almarsdottir AB, Sørensen HT. The Nordic countries as a cohort for pharmacoepidemiological research. *Basic Clin Pharmacol Toxicol*. 2010;106(2):86–94. doi:10.1111/j.1742-7843.2009.00494.x
2. 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. doi:10.2147/CLEP.S91125
3. Thygesen SK, Christiansen CF, Lash TL, Christensen S, Sørensen HT. Predictive value of coding of diagnoses in the Charlson comorbidity index in the Danish national registry of patients. *BMC Med Res Methodol*. 2011;11(83):2–7. doi:10.1186/1471-2288-11-83
4. Beck MK, Westergaard D, Jensen AB, Groop L, Brunak S. Temporal order of disease pairs affects subsequent disease trajectories: the case of diabetes and sleep apnea. *Biocomput*. 2017;2017:380–389. doi:10.1142/9789813207813_0036
5. Beck MK, Jensen AB, Nielsen AB, Perner A, Moseley PL, Brunak S. Diagnosis trajectories of prior multi-morbidity predict sepsis mortality. *Sci Rep*. 2016;6(July):1–9. doi:10.1038/srep36624
6. Skau Rasmussen L, Vittrup B, Ladekarl M, et al. The effect of postoperative gemcitabine on overall survival in patients with resected pancreatic cancer: a nationwide population-based Danish register study. *Acta Oncol*. 2019;58(6):864–871. doi:10.1080/0284186X.2019.1581374
7. 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(1):327–334. doi:10.2147/CLEP.S49773
8. 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. doi:10.1097/EDE.0000000000001208
9. Sundhedsdatastyrelsen. Disease classification system - SKS (in Danish). Available from: https://sundhedsdatastyrelsen.dk/da/rammer-og-retningslinjer/om-klassifikationer/sks-klassifikationer/klassifikation-sygdomme. Accessed March 3, 2021.
10. World Health Organization. ICD-10 Version:2016; 2016. Available from: https://icd.who.int/browse10/2016/en. Accessed March 26, 2020.
11. Varian Medical Systems Inc. ARIA OIS for medical oncology. Available from: https://www.siemens-healthineers.com/dk/radiotherapy/software-solutions/ariaois. Accessed November 17, 2021.
12. WHO Collaborating Centre for Drug Statistics Methodology. Anatomical Therapeutic Chemical (ATC) classification system. Available from: https://www.whocc.no/atc/structure_and_principles/. Accessed March 3, 2021.
13. Schmidt M, Pedersen L, Sørensen HT. The Danish Civil Registration System as a tool in epidemiology. *Eur J Epidemiol*. 2014;29(8):541–549. doi:10.1007/s10654-014-9930-3
14. Kluyver T, Ragan-Kelley B, Pérez F, et al. Jupyter Notebooks—a publishing format for reproducible computational workflows. Position Power Acad Publ Play Agents Agendas - Proc 20th Int Conf Electron Publ ELPUB. 2016;2016:87–90. doi:10.3233/978-1-61499-649-1-87.
15. Stupp R, Mason WP, van den Bent MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med*. 2005;352(10):987–996. doi:10.1056/NEJMoa043330