Completeness of prostate cancer staging in the Danish Cancer Registry, 2004–2009

Mary Nguyen-Nielsen 1
Trine Frøslev 1
Søren Friis 2
Michael Borre 3
Niels Harving 4
Mette Søgaard 1, 5

1 Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, 2 Danish Cancer Society Research Center, Danish Cancer Society, Copenhagen, 3 Department of Urology, Aalborg Hospital, Aarhus University Hospital, Aalborg, 4 Department of Clinical Microbiology, Aalborg Hospital, Aarhus University Hospital, Aalborg, Denmark

Objective: To investigate the completeness of TNM (Tumor-Node-Metastasis) staging for prostate cancer (PC) in the Danish Cancer Registry (DCR).

Methods: We identified 20,184 men registered with first-time PC in the DCR between 2004 and 2009. These patients were linked to the Danish National Patient Register to obtain data on comorbidity according to the Charlson Comorbidity Index (CCI). We calculated the completeness and corresponding 95% confidence intervals (CI) of TNM staging overall and by the individual components. We also defined a clinically-based algorithm classifying PC into four stage categories: localized, regional, distant, and unknown.

Results: The overall completeness of TNM staging was 34.2% (95% CI: 0.34–0.35). TNM completeness improved gradually over time reaching 41.2% in 2009. TNM completeness decreased substantially with age from 75.0% among patients 0–39 years to 11.3% among patients 80 years or older. Similarly, completeness decreased with increasing comorbidity level from 37.6% among patients with low CCI to 20.3% among those with high CCI. When classifying T1 cancer as a complete registration regardless of missing N or M stage, the overall TNM completeness increased to 48.7% (95% CI: 0.48–0.49). According to the clinically-based staging algorithm, 70.5% of PC cases could be categorized into a definite clinical stage.

Conclusion: One-third of PC patients had a complete registration of all TNM components in the DCR. Although TNM completeness improved over time, older age and high comorbidity were consistently associated with missing TNM staging. Research and monitoring based on cancer registries such as the DCR should account for missing TNM staging. Failing to do so could otherwise lead to biased results of stage-specific analyses.

Keywords: prostate cancer, neoplasm staging, TNM

Introduction

Prostate cancer (PC) is the most common nonskin cancer among men in Western countries.1,2 However, there is great variation in prognosis across countries and according to clinical stage.3 The overall five-year relative survival for PC in Denmark was 62%–68% during 2001–2006,4 whereas the corresponding figure was approximately 78% in other Nordic countries during 1999–2003.5 In the United States, the relative five-year survival for PC was estimated to be 99% overall; 100% for localized PC and 29% for metastatic PC during 2001–2007.6 With the dramatic differences in PC survival across countries and according to PC stage, knowledge of clinical staging has become increasingly important in both survival analyses and treatment strategies.5,7

Cancer staging is based on the TNM (Tumor-Node-Metastasis) classification system as outlined by the International Union against Cancer and the American Joint
Committee on Cancer. Exact T staging of PC (ie, T1–4) is particularly important because cancer grading (ie, Gleason scoring) and pre-operative prostate-specific antigen (PSA) is incorporated in the assessment of localized PC into low, intermediate, or high risk tumors. Data on TNM classification, incorporated in the assessment of localized PC into low, intermediate, or high risk tumors.10,11 Data on TNM classification, but not including Gleason scores and PSA values, have been registered in the Danish Cancer Registry (DCR) since 2004. The overall validity and completeness of cancer registration in the DCR is high,12,13 but the completeness of TNM cancer staging is unknown. Missing information in cancer registries may lead to biased results in epidemiological studies and monitoring, especially if the missing data are not distributed randomly. Thus, to ascertain the magnitude and direction of the potential bias, it is important to understand and examine the characteristics of unstaged cancer cases and how these differ from the staged cases.  The present study investigated the completeness of TNM staging for PC, stratified by age, comorbidity, and year of diagnosis in the DCR.

Material and methods

We performed this study in Denmark (population 5.5 million inhabitants). All residents in Denmark are provided with free, tax-supported medical care. Since 1968, the Danish Civil Registration System has assigned a unique 10-digit civil registration number to all Danish residents, which encodes date-of-birth and gender. This number is used in all Danish registries, allowing for unambiguous individual-level data linkage.

Ascertainment of patients with PC

From the DCR, we identified all patients who had a primary cancer diagnosis of PC [International Classification of Disease, 10th revision (ICD-10) code C61] recorded in the DCR between 2004 and 2009. The DCR has recorded information on incident cancers in the Danish population since 1943. Cancer diagnoses have been registered according to the ICD for Oncology (ICD-O) during the period 1978–2003 and the ICD-10 since 2004, but converted into ICD-10 codes for the entire period encompassing 1978–2009. From the DCR, we obtained information on date of diagnosis, age, gender, and TNM codes. We categorized cancer stage as localized, regional, distant or unknown based on a clinical algorithm using all T, N, and M components (Appendix 1). Missing data were allowed in the three clinical stage categories if the available information provided sufficient and clinically meaningful information to categorize cases. The “unknown” stage category consisted primarily of locally advanced tumors (T3 or T4) with unknown N and/or M stage.

Comorbidity

Data on comorbidities were obtained from the Danish National Patient Register (DNPR). The DNPR contains data on all admissions to nonpsychiatric hospitals in Denmark since 1977 and all outpatient contacts since 1995, including the civil registration number, date of admission/contact, and diagnosis codes. We identified pre-existing comorbidity using the Charlson Comorbidity Index (CCI). The CCI was based on all Charlson comorbidities (excluding PC) registered within 10 years preceding the date of PC diagnosis. CCI levels were categorized as 0 (low), 1–2 (medium), and 3+ (high).

Statistical analysis

We calculated the proportion of complete TNM registrations and corresponding 95% confidence intervals (CI), overall and by each component individually (ie, T, N, and M). Registration of TNM was defined as complete if the patient had received the following codes: T0–4, N0–3, and M0–1. A registration of Tx, Nx, or Mx was defined as an unknown stage (ie, information on tumor size, lymph node metastases, and distant metastases were not available or could not be assessed). We stratified completeness by gender, age (0–39 years, 40–59 years, 60–79 years, and ≥80 years), year of cancer diagnosis, and CCI level (0, 1–2, 3+). We defined T1-Nx-Mx as a complete TNM registration and repeated the analyses. In addition, we repeated the analyses restricting it to histologically verified PC diagnoses. Finally, we categorized PC patients into localized, regional, distant, or unknown stages using the above mentioned clinical algorithm (Appendix 1). Analyses were performed using Statistical Analysis Software (v 9.2; SAS Institute, Inc, Cary, NC, USA).

Results

We identified 20,184 men with primary PC between 2004 and 2009. The median age at diagnosis was 71 years (interquartile range, 65 to 77 years). The overall completeness of TNM registration was 34.2% (95% CI: 0.34–0.35), whereas 65.8% (95% CI: 0.65–0.66) had a registration of Tx, Nx, and/or Mx (Table 1). Completeness was lowest for “N” at 36.7%, compared with 88.8% for T, and 60.0% for M. Notably, the completeness of the TNM registration improved gradually over the six-year study period from 28.4% (95% CI: 0.27–0.30) in 2004 to 41.2% (95% CI: 0.40–0.43) in 2009. However, the completeness of TNM registration decreased with advancing age from 75.0% (95% CI: 0.28–0.97) among patients aged 0–39 years to 11.3% (95% CI: 0.10–0.12)
Table 1: Completeness of the TNM registration among prostate cancer patients in the Danish Cancer Registry (2004–2009)

<table>
<thead>
<tr>
<th>CCI score</th>
<th>Total (n = 20,184)</th>
<th>Year of diagnosis</th>
<th>Total (n)</th>
<th>Year of diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low (0)</td>
<td>10616 (52.6)</td>
<td>2004</td>
<td>1258 (46.4)</td>
<td>54 (2.0)</td>
</tr>
<tr>
<td>Medium (1–2)</td>
<td>10172 (50.2)</td>
<td>2005</td>
<td>1474 (48.7)</td>
<td>71 (2.3)</td>
</tr>
<tr>
<td>High (≥3)</td>
<td>1008 (4.8)</td>
<td>2006</td>
<td>1565 (50.6)</td>
<td>80 (2.6)</td>
</tr>
<tr>
<td></td>
<td>(95% CI: 0.36–0.37)</td>
<td>2007</td>
<td>2027 (54.3)</td>
<td>85 (2.3)</td>
</tr>
<tr>
<td></td>
<td>(95% CI: 0.36–0.37)</td>
<td>2008</td>
<td>2024 (54.6)</td>
<td>98 (2.6)</td>
</tr>
<tr>
<td></td>
<td>(95% CI: 0.36–0.37)</td>
<td>2009</td>
<td>2268 (58.0)</td>
<td>116 (3.0)</td>
</tr>
</tbody>
</table>

Table 2: Staging of prostate cancer patients in the Danish Cancer Registry by year of prostate cancer diagnosis, sex, age group, and comorbidity level

<table>
<thead>
<tr>
<th>Cancer stage</th>
<th>Total (n)</th>
<th>Year of diagnosis</th>
<th>Age group (years)</th>
<th>CCI score</th>
<th>Known (n)</th>
<th>Unknown (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Localized</td>
<td>10166 (52.6)</td>
<td>2004</td>
<td>0–39</td>
<td>Low (0)</td>
<td>3 (7.50)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>Regional</td>
<td>504 (2.5)</td>
<td>2005</td>
<td>40–59</td>
<td>Medium (1–2)</td>
<td>1418 (68.5)</td>
<td>87 (2.4)</td>
</tr>
<tr>
<td>Distant</td>
<td>3118 (15.4)</td>
<td>2006</td>
<td>60–79</td>
<td>High (≥3)</td>
<td>7375 (55.7)</td>
<td>383 (2.9)</td>
</tr>
<tr>
<td>Unknown</td>
<td>5946 (29.5)</td>
<td>2007</td>
<td>≥80</td>
<td>Low (0)</td>
<td>7375 (55.7)</td>
<td>383 (2.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2008</td>
<td></td>
<td>Medium (1–2)</td>
<td>2665 (48.4)</td>
<td>108 (2.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2009</td>
<td></td>
<td>High (≥3)</td>
<td>576 (39.8)</td>
<td>13 (0.9)</td>
</tr>
</tbody>
</table>

Abbreviations: CCI, Charlson Comorbidity Index; n, number.
Klassen et al investigated staging among 22,217 men with PC in the Maryland Cancer Registry. These authors found that older age, black race, and higher median income increased the likelihood of missing cancer staging. Another study of “all cancer types,” based on 1,907 nursing home residents in Michigan, USA, suggested that comorbidity was not associated with a higher proportion of unstaged cancer; however, the majority of the patients in this study were over 86 years old, thus reducing the generalizability of these findings due to selection bias. Finally, Radespiel-Troger et al investigated the completeness of PC staging in the Bavarian Cancer Registry and reported an overall registration completeness of 96%, and the proportions of missing T, N, and M were 15%, 37%, and 49%, respectively.

The strengths of our study are: (1) its population-based design; (2) the large sample size of over 20,000 PC patients; (3) the comprehensiveness of the Danish medical registries together with the ability to link data across multiple registries; and (4) the nature of the uniform health care system in Denmark that ensures equal access to free health care to all residents. Furthermore, the registration of cancer incidence is virtually complete in the DCR,12,13,16 and the positive predictive value of coding in the DNPR for comorbid diseases (CCI) was recently shown to be high.19

The main limitation of our study is that it only investigated the level of TNM completeness; we did not assess the actual validity of registration (eg, by medical record review). It would be relevant to proceed with validation studies to investigate the quality and validity of TNM staging in the DCR. It is also important to note that a given TNM stage in the DCR reflects the stage at the time of initial diagnosis/registration. Also, our stage categorization (into localized, regional, or distant PC) was based on a simplified algorithm which only took into consideration T, N, and M stage, and thus did not include Gleason scores or PSA values as otherwise outlined by the International Union against Cancer and the American Joint Committee on Cancer 2010 guidelines for PC stage grouping.5–10 A new nationwide clinical database, DaProCaData, has recently been established in Denmark and is projected to provide detailed data on Gleason scores, cancer staging, and lifestyle variables.26 Moreover, an alternative clinical staging of PC into localized, locally advanced, and distant PC might be more appropriate than our TNM-based staging into localized, regional, and distant PC since the distinction between localized and locally advanced PC correlates better with current treatment measures. For localized PC, it would have been ideal to perform stratification according to the classification defined by the D’Amico Risk Group into low, intermediate, and high level of risk.27 However, this would require additional information on pre-diagnostic PSA levels and Gleason scores, which are not available in the DCR. A recent study investigating data completeness in the Surveillance, Epidemiology, and End Results database of the National Cancer Institute and PC staging based on the D’Amico staging showed that 33% had incomplete registration of D’Amico variables (PSA, Gleason, and/or T stage).28 Similar to our findings, elderly PC patients were also found to have the highest level of missing data when using the DARG classification.24

The reasons for the low completeness of TNM staging for PC in the DCR are most likely multifactorial and may be related to age, comorbidity, patient/doctor preferences, and current clinical guidelines. It has been suggested that with increasing age and comorbidity, patients increasingly refrain from further diagnostic measures.20,22 The limited registration of “N” likely results from the current clinical guidelines for PC and the diagnostic process related hereto. Staging of “N” for PC requires either surgical procedures or advanced imaging of the lymph nodes, which are not performed in all patients. The European Association of Urology guidelines on PC only recommend N-staging when “potentially curative treatment is planned for.”29 Thus, if we had designed our staging algorithm in accordance with this, we would have found a lower proportion of patients with an unknown cancer stage since missing N-stage accounted for the majority of the missing T-N-M components. Furthermore, prostate biopsies are associated with a relative high risk of bleeding and infection and CT-imaging has recently been shown to have low sensitivity and specificity in detecting dissemination to the pelvic lymph nodes.30 Likewise, extended pelvic lymph node dissection and its therapeutic benefit in the staging and management of PC is much debated.31 By contrast, standard diagnostic procedures for breast cancer include “sentinel node” dissection, resulting in a higher proportion of N registrations for breast cancer patients simply due to standard clinical practice.

When we defined T1-Nx-Mx as a complete TNM registration, the overall TNM completeness improved to 48.7%. This proportion more accurately reflects TNM completeness in the DCR. In the case of a T1 PC tumor with no clinical signs of metastatic spread to the lymph nodes or to distant sites, the clinician will typically not proceed with further diagnostic testing. For such a T1 tumor, it is inferred that there is no spread outside the solid tumor, but Nx and Mx
must be registered since N0 and M0 status has not technically been confirmed. Consequently, the registration of T1-Nx-Mx typically represents a complete TNM staging based on clinical evaluation. Likewise, when using the algorithm to classify PC cases into localized, regional, or distant stage, 70.5% of all PC patients could be staged in the DCR. These results are clinically meaningful and perhaps better reflect the true completeness of cancer staging in the DCR.

Exclusion of nonhistologically verified PC had little impact on the overall results. Compared to other common cancers registered in the DCR (ie, breast, lung and colon cancer), PC had the highest proportion of nonhistologically verified cancer diagnosis (19.3%). This is expectable, given current clinical practice, the typical slow progression of the majority of PC cases, and the potential risks associated with invasive diagnostic procedures. The diagnosis of PC relies on a combination of digital rectal examination, PSA, and eventually histology or imaging (eg, MRI or bone scintigraphy). The benefit of histological verification and complete TNM cancer staging may simply not outweigh the potential harms.

In conclusion, one-third of PC patients had complete TNM staging in the DCR. When applying our clinical staging algorithm, 70.5% of all PC cases could be staged as localized, regional or distant metastatic. However, TNM completeness was consistently poorer with older age and high comorbidity. Our findings from the DCR highlight the potential for selection bias when data is missing. Missing or unknown data should therefore be accounted for (eg, by multiple imputation methods) when conducting stage-specific analyses in the DCR or other cancer registries.

Acknowledgments
This project was funded by the Regional Clinical Epidemiological Monitoring Initiative for Central and North Denmark Regions. A sincere thanks to Christina Gade Jespersen, MD, for her input on the clinical staging algorithm.

Disclosure
The authors declare no conflicts of interest.

References


Appendix 1 Algorithm for prostate cancer staging according to the TNM classification in the Danish Cancer Registry

<table>
<thead>
<tr>
<th>Stage</th>
<th>TNM-codes</th>
</tr>
</thead>
</table>
| Localized | T1–4, x N0 M0  
T1–2 N0 Mx  
T1–2 Nx M0, x |
| Regional | T1–4, x N1, M0  
T1–4, x N1, M1  
T1–4, x N0 M1  
T1–4, x Nx M1 |
| Distant  | T3–4, x Nx M0, x  
T3–4, x N0 Mx  
T1–4, x N1 Mx |

Notes: According to UICC/AJCC guidelines, Tis, Ta, N2, and N3 are not used for PC.9–11 We identified 144 PC cases who were registered with these classifications; these cases were categorized as “unknown.” Likewise eight cases registered with “T0” were also categorized as “unknown.”

Abbreviations: PC, prostate cancer; TNM, tumor-node-metastasis; UICC/AJCC, International Union against Cancer and the American Joint Committee on Cancer.