ORIGINAL RESEARCH

Registrations of Patients with Renal Cell Carcinoma in the Nationwide Danish Renal Cancer Database versus the Danish Cancer Registry: Data Quality, Completeness and Survival (DaRenCa Study-3)

Bolette Danckert[1]
Trine Allerslev Horsbøl[2]
Ole Andersen [iD][1]
Susanne Oksbjerg Dalton [iD][2,3]
Jane Christensen[4]
Margit Rasted[5]
Astrid Petersen[6]
Mette Nørgaard [iD][7]
Nessn Azawi [iD][8]
Lars Lund[9]
Frede Donskov [iD][10]

[1]Director´s Office, Danish Cancer Society Research Center, Copenhagen, Denmark; [2]Survivorship and Inequality in Cancer, Danish Cancer Society Research Center, Copenhagen, Denmark; [3]Department of Clinical Oncology & Palliative Care, Zealand University Hospital, Næstved, Denmark; [4]Statistics and Data Analysis, Danish Cancer Society Research Center, Copenhagen, Denmark; [5]Danish Health Data Authority, Copenhagen, Denmark; [6]Department of Pathology, Aalborg University Hospital, Aalborg, Denmark; [7]Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark; [8]Department of Urology, Zealand University Hospital, Roskilde, Denmark; [9]Department of Urology and Department of Clinical Research, University of Southern Denmark, Odense, Denmark; [10]Department of Oncology, Aarhus University Hospital, Aarhus, Denmark

Correspondence: Frede Donskov
Department of Oncology, Aarhus University Hospital, Palle Juul-Jensens Boulevard 99, Aarhus N 8200, Denmark
Tel +45 4046 5234
Email frede.donskov@auh.rm.dk

Background: The Danish multidisciplinary renal cancer group (DaRenCa) established the nationwide database DaRenCaData in 2010. The Danish Cancer Registry (DCR) has been considered the golden standard. In contrast to DCR, DaRenCaData required the diagnosis to be histologically or cytologically verified. DaRenCaData and DCR have not previously been compared.

Patients and Methods: We included patients with renal cell carcinoma registered in DaRenCaData and/or DCR from August 1st 2010 to December 31st 2015. We computed completeness and positive predictive value (PPV) of a diagnosis in DaRenCaData compared with DCR, 1-year, 3-year and 5-year mortality rate ratios, and relative survival.

Results: We identified 4890 patients in the two registries. Of these, 4326 were registered in DaRenCaData and 4714 in DCR. Completeness of DaRenCaData was 88% [95% CI, 87–89%] and increased during the period from 82% to 94%. The PPV was 96% [95% CI, 95–97%]. A total of 4150 patients (85%) were found in both registries, 4% (176 patients) in DaRenCaData only, and 12% (564 patients) in DCR only. The relative survival was higher for patients in DaRenCaData vs DCR; the 1-year and 5-year relative survival was 85% vs 81% and 65% vs 59%, respectively. Compared with patients registered in both registries, the mortality rates were higher in patients registered in DaRenCaData only (1-year hazard ratio (HR)=2.84 [95% CI, 2.20–3.68]) or DCR only (1-year HR=4.29 [95% CI, 3.72–4.93]). Observed in both registries, survival improved over time with a 7% yearly reduction in death based on estimations of 1-year mortality rate ratios.

Conclusion: DaRenCaData had high and increasing completeness and high PPV, establishing it as a high-quality research database. Observed in both registries, renal cell carcinoma mortality declined over time; patients only registered in DCR or DaRenCaData had poorer outcomes. This study points to the importance of assessing the inclusion criteria when interpreting registry-based studies.

Keywords: kidney cancer, national registries, survival, assessment of database

Introduction

Denmark has several population-based clinical databases comprising routinely collected health-related data.[1,2] Among these databases, the multidisciplinary Danish Renal Cancer Group (DaRenCa) established a nationwide database (DaRenCaData) in 2010 with the primary objective to monitor the clinical quality of renal cell carcinoma diagnostics and treatment in Denmark.[3] So far, the data

quality of DaRenCaData has not been formally assessed. The Danish Cancer Registry (DCR) is a well-established nationwide registry that has monitored incidence and survival of all Danish cancer cases since 1943.[4] Both registries contain details of new cases obtained through automatized linkage between national registries, ie, the Danish National Patient Register (DNPR), the Danish Pathology Register (DPR), and the Civil Registration System; however, inclusion criteria vary slightly. DaRenCaData focuses only on histologically or cytologically verified kidney cancer in patients with a diagnosis in the hospital system; the registry contains data on diagnosis, treatments, outcomes, and a few variables registered online in a web-based form by the treating clinicians.[3] The DCR contains data on cancer diagnosis and outcomes; the registry additionally receives notifications from the National Causes of Death Registry and from general practitioners and practising specialists outside the hospitals.[4] The existence of two nationwide Danish registries on renal cell carcinoma made it possible to assess the data quality in DaRenCaData in comparison to DCR. To perform an evaluation of registration overlap and impact of differences in inclusion criteria on survival measures is crucial for adequate use and understanding of research employing either of the two data sources.[5–7]

The present study was established as a collaboration between the Danish Cancer Registry, DaRenCa, and the Danish Cancer Society Research Center and aimed to: 1) assess the quality of DaRenCaData by completeness of data and the positive predictive value (PPV) compared with the DCR, 2) identify discrepancies between DaRenCaData and DCR, and 3) quantify the impact of potential discrepancies on the survival estimates. Thereby, the study may contribute to the growing body of literature regarding assessments of the quality of data from clinical databases[5–7] presenting DaRenCaData as a data source, which can be useful for future clinical research.

Patients and Methods
Data Sources
DaRenCaData included all persons with a first-time diagnosis of renal cancer in Denmark since August 2010. The renal cell carcinoma diagnosis was required to be histologically or cytologically verified. Patients were identified if they had a first-time registration of a renal cancer in the DPR, which contains information using SNOMED codes on all samples examined at Danish departments of

pathology.[8] Patients were included in DaRenCaData if they had either (a) a code indicating a tumour in the kidney (SNOMED code "T71*") followed immediately on the same material by a code indicating a malignant carcinoma (an M-code ending on "3" in the interval M80103-M958*3 – excluding nephroblastoma (M89603) and urothelial carcinoma (M81203)) – or (b) a code indicating a metastasis from a renal tumour (code "ÆF4510" unless the ÆF4510-code was proceeded by "M81206"). Patients were not included in DaRenCaData if the diagnosis was tentative (if the relevant M-code or ÆF4510-code was followed by "ÆYYY00"), or if the cancer was a relapse. The diagnosis was considered a relapse if the patient had a prior diagnosis of renal cell carcinoma identified either from notes in DaRenCaData or from registrations in DPR or in the DNPR, which recorded information on diagnoses and treatment in Danish somatic hospitals since 1977.[9] DaRenCaData did not include patients whose diagnoses were solely recorded from death certificates (death certificate only; DCOs). The date of diagnosis in DaRenCaData was defined as the date the biopsy or surgical treatment was performed.

DCR is a nationwide cancer registry that was established in 1943, administered by the Danish Health Data Authority. It is considered to hold data for all incident cancer cases in Denmark, including information on tumour characteristics, eg, ICD10 codes, topology, morphology, laterality, stage, grade, and date of diagnosis.[4] For DCR, the primary data sources were registrations in DNPR and death certificates. DCR considered all patients diagnosed with ICD-10 code C64 as patients with renal cell carcinoma. Following identification of a patient in DNPR, DCR looked for supplementary information from DPR; however, the diagnosis did not have to be histologically or cytologically verified. Thus, DCR may identify patients with renal cell carcinoma if they were only diagnosed by imaging and not by biopsy or fine needle aspiration, and as such not found in DPR. The date of diagnosis in DCR was defined as the date of the first hospital contact where a diagnosis of renal cancer was registered in DNPR.

Since 1968 all residents in Denmark have been equipped with a unique individual ten digits personal code (a CPR-number). This number, unique to each Danish resident, is used in all Danish registries, allowing unambiguous individual-level data linkage. From the Civil Registration System, we obtained information on vital status (dead or alive), date of death, and residence for all cancer patients.[10]

This study was approved by the Danish Data Protection Agency (2015-41-3726) and locally at the Danish Cancer Society Research Center (2019-DCRC-0059). According to Danish legislation, this registry-based study did not require further ethical approval because it did not involve any patient contact or intervention.

Patient Cohort

The present study included all patients with renal cell carcinoma who were permanent residents in Denmark and registered in DCR or DaRenCaData during the period August 1st 2010 to December 31st 2015. The study period ended on December 31st 2015, because a new IT system was introduced in two Danish regions during 2016 and 2017 which may have caused irregularities in registrations. A calendar year for DaRenCaData spanned from August 1st to July 31th the subsequent year.

Statistical Methods

The completeness of data in the DaRenCaData relative to DCR was estimated, indicating the number of patients with renal cell carcinoma in DCR who were also registered in DaRenCaData.[11] Furthermore, the positive predictive value (PPV) of the data in DaRenCaData was calculated and defined as the number of patients in DaRenCaData who were also registered with renal cell carcinoma in DCR.[11] Registrations in DCR and DaRenCaData were considered identical if the patients had the same CPR-number and were recorded in both registries with no more than 120 days between the registered dates of diagnosis.

Data in the two registries were merged, and agreement in the data sets was assessed by inspecting the size of the intersection (denoting patients found in both registries), while differences were assessed by inspecting the DaRenCaData set difference (observations only found in DaRenCaData) and the DCR set difference (observations only found in DCR). For patients in the set differences, it was investigated whether the patient appeared with a registered first incidence of renal cancer in either DNPR or DPR (or both) within 120 days from the date of diagnosis recorded in the set difference.

Differences in survival and mortality between patients in both registries, in the intersection, and the set differences were investigated by estimating 1-, 3- and 5-year survival relative to survival of the background population, and 1-, 3- and 5-year mortality rate ratios (MRRs) for which the end of follow-up was December 31st 2017.

Relative survival (RS) estimates were obtained based on the Ederer II weighing method.[12] The estimates were adjusted for age using the International Cancer Survival Standard population weighs (specifically ICSS1).[13] However, contrary to the ICSS standard, a few patients younger than age 15 were included in DCR and DaRenCaData, implying that the age groups were 0–44, 45–64, 65–74 and 75+ years old. The estimates of expected survival were based on population mortality rates stratified by sex, age, and calendar time (in 1-year intervals up to an age at 98 years after which all observations were grouped to age 99+ years). Note that the sizes of 1-, 3- and 5-year relative survival estimates should not be compared to each other since they were not based on the same sample (1-year follow up was available for all patients which was not the case for all patients with regard to 3- and 5-year follow up).

MRRs were calculated as hazard ratios (HRs) using Cox Proportional Hazard models for which the underlying timescale was time since diagnosis. The estimations included sex and age at the time of diagnosis as strata (age was grouped similarly to the relative survival estimations), and adjustments were made for T-stage, N-stage, M-stage, and combined TNM-stage. The proportional hazard assumptions were checked, and did not give rise to any concern. A total of 38 observations were excluded in the survival and mortality analyses, because the dates of death preceded the dates of diagnosis in the registration. Kaplan-Meier curves were generated to visualise survival across different groups.

For all statistical analyses except relative survival, SAS statistical software was used. For relative survival, the statistical software RStudio was used (specifically the package "popEpi").[14]

Results
Patient Cohort

Between August 1st 2010 and December 31st 2015, a total of 4890 patients with renal cell carcinoma were registered in the two registries; 4326 patients were recorded in DaRenCaData, and 4714 patients were identified in DCR (6 registrations in DaRenCaData were excluded because they lacked a valid CPR-number).

Of the 4326 patients registered in DaRenCaData, 4150 were registered in DCR as well, yielding a positive predictive value of being registered in DaRenCaData of 96% [95% CI, 95–97%]. Of the 4714 patients registered in

DCR, 4150 were captured in DaRenCaData as well, giving a completeness of DaRenCaData at 88% [95% CI, 87–89%] (Table 1).

The number of patients per year captured in both registries increased from 649 patients (82%) in 2010–2011 to 865 patients (94%) in 2014–2015. In all, 564 patients (12%) were found only in the DCR, and this number decreased over time from 145 patients in 2010–2011 to 57 patients in 2014–2015. Patients only found in DaRenCaData comprised 4% (n=176), and the number per year was rather stable during the study period.

Patients Found Only in DaRenCaData

Of the 176 patients registered only in the DaRenCaData (Table 1), 24 (14%) patients were also registered in DCR, but before or after the specified study time period. Of the 176 patients, 148 (84%) patients were registered in DPR with a diagnosis of renal cell carcinoma, and therefore seemed to be candidates also for inclusion in DCR. For 4 (<0.1%) patients registered only in DaRenCaData, there was no renal cancer diagnosis recorded based on the SNOMED codes in DPR with a date of diagnosis within 120 days before or after the DaRenCaData date of diagnosis. The source of these records was unknown.

Patients Found Only in DCR

Of the 564 patients registered only in DCR (Table 1), 16 (3%) patients were also registered in DaRenCaData, but after the study closure time December 31st 2015. Nineteen patients, corresponding to 0.4% of all patients in DCR, were likely to be DCOs, as the registered date of diagnosis was identical to the registered date of death. For 446 (79%) patients, a first incident renal cell carcinoma was

recorded in DNPR within 120 days between the DCR and DNPR dates of diagnosis, but the diagnosis was not verified histologically or cytologically as no record appeared in the DPR. For 45 (8%) patients, DPR confirmed renal cell carcinoma, thus fulfilling the DaRenCaData inclusion criteria. For the remaining 38 (7%) patients registered only in DCR, no coding of the first instance of renal cancer diagnosis was recorded neither in DNPR nor in DPR within 120 days from the DCR dates of diagnosis.

Patients registered in DCR only were older (median; 5–95 percentiles), with age 74 years (48–90) compared with both registries 66 years (45–82), or DaRenCaData only 68 years (49–84). There was no significant difference in sex distribution between patients in DCR only, DaRenCaData only, or both registries.

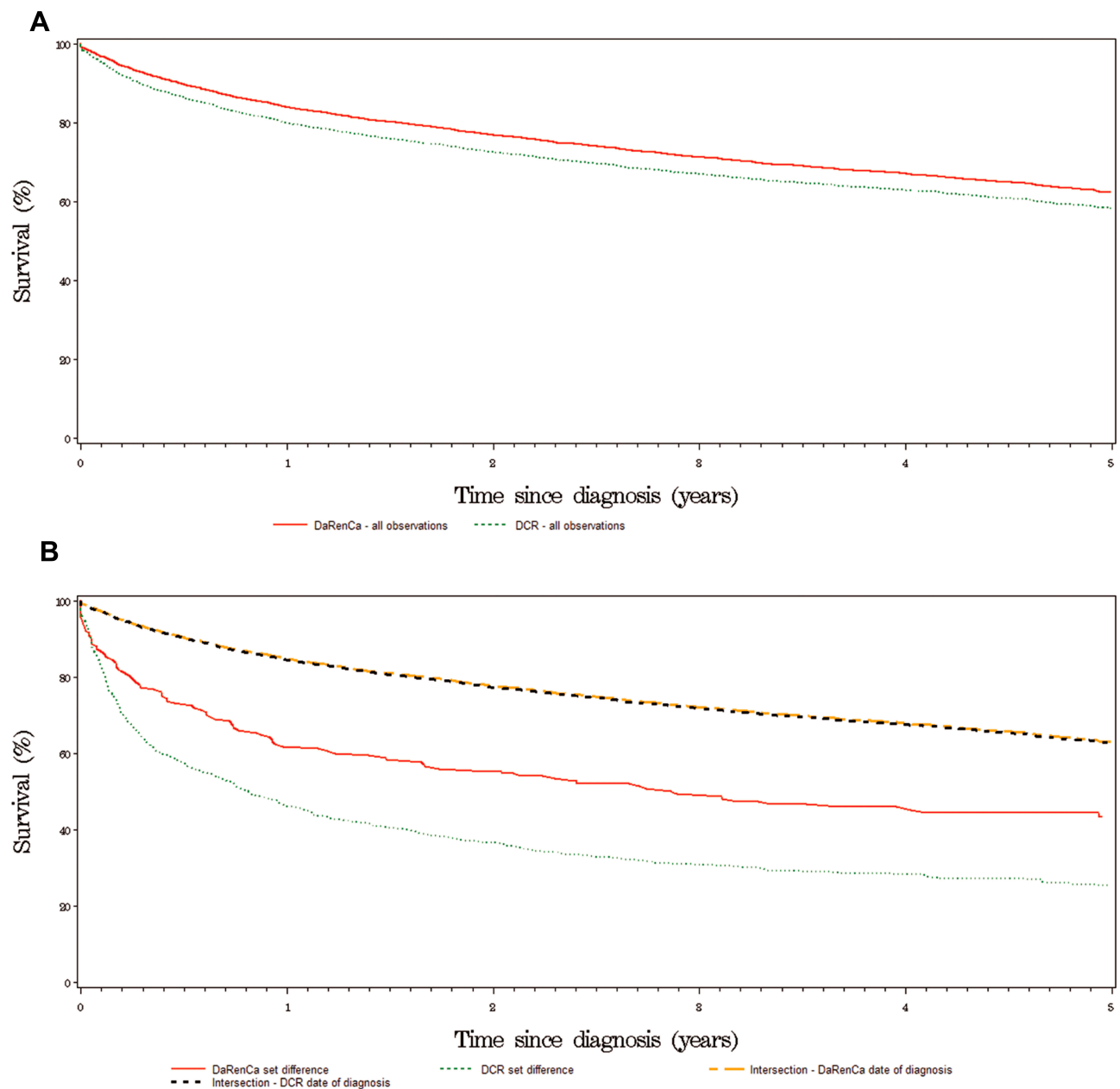## Comparison of Relative Survival Estimates

The relative survival of patients registered in DaRenCaData was higher compared to patients registered in DCR: The 1-year relative survival was 85% [95% CI: 84–86%] vs 81% [95% CI: 79–82%], 3-year relative survival was 74% [95% CI: 73–76%] vs 69% [95% CI: 67–70%], and 5-year relative survival was 65% [95% CI: 63–67%] vs 59% [95% CI: 57–61%] (Figure 1A, Table 2).

The 1-year relative survival of patients registered in both DaRenCaData and DCR was 86%, and was similar whether the DaRenCaData or DCR date of diagnosis was used as entry point (Figure 1B). In contrast, the 1-year relative survival was 64% [95% CI, 56–70%] for patients in the DaRenCaData only and 52% [95% CI 48–56%] for patients in DCR only (Figure 1B, Table 2).

Table 1 Completeness and Positive Predictive Value (PPV) of Registrations in DaRenCaData Compared to Registrations in DCR

	DaRenCaData – DCR Patients in Both Registries		DaRenCaData Only	DCR Only	Total	Completeness	PPV
	DaRenCaData N	DCR N	DaRenCa N	DCR N		% (95% CI)	% (95% CI)
01/08/2010–31/07/2011*	649	665	50	145		82 (79–84)	93 (91–95)
01/08/2011–31/07/2012	694	695	26	129		84 (82–87)	96 (95–98)
01/08/2012–31/07/2013	725	715	38	92		89 (87–91)	95 (93–97)
01/08/2013–31/07/2014	849	856	21	93		90 (88–92)	98 (97–99)
01/08/2014–31/07/2015	865	861	24	57		94 (92–95)	97 (96–98)
01/08/2015–31/12/2015	368	358	17	48		88 (85–92)	96 (94–98)
Total	**4150 (85%)**		**176 (4%)**	**564 (12%)**	**4890 (100%)**	**88 (87–89)**	**96 (95–97)**

Notes: *The time periods correspond to the annual reports of DaRenCaData going from August to July; DCR, Danish Cancer Registry; DaRenCaData, Clinical database of the multidisciplinary renal cancer group; PPV Positive Predictive Value: calculated as patients in the intersection in relation to all patients in DaRenCaData; Completeness: Calculated as patients in the intersection in relation to all patients in DCR.

Figure 1 (**A**) Relative survival for two nationwide renal cell carcinoma registries. Kaplan-Meier curves showing the outcomes for patients registered in DaRenCaData and Danish Cancer Registry. (**B**) Relative survival based on registry registration. Kaplan-Meier curves showing survival of patients found in both registries (intersection) with DaRenCaData's date of diagnosis, or with DCR's date of diagnosis, registration in DaRenCaData only (DaRenCa set difference) and registration in Danish Cancer Registry only (DCR set difference).

Further analyses of 1- and 3-year relative survival indicated that relative survival improved during the study period assessed in both DCR and DaRenCaData ([Supplementary Table 1](#)).

Comparing Mortality

The 1-year MRR for all patients in DaRenCaData was significantly lower than the MRR for all observations in DCR (HR = 0.79 [95% CI, 0.72 to 0.88]). Similar findings were obtained with regard to 3- and 5-year MRRs ([Table 3](#)).

Adjusting for stage did not impact findings. Compared with DCR patient population, 1-year MRR (HR; 95% CI) in DaRenCaData patient population adjusted for T-stage was 0.88 (0.73–0.89), N-stage 0.78 (0.71–0.86), M-stage 0.81 (0.73–0.89), and combined TNM-stage 0.83 (0.75–0.92); similar findings were obtained with 5-year MRR as endpoint. For patients included in both registries, the 1-year MRR was similar using either the date of diagnosis registered in DaRenCaData or in DCR as entry point (HR = 0.98 [95% CI, 0.88 to 1.10]). Patients in the DaRenCaData only and in

Table 2 Relative Survival Based on Registry Registration for DCR and DaRenCaData Patient Populations (Model 1), and Relative Survival for Patients in the Intersection (with Either DCR or DaRenCaData Date of Diagnosis as Entry Point), or Only Registered in DCR or DaRenCaData (Model 2)

	Relative Survival (95% CI)		
	1-Year	3-Year	5-Year
Model 1			
DCR patient population	81 (79–82)	69 (67–70)	59 (57–61)
DaRenCaData patient population	85 (84–86)	74 (73–76)	65 (63–67)
Model 2			
Intersection with DCR date of diagnosis	86 (85–87)	75 (73–77)	66 (64–68)
Intersection with DaRenCaData date of diagnosis	86 (85–88)	76 (74–77)	67 (65–69)
Patients only registered in DCR	52 (48–56)	29 (25–34)	21 (17–25)
Patients only registered in DaRenCaData	64 (56–70)	47 (39–55)	38 (30–46)

the DCR only had significantly higher MRRs (HR = 2.84 [95% CI, 2.20 to 3.68] and 4.29 [95% CI, 3.72 to 4.93], respectively), than patients registered in both registries using the DCR date of diagnosis as entry point, with the highest MRR observed among patients only found in the DCR registry (Table 3).

The MRRs for patients registered in both DaRenCaData and DCR declined significantly from 2012–2013 to 2014–2015 with lowest HR for the 1-year mortality rate at 0.74 [95% CI, 0.63 to 0.87] when compared to patients diagnosed in 2010–2011 (Table 4), and HR for the 3-year mortality rate at 0.77 [95% CI, 0.68 to 0.87] (Supplementary Table 2). Assessed with time as a continuous variable, the 1-year MRR was 0.93 [95% CI, 0.90 to 0.96] (Table 4) (corresponding to a 7% yearly reduction in the death rate), and the 3-year MRR was 0.94

[95% CI, 0.92 to 0.96] (Supplementary Table 2). All results were substantially the same when DCOs were excluded from the DCR data (results not shown).

Discussion

This study showed that the clinical database DaRenCaData had a high completeness of 88% and a positive predictive value at 96% in the study period from August 2010 to December 2015, when using the Danish Cancer Registry as reference. Furthermore, the completeness increased during the period from 82% to 94%, suggesting that an increasing number of patients with renal cell carcinoma have their suspected diagnosis confirmed histologically or cytologically. DaRenCaData made histological or cytological diagnosis of renal cell carcinoma mandatory by August 2012, and the observed increase in completeness may be a reflection of this request. Most biopsies performed were core biopsies. The main reason for the disagreement between the two registries was that DCR included a large number of patients without a histologically or cytologically verified diagnosis registered in DPR. Indeed, 446 patients who were found only in the DCR – corresponding to 60% of the total number of patients with disagreement – did not appear in DPR. These patients were older and only had a diagnosis based on imaging, representing either small indolent lesions, subject to watchful waiting, or aggressive lesions in frail elderly patients, unfit for treatment, non-renal cancer carcinoma neoplasia, or non-malignant lesions. The 24 patients (14%) found in DaRenCaData only, may have been cases identified by DCR as relapses and therefore not registered in DCR. In addition to this, 0.4% of the total DCR patient population were likely to be patients registered from death certificate only (the so-called DCOs), because the date of

Table 3 Mortality Rate Ratios for DCR and DaRenCa Patient Populations (Model 1), and Mortality Rate Ratios for Patients in the Intersection (with Either DCR or DaRenCa Date of Diagnosis as Entry Point), or Only Registered in DCR or DaRenCaData (Model 2)

	HR (95% CI)		
	1-Year	3-Year	5-Year
Model 1			
DCR patient population	1	1	1
DaRenCaData patient population	0.79 (0.72–0.88)	0.84 (0.78–0.91)	0.86 (0.80–0.92)
Model 2			
Intersection with DCR date of diagnosis	1	1	1
Intersection with DaRenCaData date of diagnosis	0.98 (0.88–1.10)	0.98 (0.91–1.07)	0.99 (0.92–1.06)
Patients only registered in DCR	4.29 (3.72–4.93)	3.58 (3.18–4.03)	3.27 (2.92–3.66)
Patients only registered in DaRenCaData	2.84 (2.20–3.68)	2.23 (1.78–2.78)	2.03 (1.64–2.51)

Table 4 1-Year Mortality Rate Ratios with Time Variations for Patients Registered in Both DCR and DaRenCaData (Model 2)

	Number of Deaths	1-Year Mortality Rate Ratio (95% CI)
Time measured as a categorical variable		
Period 1 (August 2010 – July 2011)	124	[Reference category]
Period 2 (August 2011 – July 2012)	127	1.04 (0.89–1.22)
Period 3 (August 2012 – July 2013)	120	0.84 (0.71–0.99)
Period 4 (August 2013 – July 2014)	127	0.82 (0.70–0.96)
Period 5 (August 2014 – July 2015)	133	0.74 (0.63–0.87)
Period 6 (August 2015 – December 2015)	58	0.74 (0.60–0.91)
Time measured as a continuous variable		
Years between date of diagnosis and start of observation period (August 1st 2010)	–	0.93 (0.90–0.96)

diagnosis was equal to the date of death. Actually, DCR reports to register in general 0.4% of all incident cancer cases as DCOs.[15] In essence, the main differences between DaRenCaData and DCR related to diagnoses that had not been histologically or cytologically verified, as a consequence of the differences in DaRenCaData and DCR inclusion criteria. Thus, the high completeness of DaRenCaData observed in the study period is indeed an underestimate, as not all patients in DCR should be included in DaRenCaData.

However, both DaRenCaData and DCR comprised patients that appeared to be candidates for the other registry as well. While this may partly be due to minor errors in their algorithms, some patients may have been excluded from DCR, because DCR had some observations under further investigation before decision of inclusion or exclusion (ie, the registration may have been considered incomplete). Also, both registries may have excluded some patients if they were found to have a prior diagnosis of renal cancer before the present study period. Such patients may have been registered with a registration of relapse in DaRenCaData or DNPR prior to the inclusion date.

In general, patients registered in DaRenCaData had higher relative survival and lower mortality rate ratios than patients registered in DCR. This finding may reflect that DaRenCaData used DPR as their primary source, meaning that only patients with histologically or cytologically verified diagnosis were included. This may indicate that patients in DCR without histologically or cytologically verified diagnoses primarily are patients who were considered to have a poor prognosis and consequently not being offered detailed diagnostic work-up; but on the other hand, some of them may also comprise patients with good prognosis and a minor lesion too small for biopsy. Actually, patients in both set differences displayed significantly higher MRRs and lower relative survival estimates than patients in the intersection; controlling for stage did not change our findings. Neither the different definitions of dates of diagnosis of the two registries nor DCOs registered in DCR appeared to affect survival measures to any notable extent. Intriguingly, for both registries, improved survival for patients with renal cell carcinoma was observed during the observation period, corresponding to a 7% yearly reduction in death.

The findings in this study have documented that the definition of patient populations in DCR and DaRenCaData has important implications for the survival measures. This applies to findings in Swedish RCC national dataset that uses similar definitions.[16] The main reason was that DCR also included patients that have not had their renal cancer diagnosis histologically or cytologically verified. While there may be good reasons for choosing either of the two inclusion criteria of the registries, it is important to be very explicit about the definitions, and sensitivity analysis using alternative criteria could be performed in order to provide a fuller understanding of the findings and enable dialogues about the results. The present study contributes to the growing literature regarding assessments of medical databases and their quality, providing better backgrounds for epidemiological research.

In conclusion, DaRenCaData had a high and increasing completeness at an average of 88% and high PPV at 96% compared to DCR, establishing DaRenCaData as a high-quality research database. In both registries, renal cell carcinoma mortality declined over time. Patients with renal cell carcinoma only registered in DCR or DaRenCaData have poorer outcomes than patients found in both registries. This study points to the importance of assessing the inclusion criteria when interpreting registry-based studies.

Funding

Disclosure

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. doi:10.2147/CLEP.S179083

2. Norgaard M, Johnsen SP. How can the research potential of the clinical quality databases be maximized? The Danish experience. *J Intern Med*. 2016;279(2):132–140. doi:10.1111/joim.12437

3. Petersen AC, Sogaard M, Mehnert F, et al. The database of the Danish Renal Cancer Group. *Clin Epidemiol*. 2016;8:725–729. doi:10.2147/CLEP.S106042

4. Gjerstorff ML. The Danish Cancer Registry. *Scand J Public Health*. 2011;39(7 Suppl):42–45. doi:10.1177/1403494810393562

5. Christensen J, Hojsgaard Schmidt LK, Tranberg Kejs AM, et al. Agreement between the Danish Cancer Registry and the Danish Colorectal Cancer Group Database. *Acta Oncol* .2020 Jan;59(1):116–123.

6. Pedersen SA, Schmidt SAJ, Klausen S, et al. Melanoma of the Skin in the Danish Cancer Registry and the Danish Melanoma Database: a validation study. *Epidemiology (Cambridge, Mass)*. 2018;29 (3):442–447. doi:10.1097/EDE.0000000000000802

7. Bakken IJ, Gystad SO, Christensen OO, et al. Comparison of data from the Norwegian Patient Register and the Cancer Registry of Norway. *Tidsskr nor Laegeforen*. 2012;132(11):1336–1340. doi:10.4045/tidsskr.11.1099

8. Bjerregaard B, Larsen OB. The Danish Pathology Register. *Scand J Public Health*. 2011;39(7 Suppl):72–74. doi:10.1177/1403494810393563

9. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. *Scand J Public Health*. 2011;39(7 Suppl):30–33. doi:10.1177/1403494811401482

10. Schmidt M, Pedersen L, Sorensen 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

11. Norgaard M, Skriver MV, Gregersen H, Pedersen G, Schonheyder HC, Sorensen HT. The data quality of haematological malignancy ICD-10 diagnoses in a population-based hospital discharge registry. *Eur J Cancer Prev*. 2005;14(3):201–206. doi:10.1097/00008469-200506000-00002

12. Hakulinen T, Seppa K, Lambert PC. Choosing the relative survival method for cancer survival estimation. *Eur J Cancer*. 2011;47 (14):2202–2210. doi:10.1016/j.ejca.2011.03.011

13. Corazziari I, Quinn M, Capocaccia R. Standard cancer patient population for age standardising survival ratios. *Eur J Cancer*. 2004;40 (15):2307–2316. doi:10.1016/j.ejca.2004.07.002

14. Miettinen J, Rantanen M, Seppa K. *Package 'popEpi'*.

15. The Danish Cancer Society & Statens Serum Institut. Validering af Cancerregisteret og udvalgte kliniske cancerdatabaser [Validation of The Danish Cancer Registry and selected clinical databases]. [report]; 2014. Available from: http://sundhedsdatastyrelsen.dk/-/media/sds/filer/registre-og-services/nationale-sundhedsregistre/sygedomme-laegemidler-og-behandlinger/cancerregisteret/valideringsrapport-cancerregisteret.pdf?la=da. Accessed April 15, 2020. Danish.

16. Thorstenson A, Bergman M, Scherman-Plogell AH, et al. Tumour characteristics and surgical treatment of renal cell carcinoma in Sweden 2005⁻2010: a population-based study from the national Swedish kidney cancer register. *Scand J Urol*. 2014;48(3):231–238. doi:10.3109/21681805.2013.864698