Scand J Work Environ Health 1987;13(1):1-91
Issue date: 1987

Occupation and risk of cancer in Denmark. An analysis of 93,810 cancer cases, 1970-1979.

by Olsen JH, Jensen OM

Affiliation: Institute of Cancer Epidemiology, Danish Cancer Society, Copenhagen, Denmark.

This article in PubMed: www.ncbi.nlm.nih.gov/pubmed/3659854

Print ISSN: 0355-3140 Electronic ISSN: 1795-990X

Scand J Work Environ Health 13 (1987): suppl 1, 91 p

Occupation and risk of cancer in Denmark

An analysis of 93 810 cancer cases, 1970—1979

by Jørgen H Olsen, MD, Ole M Jensen, MD

Danish Cancer Registry
Institute of Cancer Epidemiology
Danish Cancer Society
Copenhagen Denmark

PREFACE

As in most countries of the industrialized part of the world, cancer is one of the leading causes of death in Denmark. Some 25—30 % of all its inhabitants develop cancer before they reach the age of 75. There is therefore need for research to disentangle the causes of cancer so that preventive measures can be instituted.

For more than 200 years it has been known that exposures incurred during worklife may lead to an increased risk of cancer. Early observations were, for the most part, dependent on the clustering of rare cancers in certain occupations. The advance of medical information systems based on record linkage, in particular in the Nordic countries, has made it possible to undertake a comprehensive and systematic description of cancer risk in different occupations.

As part of the Danish Cancer Society's initiative in environmental cancer research, the Danish Cancer Registry has embarked on two large descriptive studies on occupational cancer based on linkage between the cancer register and the census and the Supplementary Pension Fund. The present publication concerns the study of the occupational history of persons with cancer. The publication is based on the linkage of cancer cases notified to the Danish Cancer Registry with occupational information from two computer-based national registries, the Supplementary Pension Fund and the Central Population Registry. The purpose of the linking of data has been to create a basis for the continuous monitoring of known occupational cancers in various groups or industries in Denmark, to search for previously unsuspected hazards, and to provide a basis for hypothesis-testing case-referent studies.

This first systematic analysis of the linked data set offers an overview of cancer risks in occupations held longest by the patient. Additional risk calculations are under preparation that will take into consideration latency time, "ever versus never" employment in a given industry, and various job titles occurring within each industrial activity. Recently, the data set was up-dated with another 100 000 cancer cases notified in the period 1980—1984, which, combined with occupational information on a random sample of 120 000 Danish inhabitants, will form the basis for analyses in the years to come.

This project would not have been feasible without the wholehearted assistance of the Board of Directors and staff of the Supplementary Pension Fund. The advice from our colleagues Ms M Ewertz, Ms E Lynge, Ms A Prener and Mr HH Storm at the Danish Cancer Registry is gratefully acknowledged. The programming was performed by Mr CD Møller, and the graphic presentation of the study results was prepared by Mr H Møller. Our secretaries, Ms V Clemmensen and Ms M Harnek, assisted with the preparation of the manuscript. The Danish Cancer Registry is a research institute of the Danish Cancer Society, which, in addition, provided special funds for the present investigation (grant no M-1/81).

CONTENTS

ABBREVIATIONS

The following abbreviations appear in the text, tables, and appendix:

amp Vater	= ampulla of Vater or hepatopancreatic ampulla
CIR	= Danish employer's identification number
CPR	= Central Population Registry
DCR	= Danish Cancer Registry
IARC	= International Agency for Research on Cancer
ICD-7	= seventh revision of the International Classification of Diseases
O	= observed number
SIR	= standardized incidence ratio
SMR	= standardized mortality ratio
SPF	= Supplementary Pension Fund
SPIR	= standardized proportional incidence ratio
95 % CI	= 95 % confidence interval

Key to figures:

No	= number
SPIR	= standardized proportional incidence ratio
——	= 95 % confidence interval
100	= expected cancer risk in the branches of industry shown, derived from the site-specific distribution of cancer cases among all employees in Denmark
Hatched bar	= significant deviation of the SPIR value from expectancy
RESTAUR	= restaurants
Prod	= products
Adm	= administration
Serv	= services
Transp	= transport
Manufact	= manufacture, manufacturing
Distrib	= distribution

INTRODUCTION

As stated by Pierre Decoufle (50), epidemiologic studies of occupational groups provide opportunities to investigate the consequences of continued, intimate contact of humans with a host of potentially hazardous substances. Exposure to physical and chemical agents at the workplace is usually heavier than elsewhere and often extends over relatively long periods of time. Thus, if a substance is carcinogenic to humans, the best opportunity of observing it may be in an occupational setting (50). In addition, the results of such investigations are of direct relevance to the occupational groups in question.

Cancer epidemiology is generally an exceedingly time-consuming research activity, and the setting of priorities is therefore a particular, recurrent problem. The difficulties in setting priorities derive from the identification and follow-up of study groups in cohort studies and from case ascertainment and risk assessment in case-referent studies. Data on disease(s) and exposure(s) must be extracted from archives, which have often been established to serve purposes other than research, or through questionnaires, which the researcher has to design carefully so that they will be well understood and will give a true reflection of past exposures. The restraints imposed on research output by limited resources can be partly overcome by the use of large-scale data linkage. Linkage studies have been carried out particularly in Scandinavia, due to the existence of personal identification numbers and of high-quality, long-term national cancer registries. In most Scandinavian countries, census data have been linked with cancer registries (115, 138, 168, 203) as a means to screen for occupational cancer risks.

The introduction of a personal identification number in Denmark in 1968 greatly facilitated large-scale epidemiologic cancer studies based on record linkage. Two nationwide linkage studies were initiated in Denmark in 1981. One was based on a linkage of cancer registry records from 1970—1980 with individual records from the 1970 census, the tabulation of cancer incidence thus being allowed for a 10-year period by occupational groups (115). The other, described in the present report, was based on the linkage of cancer cases notified to the Danish Cancer Registry during the period 1970—1979 with information on *occupational histories* from computer-based national registries, ie, the Supplementary Pension Fund and the Central Population Registry. The aims of the linked data set are twofold, ie, to describe the occurrence of cancer in occupational groups in the country at large and to form a common source of data for the testing of specific hypotheses about occupational cancer. This supplement describes the linked data set and reviews extensively all results of potential importance.

BACKGROUND

Alert clinicians were the first to detect cancer risks in the occupational environment in several situations, many of which were geographic or occupational clusterings of high incidences of an otherwise rare form of tumor. Thus, Sir Percival Pott, in 1775, related the unusually high frequency of scrotal cancer among London chimney sweeps to exposure to soot (156). Bladder cancer, found in excess among workers in the German dyestuffs industry in 1895 (164), was long thought to be caused by aniline dyes but was later associated with exposure to beta-naphthylamine and benzidine. More recent examples are observations of associations between nasal cancer and woodworking (3), between pleural mesothelioma and asbestos (128), and between the risk of hemangiosarcoma of the liver and vinyl chloride production (43).

It is unlikely that clinicians will identify such environmental risks in the case of common cancers when more than a few competitive causes are involved or when the increase in risk associated with the environmental exposure is modest. For the detection of exposure-effect associations under these circumstances, it is usually necessary to undertake regular epidemiologic investigations in a well-defined study population with a high degree of case ascertainment. The existence of a cancer registry greatly enhances the practicability of studies on cancer occurrence (89). The registration of cancer is, however, rarely matched by ex-

posure information of similar quality, and therefore ad-hoc epidemiologic studies are usually needed for the detection of occupational cancer risks.

The present study is the result of a unique opportunity for linking the cases in a nationwide cancer registry with the records of a pension fund, which provide occupational histories of all employees in the country. This linked Supplementary Pension Fund-Danish Cancer Registry data file is significant because it makes possible: (i) the detection of previously unsuspected hazards through the notification of unusual cancer patterns in occupational groups; (ii) the monitoring of known occupational cancer risks in different occupational groups and industries, and (iii) easy access to information on cancer and occupation as a basis for case-referent studies (151).

MATERIAL AND METHODS

The linkage of records in the Danish Cancer Registry (1970—1979), in the Supplementary Pension Fund (1964—1979) and in the Central Population Registry (1968—1979), comprising approximately 150 000 cancer patients, was accomplished in 1982.

Danish Cancer Registry

The Danish Cancer Registry was founded in May 1942 as part of the first nationwide program to register all cases of cancer arising in a defined population (36, 94). Although it operates under the auspices of the Danish Cancer Society, assistance has also been received from the Danish National Board of Health, which has provided full access to death certificates, and from the Danish Medical Association, which has urged its members to report cancer patients to the Registry on a voluntary basis. Thus the Registry regularly receives reports of diagnoses of malignant and related diseases and of changes in initial diagnoses and treatment. In addition reports are received from departments of pathology and forensic medicine on the results of autopsies of cancer patients. Cancers first recognized at autopsy are also reported and are included in the material of the Registry, irrespective of whether the cancer was suspected before death or was an incidental finding at autopsy. Finally, linkage with all death certificates issued in Denmark is undertaken annually. Before a case of cancer appearing only on a death certificate is included in the data base, it is thoroughly investigated through contact with the certifying hospital department; cancer cases are not included if the diagnosis is not verified. When, occasionally, verification is not available from the certifying physician, the information on the death certificates is included, even if the reliability of the diagnosis is questionable.

Reports are requested for all cases of carcinoma, sarcoma, leukemia, Hodgkin's disease, non-Hodgkin's lymphoma, multiple myeloma, and mycosis fungoides. In addition notification of the following lesions is also requested: (i) papillomas of the lower urinary tract (renal pelvis, ureter, urinary bladder and urethra), which are included with lower urinary tract cancers; (ii) histologically benign tumors of the central nervous system and meninges, which are included with malignant tumors at those sites; and (iii) carcinoma in situ and epithelial dysplasia of the uterine cervix, which are reported as precancerous lesions and are not tabulated with invasive cancer of the uterine cervix (and not included in the present linkage study). Tumors are classified according to the seventh revision of the International Classification of Diseases (ICD-7). However, the coding system of the Danish Cancer Registry includes some additions to the ICD-7, giving more detail. For example sarcomas and mesotheliomas can be separated as distinct groups of malignancies (36, 47, 48).

Cases included in the register are identified by means of the unique personal identification number assigned by the national Central Population Registry to each individual residing in Denmark. The use of this personal identification number by the Registry reduces the risk of duplicate registrations of the same tumor, and it allows the Cancer Registry to be linked with other registries, which also use this number.

Since the study population originates from the Cancer Registry, all persons in the study have, by definition, a tumor, and the following information is available: (i) date of birth, (ii) sex, (iii) personal identification number, (iv) tumor diagnosis (ICD-7), (v) date of diagnosis, and (vi) evidence of tumor diagnosis.

Supplementary Pension Fund

The Supplementary Pension Fund (ie, a supplement to the state pension, which is a priviledge of all inhabitants of the country) was established on 1 April 1964. The pension scheme (183) is *compulsory* for wage earners 18 to 66 years of age and, since September 1978, those aged 16 to 66 years of age. The pension supplement is financed by monthly contributions from the wage earners themselves and from their employers.

Both contributions are paid quarterly by the employers to a nationwide central fund, which retains all information on contributors, pensioners, and deceased persons.

For the correct recognition of an employer (company), a unique seven-digit employer's identification number (not to be confused with the employer's personal identification number) has been allocated by the tax authorities. This so-called CIR number is used by the Supplementary Pension Fund to register payment of contributions. At the same time the identification number of the benefiting employee is registered to insure that the contribution is paid into the account of the correct person at retirement age. The calendar period of the contribution is also included. The occupational history of each employee can thus be reconstructed from the start of the Pension Fund on 1 April 1964. Figure 1 shows examples for three persons, all of whom were cancer patients.

Since 1975, the Pension Fund has not only registered the CIR number but also the industrial category of the employer. The coding is performed by the Danish Central Bureau of Statistics on the basis of a standard registration form filled out by the employer, who specifies the type of industry and describes the most important activities of the company (employer). These registration forms are collected by the Danish customs authorities, who are responsible for collecting value added tax (VAT) in the country. The codes for the industrial category of the employers are transferred regularly to the Supplementary Pension Fund. This procedure insures that all CIR numbers which were not abolished before the introduction of the industry coding appear with a code for the industrial category attached. This is a five-digit, hierarchically formed number that corresponds closely to the International Standard Industrial Classification of All Economic Activities (189). Figure 2 gives an example of the structure of the industrial classification, and figure 3

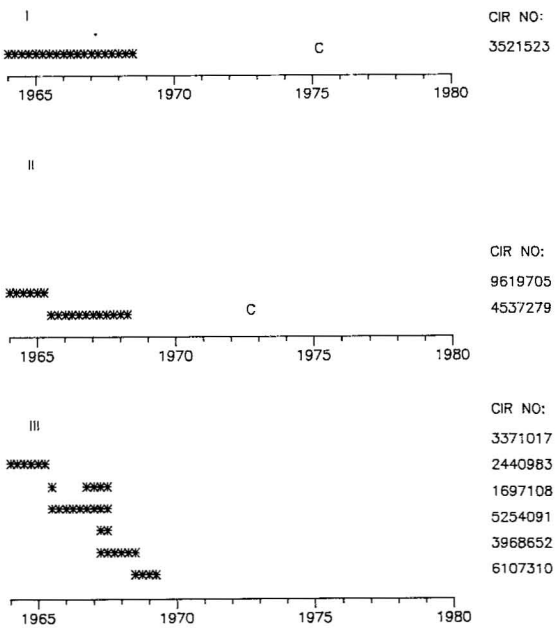

Figure 1. Occupational histories (***) of three persons (I, II and III) all born on 23 January 1900, and time of diagnosis of cancer (C).

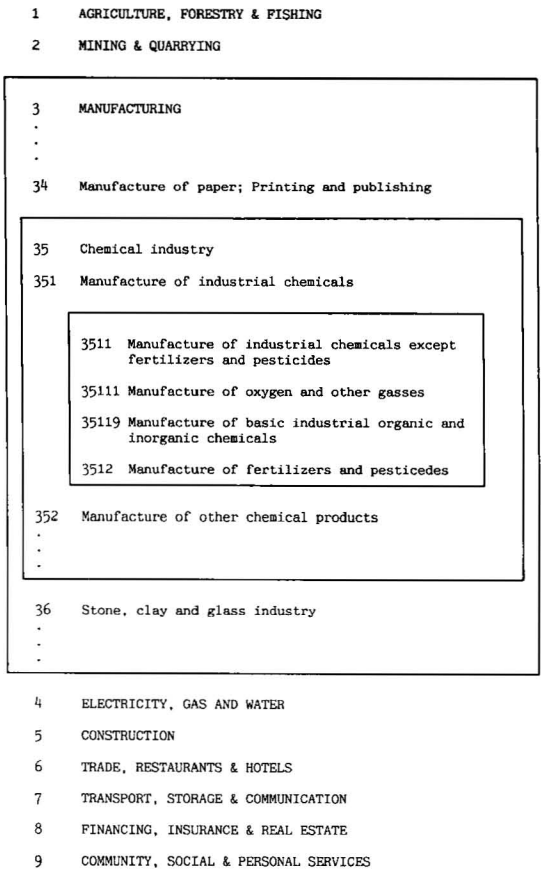

Figure 2. Typical section of the industry code grouping used in the present analysis.

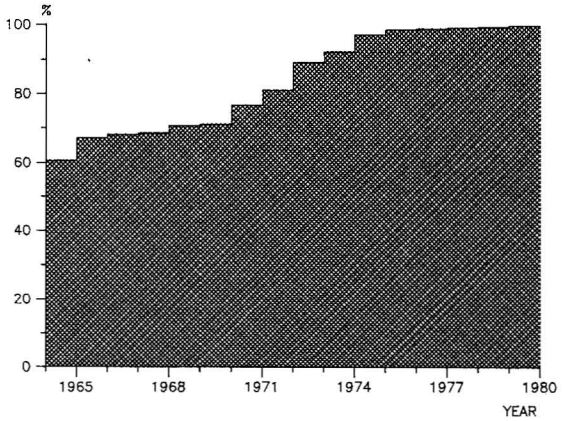

Figure 3. Proportion of firms with an industry code (International Classification Standard) added to the CIR number, 1964—1979.

shows the proportion of firms for which an industry code has been added to the CIR number since the start of the Pension Fund scheme.

The following data from the Supplementary Pension Fund are available on the linked file: (i) personal identification number of the employee, (ii) CIR number of the firm(s), (iii) code for the industry represented by the firm, (iv) date of commencement of each employment, and (v) date for termination of each employment.

Table 1. Source of information included in the linked data file, (+ = information obtained, − = information not obtained)

Item	Cancer Registry	Supplementary Pension Fund	Central Population Registry
Identification number[a]	+	+	+
Company name	−	+	−
Date of start of employment	−	+	−
Date of ending employment	−	+	−
Industry	−	+	−
Job title	−	−	+
Cancer diagnosis	+	−	−
Date of diagnosis	+	−	−
Evidence of diagnosis	+	−	−
Vital status	−	−	+
Date of death	−	−	+

[a] Includes information on date of birth and sex.

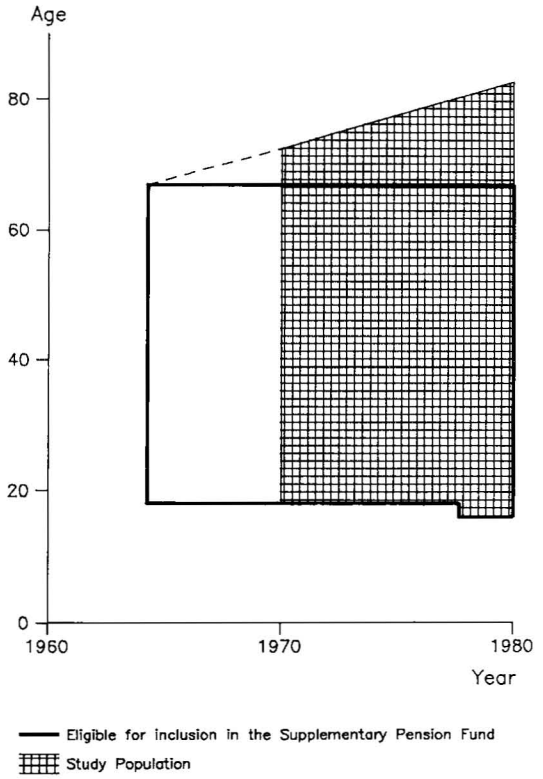

Figure 4. Study population defined by year and age at diagnosis, and potential membership of the Supplementary Pension Fund, 1964—1979.

Central Population Registry

The Central Population Registry was established on 1 April 1968 with the purpose of storing commonly used personal data on each inhabitant of the country. Its register acts as source material for the administrative system of Denmark. The Central Population Registry is administered by the Ministry for the Interior (32).

The Registry collects information on all persons who have resided in Denmark since 1 April 1968, and each inhabitant is allocated a unique, 10-digit identification number composed of date of birth and a four-digit running number, the last of which is a check digit.

Among other personal information, the Central Population Registry retains the job title, as supplied by the citizen himself on the annual income tax form. Although historical job titles are not stored, information on current occupation is still of potential importance for occupational studies, as it indicates whether an employee is likely to be (or have been) exposed in the process of production (blue-collar worker) or not (white-collar worker).

For the present linkage study the Central Population Registry has thus supplied (i) the personal identification numbers of the persons under study, (ii) their job titles, (iii) their vital status, and (iv) their dates of death.

Record linkage and study population

All cases of cancer in the decade 1970—1979 were included in the data linkage, provided the patients resided in Denmark and were of an age compatible with inclusion in the Supplementary Pension Fund. A computer tape was produced from the Cancer Registry file, giving the personal identification numbers of the patients, and the items listed in table 1 were procured from the three registries by computerized record linkage, the personal identification number being used as the key identifier. This linkage gave rise to a study population which is illustrated graphically in figure 4.

After the exclusion of 134 persons for whom the personal identification data were not verifiable, 153 427 cancer patients (76 985 men and 76 442 women) were eligible for the linkage study, corresponding to 78 % of all persons diagnosed with a malignancy during the decade 1970—1979. For these patients, more than 2.8 million items of information on employment were recorded in the Supplementary Pension Fund register. Since status as an employee is required for inclusion in the Supplementary Pension Fund, not all of the 153 427 cancer patients of the study population benefit from the Fund, even if they fall within the age limits for membership (figure 4). Some patients had never been employed (eg, self-employed persons, housewives, students), others were unemployed, while other patients had had up to 152 different employers since 1964.

As can be seen in figure 1, which shows the occupational histories of three of the cancer patients, some inaccuracies apparently exist in the files of the Pension Fund in the determination of exact employment periods. Overlapping periods of employment by two or more employers were detected regularly, and short discontinuations of registration in the Fund were observed. These irregularities were thought to be due mainly to faulty registration (eg, the first employer continued to contribute to the fund for some weeks or months after termination of the employment), or to temporary discontinuation of payment during holidays. Thus it was decided to edit the employment in-

Table 2. Proportions of the study population recovered from the files of the Supplementary Pension Fund, with and without job title as stated in the files of the Central Population Registry. ("Housewife" was regarded as a job title.)

	Members of Fund[a]		Non-members of Fund		Total	
	N	%	N	%	N	%
With job title						
Men	45 979	60	17 232	22	63 211	82
Women	27 742	36	23 471	31	51 213	67
Without job title						
Men	8 025	10	5 749	8	13 774	18
Women	8 905	12	16 324	21	25 229	33
Total						
Men	54 004	70	22 981	30	76 985	100
Women	36 647	48	39 795	52	76 442	100
Both sexes	90 651		62 776		153 427	

[a] Included in the analysis.

Table 3. Record description for data file compiled from the linkage of the Supplementary Pension Fund (SPF) and the Danish Cancer Registry (DCR).[a] (CPR = Central Population Registry)

Item on record	Source
1. Personal identification number (unique 10-digit identification number of a person)	CPR
2. Sex	CPR
3. Day of birth	CPR
4. Job title (exists for only 75 % of the study population)	CPR
5. Status on 1 January 1980 (alive, dead, emigrated, disappeared)	CPR
6. Date of last change in status	CPR
7. Number of cancer diagnoses	DCR
8. Number of actual cancers	DCR
9. Cancer diagnosis, ICD-7 code (1970—1977)	DCR
10. Cancer site, ICD-0 (1978—1979)	DCR
11. Histology, ICD-0 (1978—1979)	DCR
12. Behavior code for the tumor, ICD-0 (1978—1979)	DCR
13. Method of verification (two-digit code for the basis of the cancer diagnosis)	DCR
14. Age at diagnosis	DCR
15. Day of diagnosis (date of first hospital admission)	DCR
16. Municipality (three-digit code for municipality of residence at time of diagnosis)	DCR
17. Job (two-digit code based on job title from hospital record)	DCR
18. Number of different employers before 1 January 1980	SPF
19. Number of different employers before cancer diagnosis	Calculated
20. Days at work (total number of days of employment from 1 April 1964 until day of diagnosis)	Calculated
21. Days of work, five years before cancer (total number of days of employment from 1 April 1964 until five years before day of diagnosis)	Calculated
22. Days of work, ten years before cancer (total number of days of employment from 1 April 1964 until ten years before day of diagnosis)	Calculated
23. Days of work, 15 years before cancer (total number of days of employment from 1 April 1964 until 15 years before day of diagnosis)	Calculated
24. Time since April 1964 (total number of days from 1 April 1964 until day of cancer diagnosis)	Calculated
25. Time since April 1964, five years before cancer (total number of days from 1 April 1964 until five years before day of diagnosis)	Calculated
26. Time since April 1964, ten years before cancer (total number of days from 1 April 1964 until ten years before day of diagnosis)	Calculated
27. Time since April 1964, 15 years before cancer (total number of days from 1 April 1964 until 15 years before day of diagnosis)	Calculated
28. Employers' identification number (seven-digit identification number, so-called CIR number, of the company)	SPF
29. Industry (five-digit code for the industry represented by the firm) — The industrial code is hierarchical, with the following nine main classes: 1. agriculture, hunting, forestry and fishing; 2. mining and quarrying; 3. manufacturing; 4. electricity, gas and water; 5. construction; 6. wholesale and retail, restaurants and hotels; 7. transport, storage and communication; 8. financing, insurance, real estate and business services; 9. community, social and personal services	SPF
30. Start of work (day of commencement of the employment period)	SPF
31. End of work (day of termination of employment period)	SPF
32. Days of employment period (total)	Calculated
33. Days of employment period, five years before cancer (until five years before the day of cancer diagnosis)	Calculated
34. Days of employment period, ten years before cancer (until ten years before the day of cancer diagnosis)	Calculated
35. Days of employment period, 15 years before cancer (until 15 years before the day of cancer diagnosis)	Calculated
36. Employment priority (order of the employment period in relation to length of all employment periods)	Calculated
37. Employment priority, five years (order of the employment period, allowing for a five-year latency period)	Calculated
38. Employment priority, ten years (order of employment period, allowing for a 10-year latency period)	Calculated
39. Employment priority, 15 years (order of employment period, allowing for a 15-year latency period)	Calculated

[a] The file contains 98 810 records, ie, one record per tumor, 1970—1979.

formation along the following lines:

1. Discontinuation of payment for up to one month was ignored if the employment periods in question were with the same employer.

2. Two or more simultaneous employments with the same employer were counted as one employment.

3. Two or more simultaneous employments for less than three months with different employers were counted as one employment with the last employer.

4. Overlapping employments for more than three months were accepted as multiple employments.

One job title per patient was available from the Central Population Registry. Table 2 shows that 70 % of the men and 48 % of the women had been employed at some time and were listed in the Pension Fund register with information on one or more jobs. For 82 % of the men and 67 % of the women a job title was attached to their record in the Central Population Registry. Finally, the table gives the frequency of the combined information on occupation from both the Fund and the Registry, leaving 8 % of the men and 21 % of the women for whom no information on occupation was available. However, even though the job title was included in the linked data set and thus available for analysis — if possible, in combination with the CIR number of a company or a code for branch of industry — the information contained in this job title was not used in the present analysis of the material.

This report is based on data on 90 651 cancer patients, corresponding to 93 810 cancer cases *for which information existed in the Supplementary Pension Fund* on employment between 1964 and date of cancer diagnosis, irrespective of the existence of a job title. The tumor is the key unit of the linked data file to which a varying number of items of employment information was attached. One patient may have more than one tumor. A full list of the record variables of the linked data set is given in table 3.

Analysis

For 3.5 % of the study population on record at the Cancer Registry, two (or more) primary malignancies were noted in the period 1970—1979. As the *tumor* is the key unit of the analysis, 158 856 primary tumors constitute the linkage file created between the Supplementary Pension Fund and the Danish Cancer Registry for the investigation of occupation and cancer risk. Because of the complexity of the number of employments relative to cancer patients and tumors, the linked data file was arranged hierarchically, with the diagnosis as the standard entity of entrance into the file, to which personal data and an appropriate number of occupational records were added.

The fact that only cancer cases were included in the study had to be considered in the analysis; thus, in the absence of population denominators, a proportional incidence analysis was performed. The risk of cancer in a defined company or branch of industry was estimated as the standardized proportional incidence ratio (SPIR). This value approximates the conventional standardized incidence ratio (SIR) or standardized mortality ratio (SMR) when the cancer under investigation constitutes a minor part of all the malignancies included in the study and when exposure has no effect on cancer risk in general (133).

In the present analysis, the observed number of cancer cases was allocated to the *industry in which the cancer patient was employed for the longest period of time.*

For calculation of the expected number of a defined tumor in a given industry, the following formula was used:

$$\sum_{i=15-19}^{(80+)} \sum_{j=1970}^{1979} (T_f \times t/T)_{ij},$$

where t is the observed number of the defined tumor in the material, T is the total number of tumors, and T_f is the total number of tumors in the selected industry, f. In this calculation adjustment was made for calendar time and age; thus i is the age groups 15—19, 20—24, ..., 75—79, ≥80, and j the calendar periods 1970, 1971, ..., 1979.

On the basis of the observed number of cancer cases, corresponding 95 % confidence limits were calculated from an approximation to the normal distribution:

$$\text{lower confidence limit (LL)} = e^{\frac{-1.96}{\sqrt{obs}}} \times \text{SPIR}$$

and

$$\text{upper confidence limit (UL)} = e^{\frac{1.96}{\sqrt{obs}}} \times \text{SPIR},$$

where e is the basis of the natural logarithm, *obs* is the observed number of cancer cases, and SPIR is the standardized proportional incidence ratio (4). However, when the number of observed incident cancer cases in a given industry is less than 20, the confidence limits must be interpreted with caution due to the approximation undertaken.