

# Completeness of TNM staging of small-cell and non-small-cell lung cancer in the Danish Cancer Registry, 2004–2009

Thomas Deleuran<sup>1,2</sup>  
Mette Søgaard<sup>1,3</sup>  
Trine Frøslev<sup>1</sup>  
Torben Riis Rasmussen<sup>4</sup>  
Henrik Kirstein Jensen<sup>5</sup>  
Søren Friis<sup>6</sup>  
Morten Olsen<sup>1</sup>

<sup>1</sup>Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark; <sup>2</sup>Department of Medicine V (Hepatology and Gastroenterology), Aarhus University Hospital, Aarhus, Denmark; <sup>3</sup>Department of Clinical Microbiology, Aalborg Hospital, Aarhus University Hospital, Aalborg Denmark; <sup>4</sup>Department of Respiratory Medicine, Aarhus University Hospital, Aarhus, Denmark; <sup>5</sup>Department of Respiratory Medicine, Aalborg Hospital, Aarhus University Hospital, Denmark; <sup>6</sup>Danish Cancer Society Research Center, Danish Cancer Society, Copenhagen, Denmark

**Objective:** We examined the completeness of TNM staging of small-cell (SCLC) and non-small-cell (NSCLC) lung cancer in the national Danish Cancer Registry (DCR) and whether staging varied by year of diagnosis, gender, age, degree of comorbidity, or presence of histopathological diagnosis.

**Methods:** We identified all patients with SCLCs and NSCLCs registered in the DCR during 2004–2009 and examined the completeness of their TNM registrations. Completeness was defined as the number of recorded individuals with TNM divided by the total number of patients. Completeness was calculated for TNM, T, N, and M individually, overall, and by year of diagnosis, gender, age at diagnosis, and comorbidity. Data regarding comorbidity was obtained from the Danish National Patient Register (DNPR). We performed separate analyses for patients with a histopathologically verified diagnosis of NSCLC. Finally, we designed an algorithm to categorize tumors with missing TNM components as limited, extensive, or distant disease.

**Results:** Overall TNM staging completeness was 77.5% (95% confidence interval (CI): 76.1%–78.8%) for SCLC and 77.9% (95% CI: 77.3%–78.4%) for NSCLC. Completeness did not vary by gender and increased during the study period. The proportion of staged patients was lower among patients above 80 years of age or with medium to high levels of comorbidity.

**Conclusion:** Overall TNM completeness for SCLC and NSCLC in the Danish Cancer Registry is high, but decreases with increasing levels of comorbidity and at ages greater than 80 years. Researchers should be aware of these potential sources of bias.

**Keywords:** cancer registers, cancer staging, completeness, lung cancer

## Introduction

Lung cancer is one of the most common malignancies and is the leading cause of cancer-related death in Denmark, as well as worldwide.<sup>1–3</sup> In Denmark, incidence rates of lung cancer in 2009 were 83 (men) and 64 (women) per 100,000 person-years. Lung cancer is classified into two subtypes, including small cell (SCLC, 15%–18% of incident cases) and non-small cell lung cancer (NSCLC, 82%–85% of incident cases) based on histopathology and differences in prognosis and treatment. Stages of SCLC (limited and extensive) and NSCLC (I–IV) are assigned according to TNM classification, which addresses tumor size and growth into neighboring organs (T), lymph node involvement (N), and distant metastases (M). Prognosis and treatment choices vary according to cancer stage, but prognosis is also affected by age and level of comorbidity.<sup>4–8</sup> The five-year survival rate for SCLC patients is approximately 20% for limited disease and 0% for extensive disease.<sup>9</sup> Corresponding rates for NSCLC patients are 50% and approximately 5% for stage IA and IV, respectively.<sup>10,11</sup> Only two studies

Correspondence: Thomas Deleuran  
Department of Clinical Epidemiology,  
Aarhus University Hospital, Olof Palmes  
Allé 43-45, 8200 Aarhus N, Denmark  
Tel +4587134797  
Fax +45 87167215  
Email thomas.deleuran@dce.au.dk

have been conducted regarding the registration of TNM staging for lung cancer in population-based registries.<sup>12,13</sup> The proportion of patients with incomplete lung cancer stage in these registries increased with older age, as well as those that lived alone and were of black ethnicity.

The Danish Cancer Registry (DCR)<sup>14,15</sup> is a nationwide registry that has recorded cancer incidence in the Danish population since 1943. Cancer diagnoses have been recorded according to the International Classification of Disease, 10th revision (ICD-10), since 2004. Stage at diagnosis has been recorded according to the 6th edition of TNM classification<sup>16</sup> since 2004 and the 7th edition since 2009.

Completeness of TNM staging in the DCR is unknown. We therefore studied TNM completeness of SCLC and NSCLC, as well as conducted stratified analyses by year of diagnosis, gender, age, and level of comorbidity. Additionally, we suggest a method for defining stage categories in the presence of specific missing T, N, or M components.

## Material and methods

We performed this study in Denmark, which has a population of approximately 5.4 million. All residents are provided with tax-supported medical care which is free of charge. Since 1968, the Danish Civil Registration System has assigned a unique 10-digit personal identification number to all Danish residents.<sup>17</sup> This number is used in all Danish registers and allows unambiguous individual-level data linkage.

## Identification and categorization of lung cancer patients

We identified all patients with a first-time lung cancer diagnosis (ICD-10 code C33-C34) recorded between January 1, 2004 and December 31, 2009 in the DCR. Patients were categorized according to histopathological (morphological) diagnosis in the DCR, which is derived from the Danish Pathology Register and recorded according to the International Classification of Diseases for Oncology, 3rd version (ICD-0-3) and combined with a topographic diagnosis code.<sup>14,18</sup> We defined SCLC as cases with morphology codes 80413–80459, while NSCLC was defined as any other morphology code including those with no available histopathological diagnosis.

We designed an algorithm allowing assignment of stage categories, despite specific missing T, N, or M values. We categorized SCLC tumors as limited (if tumor was M0), extensive disease (if tumor was M1), or unknown (if tumor was Mx), regardless of the values, known or unknown, of other components (T and N) (Appendix 1). We categorized

NSCLC tumors into limited, extensive, distant disease, and unknown (Appendix 1). If TNM stage included component values of T4, N3, or M1, tumor stage was categorized as IIIb or IV, and these tumors were categorized as distant. This was done regardless of the stage, known or unknown, of other components. All other cancers with missing components were categorized as unknown.

## Comorbidity

We obtained data regarding comorbidity from the Danish National Patient Register (DNPR).<sup>19</sup> This register contains data regarding all admissions to non-psychiatric hospitals in Denmark since 1978 and outpatient settings since 1995. The data includes dates of admission/contact, discharge, and diagnosis codes according to the ICD-10. The presence of comorbidity was defined according to the Charlson Comorbidity Index<sup>20,21</sup> (CCI), and categorized as low (CCI = 0), medium (CCI = 1–2), or high (CCI > 2) level of comorbidity. The CCI was based on all hospital diagnoses recorded in the DNPR within 10 years preceding the date of cancer diagnosis.

## Statistical analysis

We computed completeness and corresponding 95% confidence intervals (CI) for the TNM registration overall and for each component (ie, T, N, and M). Completeness was defined as the number of individuals with TNM recordings divided by the total number of patients. We stratified completeness according to year of cancer diagnosis, gender, age at diagnosis (0–39 years, 40–59 years, 60–79 years, or ≥80 years), and level of comorbidity. We computed completeness separately for SCLC and NSCLC and, for NSCLC, for histologically verified disease. SCLC diagnosis had been verified histologically. We repeated the analyses using the algorithm for stage assignment in the presence of specific missing T, N, or M values. Analyses were performed using SAS (v 9.2; SAS Institute, Inc, Cary, NC, USA).

## Results

We identified 24,734 patients (50% males) that had lung cancer in Denmark between 2004 and 2009. Of these, 3,658 (15%) had SCLC (Table 1) and 21,076 (85%) had NSCLC (Table 2). Among those with NSCLC, 17,972 (85%) cases had been histologically verified. Median age at diagnosis was 69 years for all patients combined. The proportion of patients with preexisting hospital recorded comorbidity was similar between the two patient categories (47% among patients with SCLC, 50% among patients with NSCLC).

**Table 1** Completeness of TNM registration for 3658 patients with small cell lung cancer according to year of diagnosis, gender, age, and level of comorbidity

	TNM completeness		T completeness		N completeness		M completeness		Total N
	% (N)	95% CI	% (N)	95% CI	% (N)	95% CI	% (N)	95% CI	
Year of diagnosis									
2004	75.9 (421)	72.2–79.3	87.7 (487)	84.8–90.3	80.9 (449)	77.5–84.0	92.6 (514)	90.2–94.6	555
2005	76.0 (469)	72.5–79.3	85.7 (529)	82.8–88.3	81.5 (503)	78.3–84.4	92.2 (569)	89.9–94.1	617
2006	75.5 (468)	72.0–78.8	85.5 (530)	82.6–88.1	81.6 (506)	78.4–84.5	91.9 (570)	89.6–93.9	620
2007	75.3 (484)	71.8–78.5	84.6 (544)	81.7–87.2	80.7 (519)	77.5–83.6	92.4 (594)	90.1–94.2	643
2008	78.7 (474)	75.3–81.9	90.4 (544)	87.8–92.5	85.4 (514)	82.4–88.0	93.5 (563)	91.3–95.3	602
2009	83.4 (518)	80.3–86.2	91.0 (565)	88.5–93.1	89.0 (553)	86.4–91.3	95.8 (595)	94.0–97.2	621
Gender									
Female	76.9 (1359)	74.9–78.8	87.5 (1546)	85.9–89.0	83.1 (1468)	81.3–84.8	92.5 (1635)	91.2–93.7	1767
Male	78.0 (1475)	76.1–79.8	87.4 (1653)	86.0–88.9	83.3 (1576)	81.6–85.0	93.6 (1770)	92.4–94.6	1891
Age									
0–39 years	85.7 (6)	49.9–98.4	100 (7)	70.8–100	85.7 (6)	49.9–98.4	100 (7)	70.8–100	7
40–59 years	78.7 (548)	75.6–81.7	88.2 (614)	85.7–90.5	83.6 (582)	80.7–86.2	95.5 (665)	93.8–96.9	696
60–79 years	78.5 (2043)	76.9–80.1	88.0 (2290)	86.7–89.2	84.2 (2192)	82.8–85.6	93.7 (2437)	92.7–94.6	2602
≥80 years	67.1 (237)	62.1–71.9	81.6 (288)	77.3–85.4	74.8 (264)	70.1–79.1	83.9 (296)	79.8–87.4	353
Level of comorbidity									
0	80.5 (1576)	78.7–82.2	89.7 (1755)	88.3–91.0	85.6 (1675)	84.0–87.1	94.7 (1854)	93.7–95.7	1957
1–2	75.6 (961)	73.2–77.9	86.5 (1099)	84.5–88.3	81.7 (1038)	79.5–83.7	92.0 (1169)	90.4–93.4	1271
≥3	69.1 (297)	64.6–73.3	80.2 (345)	76.3–83.8	77.0 (331)	72.8–80.8	88.8 (382)	85.6–91.6	430
Total	77.5 (2834)	76.1–78.8	87.5 (3199)	86.4–88.5	83.2 (3044)	82.0–84.4	93.1 (3405)	92.2–93.9	3658

Overall completeness of the TNM staging was 77.5% (95% confidence interval (CI): [76.1%–78.8%]) for SCLC and 77.9% (95% CI: [77.3%–78.4%]) for NSCLC. Overall completeness for NSCLC with histological verification was 83.3% (95% CI: [82.8%–83.9%]) (data not shown).

During the study period, completeness increased from 75.9% (95% CI: [72.2%–79.3%]) to 83.4% (95% CI: [80.3%–86.2%]) for patients with SCLC and from 77.2% (95% CI: [75.8%–78.6%]) to 82.4% (95% CI: [81.2%–83.6%]) for patients with NSCLC (Tables 1 and 2).

**Table 2** Completeness of TNM registration for 21,076 patients with non-small cell lung cancer according to year of diagnosis, gender, age, and level of comorbidity

	TNM completeness		T completeness		N completeness		M completeness		Total N
	% (N)	95% CI	% (N)	95% CI	% (N)	95% CI	% (N)	95% CI	
Year of diagnosis									
2004	77.2 (2594)	75.8–78.6	87.2 (2927)	86.0–88.3	80.6 (2707)	79.3–81.9	88.0 (2956)	86.9–89.1	3358
2005	75.6 (2543)	74.1–77.0	86.6 (2913)	85.4–87.7	79.7 (2681)	78.3–81.0	88.9 (2993)	87.9–90.0	3365
2006	74.0 (2549)	72.5–75.4	84.8 (2922)	83.6–86.0	78.1 (2692)	76.7–79.5	87.1 (3003)	86.0–88.2	3446
2007	76.7 (2797)	75.3–78.0	89.0 (3246)	87.9–89.9	81.5 (2974)	80.2–82.7	90.2 (3290)	89.2–91.1	3649
2008	81.0 (2903)	79.7–82.2	91.1 (3265)	90.1–92.0	85.9 (3078)	84.7–87.0	92.2 (3304)	91.3–93.0	3585
2009	82.4 (3027)	81.2–83.6	91.9 (3376)	91.0–92.8	86.8 (3190)	85.7–87.9	92.1 (3383)	91.2–92.9	3673
Gender									
Female	77.6 (7605)	76.8–78.4	88.4 (8661)	87.7–89.0	82.0 (8035)	81.2–82.7	89.5 (8775)	88.9–90.1	9800
Male	78.1 (8808)	77.3–78.9	88.6 (9988)	88.0–89.2	82.4 (9287)	81.7–83.1	90.0 (10154)	89.5–90.6	11276
Age									
0–39 years	83.7 (82)	75.4–90.0	89.8 (88)	82.7–94.6	87.8 (86)	80.2–93.1	95.9 (94)	90.6–98.6	98
40–59 years	82.9 (3224)	81.7–84.1	91.4 (3553)	90.5–92.3	87.1 (3384)	86.0–88.1	94.8 (3683)	94.0–95.4	3887
60–79 years	80.8 (11085)	80.2–81.5	90.5 (12413)	90.0–91.0	84.8 (11630)	84.2–85.4	91.9 (12607)	91.5–92.4	13715
≥80 years	59.9 (2022)	58.2–61.5	76.9 (2595)	75.4–78.3	65.8 (2222)	64.2–67.4	75.4 (2545)	73.9–76.8	3376
Level of comorbidity									
0	81.1 (8535)	80.4–81.9	90.5 (9524)	90.0–91.1	85.1 (8953)	84.4–85.8	92.6 (9744)	92.1–93.1	10521
1–2	76.4 (5868)	75.5–77.4	87.6 (6729)	86.9–88.4	80.8 (6203)	79.9–81.7	88.2 (6775)	87.5–89.0	7678
≥3	69.9 (2010)	68.2–71.5	83.3 (2396)	81.9–84.6	75.3 (2166)	73.7–76.8	83.8 (2410)	82.4–85.1	2877
Total	77.9 (16413)	77.3–78.4	88.5 (18649)	88.1–88.9	82.2 (17322)	81.7–82.7	89.8 (18929)	89.4–90.2	21076

**Table 3** Patients with limited, extensive, or unknown stage of small cell lung cancer according to year of diagnosis, gender, age, and level of comorbidity

	Limited disease		Extensive disease		Unknown	
	N	%	N	%	N	%
Year of diagnosis						
2004	153	27.6	361	65.0	41	7.4
2005	176	28.5	393	63.7	48	7.8
2006	149	24.0	421	67.9	50	8.1
2007	168	26.1	426	66.3	49	7.6
2008	160	26.6	403	66.9	39	6.5
2009	152	24.5	443	71.3	26	4.2
Gender						
Female	467	26.4	1168	66.1	132	7.5
Male	491	26.0	1279	67.6	121	6.4
Age						
0–39 years	2	28.6	5	71.4	0	0
40–59 years	191	27.4	474	68.1	31	4.5
60–79 years	696	26.7	1741	66.9	165	6.3
≥80 years	69	19.5	227	64.3	57	16.1
Level of comorbidity						
0	566	28.9	1288	65.8	103	5.3
1–2	297	23.4	872	68.6	102	8.0
≥3	95	22.1	287	66.7	48	11.2
Total	958	26.2	2447	66.9	253	6.9

Completeness did not vary by gender. However, among patients with SCLC, completeness declined from 78.5% (95% CI: [76.9%–80.1%]) among those aged 60–79 years to 67.1% (95% CI: [62.1%–71.9%]) among those aged 80 years or more. For patients with NSCLC, completeness declined

from 80.8% (95% CI: [80.2%–81.5%]) among those aged 60–79 years to 59.9% (95% CI: [58.2%–61.5%]) among those aged 80 years or more. Completeness did not vary significantly between age groups 0–39, 40–59 or 60–79 years for SCLC or NSCLC.

Completeness also decreased with comorbidity level. The proportions of complete TNM staging were 81%, 76%, and 69% for low, medium, and high levels of comorbidity among patients with SCLC. The corresponding proportions were 81%, 76%, and 70% for patients with NSCLC.

Finally, we evaluated stage completeness for SCLC and NSCLC, allowing for the inclusion of missing TNM components as described above. When using this algorithm, overall stage completeness was 93.1% for SCLC and 91.1% for NSCLC. However, variations in completeness according to gender, age, year of diagnosis, and comorbidity showed a pattern similar to that of the original data (Tables 3 and 4).

## Discussion

We found that the completeness of TNM staging for SCLC and NSCLC in the DCR was lower for elderly patients and/or those with severe comorbidity. Completeness increased during the study period (2004–2009) and did not vary by lung cancer subtype or gender. Among patients with NSCLC, completeness was higher for patients with a histologically verified diagnosis than for those without such diagnosis.

**Table 4** Patients with limited, extensive, disseminated, or unknown stage of non-small cell lung cancer, according to year of diagnosis, gender, age, and level of comorbidity

	Limited disease		Extensive disease		Disseminated disease		Unknown	
	N	%	N	%	N	%	N	%
Year of diagnosis								
2004	519	15.5	294	8.8	2171	64.7	374	11.1
2005	521	15.5	337	10.0	2157	64.1	350	10.4
2006	447	13.0	309	9.0	2303	66.8	387	11.2
2007	560	15.3	316	8.7	2467	67.6	306	8.4
2008	521	14.5	331	9.2	2510	70.0	223	6.2
2009	551	15.0	393	10.7	2488	67.7	241	6.6
Gender								
Female	1526	15.6	866	8.8	6524	66.6	884	9.0
Male	1593	14.1	1114	9.9	7572	67.2	997	8.8
Age								
0–39 years	18	18.4	7	7.1	68	69.4	5	5.1
40–59 years	542	13.9	373	9.6	2816	72.4	156	4.0
60–79 years	2177	15.9	1344	9.8	9252	67.5	942	6.9
≥80 years	382	11.3	256	7.6	1960	58.1	778	23.0
Level of comorbidity								
0	1396	13.3	1005	9.6	7465	71.0	655	6.2
1–2	1242	16.2	709	9.2	4919	64.1	808	10.5
≥3	481	16.7	266	9.2	1712	59.5	418	14.5
Total	3119	14.8	1980	9.4	14096	66.9	1881	8.9



Some SCLC as NSCLC cases without morphological information may have been misclassified. However, this potential misclassification likely had minor impact on the results, since we observed nearly the same completeness for SCLC as for NSCLC.

Completeness of lung cancer staging in the National Cancer Institute's (US) Surveillance, Epidemiology, and End Results (SEER) Program<sup>12</sup> and completeness of the Swedish Cancer Register<sup>13</sup> are higher than that of the DCR. Institutions participating in the SEER are financially rewarded, and SEER promotes education and conducts regular audits for cancer registration,<sup>22</sup> which may explain this higher completeness. However, neither report presented overall estimates nor information regarding which components of TNM were missing or the impact of cancer type (SCLC or NSCLC), diagnosis year, or comorbidity on the likelihood of proper TNM staging.

Our findings are consistent with those of previous studies examining unstaged malignant disease which showed that fragile, elderly and/or those requiring nursing care are more likely to have unstaged disease.<sup>13,23,24</sup> This is the first study to present TNM staging according to comorbidity levels.

In conclusion, overall completeness of the TNM registration of SCLC and NSCLC in the DCR is high, and the amount of incomplete data is likely to have a minor impact on studies examining prognosis or surveillance that rely on DCR data, particularly when evaluating potential sources of bias such as age and level of comorbidity. We further showed that overall stage completeness can be increased by including patients for which the TNM components provide unambiguous prognostic information, despite missing or misclassified information of individual staging components. Completeness of both SCLC and NSCLC registration decreased with increasing age and comorbidity level, but was not affected by gender or lung cancer subtype. Future studies examining treatment, prognosis, or screening policies that rely on TNM data from the DCR should be conducted to address these potential sources of bias.

## Acknowledgments

This study received financial support from the Regional Clinical Epidemiological Monitoring Initiative for Central and North Denmark Regions.

## Disclosure

The authors declare no conflicts of interest in this work.

## References

1. Danish Board of Health. The Danish Cancer Registry 2010. Copenhagen, Danish Board of Health, 2011. <http://www.sst.dk/publ/Publ2011/DAF/Cancer/Cancerregisteret2010.pdf>.
2. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*. 2010;127:2893–2917.
3. Danish Board of Health. The Danish Register of Causes of Death 2010. Copenhagen, Danish Board of Health, 2011. <http://www.sst.dk/publ/Publ2011/DAF/Doedsaarsagertal2010.pdf>.
4. Kuo YW, Jerng JS, Shih JY, Chen KY, Yu CJ, Yang PC. The prognostic value of the simplified comorbidity score in the treatment of small cell lung carcinoma. *J Thorac Oncol*. 2011;6(2):378–383.
5. Birim O, Kappetein AP, Bogers AJ. Charlson comorbidity index as a predictor of long-term outcome after surgery for nonsmall cell lung cancer. *Eur J Cardiothorac Surg*. 2005;28(5):759–762.
6. Janssen-Heijnen ML, Smulders S, Lemmens VE, Smeenk FW, van Geffen HJ, Coebergh JW. Effect of comorbidity on the treatment and prognosis of elderly patients with non-small cell lung cancer. *Thorax*. 2004;59(7):602–607.
7. Blanco JA, Toste IS, Alvarez RF, Cuadrado GR, Gonzalez AM, Martin IJ. Age, comorbidity, treatment decision and prognosis in lung cancer. *Age Ageing*. 2008;37(6):715–718.
8. Tammemagi CM, Neslund-Dudas C, Simoff M, Kvale P. Impact of comorbidity on lung cancer survival. *Int J Cancer*. 2003;103(6):792–802.
9. Sorensen M, Pijls-Johannesma M, Felip E. Small-cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2010;21(5 Suppl):120–125.
10. Crinò L, Weder W, van Meerbeeck J, Felip E. Early stage and locally advanced (non-metastatic) non-small-cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2010;21(5 Suppl):103–115.
11. Asamura H, Goya T, Koshiishi Y, et al. A Japanese Lung Cancer Registry study: prognosis of 13,010 resected lung cancers. *J Thorac Oncol*. 2008;3(1):46–52.
12. Merrill RM, Sloan A, Anderson AE, Ryker K. Unstaged cancer in the United States: a population-based study. *BMC Cancer*. 2011;11:402.
13. Barlow L, Westergren K, Holmberg L, Talbäck M. The completeness of the Swedish Cancer Register: a sample survey for year 1998. *Acta Oncol*. 2009;48(1):27–33.
14. Gjerstorff ML. The Danish Cancer Registry. *Scand J Public Health*. 2011;39(7 Suppl):42–45.
15. Storm HH, Michelsen EV, Clemmensen IH, Pihl J. The Danish Cancer Registry – history, content, quality and use. *Dan Med Bull*. Nov 1997;44(5):535–539.
16. AJCC. *Cancer Staging Manual*. 6th ed: Springer-Verlag; 2002.
17. Pedersen CB. The Danish Civil Registration System. *Scand J Public Health*. 2011;39(7 Suppl):22–25.
18. Bjerregaard B, Larsen OB. The Danish Pathology Register. *Scand J Public Health*. 2011;39(7 Suppl):72–74.
19. Lyng E, Sandegaard JL, Rebolj M. The Danish National Patient Register. *Scand J Public Health*. 2011;39(7 Suppl):30–33.
20. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373–383.
21. Thygesen SK, Christiansen CF, Christensen S, Lash TL, Sorensen HT. The predictive value of ICD-10 diagnostic coding used to assess Charlson comorbidity index conditions in the population-based Danish National Registry of Patients. *BMC Med Res Methodol*. 2011;11:83.
22. Yu JB, Gross CP, Wilson LD, Smith BD. NCI SEER public-use data: applications and limitations in oncology research. *Oncology (Williston Park)*. 2009;23(3):288–295.
23. Klassen AC, Curriero F, Kulldorff M, Alberg AJ, Platz EA, Neloms ST. Missing stage and grade in Maryland prostate cancer surveillance data, 1992–1997. *Am J Prev Med*. 2006;30(2 Suppl):S77–S87.
24. Koroukian SM, Xu F, Beaird H, Diaz M, Murray P, Rose JH. Complexity of care needs and unstaged cancer in elders: a population-based study. *Cancer Detect Prev*. 2007;31(3):199–206.

## Appendix I

### Stage categories for SCLC and NSCLC allowing for specific missing T, N, or M values

#### Small cell lung cancer

Limited disease

T0–4,x N0–3,x M0

Extensive disease

T0–4,x N0–3,x M1

Unknown

T0–4,x N0–N3,x Mx

#### Non-small cell lung cancer

Limited disease

T0–2 N0–1 M0 = Stadium IA–IIB

Extensive disease

T0–2 N2 M0 = Stadium IIIA

T3 N0–2 M0 = Stadium IIIA (& IIB)

Distant disease

T4 N0–3,x M0,x = Stadium IV

T0–3,x N3 M0,x = Stadium IIIB

T0–4,x N0–3,x M1 = Stadium IV

Unknown

Tx N0–2,x M0,x

T0–3 Nx M0

T0–3 N0–2,x Mx

### Clinical Epidemiology

#### Publish your work in this journal

Clinical Epidemiology is an international, peer-reviewed, open access journal focusing on disease and drug epidemiology, identification of risk factors and screening procedures to develop optimal preventative initiatives and programs. Specific topics include: diagnosis, prognosis, treatment, screening, prevention, risk factor modification, systematic

Submit your manuscript here: <http://www.dovepress.com/clinical-epidemiology-journal>

reviews, risk & safety of medical interventions, epidemiology & biostatistical methods, evaluation of guidelines, translational medicine, health policies & economic evaluations. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use.

Dovepress