

# Population-based Aarhus Sarcoma Registry: validity, completeness of registration, and incidence of bone and soft tissue sarcomas in western Denmark

Katja Maretty-Nielsen<sup>1,2</sup>  
Ninna Aggerholm-Pedersen<sup>1–3</sup>  
Johnny Keller<sup>1,4</sup>  
Akmal Safwat<sup>1,3</sup>  
Steen Baerentzen<sup>1,5</sup>  
Alma B Pedersen<sup>6</sup>

<sup>1</sup>Sarcoma Centre of Aarhus University Hospital, <sup>2</sup>Department of Experimental Clinical Oncology, <sup>3</sup>Department of Oncology, <sup>4</sup>Department of Orthopedic Surgery, <sup>5</sup>Department of Pathology, <sup>6</sup>Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark

**Background:** The aim of the present study was to validate the data in the Aarhus Sarcoma Registry (ASR), to determine if this registry is population-based for western Denmark, and to examine the incidence of sarcomas using validated, population-based registry data.

**Methods:** This study was based on patients with bone and soft tissue sarcoma treated at the Sarcoma Centre of Aarhus University Hospital between January 1, 1979 and December 31, 2008. The validation process included a review of all medical files by two researchers using a standardized form. The Danish Cancer Registry was used as a reference to assess the completeness of registration of patients in the ASR. Crude and World Health Organization age-standardized incidence, as well as age-, gender-, and year-specific incidences were estimated.

**Results:** The validation process added 385 to the 1442 patients who were registered in the ASR. Before validation, on average, 70.5% of the data for the variables was correct. Validation improved the average completeness of the registered variables from 83.7% to 99.3%. The 1827 patients in the ASR after validation include 85.3% of the patients registered in the Danish Cancer Registry. The overall World Health Organization age-standardized incidence of sarcoma in the trunk or extremities in western Denmark in the period 1979–2008 was 2.2 per 100,000, being 0.8 for bone sarcomas and 1.4 for soft tissue sarcomas.

**Conclusion:** The validation process significantly improved the completeness of the variables and the quality of the ASR data. ASR is now a valuable population-based tool for epidemiological research and quality improvement in the treatment of sarcoma. It is our recommendation that documented validation of registries should be a prerequisite for publishing studies derived from them.

**Keywords:** clinical databases, cancer registry, sarcoma, incidence, completeness, validity

## Introduction

Sarcomas comprise a heterogeneous group of rare tumors, and detailed data are important for predicting prognosis as well as monitoring and improving quality of treatment. These data can be obtained from appropriately designed clinical databases.

Clinical databases hold data on disease characteristics, possible prognostic factors, treatment, and follow-up, as well as the indicators used to monitor quality of treatment. Use of such databases ensures large sample sizes, long follow-up periods, and high external validity. Validation of data in the few existing sarcoma databases worldwide is either not reported or not done, although it is a crucial factor determining the quality and value of the results reported. Even when it is reported to be done, the

Correspondence: Katja Maretty-Nielsen  
Department of Experimental Clinical  
Oncology, Aarhus University Hospital,  
Noerrebrogade 44,  
8000 Aarhus C, Denmark  
Tel +45 7846 2621  
Fax +45 8619 7109  
Email [katja.maretty-nielsen@ki.au.dk](mailto:katja.maretty-nielsen@ki.au.dk)

level of validation may be unclear and is suspected to be on a group level.<sup>1-3</sup>

In Denmark there is a unique opportunity to validate data on the individual level. Further, use of clinical databases can be extended by unambiguous linkage of data with other Danish population-based data sources in order to achieve complete follow-up of patients and detailed data on, eg, incidence of diseases and changes over time. A preliminary study from the Sarcoma Centre of Aarhus University Hospital reported an incidence of 16 soft tissue sarcomas per million; however, the incidence of sarcoma has not previously been determined systematically in Denmark.<sup>4,5</sup>

The aim of the present study was to validate data in the Aarhus Sarcoma Registry (ASR), to assess if the ASR is population-based for western Denmark, and to examine the incidence of sarcoma using validated population-based registry data.

## Materials and methods

### Danish health care system

The population of Denmark is approximately 5.5 million.<sup>6</sup> The health care system provides tax-supported health care for all residents, allowing free access to public and private general practitioners and hospitals. Since 1968, all citizens in Denmark have been assigned a unique 10-digit civil personal registration (CPR) number, which goes throughout all the Danish administrative registries and clinical databases, allowing for unambiguous linkage on an individual level and tracking of patients who have died, emigrated, or been admitted to different hospitals.<sup>7</sup>

### Data sources

Since 1979, the treatment of patients with sarcoma in western Denmark (population 2.5 million) has been carried out at the Sarcoma Centre of Aarhus University Hospital, resulting in initiation of the ASR. All patients treated for soft tissue sarcoma, bone sarcoma, some borderline (World Health Organization [WHO] classification: intermediate malignancy, locally aggressive and rarely metastasizing)<sup>8</sup> and benign tumors, eg, desmoid-type fibromatosis and aneurysmal bone cyst, have been registered in the ASR. The ASR collects basic patient data, including CPR number, gender, county of residence, date of diagnosis, specific data on tumor characteristics and treatment, including tumor size, localization, histological type, tumor grade, stage of disease, date and type of treatment, as well as data on follow-up examinations, local recurrence, distant metastases, and death. From 1993, data were registered prospectively.

The Danish Cancer Registry has recorded all incident cases of cancer in Denmark since 1943. The main variables are CPR number, date of diagnosis, clinical stage, initial treatment, topography codes according to the tenth version of the International Classification of Diseases (ICD-10), and morphology codes according to the third version of the International Classification of Diseases for Oncology (ICD-O-3). Until 1987, reporting to the Danish Cancer Registry was voluntary, but then became mandatory for doctors in hospital departments and private medical specialists.<sup>9,10</sup> Since 2004, registration has been done electronically and the completeness of patient registration and validation of data has been ensured by cross-referencing data from the Danish Cancer Registry to data from three national registries: the National Patient Registry, which contains data on all patients admitted to any hospital department, including discharge diagnoses; the Danish Pathology Registry, which contains histological diagnosis, anatomical localization, and date at diagnosis for both benign and malignant pathology; and the Danish Cause of Death Registry, which contains date and immediate and underlying cause of death.

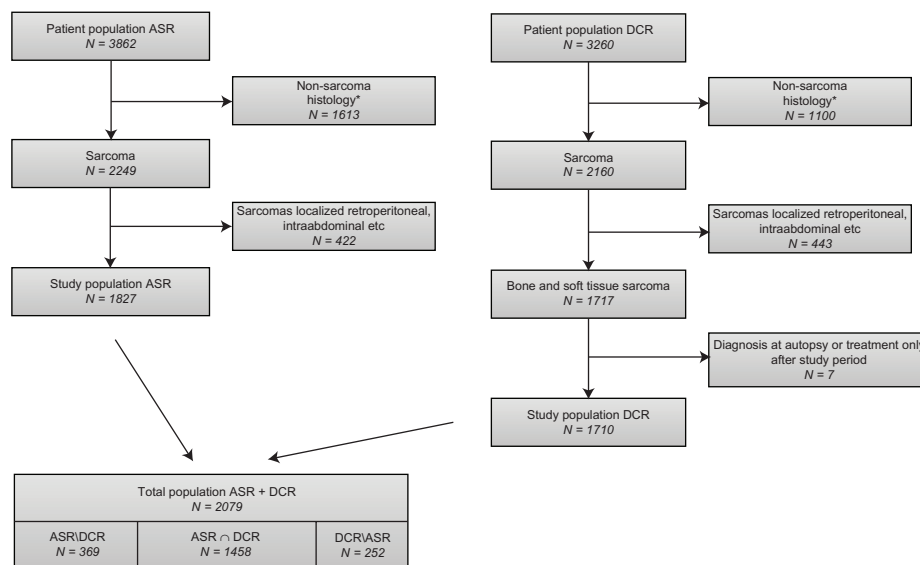
Data on the population size in western Denmark was obtained from StatBank Denmark, a database containing detailed statistical information on Danish society, including number of citizens per calendar year.<sup>11</sup>

### Study population

Patients were included in the study if they fulfilled the following inclusion criteria: residence in western Denmark, a diagnosis of sarcoma, treatment at the Sarcoma Centre of Aarhus University Hospital in the period between January 1, 1979 and December 31, 2008, and/or registration in the ASR. Patients with nonsarcoma pathology, or borderline or benign tumors, were excluded from our analyses (Appendix A). Furthermore, we excluded patients with tumors not located in the extremities or trunk (Appendix B). The remaining 1827 (47.3%) patients with bone sarcoma and soft tissue sarcoma in the trunk or extremities formed our study population (Figure 1).

### Validation of ASR

Initially, a revision of the variables registered and the registration forms used in the ASR was made. A random sample was selected using the first two digits of the CPR numbers, indicating the day of birth, and choosing the 100 patients with the lowest numbers. Based on a revision of these patients and a review of the literature regarding



**Figure 1** Flow chart for patients registered in the ASR and DCR in the period 1979–2008, with number of patients (N), reasons for exclusion, and distribution across registries. Includes patients only registered in the ASR (ASR/DCR), patients only registered in the DCR (DCR/ASR), and patients registered in both registries (ASR ∩ DCR). **Notes:** \*Based on the WHO ICD-03 codes. Non-sarcoma pathology includes benign, borderline and some malignant tumors (Appendix A). This population is not complete. **Abbreviations:** ASR, Aarhus Sarcoma Registry; DCR, Danish Cancer Registry; WHO, World Health Organization.

prognostic factors, standardized registration forms were designed in cooperation with a surgeon, an oncologist, and a pathologist with expertise in sarcoma.

Medical files, considered to be the gold standard for this type of research, were reviewed for all patients registered in the ASR by two independent researchers not involved in treatment using the standardized forms. Based on medical files from both the Department of Orthopedic Surgery and the Department of Oncology, data that were missing or incorrect in the ASR were added or corrected. All uncertain cases were discussed with the same team of sarcoma experts and a consensus was reached.

Finally, 385 patients with sarcoma were identified either through the Department of Pathology or in the Danish Cancer Registry (DCR)<sup>10</sup> as treated at the Sarcoma Centre of Aarhus University Hospital in the study period but yet not registered in the ASR. These medical records were reviewed in the same manner and entered in the ASR.

To assess whether or not the validation process has resulted in substantial changes in the quality of the ASR, a data extract including important variables selected from the ASR before and after validation was used. The proportion of correctly registered variables before validation was assessed. Correctness of the data was defined as the amount of registered data not changed during validation divided by the number of registered data before validation. Completeness of data in each selected variable before and after validation was assessed and compared. Completeness of data was defined

as the amount of registered data divided by the number of registered patients.

Completeness > 90% was considered satisfactory for the important variables selected as well as patient registration (described below), with ≤3% missing values in the demographic data, in agreement with the requirements for clinical databases approved by the Danish National Board of Health and the North American Association of Central Cancer Registries.<sup>12,13</sup>

## Completeness of patient registration in ASR

The completeness of patient registration in the ASR was assessed using the DCR as a reference. Data were retrieved from the DCR, including 2260 patients living in western Denmark at the time of diagnosis with a tumor located in soft tissue or bone, based on the following ICD-10 codes: C40, C41, C47, and C49. These data were reviewed by one pathologist with expertise in sarcoma, and 1100 patients with nonsarcoma pathology (Appendix A) were excluded from the DCR.<sup>10,14</sup> Furthermore, based on the Danish Pathology Registry, patients with tumors not located in the extremities or trunk were excluded, as in the ASR (Appendix B). Finally, patients in whom sarcoma was found incidentally at autopsy and patients treated only after the study period were excluded from the DCR (Figure 1). Data from the DCR was linked on an individual level with data from the ASR. The completeness of patient registration in the ASR (P) was estimated as the

number of the patients registered in both the ASR and DCR ( $ASR \cap DCR$ ) divided by the total number of patients registered in the DCR (DCR):

$$P = \frac{ASR \cap DCR}{DCR}$$

Differences in completeness of patient registration according to gender, age group (<20, 20–39, 40–59, and  $\geq 60$  years), region (south, central, and north), and 5-year periods (1979–1983, 1984–1988, 1989–1993, 1994–1998, 1999–2003, and 2004–2008) were examined using the Chi-squared test, with a *P* value less than 0.05 considered to be statistically significant.

The number of expected patients with bone and soft tissue sarcoma not registered in either the ASR or the DCR (*N*) was determined using the capture-recapture method,<sup>13,14</sup> and defined as the number of patients only registered in the ASR ( $ASR/DCR$ ) multiplied by the number of patients only registered in the DCR ( $DCR/ASR$ ) divided by the number of patients registered in both the ASR and DCR:

$$N = \frac{ASR/DCR * DCR/ASR}{ASR \cap DCR}$$

## Incidence

Incidence was calculated using data from the ASR and StatBank Denmark, expressed per 100,000 inhabitants per year with 95% confidence intervals (CI). Crude and age-standardized incidence rates using the WHO standard population were calculated as overall, and separately for bone sarcoma and soft tissue sarcoma.<sup>15</sup>

The changes in age-standardized 5-year incidence rate for gender and age were assessed as an incidence rate ratio with 95% CI using Poisson regression, and a *P* value of 0.05 was considered to be statistically significant. In addition, the age-specific and gender-specific incidence rate was calculated as overall and sarcoma-specific.<sup>11</sup>

## Ethical approval

This study was approved by the Danish Data Protection Agency (2007-58-0010) and the Danish Health and Medicines Authority (7-604-04-2/262/KWH).

## Results

### ASR patient demographics

In total, 1827 patients with soft tissue sarcoma or bone sarcoma were treated at the Sarcoma Centre of Aarhus

University Hospital in the period 1979–2008. The majority of patients (824, 45.0%), were diagnosed in the 1999–2008 period compared with 446 (24.4%) and 557 (30.5%) in 1979–1988 and 1989–1998, respectively. The median age at diagnosis was 53.4 years (range 0.1–95.5), with 1013 (55.5%) males compared with 814 (44.5%) females. There were 1275 patients (70.0%) with soft tissue sarcoma and 552 (30.0%) with bone sarcoma.

### Validity of ASR

Validation increased the number of patients with bone sarcoma and soft tissue sarcoma in the ASR from 1442 to 1827 (an increase of 26.7%), and completeness of the registered variables increased from 83.7% (95% CI 83.1–84.2) to 99.3% (95% CI 99.2–99.4), as shown in Table 1. Before validation, the proportion of correctly registered data for the variables varied from 25.3% (95% CI 22.4–28.2) to 97.7% (95% CI 96.7–98.5), with an average of 70.5% (95% CI 69.8–71.2), Table 1.

### Completeness of patient registration in ASR

We identified 2078 patients with bone sarcoma or soft tissue sarcoma in western Denmark in the period 1979–2008 who were registered in both the ASR and DCR (Figure 1), giving an overall completeness of registration in the ASR of 85.3% (95% CI 83.5–86.9). A total of 369 patients were registered in the ASR but not in the DCR. There was no difference in completeness according to gender, but completeness was significantly greater for those younger than 60 years than for those older than 60 years, in patients from the central region compared with the north and south regions and in patients diagnosed after 1994 compared with those diagnosed before 1994 (Table 2).

The expected number of patients with bone sarcoma or soft tissue sarcoma in the trunk or extremities in western Denmark in 1979–2008 not registered in either the ASR or DCR was 64. Of these, 40 patients would be expected in 1979–1988, while 18 and nine would be expected in 1989–1998 and 1999–2008, respectively.

### Overall incidence

The crude overall incidence rate of sarcoma in the ASR for the period 1979–2008 was 2.5 (95% CI 2.4–2.6) per 100,000 inhabitants per year. The WHO age-standardized overall incidence rate was 2.2 (95% CI 2.1–2.3). The crude and WHO age-standardized incidence rate was 0.8 (95% CI 0.7–0.9) and 0.8 (95% CI 0.7–0.8), respectively, for bone

**Table I** Completeness and correctness of data by selected variables in the Aarhus Sarcoma Registry

Variables	Completeness of data % (95% CI), n		Correct data % (95% CI), n
	Prevalidation	Postvalidation	Prevalidation
Date of admission	81.8 (79.7–83.8) 1180/1442	100.0 (99.8–100.0) 1827/1827	43.2 (40.4–46.1) 510/1180
Duration of symptoms	88.1 (86.3–89.7) 1270/1442	95.9 (94.9–96.8) 1752/1827	50.2 (47.4–53.0) 638/1270
Cause of referral	81.1 (78.9–83.1) 1169/1442	100.0 (99.8–100.0) 1827/1827	93.7 (92.1–95.0) 1095/1169
Tumor size	79.1 (76.9–81.2) 1141/1442	95.5 (94.4–96.4) 1744/1827	42.7 (39.8–45.6) 487/1141
Anatomic localization			
Bone sarcoma	77.6 (72.3–81.7) 287/370	100.0 (99.3–100.0) 552/552	94.1 (90.7–96.5) 270/287
Soft tissue sarcoma	81.8 (79.4–84.1) 877/1072	100.0 (99.7–100.0) 1275/1275	86.1 (83.6–88.3) 755/877
Compartmental localization	79.3 (77.1–81.3) 1143/1442	100.0 (99.8–100.0) 1827/1827	73.0 (70.3–75.5) 834/1143
Histological subtype			
Bone sarcoma	77.8 (73.3–82.0) 288/370	100.0 (99.3–100.0) 552/552	64.2 (58.4–69.8) 185/288
Soft tissue sarcoma	81.0 (78.5–83.3) 868/1072	100.0 (99.7–100.0) 1275/1275	79.4 (76.5–82.0) 689/868
Grade of malignancy	95.6 (94.4–96.6) 1378/1442	100.0 (99.8–100.0) 1827/1827	53.4 (50.7–56.1) 736/1378
Surgery <sup>a</sup>			
Date	68.1 (65.5–70.6) 895/1315	100.0 (99.8–100.0) 1564/1564	25.3 (22.4–28.2) 226/895
Type	80.7 (78.4–82.8) 1061/1315	100.0 (99.8–100.0) 1564/1564	97.7 (96.7–98.5) 1037/1061
Margin	79.9 (77.7–82.1) 1051/1315	99.9 (99.6–100.0) 1563/1564	76.0 (73.3–78.6) 799/1051
Radiotherapy <sup>b</sup>	95.6 (94.4–96.6) 1379/1442	100.0 (99.8–100.0) 1827/1827	90.5 (88.8–92.0) 1248/1379
Chemotherapy <sup>b</sup>	95.6 (94.4–96.6) 1379/1442	100.0 (99.8–100.0) 1827/1827	95.9 (94.8–96.9) 1323/1379

**Notes:** <sup>a</sup>Subgroup includes only patients treated with surgery (prevalidation, n = 1315; postvalidation, n = 1564); <sup>b</sup>data on whether or not the patient was treated with radiotherapy and chemotherapy.

**Abbreviation:** CI, confidence interval.

sarcoma and 1.8 (95% CI 1.7–1.9) and 1.4 (95% CI 1.3–1.5) for soft tissue sarcoma, respectively.

The incidence rate for bone sarcoma in both genders was highest at age 10–24 years and at age 75–79 years. The highest incidence rate for soft tissue sarcoma was found at age 70–74 years for females and age 85–90 years for males (Figure 2).

The WHO age-standardized incidence rate for bone sarcoma decreased with increasing age, with an incidence rate ratio of 0.30 (95% CI 0.31–0.62,  $P < 0.001$ ) for females and 0.59 (95% CI 0.42–0.82,  $P = 0.002$ ) for males older than 60 years versus younger than 20 years. The opposite was found for soft tissue sarcoma, for which the WHO age-standardized incidence rate increased with increasing age, with an incidence rate ratio of 4.46 (95% CI 3.09–6.45,  $P = 0.000$ ) for females and 4.55 (95% CI 3.47–5.97,

$P < 0.001$ ) for males (Figure 2). The overall incidence rate ratio was 1.3 (95% CI 1.2–1.5,  $P < 0.001$ ) for males compared with females.

## Change in incidence rate over time

The WHO age-standardized overall incidence rate increased over the entire study period, with an incidence rate ratio of 1.80 (95% CI 1.51–2.15,  $P < 0.001$ ) in 2004–2008 compared with 1.00 (reference) in 1979–1983 (Figure 3). A steady increase in incidence rate was seen in soft tissue sarcoma, whereas the incidence rate of bone sarcoma only increased significantly in the last time period compared with the first time period (Figure 3). The increase in incidence rate over time was highest in the age group 40–59 years in males (incidence rate ratio 2.9; 95% CI 1.8–4.6,  $P < 0.001$ )



**Table 2** Completeness of patient registration in the Aarhus Sarcoma Registry by gender, age, region, and time period

	%	(95% CI)	n	P <sup>a</sup>
Gender				
Male	84.5	(82.1–86.7)	829/981	0.31
Female	86.3	(83.6–88.7)	629/729	
Age (years)				
<20	90.2	(85.3–93.9)	184/204	<0.0001
20–39	90.5	(86.5–93.6)	266/294	
40–59	89.6	(86.4–92.3)	396/442	
≥60	79.5	(76.5–82.3)	612/770	
Regions				
South	76.9	(73.2–80.4)	423/550	<0.0001
Central	94.2	(92.2–95.8)	634/673	
North	82.3	(78.7–85.6)	401/487	
Time periods				
1979–1983	68.0	(61.5–73.9)	157/231	<0.0001
1984–1988	76.8	(70.9–81.9)	185/241	
1989–1993	77.4	(71.9–82.3)	206/266	
1994–1998	89.8	(85.7–93.1)	255/284	
1999–2003	95.9	(93.1–97.8)	305/318	
2004–2008	94.6	(91.8–96.7)	350/370	
Total	85.3	(83.5–86.9)	1458/1710	

**Note:** <sup>a</sup>P values based on the Chi-square test.

**Abbreviation:** CI, confidence interval.

and 20–39 years in females (incidence rate ratio 3.5; 95% CI 1.8–6.8,  $P < 0.001$ ) in 2004–2008 compared with 1979–1983 (Table 3).

## Discussion

### Validity of ASR

The validation process eliminated incorrect registration and values missing from the ASR. This emphasizes the importance of validation, as well as ensuring the completeness of patient registration by continuous cross-referencing to other data sources. The overall completeness of the variables registered in the ASR is now considered to be very satisfactory.

### Completeness of patient registration in ASR

Validation increased the number of patients registered in the ASR by 26.7%. The overall completeness of patient registration in the ASR of 85.3% appears to be in agreement with the completeness reported for the few other sarcoma registries. For example, the Scandinavian Sarcoma Group Register, a registry including sarcoma patients from Sweden, Norway, and Finland, reported a completeness of at least 90%.<sup>1–3</sup> However, it was not possible to determine whether the analyses of patient completeness for other sarcoma registries were based on individual or group levels. A comparison on the group

level may result in misleading estimates, because it would be unclear whether an individual patient was registered in both data sources.

Completeness was improved after 1993, which is probably attributable to retrospective registration of patients before 1993 and prospective registration afterwards. Some of the old medical files were lost, owing to the legal obligation in Denmark to store medical files for 10 years only, making them difficult to register and validate in the ASR. Furthermore, it is only recently that a tradition of centralizing the treatment of rare diseases, such as sarcomas, has been established in Denmark.

The lesser completeness in patient  $\geq 60$  years may be explained by the fact that comorbidity and/or metastasis in this age group would mean no treatment strategy, and thus not being referred to the Sarcoma Centre of Aarhus University Hospital. The lesser completeness among patients 60 years and older could be a source of selection bias in this study in relation to estimates of incidence as well as in future survival studies based on ASR patients. However, the size of the possible selection bias would be expected to be low.

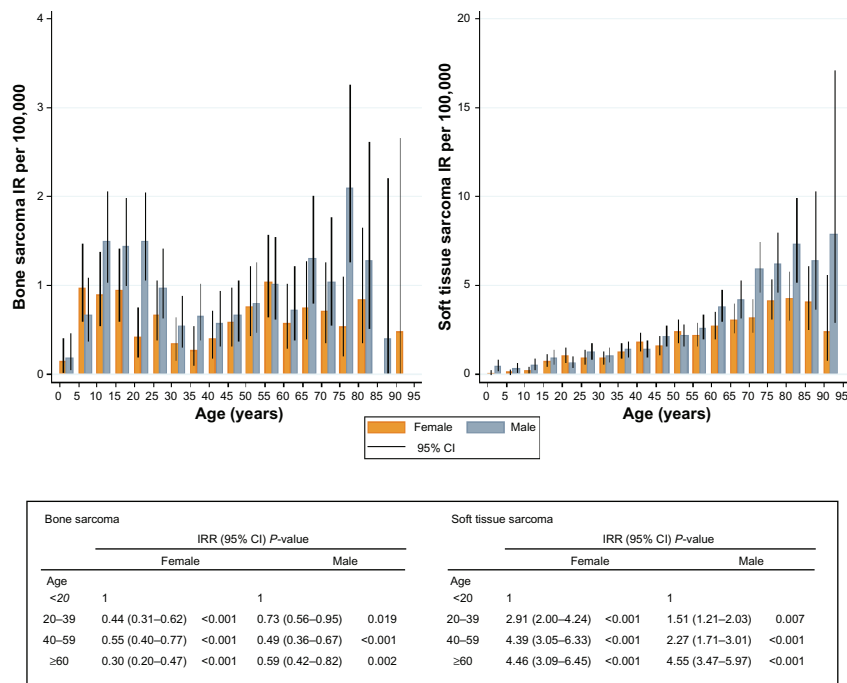
The lesser completeness of data in the southern region is most likely explained by the fact that some patients in this region were referred incorrectly to a nearer sarcoma centre in Odense. The number of patients not registered in either the ASR or DCR was estimated to be approximately two patients per year. Although ideally this number should be 0, we consider it acceptable and conclude that ASR is a population-based registry for western Denmark.

### Incidence

The estimated crude and age-standardized incidence rates of 2.5 and 2.2 per 100,000 were in general agreement with results from other studies; however, comparability is difficult because of differences in the methods of reporting data and inclusion criteria, such as age, histological subtypes, and anatomical localization.<sup>16</sup> Two European studies that also included visceral sarcomas reported an overall incidence rate of 2.9 and 6.2 sarcomas per 100,000 per year in 1982–1984 and 2005–2007, respectively.<sup>17,18</sup> A Swedish study of soft tissue sarcoma excluding visceral sites reported an incidence of 1.8 per 100,000 per year in the period 1964–1989, consistent with our findings.<sup>19</sup>

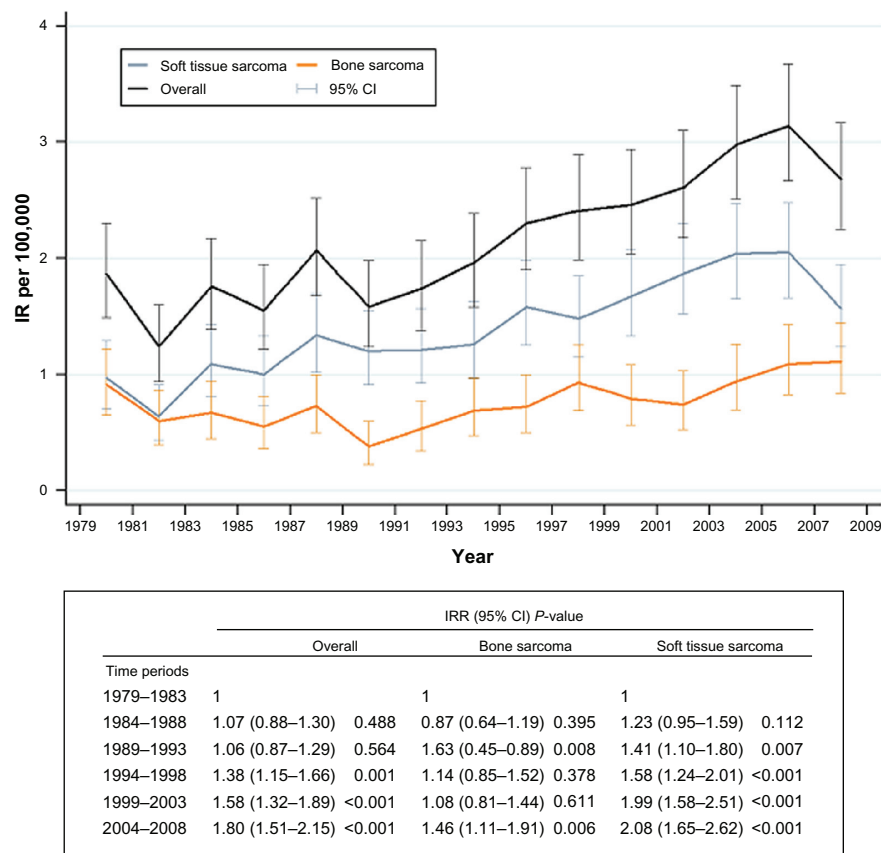
The increase in incidence found with higher age and male gender was also in accordance with results from other research.<sup>1,17,18,20,21</sup>

Other studies<sup>1,20</sup> have shown an increase in incidence in soft tissue sarcoma over time, consistent with our findings.



**Figure 2** Crude IR per 100,000 inhabitants and World Health Organization age-standardized IRR, with 95% CI and *P* values, for bone and soft tissue sarcoma in the Aarhus Sarcoma Registry from 1979 to 2008 according to age (years) and gender (*n* = 1827).

**Abbreviations:** CI, confidence intervals; IR, incidence rate; IRR, incidence rate ratios.



**Figure 3** World Health Organization age-standardized IR per 100,000 inhabitants and IRR, with 95% CI and *P* values for overall, bone, and soft tissue sarcoma in the Aarhus Sarcoma Registry from 1979 to 2008 (*n* = 1827).

**Abbreviations:** CI, confidence intervals; IR, incidence rate; IRR, incidence rate ratios.

**Table 3** Age-specific and gender-specific incidence rates per 100,000 inhabitants per year by time period

Age, years	Incidence rate (95% CI)					
	1979–1983	1984–1988	1989–1993	1994–1998	1999–2003	2004–2008
Male						
<20	1.3 (0.8–2.0)	1.5 (0.9–2.2)	0.8 (0.4–1.4)	2.0 (1.3–2.8)	1.9 (1.2–2.7)	1.7 (1.1–2.4)
20–39	2.0 (1.4–2.7)	1.7 (1.2–2.4)	1.5 (1.0–2.2)	2.4 (1.8–3.3)	2.2 (1.6–3.1)	2.1 (1.5–3.0)
40–59	1.6 (1.0–2.5)	1.9 (1.3–2.8)	2.0 (1.4–2.9)	2.6 (1.9–3.5)	3.2 (2.4–4.1)	4.7 (3.7–5.8)
≥60	4.3 (3.1–5.8)	5.3 (4.0–6.9)	4.8 (3.6–6.4)	6.0 (4.6–7.7)	7.3 (5.8–9.1)	9.2 (7.6–11.1)
Total	2.1 (1.7–2.5)	2.3 (2.0–2.8)	2.1 (1.7–2.5)	3.0 (2.6–3.5)	3.3 (2.9–3.8)	4.1 (3.6–4.7)
Female						
<20	1.1 (0.6–1.7)	1.0 (0.6–1.7)	0.9 (0.4–1.5)	0.6 (0.3–1.2)	1.3 (0.7–2.0)	1.3 (0.8–2.0)
20–39	0.6 (0.3–1.2)	0.9 (0.5–1.5)	1.3 (0.8–1.9)	1.9 (1.3–2.7)	1.8 (1.2–2.5)	2.2 (1.6–3.1)
40–59	2.0 (1.3–2.9)	1.9 (1.3–2.8)	2.6 (1.9–3.6)	2.4 (1.7–3.3)	3.2 (2.4–4.1)	3.4 (2.6–4.4)
≥60	2.9 (2.0–3.9)	2.9 (2.1–4.0)	3.6 (2.7–4.8)	3.7 (2.8–4.9)	4.9 (3.8–6.3)	5.5 (4.4–6.9)
Total	1.5 (1.2–1.9)	1.6 (1.3–2.0)	2.1 (1.7–2.4)	2.2 (1.8–2.6)	2.7 (2.3–3.2)	3.1 (2.7–3.6)
Overall	1.8 (1.6–2.1)	2.0 (1.7–2.3)	2.1 (1.8–2.3)	2.6 (2.3–2.9)	3.0 (2.7–3.3)	3.6 (3.3–4.0)

**Abbreviation:** CI, confidence interval.

However, it can be questioned whether our results represent a true change in incidence or rather reflect an increase in completeness of patient registration in the ASR. The latter can be supported by the fact that the expected number of patients not registered in either the ASR or the DCR dropped by more than four-fold when comparing the first with the last time period. However, even when accounting for the expected number of non registered patients in the three 10-year time periods (1979–88, 1989–98 and 1999–2008), respectively, the total number of patients increased over time (579 + 40 versus 642 + 18 versus 858 + 9), which could indicate a real change in incidence. The apparent decrease in incidence in the last 2 years of the study period could be explained by common variation or possibly a small diagnostic delay, ie, patients diagnosed with sarcoma retrospectively at revision of histopathology later on.

## Methodological considerations

The risk of information bias in this study was low, given that the medical files were systematically reviewed using standardized forms and by only two persons in close collaboration. A manual with definitions of variables was elaborated, and all uncertain cases were discussed by a multidisciplinary team, and consensus was reached. Medical files were used to validate data in the ASR. Even though there is a possibility that correct data in the ASR were replaced by incorrect data from the medical files, this is considered very unlikely given the fact that medical files are generally dictated immediately after the consultation, whereas registration in the ASR is more likely to be done later on.

Although DCR was considered to be the gold standard for assessing completeness of patient registration in ASR,

we found 369 cases in the ASR which were not registered in the DCR. Differences in patient registration in the ASR and DCR may be attributable to various factors: firstly, incorrect diagnoses, where the patient is initially diagnosed with sarcoma and registered, while not deleted when a later revision shows nonsarcoma histology; secondly, differences in inclusion criteria, where some borderline or low-grade sarcoma types are included and registered in, eg, the ASR but not in the DCR; thirdly, missing registrations, where it is especially important that reporting to the DCR until 1987 was voluntary, resulting in less patient completeness in the DCR for that period. However, even though numerous studies have shown high quality of the data in the DCR, our results show that the DCR is not a perfect reference for sarcoma, and validation of the sarcoma data in the DCR is warranted.<sup>22–25</sup>

## Conclusion

The incidence of bone sarcoma and particularly soft tissue sarcoma increased substantially during the period 1979–2008. The increase in incidence rate over time is gender-dependent and age-dependent. The validation process significantly improved the completeness of variables and the quality of the ASR data. The ASR is now a population-based, validated, and valuable tool for epidemiological research and improvement in quality of treatment for sarcoma. It is our recommendation that documented validation of registries should be a prerequisite for publishing studies derived from them.

## Acknowledgments

The study was supported by grants from the “Frits, Georg and Marie Cecilie Gluds legat”, the “Max and Inge Wørzners mindelegat”, the Danish Council for Independent Research/



Medical Sciences, and Aarhus University. We extend our sincere thanks to the orthopedic surgeons and oncologists from the Aarhus Sarcoma Centre who register data in the Aarhus Sarcoma Registry, as well as the staff from the archives for locating the medical files reviewed in this study.

## Disclosure

The authors report no conflicts of interest in this work.

## References

- Gutierrez JC, Perez EA, Franceschi D, Moffat FL Jr, Livingstone AS, Koniaris LG. Outcomes for soft-tissue sarcoma in 8249 cases from a large state cancer registry. *J Surg Res*. 2007;141(1):105–114.
- Bauer HC, Ahlén J, Alvegard TA, et al. The Scandinavian Sarcoma Group register 1986–2008. *Acta Orthop Scand Suppl*. 2009;80(334):13–14.
- Bauer HC, Trovik CS, Alvegard TA, et al. Monitoring referral and treatment in soft tissue sarcoma: study based on 1,851 patients from the Scandinavian Sarcoma Group register. *Acta Orthop Scand*. 2001;72(2):150–159.
- Vraa S, Keller J, Nielsen OS, Sneppen O, Jurik AG, Jensen OM. Prognostic factors in soft tissue sarcomas: the Aarhus experience. *Eur J Cancer*. 1998;34(12):1876–1882.
- Jensen OM, Hogh J, Ostgaard SE, Nordentoft AM, Sneppen O. Histopathological grading of soft tissue tumours. Prognostic significance in a prospective study of 278 consecutive cases. *J Pathol*. 1991;163(1):19–24.
- Statistics Denmark. Population Denmark. Available from: <http://www.dst.dk/en/Statistik/emner/befolkning-og-befolkningsfremskrivning/folketal.aspx>. Accessed February 3, 2013.
- The Danish Civil Registry. Danish Act on the Civil Registration System. <http://cpr.dk/cpr/site.aspx?p=194&ArticleID=4326>. Accessed February 3, 2013.
- Fletcher CDM, Unni KK, Mertens F; World Health Organization, International Academy of Pathology. *Pathology and Genetics of Tumours of Soft Tissue and Bone*. Lyon, France: IARC Press; 2002.
- Storm HH, Michelsen EV, Clemmensen IH, Pihl J. The Danish Cancer Registry – history, content, quality and use. *Dan Med Bull*. 1997;44(5):535–539.
- The Danish Pathology Registry. SNOMED codes. Available from: <http://patobank.dk/>. Accessed February 3, 2013.
- StatBank Denmark. Available from: <http://statbank.dk/statbank5a/default.asp?w=1280>. Accessed February 3, 2013.
- Hansen MT. Basic requirements for national clinical databases. Available from: [http://www.regioner.dk/Sundhed/Fonde+vejledninger+og+databaser/Databaser/~media/Filer/IT%20og%20Kvalitet/Kliniske%20Kvalitetsdatabaser/11.05.2007\\_basiskrav\\_for\\_landsdaekkende\\_kliniske\\_kvalitetsdatabaser.pdf.ashx](http://www.regioner.dk/Sundhed/Fonde+vejledninger+og+databaser/Databaser/~media/Filer/IT%20og%20Kvalitet/Kliniske%20Kvalitetsdatabaser/11.05.2007_basiskrav_for_landsdaekkende_kliniske_kvalitetsdatabaser.pdf.ashx). Accessed February 3, 2013.
- Hofferkamp J. North American Association of Central Cancer Registries. Available from: <http://www.naaccr.org/LinkClick.aspx?fileticket=hvFzJKUcRM8%3d&tabid=134&mid=474>. Accessed February 3, 2013.
- National Cancer Institute. ICD-O-3 histology codes. Available from: <http://seer.cancer.gov/icd-o-3/>. Accessed February 3, 2012.
- Ahmad O, Boschi-Pinto C, Lopez A, Murray C, Lozano R, Inoue M. Age standardization of rates: a new WHO standard. Available from: <http://www.who.int/healthinfo/paper31.pdf>. Accessed February 3, 2013.
- Howlader N, Noone AM, Krapcho M, et al. SEER cancer statistics review. Available from: [http://seer.cancer.gov/csr/1975\\_2009\\_pops09/](http://seer.cancer.gov/csr/1975_2009_pops09/). Accessed February 3, 2013.
- Hartley AL, Blair V, Harris M, et al. Sarcomas in North West England: II. Incidence. *Br J Cancer*. 1991;64(6):1145–1150.
- Ducimetiere F, Lurkin A, Ranchere-Vince D, et al. Incidence of sarcoma histotypes and molecular subtypes in a prospective epidemiological study with central pathology review and molecular testing. *PLoS One*. 2011;6(8):e20294.
- Gustafson P. Soft tissue sarcoma. Epidemiology and prognosis in 508 patients. *Acta Orthop Scand Suppl*. 1994;259(1):1–31.
- Toro JR, Travis LB, Wu HJ, Zhu K, Fletcher CD, Devesa SS. Incidence patterns of soft tissue sarcomas, regardless of primary site, in the surveillance, epidemiology and end results program, 1978–2001: an analysis of 26,758 cases. *Int J Cancer*. 2006;119(12):2922–2930.
- Mastrangelo G, Coindre JM, Ducimetiere F, et al. Incidence of soft tissue sarcoma and beyond: a population-based prospective study in 3 European regions. *Cancer*. 2012;118(21):5339–5348.
- Storm HH. Completeness of cancer registration in Denmark 1943–1966 and efficacy of record linkage procedures. *Int J Epidemiol*. 1988;17(1):44–49.
- Jensen AR, Overgaard J, Storm HH. Validity of breast cancer in the Danish Cancer Registry. A study based on clinical records from one county in Denmark. *Eur J Cancer Prev*. 2002;11(4):359–364.
- Gjerstorff ML. The Danish Cancer Registry. *Scand J Public Health*. 2011;39(Suppl 7):42–45.
- Kjaerbye-Thygesen A, Huusom LD, Frederiksen K, Kjaer SK; Danish Cancer Registry and the Pathology Data Bank. [Primary ovarian cancer. A comparison of registrations in the Danish Cancer Registry and the Pathology Data Bank]. *Ugeskr Laeger*. 2007;169(1):50–54. [Danish.]

## Appendices

**Appendix A** Histology subtypes, based on ICD-O-3 World Health Organization codes, excluded from both the Aarhus Sarcoma Registry and the Danish Cancer Registry in this study

Code	Text
80003	Neoplasm, malignant
80009	Neoplasm, malignant, uncertain whether primary or metastasis
80013	Tumor cells, malignant
80019	Tumor cells, uncertain whether primary or metastasis
80043	Malignant tumor, spindle cell type
80103	Carcinoma, NOS
80109	Carcinoma, uncertain whether primary or metastasis
80123	Large cell carcinoma, NOS
80203	Carcinoma, undifferentiated type, NOS
80223	Pleomorphic carcinoma
80413	Small cell carcinoma, NOS
80703	Squamous cell carcinoma, NOS
81403	Adenocarcinoma, NOS
81409	Adenocarcinoma, uncertain whether primary or metastasis
82403	Carcinoid tumor, NOS
82409	Carcinoid, malignant, uncertain whether primary or metastasis
82463	Neuroendocrine carcinoma
84413	Serous cystadenocarcinoma, NOS
84419	Serous cystadenocarcinoma, uncertain whether primary or metastasis
84613	Serous surface papillary carcinoma
84803	Mucinous adenocarcinoma
84813	Mucin-producing adenocarcinoma
84819	Mucin producing adenocarcinoma, uncertain whether primary or metastasis
86803	Paraganglioma, malignant
87203	Malignant melanoma, NOS
88013	Spindle cell tumor
88211	Aggressive fibromatosis, extra-abdominal desmoid tumor
88301	Atypical fibrous histiocytoma, atypical fibroxanthoma
88400	Myxoma, NOS
88500	Lipoma, NOS
89353	Stromal tumor, NOS
89363	Gastrointestinal stromal sarcoma
89513	Mesodermal mixed tumor
89803	Carcinosarcoma, NOS
90203	Phyllodes tumor, malignant
90503	Mesothelioma, malignant
90513	Fibrous mesothelioma, malignant
90523	Epithelioid mesothelioma, malignant
90533	Biphasic mesothelioma, malignant
90613	Seminoma, NOS
90619	Seminoma, uncertain whether primary or metastasis
90643	Germinoma
90649	Germinoma, uncertain whether primary or metastasis
90703	Embryonal carcinoma, NOS
90709	Embryonal carcinoma, uncertain whether primary or metastasis
90713	Yolk sac tumor
90719	Yolk sac tumor, uncertain whether primary or metastasis
90803	Teratoma, malignant, NOS
90813	Teratocarcinoma
90853	Mixed germ cell tumor
91003	Choriocarcinoma
91013	Choriocarcinoma combined with other germ cell elements
91200	Hemangioma
91403	Kaposi's sarcoma

(Continued)

**Appendix A** (Continued)

Code	Text
91800	Osteoma, NOS
91910	Osteoid osteoma, NOS
92100	Osteochondroma, cartilaginous exostosis
92200	Chondroma, NOS; enchondroma
92501	Giant cell tumor of bone, NOS
92510	Diffuse-type giant cell tumor
92511	Giant cell tumor of soft parts, NOS
93103	Ameloblastoma, malignant
93643	Peripheral neuroectodermal tumor
94903	Ganglioneuroblastoma
95033	Neuroepithelioma, NOS
95600	Neurilemmoma
95803	Granular cell tumor, malignant
95903	Malignant lymphoma, NOS
95913	Malignant lymphoma, non-Hodgkin, NOS
95919	Malignant lymphoma, non-Hodgkin, NOS, uncertain whether primary or metastasis
96503	Hodgkin lymphoma, NOS
96633	Hodgkin lymphoma, nodular sclerosis, NOS
96703	Malignant lymphoma, small B lymphocytic, NOS
96713	Malignant lymphoma, lymphoplasmacytic
96733	Mantle cell lymphoma
96753	Malignant lymphoma, mix small + large, diffuse (obsolete)
96803	Malignant lymphoma, large B-cell, diffuse, NOS
96843	Malignant lymphoma, large B-cell, diffuse, immunoblastic, NOS
96873	Burkitt lymphoma, NOS
96903	Follicular lymphoma, NOS
97023	Mature T-cell lymphoma, NOS
97313	Plasmacytoma, NOS
97323	Myelomatosis
97343	Plasmacytoma, extramedullary
97403	Mast cell sarcoma
97503	Malignant histiocytosis
97563	Langerhans cell sarcoma
97583	Follicular dendritic cell sarcoma
99303	Myeloid sarcoma
99903	No microscopy, clinical malignant tumor (cancer)
99909	No microscopy, uncertain whether primary or metastasis
99993	NA morphology, malignant
99999	NA morphology, uncertain whether primary or metastasis
	Aneurysmal bone cyst
	Eosinophilic granuloma
	Neoplasm with perivascular epithelioid cell differentiation
	Nodular fasciitis
	Nonossifying fibroma
	Simple bone cyst
	Small cell anaplastic sarcoma in patients > 30 years

**Abbreviations:** NOS, not otherwise specified; NA, not available.

**Appendix B** Topography in the Aarhus Sarcoma Registry and the Danish Cancer Registry excluded from the study

---

Adrenal gland  
 Cauda equina  
 Cerebrum  
 Cordis  
 Pouch of Douglas  
 Iliac fossa  
 Gaster  
 External genitalia  
 Thyroid gland  
 Head and neck  
 Hepatic  
 Intra-abdominal  
 Kidney  
 Lung  
 Mediastinum  
 Spinal medulla  
 Iliopsoas muscle  
 Omentum majus  
 Pelvis minor  
 Pericardium  
 Pleura  
 Prostate  
 Retroperitoneum  
 Sigmoid mesentery  
 Small intestine  
 Small intestine mesentery  
 Spleen  
 Ureter  
 Urethra  
 Uterus  
 Vena cava inferior  
 Vesica urinaria

---

**Clinical Epidemiology**

Dovepress

**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.