HSE

Health and Safety Executive

The burden of occupational cancer in Great Britain

Methodology

Prepared by **Imperial College London**
for the Health and Safety Executive 2012

RR927
Research Report

Health and Safety Executive

The burden of occupational cancer in Great Britain

Methodology

Sally Hutchings, Lesley Rushton
Department of Epidemiology and Biostatistics
Imperial College London
Norfolk Place
London W2 1PG

This project aimed to estimate the current burden of cancer for Great Britain due to occupational exposure to carcinogenic agents/exposure circumstances classified by the International Agency for Research on Cancer as definite (Group 1) or probable (Group 2A) human carcinogens.

The measure of the burden of cancer used was the attributable fraction (AF) ie the proportion of cases that would not have occurred in the absence of exposure. For each cancer/exposure pair, risk estimates, adjusted for confounders, were obtained from key industry-based studies, meta-analyses or reviews. The period of exposure relevant to the development of the cancer in the target year 2005 was defined as the risk exposure period (REP) and was assumed to be 1956-1995 for solid tumours and 1986-2005 for haematopoietic neoplasms. National data were used to derive the proportion of the population ever exposed to each carcinogenic agent or occupation in the REP ie the ratio of the numbers ever exposed within GB over the total number of people ever employed. Adjustment was made for staff turnover, life expectancy and industry-employment trends over the REP. Exposed numbers were obtained for high and low exposures matched to appropriate risk estimates from the literature. The AFs were applied to national cancer deaths and registrations to give occupation attributable cancer numbers.

This report and the work it describes were funded by the Health and Safety Executive (HSE). Its contents, including any opinions and/or conclusions expressed, are those of the authors alone and do not necessarily reflect HSE policy.

HSE Books

ACKNOWLEDGEMENTS

Funding was obtained from the Health and Safety Executive (HSE). Andrew Darnton from the HSE was responsible for the work on mesothelioma. The contributions to the project and advice received from many other HSE and Health and Safety Laboratory staff is gratefully acknowledged. Two workshops were held during the project bringing together experts from the UK and around the world. We would like to thank all those who participated and have continued to give advice and comment on the project. We would also like to thank Helen Pedersen and Gareth Evans for their help in editing and formatting the reports.

CONTENTS

1 INTRODUCTION

The currently used estimate of the proportion of cancer deaths in Great Britain due to occupational causes is 4%, with an uncertainty range of from 2% to 8%, based on the study of Doll and Peto (1981). This equates to approximately 6,000 deaths per annum (with a range of 3,000 to 12,000). The aim of this project was to update these figures and to estimate the burden of cancer due to occupation in Great Britain. Specific aims were:

(1) To estimate a current overall attributable fraction and attributable number of cancers due to occupation.
(2) To indicate the relative contributions of the occupational carcinogens to the current total burden of occupational cancer in Great Britain, thereby providing evidence on which to prioritise intervention.
(3) To estimate future burden resulting from current exposures, in order to indicate where to prioritise future intervention.
(4) To suggest areas for future data collection, to improve the quality of the estimates.

This technical report describes the methodology for aims (1) and (2).

2 ALTERNATIVE APPROACHES TO CALCULATING THE OCCUPATIONAL ATTRIBUTABLE BURDEN

There are a number of basic approaches that have been used to calculate the occupation attributable cancer burden. These are:

1. For each cancer/exposure pairing, a relative risk estimate is obtained from epidemiological studies, and an estimate of the proportion of the population exposed is obtained from census or national employment data or from a job-exposure matrix (JEM), or from the CARcinogen EXposure (CAREX) database (Pannett *et al,* 1998).

Examples of this approach are:

a) Dreyer *et al*. (1997), who calculated the attributable fraction for cancers due to occupation in the Nordic countries, using a common risk exposure period of 1970-84 for a target year of 2000 and a common of employment turnover (15% for those employed for over one year).
b) World Health Organisation (WHO) Global Burden of Disease (GBD) Study;
c) Steenland *et al.* (2003), for the USA;
d) Nurminen & Karjalainen (2001), for Finland;
e) Mathers *et al.* (2000), for Australia, using GBD methodology;
f) Driscoll *et al.* (2004) , for New Zealand, who used the attributable fractions calculated by Nurminen & Karjalianen plus Steenland's estimates for lung cancer;
g) Morabia *et al.* (1992), for the US, for lung cancer only, using relative risks from one large hospital-based case-control study;
h) Steenland *et al.* (1996), for the USA, for lung cancer only, who combined relative risk estimates by exposure agent from a review of cohort studies, and matched these to National Institute for Occupational Safety and Health (NIOSH) numbers of workers exposed.

2a). The results of a pooled analysis of (usually population-based) case control studies are used for an estimate of relative risk, with internal estimates of the proportion of the population exposed (Pr(E)) or the proportion of cases exposed (Pr(E|D)) obtained from the distribution of exposures amongst the controls, or cases.

Examples of this approach are:

a) Vineis *et al.* (1988), for the US, for lung cancer, using a pool of population-based case-control studies;
b) Gustavsson *et al.* (2000), for Sweden, lung cancer only, using a single population-based case-control study;
c) Kogevinas *et al.* (1998) for Europe, who also quoted method 5 below in a review of methodology suitable to produce an attributable fraction for Europe.

2b) Similar to the above, estimates of relative risk are obtained from a search of the literature for the 'best studies' (population-based case-control studies) for an occupational cancer, and an attributable fraction is calculated from each of these studies using internal estimates of Pr(E) (the proportion of the population exposed) or Pr(E|D) (the proportion of cases exposed (E = exposed, D = case)), to give a range of possible attributable fraction estimates. Additional lower and upper estimates are also provided that are based on inclusion of only well-established exposures for the lower estimates, plus 'possible' and 'uncertain' exposures for the upper estimates.

An example of this approach is:

d) Vineis and Simonato (1991), for the USA and Europe

Normally, exposure assessments will not be detailed enough from population or hospital based case-control studies to estimate a dose-response relationship. To estimate attributable fraction, 'ever/never' exposed categories are needed, but it is necessary to have a defined ever/never boundary level. This will normally be less precise from population or hospital based case-control studies, with exposure data generally collected by postal questionnaire or interview rather than expert hygiene assessment of the exposures scenario.

3. Absolute risk measures rather than relative risks are appropriate for a few cancer/exposure pairings for which the attributable fraction is associated with an occupational exposure thought to account for 100% of the risk, for example mesothelioma that is uniquely caused by exposure to asbestos.

Examples of this approach are:

a) WHO GBD study (Nelson et al, 2005)
b) Hodgson & Darnton (2000)

In these cases an age period cohort method is also used to predict future numbers.

4a) Proportional Mortality Ratios (PMRs), which can substitute for relative risks in the calculation of attributable fraction, are available for the UK for most cancers by occupation from the Occupational Health Decennial Supplement (1995). The proportion of the population exposed (Pr(E)) in a risk exposure period would come from national employment data (LFS for occupations), although the reliability of occupation as recorded on the death certificate is known to be problematic, as is matching with other sources of routinely collected data on occupation (Carpenter and Roman, 1999). This method could be used as a 'reality check' on methods (1) and (2) above, where UK studies are not available, and to contribute to confidence intervals. Again an age period cohort method is used to predict future numbers.

4b) Similar to the above, Bouchardy *et al.* (2002) used cancer registry data alone to estimate occupational odds ratios for Switzerland, using a case-referent method, with cases defined for a particular cancer and all other cancers as controls, and job history as recorded by the registry. To extend this approach to obtain an attributable fraction, an external source of employment data would be required.

5. Linkage analysis of census and cancer registry data can be used if national databases permit. For example a population based (often a national census-based) cohort can be established with cancer registration or death certificate follow-up.

Examples of this approach are:

a) Andersen *et al.* (1999), in Norway, Sweden, Finland and Denmark, linked cancer registry and census data to produce SIRs by occupation (but not attributable fractions);
b) van Loon *et al.* (1997) in the Netherlands used a population-based cohort study but only for lung cancer.

6. The 'Delphic principle' (Morrell et al, 1998) of Doll and Peto, which uses panels of experts to estimate attributable fraction.

An example of this approach is that of Landrigan *et al.* (2002), who examined environmental attributable burden of disease in US children, using panels of experts to arrive at an estimate of attributable fraction by a consensus process of meetings and ballots.

7. A proportionate attributable risk approach, which is an extension of the Delphic principle of Doll and Peto, can be used.

Examples of this approach are:

a) Morrell et al (1998), for Australia, used estimates of attributable percentages for all hazardous substances from countries considered to have similar exposure conditions. Estimates from the United States (Markowitz *et al,* 1989) and Israel (Blanc *et al,* 1992), plus the Doll and Peto estimates, were applied to numbers of deaths in Australia, using age at death to calculate person-years of life lost (PYLL).
b) Leigh et al. (1997) also used this method to estimate occupation attributable numbers for the USA, reviewing US-based studies for the established occupational cancers to arrive at an attributable fraction range for each cancer, then applying this to cancer mortality for one year, 1992.

8. Another approach is the descriptive study of incident cases. Deschamps *et al.* (2006) collected work histories for all newly diagnosed first cancers occurring over three years (1995-98) in a single county in France in those over 16 (excluding housewives or permanently unemployed). They considered 13 well-established cancer/exposure pairings as occupational candidate cancers. Assigning cases as occupational if the individual was exposed to the specific occupational agent for at least 5 weeks per year for a minimum three (not necessarily consecutive) years. They then calculated the proportion of cases that were occupational as a percentage of all the cancer cases in the study. The authors also give the highest and lowest attributable fraction estimates taking the highest and lowest percentages published internationally. This is a straightforward approach, but a problem with this study is that the authors had to consider the distribution of occupations in their cancer patients (in a single French county) to be the same as in the general population (which was supported by a comparison with census data). Excluding non-working cases would also have introduced bias into the results.

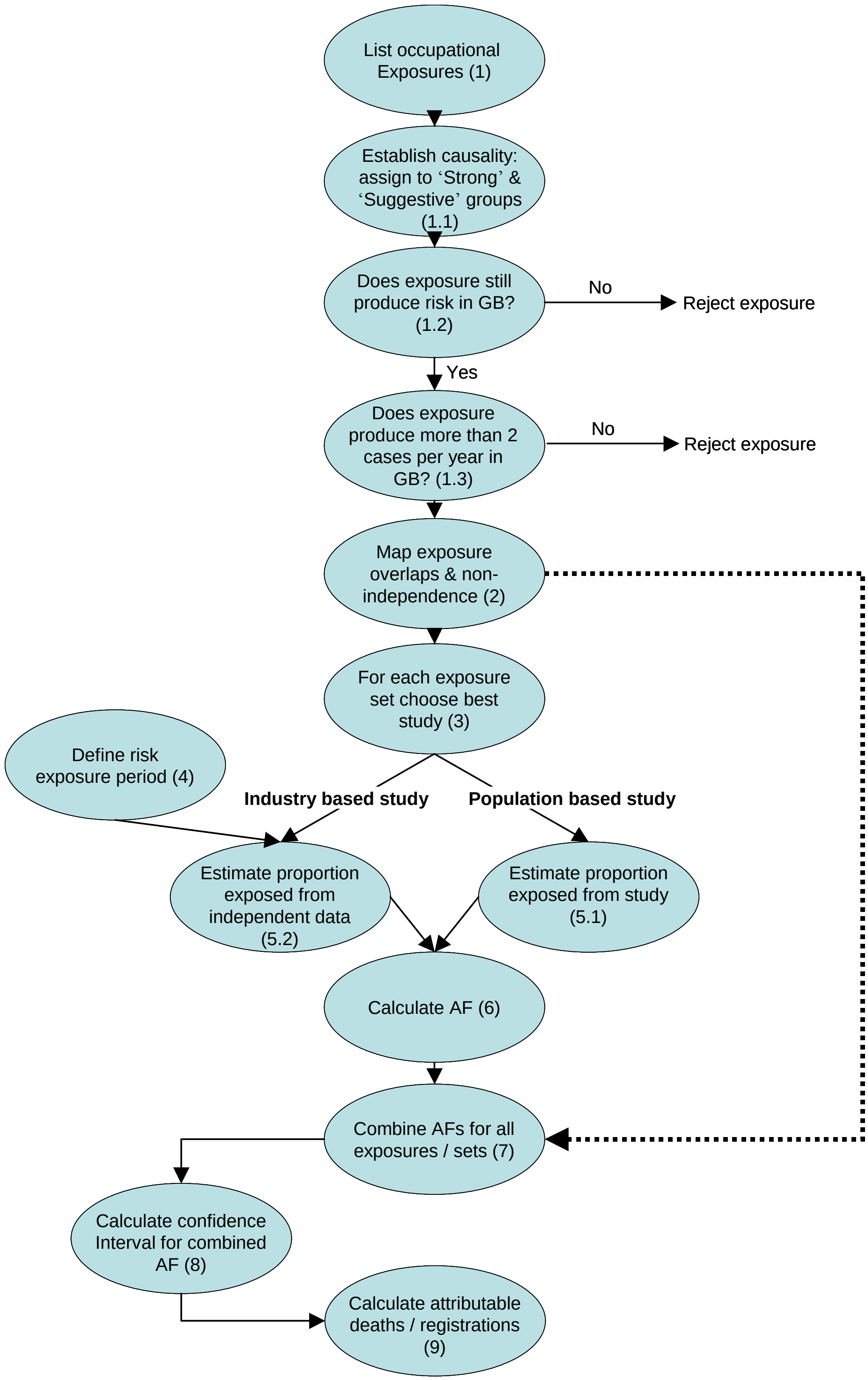
The first two of the approaches described above have been adapted for this study to estimate the burden of cancer due to occupation in GB. Elements of the 'Delphic Principle' have also been used in choosing the 'best study' using a panel of experts to derive a relative risk (RR) estimate for each cancer/exposure pairing. The first approach, and Levin's equation for the attributable fraction (see section 3.6), were used where the RR was obtained from an industry based cohort study, or a review or meta-analysis, in which case national estimates of numbers exposed were used. In practice most estimates to date have been obtained using this first approach. The second approach, and Miettinen's equation for the attributable fraction, was used where a population-based study was the source of the RR, from a population comparable in the type and distribution of exposures and confounding factors to that in Great Britain (so that the risk estimates are 'portable', see Appendix 1). In practice this approach was rarely used due to a lack of appropriate population-based studies in GB. The third approach above was used for mesothelioma caused by exposure to asbestos.

3 KEY FEATURES OF THE APPROACH ADOPTED

The overall attributable fraction was calculated on a ‘cancer by cancer’ basis. Relative risks were obtained from epidemiological studies considered the most appropriate currently available. The choice of these ‘best studies’ then determined the method by which attributable fraction was calculated.

A diagram of the calculation process for the attributable fraction is shown in Figure 1. The numbers in blue in Figure 1 refer to sub-sections in the text, which follows. Detailed methodology and the derivations of factors used in the calculations are in the appendices. A glossary of terms used in this report and accompanying papers is given at the end of the appendices.

Figure 1: Calculation process for the attributable fraction



3.1 LISTING THE OCCUPATIONAL EXPOSURES

For each cancer, major literature sources were reviewed to identify relevant occupational exposures, confounders and other non-occupational risk factors, and obtain information on latency and trends in mortality, incidence and survival (see also Section 3.3). Sources for the preliminary review for each cancer included the IARC monographs[1], NIOSH reports[2] and the US National Toxicology Program[3], the ONS DH and MB Series' for mortality and incidence data and the Decennial Supplements on Occupational Mortality. At two international workshops held as part of the project to discuss the methodology[4,5], the participants advised that priority should be given initially to International Agency for Research on Cancer (IARC) Group 1 and 2A occupationally-related carcinogens. Agents or occupations in these IARC groups were included that had either "strong" or "suggestive" evidence of carcinogenicity in humans for the specific cancer site, as defined by Siemiatycki *et al* (2004) and subsequent IARC publications. Those with "strong" evidence were defined as 'strong' carcinogens for the purposes of this study and those with "suggestive" evidence of carcinogenicity in humans were defined as 'suggestive' carcinogens. In addition there had to be substantial existing exposures in GB and/or cases of cancer still occurring due to past exposures. For each exposure, a list of industry and occupation settings (the 'exposure scenarios') was also drawn up using IARC monographs and Siemiatycki et al. (2004).

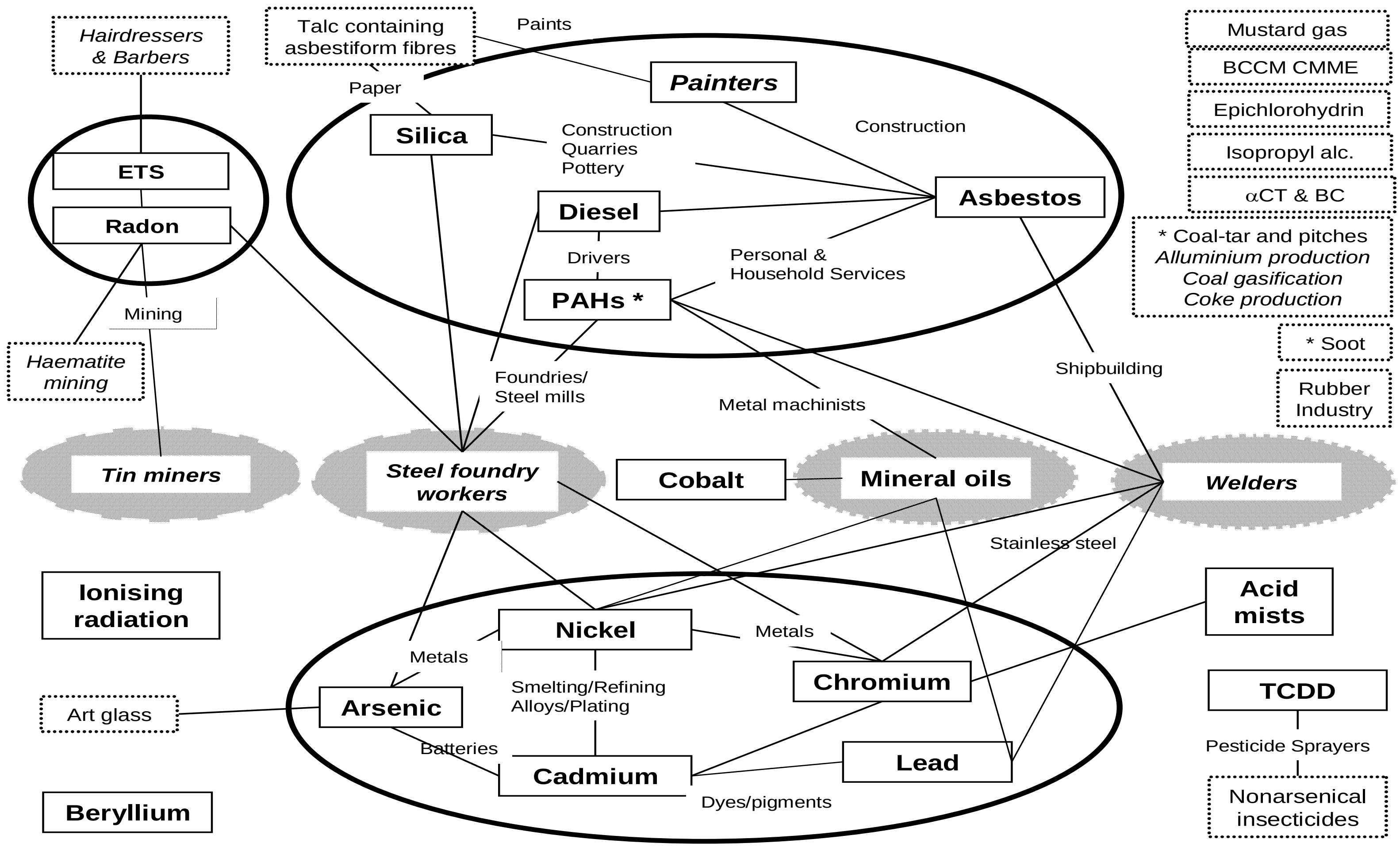
3.2 IDENTIFYING OVERLAPPING AND DOMINANT EXPOSURES

For cancer sites with several occupational carcinogens an 'exposure map' was drawn up. This is a diagram of how the exposures overlap in the working population, and could interact with one another in causing disease. An example of such a map, for lung cancer, is given in Figure 2. It illustrates the potential for double counting of the exposed population when an overall attributable fraction is calculated, and facilitates strategies to avoid this. For a given cancer, the map entries consist of either an agent (or group of agents such as PAHs), or an exposure scenario (i.e. an industry or occupation in which such exposure occurs). Agents are in plain type, exposure scenarios in italics. Lines joining boxes then indicate where overlap would occur were all the entries in the map simply considered separately – for example, if Painters and Asbestos were considered separately overlap would occur in construction (these exposure scenarios are indicated in the smaller print). For substances and occupations shown within boxes with dotted lines a separate attributable fraction was not estimated, as these exposure scenarios were included with another exposure. More explanation of the content of this map is given in Appendix 6 Section A6.3.

---

[1] http://monographs.iarc.fr/ENG/Monographs/allmonos90.php
[2] http://www.cdc.gov/niosh/homepage.html
[3] http://www.cdc.gov/niosh/homepage.html
[4] http://www.hse.gov.uk/research/hsl_pdf/2005/hsl0554.pdf
[5] http://www.hse.gov.uk/research/hsl_pdf/2007/hsl0732.pdf

Figure 2: Example of an exposure map, for lung cancer



The contribution towards the attributable fraction of these overlapping exposures could not (correctly) be estimated separately. These exposures should ideally be treated as a set, and a 'best epidemiological study' sought that has relative risk estimates for the set as a whole, or at least a study that takes account of the confounding effects and interactions within the set. In practice, however, such a study is rarely available. Therefore to take account of potential multiple exposures strategies included partitioning exposed numbers between overlapping exposures or estimating only for the 'dominant' carcinogen with the highest risk being used. The IARC Monograph process has been taking place over many years and has resulted in overlap between substances evaluated. For lung cancer, for example, there are 32 occupations or carcinogenic agents evaluated by IARC. We estimated attributable fractions for 21 of these; for example, substances such as coal-tars and pitches and processes such as coal gasification and coke production were included within our evaluation of Polycyclic Aromatic Hydrocarbons (PAH). Where exposure to multiple carcinogens remained it was assumed that the exposures were independent of one another and that their joint carcinogenic effects were multiplicative.

For each cancer/exposure pairing, the approach in general was to partition the exposed population between 'higher exposed' and 'lower' and/or 'background exposure' groups, and apply an appropriate relative risk from the source study or studies to each group separately. Ideally an exposure level threshold would be identified that separated the 'higher' and 'lower' exposure groups, but in practice the groups were usually defined by exposure scenarios.

3.3 REVIEW OF THE LITERATURE AND SELECTION OF RELATIVE RISKS

The literature was searched for relevant papers for the selection of relative risks. Standard search criteria were used, supplemented by references from review papers and textbook bibliographies.

The project team then held a review meeting to select the 'best studies'. Papers and other evidence were examined 'exposure by exposure' and standard selection criteria applied, which took into account whether the exposures in a published study represented conditions in GB, and were therefore portable to the GB situation and whether the study was of good quality (Appendix 1). Combined studies (pooled studies and meta-analyses) were usually the first choice if available, with UK studies added if necessary. Generally a best source was identified for each exposure scenario (industrial setting or occupation), but studies covering multiple exposures, especially those known to overlap in the working population, were also considered. At this stage exposures that did not contribute significantly to the study were excluded by applying a standard decision-making procedure (Appendix 2) and the reasons for exclusion were recorded. Apart from the requirement that causality was established, for an attributable fraction to be calculated for a cancer/exposure pairing, the exposure scenario needed to be present in GB during the risk exposure period for the specific cancer. There also had to be sufficient numbers exposed to produce more than two cases per year.

The relative risks (RRs) that were used in the calculation of attributable fraction were identified and extracted from the chosen studies, following the criteria set out in Appendix 3. Where only a narrative review was available giving a range of risk estimates from several relevant studies we calculated a combined estimate of the relative risks (RR) based on a random- (for heterogeneous RRs) or fixed- (for homogeneous RRs) effects model (Appendix 3, section A3.1). Formal systematic reviews and meta-analyses were carried out to estimate relative risks for laryngeal and stomach cancers related to asbestos exposure. Dose-response risk estimates were generally not available in the epidemiological literature nor were proportions of those exposed at different levels of exposure over time available for the working population in Britain. However, where possible risk estimates were obtained for an overall 'lower' level and an overall 'higher' level of exposure to the agents of concern. Where no risk estimate could be identified for very low/background levels of exposure, we estimated a RR for the 'lower exposed' group by (i) taking the harmonic mean of all the available ratios of 'higher' to 'lower' RR estimates for cancer-exposure pairs for which data were available, (see Appendix 3, section A3.2) and (ii) applying this average ratio to the 'higher' level estimates to obtain 'lower' level RR estimates. If the resulting RR estimate was less than 1, RR was set to one.

For industry-based studies, data for estimating the proportion of the population exposed was obtained from an external source (see section 3.5). In this case the exposure scenarios (the industrial settings or occupations) for which the RRs were calculated were recorded. These were matched by job or industry codes to the external source (see Appendix 3). For population-based studies, data for estimating the proportion of cases exposed (i.e. the numbers of exposed cases contributing to the relative risk estimate, and the total number of cases) was recorded from the RR source study.

3.4 THE RISK EXPOSURE PERIOD

We have defined the risk exposure period (REP) as the window of time during which exposure to an occupational carcinogen can result in a cancer being diagnosed or appearing in national mortality or cancer registration records. It is based on cancer latency and is defined in detail in Appendix 4. As there were very little data available on cancer latency, two 'standard' REPs were used for this study, to cover the solid tumours, for which a latency of between 10 and 50 years was assumed, and for the haematopoietic neoplasms such as leukaemia, for which a latency of between 0 and 20 years was assumed. For deaths in 2005 this gives a REP of 1956-1995 for solid tumours and a REP of 1986-2005 for haematopoietic neoplasms. Although strictly speaking the REP for cancer registrations recorded in 2004, the year for which estimation has been carried out, would be 1955-1994, for simplification the years 1956 to 1995 have also been used, as for deaths, as the proportion exposed will not be affected. Where possible risk estimates were extracted from studies where the exposure had occurred over a similar period to the REPs used in our estimation

Data specific to the standard REPs used to calculate the attributable fraction are given in Table 1. Note that more workers exposed during the shorter but more recent REP for haematopoietic neoplasms are still alive in 2005 than workers exposed during the solid tumour REP (see Appendix 5 Section A5.4 for the method of calculation of numbers ever of working age).

Table 1: Values of variables used in the calculation of attributable fraction specific to the standard risk exposure periods

Standard REP:	**Solid tumours**		**Haematopoietic neoplasms**	
Variable	**Men**	**Women**	**Men**	**Women**
Latency	10-50 years	10-50 years	0-20 years	0-20 years
Risk exposure period (REP)	1956-1995	1956-1995	1986-2005	1986-2005
'Peak' latency *	35 years	35 years	15 years	15 years
Number of years in REP	40	40	20	20
Total number of those ever of working age during the REP **	19.4 million	21.0 million	23.0 million	23.1 million
Ages included	25-90+	25-90+	15-84	15-79

* 'peak' latency is the latency thought to relate to the highest number of cancer cases leading to current deaths or registrations. However, a uniform distribution of cancer latencies has implicitly been assumed across the REP. This does not affect current burden results, but a lognormal distribution has been assumed for estimates of future burden (see Future Burden Methodology Technical Report)
** see Section 3.5

3.5 ESTIMATING THE PROPORTION EXPOSED

If the relative risk was extracted from a population-based study, an estimate of the proportion of the population exposed was available directly from the study. However it would be particularly important in this case that the study population was comparable and portable to GB. In practice population based studies (and therefore this source of the proportion exposed) were rarely available.

If the relative risk was extracted from an industry-based study, national data sources such as the CARcinogen EXposure database (CAREX) (Pannett *et al,* 1998), or the UK Labour Force Survey (LFS) (LFS, 2009) and Census of Employment (ONS, 2009) combined with the best available estimates of the proportions in the workplace exposed (at relevant industry sector levels), provided the estimate of the exposed population. CAREX was often the best and only source for estimates of numbers exposed. In practice for exposures not covered by CAREX there was rarely an estimate of the proportions in the workplace exposed; estimates of numbers employed for the narrowest possible definition of an occupation (from the LFS) or industry (CoE) were then used. Exposed numbers from CAREX, or national employment sources, were allocated to a 'higher' or 'lower' exposure group assuming distributions of exposure and risk that corresponded broadly to those of the studies from which the risk estimates were selected. The initial allocations were based on the judgment of an experienced human exposure scientist; each assessment was then independently peer-reviewed and if necessary, a consensus assessment agreed. Data from CAREX are not differentiated by sex; CAREX exposed numbers were split between men and women according to an estimate of relative proportions based on numbers in the appropriate occupation by industry codes from the 1991 Census (details are in Appendix 5 Section A5.6). In order to present results by separate industries and occupations, the codes and descriptions encountered in the different classification schemes used for the CAREX, LFS, Census of Employment and other sources of data on exposed numbers were mapped onto a common set of 60 industry and occupation desciptions, used for all results presented by industry (Appendix 9).

CAREX data for Britain relate only to the period 1990-93. For the LFS and CoE an available year was chosen to represent numbers employed about 35 years before the target year of 2005, as this was thought to represent a 'peak' latency for the solid tumours, and is also close to the mid-point of the REP for estimating numbers ever exposed across the period (for which a linear change in employment levels was implicitly assumed). Where the Census of Employment was used, the data are for 1971 or 1981 for solid tumours. Where the LFS was used, the first year available was 1979 for solid tumours, and 1991 was used for short latency cancers.

When CAREX data were used, adjustment factors were applied to take account of the change in numbers employed in the primary and manufacturing industry and service sectors in GB between the late 1970s and early 1990s. The factors were estimated from LFS employment data, and were grouped by main industry sector, and for men and women separately (see Appendix 5 Section A5.2). Employment turnover (TO) factors were applied to the point estimates of numbers exposed, which were specific to the same grouped main industry sectors as the CAREX adjustment factors. The turnover factors were estimated using LFS data for the length of time with current employer, which was available by length of employment category, and was used to estimate numbers of new starters per year. Ideally an estimate of staff turnover requires full national starter and leaver data across the REP for all industry sectors. In the absence of this quality of data, estimating turnover directly using new starters in years within the REPs gives the best approximation for the purpose of estimating those ever exposed. This method estimates starters in the past year as a proportion of the average number employed (Gregg and Wadsworth, 2002). As exposure in occupational epidemiological studies is usually defined as for at least one year, we have adapted this to exclude short-term labour turnover by taking new starters in the past year who are expected to remain employed for at least one year as a proportion of all those expected to be employed for at least 1 year. This is estimated as the number recorded as employed for between 1 and 2 years divided by the total employed for at least one year using LFS data averaged over the REP (see Appendix 5 section A5.3). The estimates used were for those employed for at least one year. The numbers 'ever exposed' therefore depended on the timing and length of the 'risk exposure period' described above. The basic form of the equation used to calculate numbers 'ever exposed' during the REP taking into account turnover is:

$$N_{e(REP)} = n_0 + \{n_0 * TO * t\}$$

where n_0 = numbers employed in middle year of the REP
TO = staff turnover per year
t = number of years in the REP

When this is adapted to include survivors to the target year only, the equation (in a simplified form, allowing for entry into the cohort at age 15 only) becomes:

$$N_{e(REP)} = \sum_{i=65}^{100} \{l_i * n_0/(R - 15)\} + \sum_{j=25}^{64} \{l_j * n_0 * TO\}$$

where l_i = proportion of 15-year-olds surviving to age i, summations for REP of 1956 to 1995
R = retirement age

Life table estimates of the proportions of the general population surviving to the target year (2005) were used to adjust the estimates of numbers ever exposed obtained using the turnover factors, so that only the ever exposed cohort members surviving to the target year would be counted. This part of the estimation process assumed that there was an even distribution of ages across the exposed cohort in its first year (1956 for the 'solid tumour' example REP). We also assumed that recruitment to the cohort was across the age range of 15 to 24 only.

More detail on how numbers 'ever exposed' were calculated in the case of the industry-based studies is given in Appendix 5, including the full form of the turnover equation allowing entry into the cohort at multiple ages (equation 5.4). The CAREX adjustment factors and turnover factors (TO) used are listed in Table 2. In the last two columns of Table 2 turnover factors based on our estimates of length of REP, employment turnover factors and GB life expectancy are given. The formula used to estimate these factors is given in Appendix 5, sections A5.3 and A5.5. These values are close to that of 4 used by the Global Burden of Disease project (Nelson *et al,* 2005).

Table 2: CAREX adjustment and employment turnover factors used in the calculation of attributable fraction

	Main Industry Sector		CAREX adjustment factor	TO	Turnover factors based on a 40 year (long latency) and 20 year (short latency) REP and GB life expectancy tables	
					Long latency REP	**Short latency REP**
Men	A,B	Agriculture, hunting and forestry; fishing	1	0.07	3	2
	C,D,E	Mining and quarrying, electricity, gas and water; manufacturing industry	1.4	0.09	3	3
	F	Construction	1	0.12	4	3
	G-Q	Service industries	0.9	0.11	4	3
		Total	1	0.10	4	3
Women	A,B	Agriculture, hunting and forestry; fishing	0.75	0.10	4	3
	C,D,E	Mining and quarrying, electricity, gas and water; manufacturing industry	1.5	0.14	5	4
	F	Construction	0.67	0.15	6	4
	G-Q	Service industries	0.8	0.15	6	4
		Total	0.9	0.14	6	4

Where no CAREX data were available, as is generally the case where an exposure was defined by occupation rather than substance, the LFS was the source most commonly used. Data for 1979 (the earliest year available and closest to the ‘peak’ exposure period in the early to mid 1970s) was used for the solid tumours, and 1991 was chosen for the shorter latency cancers. As tight as possible a definition of ‘occupation’ was used to minimise overestimation that would result where there was no estimate available of the proportion of workers actually exposed within the target occupation.

To estimate the proportion of the population exposed, the estimated numbers ‘ever exposed’ during the risk exposure period were then divided by the total population ever of working age during the same period (given in Table 1 above for the standard REPs). This was based on population estimates by age cohort in the target year (see Appendix 5 Section A5.4).

For some cancer/exposure pairings it was necessary to partition the risk exposure period between ‘higher’ and ‘lower’ exposed groups, if an industrial process was known to have changed during the period or controls had been put in place but the exposure has not disappeared. The ‘old’ industry-based relative risks were then applied to the earlier-exposed group, and a lower relative risk estimate to the more recently exposed group. The method for partitioning the proportion of the population exposed between two periods is detailed in Appendix 5 (Section A5.6).

3.6 CALCULATING THE ATTRIBUTABLE FRACTION

As previously described, the method for calculating attributable fraction (AF) depended on the data source. If the relative risk was extracted from a population-based study the estimate of the proportion of cases exposed was extracted directly from the study. In this case Miettinen’s equation (Miettinen, 1974) for calculating the attributable fraction was used:

$AF = Pr(E|D) * (RR – 1) / RR$
where $Pr(E|D)$ = proportion of cases exposed (E = exposed, D = case)

However in practice such studies were rarely available, and Levin's equation which can be used in the case of industry-based studies with an estimate of the proportion of the population exposed from an independent data source, was nearly always used to calculate the AF (Levin, 1953):

$AF = Pr(E) * (RR – 1) / \{1 + Pr(E) * (RR – 1)\}$
where $Pr(E)$ = proportion of the population exposed

Levin's and Miettinen's equations are equivalent, and can be derived from one another using the definition of relative risk (RR) and Bayes Theorem (Benichou, 2001).

Separate AFs were calculated by industry/occupation, and for the 'higher' and 'lower' exposed groups, and then summed to give an overall AF for the specific cancer/exposure pair (see Appendix 6 Section A6.3).

Full details of the methods for calculating AF are outlined in Appendix 6.

For a few cancer/exposure pairings, in particular mesothelioma, which is uniquely associated with asbestos exposure, absolute risk measures rather than relative risks were appropriate. The occupational AF depended only on the degree to which exposures to asbestos were considered to be non-occupational. The method used to calculate AF and attributable numbers for mesothelioma is covered in the Mesothelioma Technical Report.

In the case of lung cancer associated with asbestos exposure, the estimate of attributable numbers and hence of AF was derived directly from the ratio of lung cancer to mesothelioma cases estimated amongst asbestos exposed workers (Appendix 7 Section A7.1). Specialised methods were also used to calculate the AF for lung cancer due to radon exposure, based on estimates of numbers developing lung cancer from exposure to radon in domestic buildings (see Lung Cancer Technical Report, Annex 1).

The risk estimates for occupational exposure to ionising radiation were derived using generalized linear dose response models of excess relative risk per unit of cumulative radiation dose from the United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR, 2006), and were therefore applied to cumulative lifetime dose estimated using data from the Central Index of Dose Information (CIDI) (HSE, 1998). For aircrew, who are not covered by CIDI, the mean total lifetime radiation dose per pilot was obtained from a large cohort study of European airline pilots (Langner *et al,* 2004) and combined with numbers employed obtained from the British Airways Stewards and Stewardesses Union (BASSA, 2008) (Appendix 7 Section A7.4).

3.7 COMBINING THE ATTRIBUTABLE FRACTIONS ACROSS EXPOSURES

Attributable fractions cannot in general be summed directly, due mainly to the possibility that workers will have been exposed to more than one occupational carcinogen during their working lifetimes in the REP. There are various strategies to handle overlapping exposures, which are described in Appendix 6 Section A6.3, using examples from the analysis for lung cancer based on the exposure map illustrated in Figure 2. The aim of the strategies is to restrict the estimate of the exposed population to eliminate overlaps. However, summing even disjoint exposure AFs estimated using Levin's equation can introduce bias. If it was known that the exposures were independent (including disjoint exposures) and their joint effect on the cancer was multiplicative the AFs were combined by taking the complement of the product of complements (Appendix 6 Section A6.3):

$AF_{overall} = 1 - \Pi_k(1-AF_k)$ for the k exposures in the set.

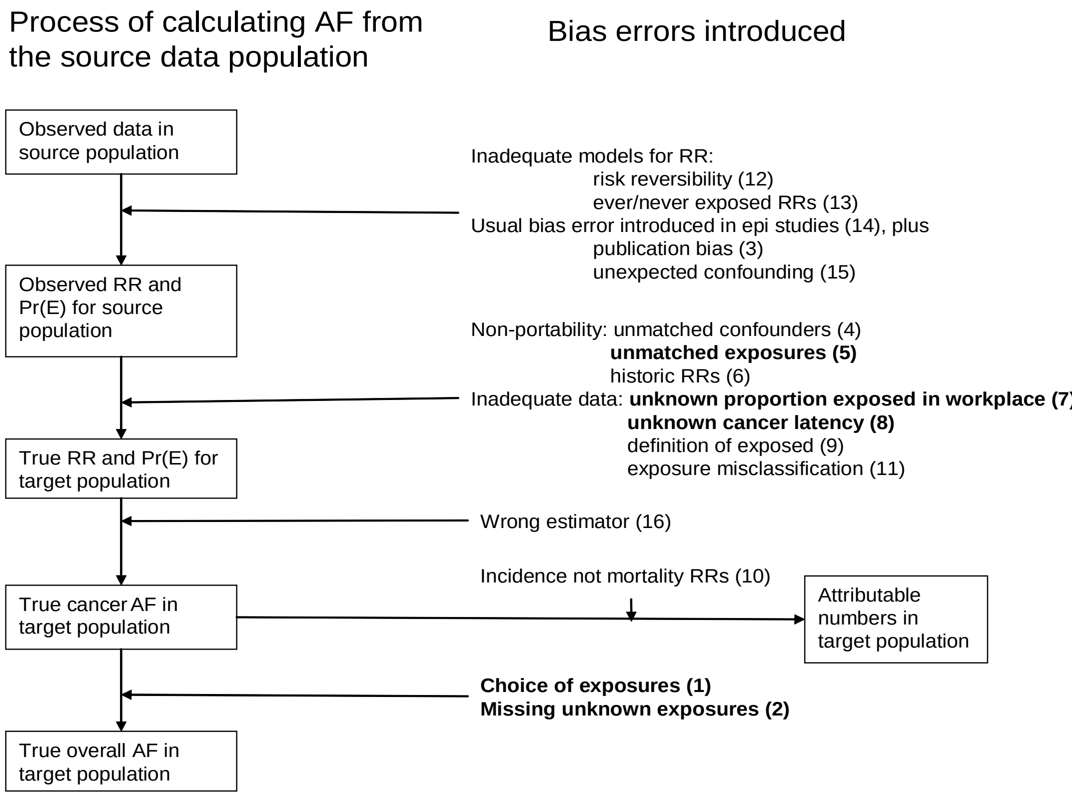
AFs for the strong, and for the strong plus suggestive, groups of exposures were estimated separately.

3.8 CONFIDENCE INTERVAL FOR THE OVERALL ATTRIBUTABLE FRACTION

A 95% random error only confidence interval was calculated for each AF, based only on the variability of the relative risk estimates used, and using Monte Carlo methods where more than one relative risk was involved (see Appendix 8).

In future work, Monte Carlo methods (Greenland, 2004) will be used to calculate a 'credibility interval' (a confidence interval covering bias as well as random error) for the combined attributable fractions. These limits are able to take into account errors due to bias at all stages of the estimation process, as well as random error in the estimate of relative risk. The details of this work will follow in a subsequent report. An indication of the sources of bias identified in the calculation of attributable fraction that need to be accounted for in the 'credibility interval' for the attributable fraction are given in Figure 3. The sources of greatest bias are in bold (Numbers 1,2,5,7 and 8) in Figure 3.

Figure 3: Sources of bias in the calculation of an overall attributable fraction



3.9 ESTIMATING ATTRIBUTABLE NUMBERS

Attributable numbers were estimated by multiplying the attributable fraction by deaths from the cause with the matching ICD code (from DH2 series mortality statistics (ONS, 2005) for England and Wales plus equivalent data for Scotland), and by cancer registrations (from the MB1 series (ONS, 2006) for England plus equivalent data for Wales and Scotland) for the target year. This was the most recent year for which mortality or registration data were available, 2005 for deaths, and 2004 for cancer registrations, for England, Wales and Scotland.

Data were not always available for some cancer sub categories (for example some leukaemia sub-types in the Welsh and Scottish data). In this case sub-type estimates were taken from proportions in English data or other sources.

For the purposes of these estimates, it was acceptable to apply attributable fractions calculated from relative risks, which were derived from either incident or mortality data, or both as in the case of pooled studies and meta-analyses. In doing this, the assumption was made that survival as a result of occupational exposures was similar to survival as a result of non-occupational exposures. If, for example, survival times were reduced amongst the (perhaps more heavily) occupationally exposed, than amongst the non-occupationally exposed cases, AN would have been underestimated if an incidence attributable fraction was applied to mortality data.

3.10 COMBINING THE ATTRIBUTABLE FRACTIONS ACROSS CANCER SITES

To obtain the overall occupation attributable fraction attributable numbers for each cancer were summed, and divided by total cancer numbers. This assumes that all the cancers being considered as caused by occupational exposure were relatively rare in the GB population. If they were common and a single exposure was known to produce cancer at more than one site in an individual at a specific time then double counting of that individual's risk could have occurred in estimating overall attributable fraction based on cancer registrations. However only one primary cancer would be recorded in the mortality figures as the underlying cause of death.

Overall attributable numbers and therefore attributable fraction has been estimated separately by exposure, or by exposure level, in the same way. Note that if attributable numbers summed across cancers by exposure are then summed the result will not normally equal the overall attributable numbers estimate because of overlapping exposures. The product sum of attributable fractions should always be used.

Different results were obtained for the combined attributable fraction depending on whether it was based on mortality and on incidence data. The difference depended on the ratios of deaths to registrations in the cancers that were included (i.e. on survival rates). An attributable fraction based on attributable deaths would have under weighted the contribution of 'good' survival cancers, such as non-melanoma skin cancer (NMSC), which is rarely fatal but occurs in relatively large numbers. On the other hand an attributable fraction based on incidence would have given all the cancers equal weighting so that the large numbers of NMSCs would have made a disproportionately large contribution. This indicates the value of using such estimates as Disability Adjusted Life Years (DALYs) in conjunction with attributable numbers to estimate an overall attributable fraction weighted according to the relative severity of the cancer outcome (especially as survival improves). Details of the calculation and use of these estimates are given in a separate report (Hutchings, 2010a).

4 ESTIMATING FUTURE BURDEN

The method used to estimate future burden of occupational cancer builds on the approach using attributable fractions to estimate the current burden of occupational cancer, by projecting risk exposure periods (REPs), which allow for cancer latencies of up to 50 years for solid tumours, forward in time, to estimate attributable fractions (attributable fractions) for a series of forecast target years. Estimates for years up to 2060 are required to allow for latency. The continuing effect of past exposure is allowed for in estimates for the earlier years. Adjustment factors are applied in estimation intervals within the REPs so that attributable fractions and attributable cancer numbers can be estimated for different forecasts of exposed numbers and exposure levels, and the effect of reducing exposure levels can be tested. The factors allow for changes in exposed numbers, and for changes in exposure levels, and the relative risks also can be adjusted. Full details are in the Future Burden Methodology Technical Report (Hutchings, 2010b).

5. APPENDICES

APPENDIX 1: CRITERIA FOR CHOOSING AN APPROPRIATE ('BEST STUDY') FOR RELATIVE RISK ESTIMATES

The study needed to be broad based and representative of the occupations or industrial exposures that were encountered in Great Britain, and had to cover women as well as men, if this was relevant to the exposure and cancer. Other criteria were those that defined a good epidemiological study. These were:

- a large sample size (so that estimates were robust);
- an appropriate comparison population e.g. national and large in number for a population-based study;
- controlled for the non-occupational confounders identified as important.
- adequate exposure assessment
- a clear and standardised case definition e.g. if the ICD codes could be identified then the results could be applied correctly to national mortality and incidence data.
- Portability of the relative risks to the GB population, e.g. for a population-based study from which the proportion of the population exposed was to be used, portability between the source (study) population and the target (GB) population was critical. Important issues concerning portability included:
  - Matching across populations of the types and levels of exposures; it would be preferable to use GB based studies or those from countries with a similar industrial heritage and level of development.
  - Matching of exposures across time, so that exposures from the past, in the source risk exposure period (REP, see Appendix 4), were similar to those in the target REP. It was, therefore, important that the source risk exposure period was not so long ago that the exposures on which the relative risk estimates were based no longer matched.
  - Less important was that the distribution of known matched confounders in the source and target populations.

Studies were required for all relevant occupations, chemicals, industries or occupational situations that were identified (the exposure scenarios). The nature of the substance, and of the tasks, jobs and processes needed to be similar to the GB situation.

Large pooled studies or meta-analyses and reviews offering cross-study estimates were usually to be the studies of first choice. The advantages of combined analyses were increased precision due to increased sample size and the opportunity to explore potential confounders, interactions and modifying effects that may explain heterogeneity among studies (Greenland, 1998). A disadvantage of combined analyses was the possible lack of compatibility of data from various studies due to differences in subject recruitment, procedures of data collection, methods of measurement and effects of unmeasured covariates that may differ among studies[6]. Heterogeneity was an important consideration in combined analyses and needed to be accounted for (see Appendix 3).

Population-based studies had the potential advantage of covering multiple exposures and so accounting for overlaps in the working population between these exposures. They were less likely to provide precise definitions of exposure as they are often based on 'self-reported' occupational histories. These were also generally limited to an 'ever/never' exposed dichotomy. Population-based studies were also unsuitable to pick up rarer occupational exposures, unless they were very large.

[6] IARC Monograph Preamble, 2006

Industry-based studies were likely to provide more precise exposure data analysed by level and duration of exposure, but required comparable independent estimates of the proportion of the GB population exposed.

Odds ratios (ORs) from case-control studies, standardised mortality ratios (SMRs) from cohort studies or proportional mortality ratios (PMRs), were all used as relative risk estimates in the calculation of AF. In the case of ORs however, the ‘rare disease’ assumption (that the probability of disease was very small) needed to be satisfied. The choice of ‘best study’ did not therefore depend on whether it was a case-control or cohort study (although the quality of exposure assessment may have been an issue here). Whether the study was population-based or industry-based, however, determined the way in which the attributable fraction was calculated.

Where a UK study was available, this was always considered for inclusion. This was achieved for example by combining the results from the UK study with a ‘best study’ review from which this study was omitted.

APPENDIX 2: DECISION-MAKING PROCESS FOR SELECTING CANCER/EXPOSURE PAIRINGS

For each cancer site, firstly all known and suspected occupational carcinogens, or exposure scenarios (usually occupations) where the causal agent may be unknown, were listed. The main source for this list was the substances and occupations classified by IARC as group 1 and 2 carcinogens, as tabulated by Siemiatycki *et al.* (2004). This list was supplemented by information from the HSE Manchester Burden of Cancer workshop (2004)[7]and from a priority list of carcinogens drawn up for a carcinogen hygiene profiling exercise developed for HSE's Disease Reduction Programme. The Occupational Health Decennial Supplement was an additional source of information.

For each exposure, a list of occupations and industries in which the exposure was known to have occurred was then drawn up. These were the exposure scenarios. The main source for this was Siemiatycki *et al.* (2004), specifically Tables 3, 4 and 6 from this paper.

Each exposure listed then had to pass three tests to be retained in the analysis, as follows:

Test 1: **Was the exposure causal?** From this test, the exposures were allocated to one of two groups based on the strength of evidence for causality. Substances or occupations in IARC group 1 and with 'strong' evidence of carcinogenicity in humans for the relevant cancer site (from Siemiatycki *et al,* 2004, Tables 3 and 6) were allocated to a 'strong' exposure group. Substances or occupations in IARC group 1 or 2A and with 'suggestive' evidence of carcinogenicity in humans for the relevant cancer site (from Siemiatycki *et al.* 2004, Tables 3, 4 and 6) were allocated to a second tier 'suggestive' group. Additional exposures identified from the other sources were allocated to the 'suggestive' group unless there was 'strong' evidence to the contrary, probably founded in UK experience.

Test 2: **Did the exposure scenario (i.e. occupation or industry situation) exist in Great Britain during the target 'risk exposure period'?** This was defined by cancer latency (see below.) For each industry or occupation listed for an exposure, if the industry did not exist in GB, or if processes had been changed by the mid 1950's for solid tumours, or by around 1985 for the shorter latency neoplasms, it was excluded. If the industry had disappeared or the processes had changed by the early 1970's for the solid tumours or by around 1990 for the shorter latency neoplasms, this was accounted for in the analysis. If effective controls were known to have been introduced before this time, but the exposure scenario still existed (or exposure levels had otherwise been substantially reduced), this was noted and accounted for in the analysis.

Test 3: Did the exposure scenario (occupation or industry) produce more than two cases per year? Cases would be the number of deaths for short survival cancers, or registrations for those cancers with better treatment outcomes. In practice this restriction was not used for the GB study, and results are quoted for all cancer exposure pairings with no minimum cut-off, unless numbers associated with an exposure were very low indeed. Figures A2.1 and A2.2 illustrate the size of the exposed population for which two deaths from cancer would be expected to occur given a range of death rates and relative risks of between 1 and 5. Numbers ever exposed for the figures are estimated as observed (2) divided by relative risk divided by Death Rate; A point estimate of numbers exposed is the numbers ever exposed/4where 4 is from Table 2 and is based on our estimates of length of REP for a long latency cancer, employment turnover factors and GB life expectancy for men.

The graphs indicate that minimum point estimates of approximately 500 exposed workers for lung cancer (male DR=513/million), and 2,000 for the less common bladder cancer (male DR=84/million),

[7] GB Burden of Occupational Cancer: Summary Report of Workshop Held on the 22nd and 23rd November 2004 in Manchester: HE/05/03

would be required for the maximum relative risks identified in the study. The CAREX tables of exposures by agent and industry[8] for GB would be the usual source for making this judgement.

Where there was a clear overlap between populations subject to individual exposure scenarios for a single cancer, the exposure scenario was excluded from the analysis if it was clear from a 'best study' that the estimated raised risk was attributable to another (dominant or higher relative risk (RR)) exposure scenario.

Figures A2.1 & A2.2 Exposed population size required to produce 2 observed deaths for a range of possible death rates and a given relative risk for a large (A2.1) and small population (A2.2)

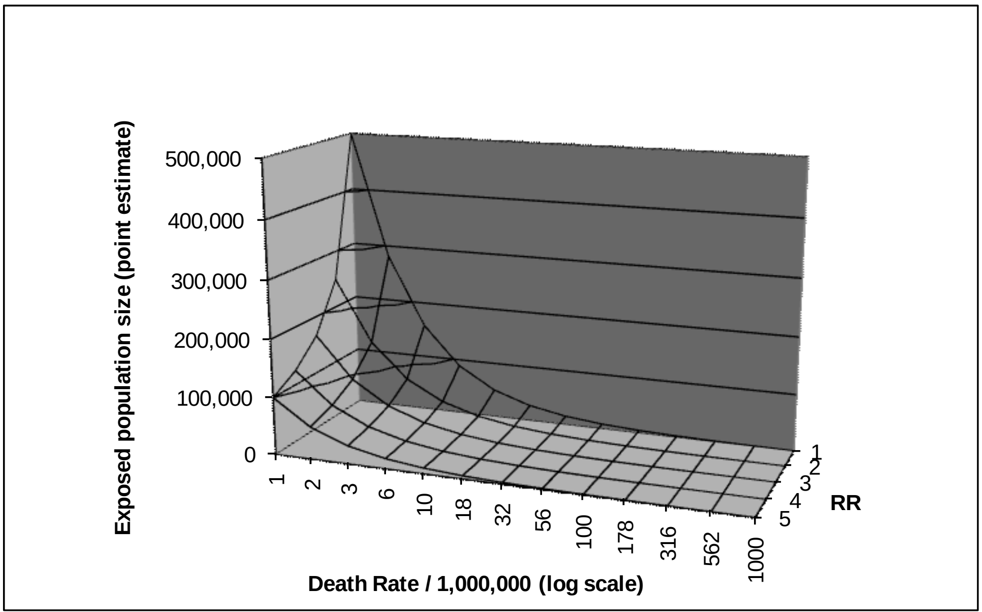
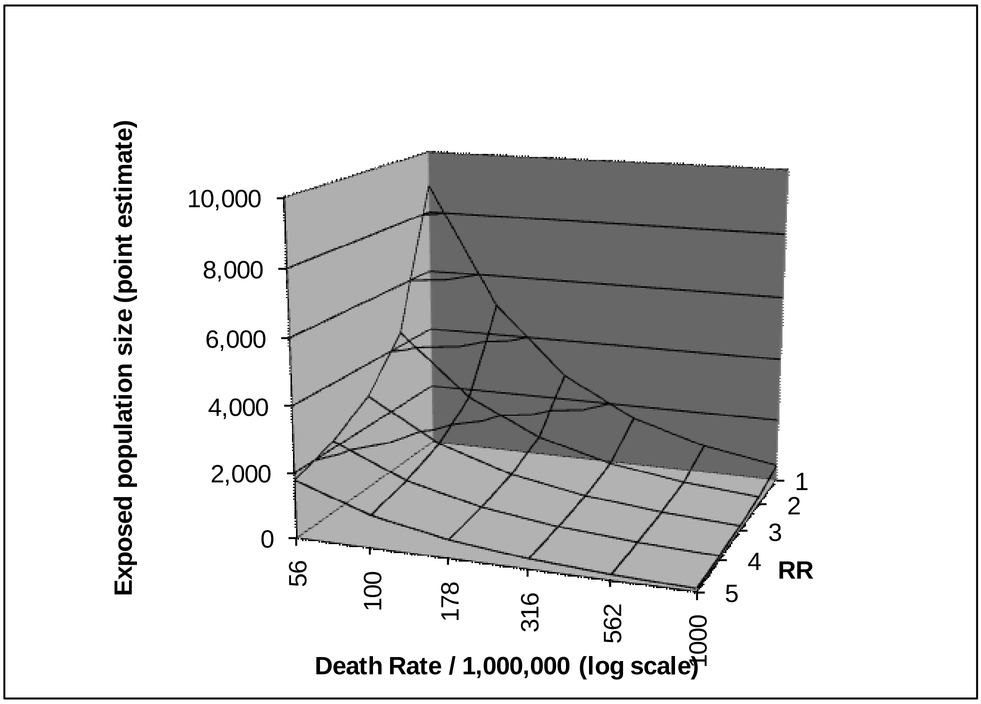


Figure A2.2



[8] CAREX: Industry specific estimates, Great Britain 1990-93. Finnish Institute of Occupational Health

APPENDIX 3: CRITERIA FOR SELECTING THE RELATIVE RISKS TO USE FROM THE 'BEST STUDY', AND MATCHING TO APPROPRIATE EXPOSURE DATA

A key relevant paper, or 'best study', was needed for each exposure scenario that had been identified. Relative risks adjusted for the effects of known confounding (adjustment) factors, and for the effects of interaction between the exposures of interest and adjustment factors, were normally preferred. This was particularly the case for 'ever/never' exposure estimates.

For each exposure a minimum level at which disease can be induced was identified explicitly. This usually came from the choice of a cut-off in the chosen 'best study' although, particularly for population-based studies relying on self-reported employment history, occupation was sometimes all the information given.

When the 'best study' was a review, or meta-analysis that did not offer a suitable pooled or combined estimate, an average overall estimate (inverse-variance weighted) of the relative risk and its (95%) confidence interval were calculated from the data. This estimate was based on a fixed effects model if the standard test for heterogeneity of the constituent relative risks (RRs) was not significant (at the 95% level), otherwise a random effects model was chosen (Section A3.1).

The exposed population was generally partitioned between 'higher exposed' and 'lower exposure' groups, and an appropriate relative risk from the source study or studies was applied to each group separately. Ideally an exposure level threshold was identified that separated the 'higher' and 'lower' exposure groups, but in practice exposure scenarios usually defined the groups.

The exposed numbers were partitioned to match the 'higher' and 'lower' RR estimates, on the basis of job/industry codes. In the case of CAREX data, exposures were coded to the UN ISIC Rev2 (1968) classification, which were matched to SIC68 codes. Exposed numbers for codes that matched the 'best study' high exposure scenarios were assigned to the 'higher exposed' group, and other codes to the 'lower exposure' group, so long as the exposure scenario passed the inclusion tests described in Appendix 2 (see also Appendix 5 Section A5.6).

Where there were no CAREX data, national employment estimates were used. Employee counts were needed for industry, or occupation codes, that matched as precisely as possible to the occupation and industries for the relative risks derived from the best study. Data were available from the Census of Employment (CoE) for the period 1971 to 1991, and from its successors (the Annual Employment Survey (AES) and the Annual Business Inquiry (ABI)) for 1991 – 1998 and 1998 – 2004 respectively, or the Labour Force Survey (LFS) from 1979 onwards. The CoE (for industry-based exposure classes), or the LFS (for occupations - earliest available date 1979) were used for exposed counts for 1973 for solid tumour cancers, and the AES and LFS for 1991 for the shorter latency cancers (see Appendices 4 and 5). These counts also were allocated to 'higher' and 'lower' exposure groups on expert judgement, unless there was any data available from HSE (the National Exposure DataBase for example) to base this on.

There were potential problems in matching RRs, which were often based on precisely defined levels of exposure to the (historic) GB working population based on occupation or industry codes without a detailed job/exposure matrix or good exposure assessment information. Expert judgement was used in these circumstances.

A3.1 Estimating a weighted average relative risk and its confidence interval

Occasionally it was necessary to calculate a weighted average across a range of SMRs or ORs where a suitable overall estimate was not offered by a 'best study'. In these circumstances a precision (inverse variance) weighted average was calculated (Rothman and Greenland, 1998). The weights were the inverse of the squared standard error:

$$w = 1/SE^2$$

so that $RR_{overall} = \Sigma_i (w_i RR_i)/ \Sigma_i w_i$ where RR_i = relative risk; w_i = weight for each estimate i

The standard error of this estimate is given by:

$$s = 1/(\Sigma_i w_i)^{1/2}$$

For RR estimated by an SMR (Breslow and Day, 1989):

$$SE(\ln(SMR)) = 1/D^{1/2} \quad \text{where} \quad D = \text{observed deaths}$$

$$w = D$$

$$\ln(SMR_{overall}) = \Sigma_i (w_i * \ln(SMR_i))/ \Sigma_i w_i$$

$$s = 1/(\Sigma_i D_i)^{1/2}$$

Therefore

$$SMR_{overall} = \exp\{\Sigma_i (w_i * \ln(SMR_i))/ \Sigma w_i\}$$

and an approximate 95% confidence interval for the weighted average estimate was given by:

$$CI = \exp(\ln(SMR_{overall}) \pm 1.96 * s)$$

For RR estimated from an OR:

$$SE(\ln(OR)) = (\ln(UCL) - \ln(LCL)) / (2 * 1.96)$$

where UCL and LCL are the upper and lower 95% confidence limits for OR; this was continued as for the SMR.

This was the estimate for a fixed effects model. The estimate for a random effects model was calculated using the methods described in DerSimonian and Laird (1986). This was programmed in the AF calculation EXCEL spreadsheets, or obtained using MIX software.

A3.2 Estimating a 'lower' exposure group relative risk where none was available

In the absence of any 'best study' data for the 'lower' exposure group, instead of applying a relative risk of 1 as a default, which was used in the preliminary estimation (Rushton *et al,* 2008), a RR for the 'lower exposed' group was estimated by taking the harmonic mean of all the available ratios of 'higher' to 'lower' RR estimates for cancer-exposure pairs for which data was available, and also separately for their upper and lower 95% confidence limits, and applying this average ratio to the 'higher level' estimates to obtain 'lower level' RR estimates and CIs. If the resulting RR estimate was less than 1, RR was set to one. This resulted in a zero AF for these exposed groups, but allowed recognition in the tabulated results of their zero contribution. Figure A3.1 shows the high and low exposure level RRs on which the average ratio estimate was based, and their inverse variance weighted and harmonic means. Of these two alternative estimates of an average ratio, the one based on inverse-variance weighted averages of the high and low RRs down-weights the higher RR-high estimates as these tend to have the wider CIs (e.g. based on lower case numbers), and therefore produces unrealistically low results where the high RRs are greater than 2. The other estimate based on harmonic means of the RRs, i.e. equal weighting, results in more realistic RR estimates. However there are more low RRs set to one using this method. The estimated relative risks used are in Table A3.1.

Figure A3.1 High and low exposure level relative risks used to estimate 'missing' low exposure level RRs

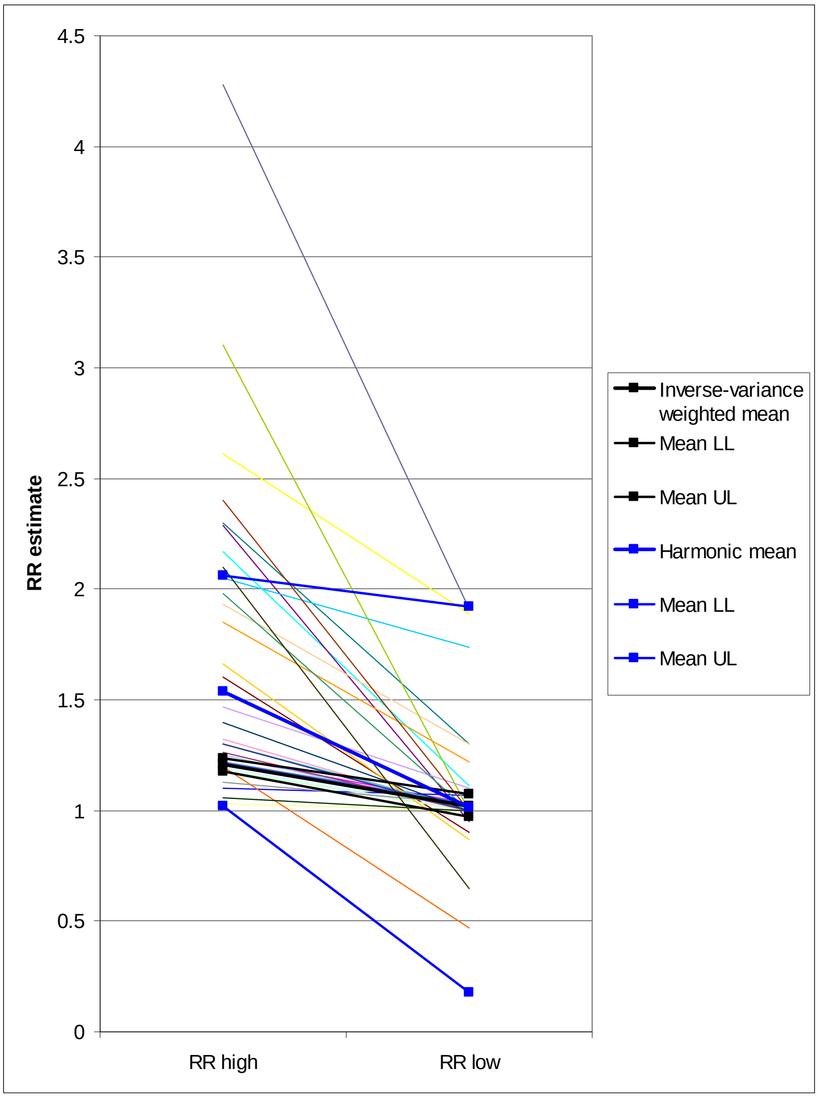


Table A3.1 'Lower exposure' relative risks estimated using a harmonic mean

Cancer	Substance	'High' exposed			'Low exposed' (harmonic mean)		
		RR	LL	UL	RR	LL	UL
Bladder	mineral oils	1.39	1.2	1.61	1.0		
Lung	beryllium	1.12	0.99	1.26	1.0		
	ETS	1.16	1.05	1.28	1.0		
	mineral oils	1.08	1.04	1.11	1.0		
	lead	1.14	1.04	1.73	1.0		
	strong inorganic acid mists	1.36	0.97	1.94	1.0		
NMSC	mineral oils	1.21	0.48	3.06	1.0		
Sinonasal	chromium	5.18	2.37	11.3	3.4	0.4	10.5
	mineral oils	2.8	1.4	5.7	1.8	0.2	5.3
	wood dust	3.1	1.6	5.6	2.0	0.3	5.2
Oesophagus	tetrachloroethylene	2.47	1.35	4.14	1.6	0.2	3.9
Brain	lead	1.06	0.8	1.4	1.0		
Cervix	tetrachloroethylene	1.95	1	3.4	1.29	0.18	3.16
Larynx	asbestos	1.37	1.17	1.6	1.0		
Stomach	inorganic lead	1.5	1.8	1.49	1.0		
NHL	tetrachloroethylene	1.39	0.56	2.86	1.0		
NHL	trichloroethylene	1.29	1	1.66	1.0		
Liver	VCM	2.86	1.83	4.25	1.9	0.3	4.0

APPENDIX 4: DEFINING THE RISK EXPOSURE PERIOD

The 'risk exposure period' (REP) was the period during which exposure to a causal factor resulted in a cancer mortality or registration. This depended on the cancer latency.

Below is an example of an individual case time line for a specific cause of death related to an occupational exposure:

```
1        2        3                    4            5             6  7  8
|________|________|____________________|____________|_____________|__|__|
                  LL                   UL
```

Key:

1	age 16
2	start of employment in a relevant industry or occupation
3	LL = start of REP
4	UL = end of REP
5	control measures introduced /process closed
6	start of risk period for development of disease (RP)
7	death from the specific cause
8	end of RP

where	for the case	(LL) to (7) is the maximum latency (l_{max})
		(UL) to (7) is the minimum latency (l_{min})
	In general	(LL) to (6) is l_{max}
		(UL) to (8) is l_{min}

If points 1 and /or 2 occurred after LL (for example if someone started work after the start of the risk exposure period), 2 (or if unavailable then 1) replaced LL as the start of the individual's REP.

Where an exposure occurred over a defined period of time only (for example prior to the introduction of effective control measures, or closure of the industry/process), the beginning and end of the REP was adjusted to correspond to this. For example if point 5 occurred before UL, 5 replaced UL as the end of the REP. If 5 preceded LL, the REP disappeared. In this case the exposure failed inclusion test 2 (see Appendix 2).

Distinction should be drawn between the source REP and target REP. The source REP is the period for a source study during which cases contributing to the results of that study (in particular the estimates of relative risk) could have been initiated. It is relevant to the calculation of AF only in determining the portability of the estimate of relative risk between the source and target populations (see Appendix 1). The target REP is the period during which cancers appearing in the national mortality and registration statistics for the target year may have been initiated. It is defined on the estimated maximum and minimum latency for each cancer, and is used in the estimation of the numbers and therefore the proportion of the population that has ever been exposed to the occupational carcinogen being considered.

The *source REP*, for the study chosen for the estimate of relative risk, was calculated as:

Source REP =
[*source risk period* minus maximum cancer latency] to [*source risk period* minus minimum latency].

The source study risk period is the period of case collection for a case-control study or follow-up for a cohort study. The *target REP*, used to estimate AF was calculated as:

Target REP = [*target year* minus maximum latency] to [*target year* minus minimum latency].

The target year was the year for which an AF was to be estimated. This was the most recent year for which cancer mortality, or registration data, was available (currently 2005 for mortality and 2004 for registrations), or it can be set in the future, if predictions based on current exposures are needed.

For example, to estimate the burden of cancer in the year 2005 for a solid tumour for which latency of between 10 and 50 years has been assumed, the target REP would be 1956 to 1995. However, the best source study may have been carried out on cases collected around 1990 representing a source REP of 1941 to 1980. If the proportion of those exposed (if it is a population-based study), and the level of that exposure had remained constant over time this would not be a problem. If this was not the case adjustments to the estimates of the proportions exposed (overall and at the 'higher' and 'lower' exposure levels) would be needed.

Where external estimates of proportions exposed were used, such as from numbers employed in an occupation or industry in GB, these were chosen to apply to the risk exposure period or, in the case of the CAREX estimates, adjusted to apply to this period (Appendix 5 Section A5.2).

In practice, very little information was available on latency for individual cancer sites and cancer/exposure pairings. Therefore for the sake of clarity, a single target risk exposure period for the solid tumours, and a separate single target REP for the haematopoietic cancers were used throughout, unless the period was shortened due to specific changes to the individual exposures. For the target year of 2005 these were as follows:

Solid tumours: The REP was defined as 1956 – 1995, based on a maximum latency of 50 years and a minimum of 10 years, and with a 'peak' at about 35 years (i.e. the highest number of solid tumours leading to current deaths or registrations related to exposures in the early 1970s).

Haematopoietic neoplasms: The REP was defined as 1986 – 2005, based on a maximum latency of 20 years and a negligible minimum, with a 'peak' at about 15 years (i.e. the highest number of haematopoietic leading to current deaths or registrations relate to exposures in the early 1990s).

Data used in the calculation of AF specific to each standard REP are given in Table 1 in Section 3.5. Although strictly speaking the REP for cancer registrations recorded in 2004, the year for which estimation has been carried out, would be 1955-1994 for solid tumours and 1985-2004 for haematopoetic neoplasms, for simplification the years 1956 to 1995 and 1986-2005 respectively have also been used, as for deaths, as the proportion exposed will not be affected.

APPENDIX 5: ESTIMATING THE PROPORTION IN THE POPULATION EXPOSED

For industry-based studies, the proportion of the population exposed was calculated from independent data. Here the industry study members were not considered representative of the general population as a whole. Indeed, an industry cohort was in certain circumstances considered 100% exposed, depending on the definition of exposed versus unexposed.

A5.1 Data used for the estimates of numbers exposed

Two estimates are needed from independent data to arrive at the proportion of the population exposed, namely an estimate of the numbers ever employed in exposed industries or jobs during the risk exposure period, and the proportion of these actually exposed to the carcinogen in the workplace. The proportion in the workplace exposed to the substance of interest [$Pr_{wk}(E)$] is the proportion above a defined baseline level, to match the definition of exposure used for the source study estimate of relative risk (RR). This was in practice very difficult to obtain, as it required numbers exposed and unexposed in the workplace, as well as substance level measurements to define the exposure. Apart from the possibility of sporadic coverage in the National Exposure Database (NEDB), the Carcinogen Exposure Database (CAREX) was the only available comprehensive source for this part of the estimate. CAREX gave numbers exposed in 1990-93 as a point estimate by industry sector (n_0), as well as numbers employed (N_e), so that

$$n_0 = N_e * Pr_{wk}(E) \quad (5.1)$$

Where CAREX data were not available, the numbers ever employed in that industry sector (or occupation category) during the risk exposure period had to suffice, and the proportion exposed in the workplace [$Pr_{wk}(E)$] was then set to one in the absence of any alternative estimate, but it was likely to have overestimated actual substance exposures. The important issue was how close the match was between the definition of the exposed versus the non-exposed used in the study from which the RR estimate was extracted, and the nearest approximation that can be obtained from the independent SIC code based GB national employment estimates. An estimate of the numbers employed in exposed industries or occupations (N_e) was obtained from national employment data (usually the Census of Employment (CoE), or the Labour Force Survey (LFS)), using where possible the relevant Standard Industrial Classification (SIC) or Standard Occupational Classification (SOC) codes.

In the case of LFS and CoE data, data were used for a year as close as possible to the mid-point of the REP, usually 1979 (the earliest available) for LFS data or 1981 (CoE) for long latency cancers, or 1995 for the short latency REP. This was so that the point estimate would best represent average employment levels during the REP.

A5.2 Adjusting CAREX exposed numbers to the risk exposure period [$N_{e(adj)}$]

There have been some significant changes in the structure of industry and therefore patterns of employment in GB in the last 35 years. Manufacturing industry and employment in mining and agriculture in particular have declined, and numbers employed in service industries have increased. As most cancers have a relatively long latency period, these changing patterns of employment needed to be taken into account when estimating the numbers exposed to occupational carcinogens.

Figures A5.1 and A5.2 illustrate the changes in employment over the time periods for which we had data, for grouped main industry sectors which were comparable across the different classifications used over the years, for the LFS and industry-based surveys. Due to changes in the way the broad industry groups were defined between SIC68 and SIC80, mining and quarrying and energy and water industries were grouped with manufacturing. Data were used for 1971 to 2004 based on the CoE and its successors, and for 1979 to 1997 based on the LFS, which includes estimates for the self employed. Employment patterns are thought not to have changed significantly between the mid 1950s and early 1970s.

Both data sources indicated a sharp decline in employment in manufacturing industry and mining between the late 1970s and early 1990s. The LFS included the self-employed, and was therefore a better source for overall employment and for the construction industry. The LFS data for agriculture also looked more realistic when matched to numbers for farmers and their employees from the June Agricultural Census.

Figure A5.1

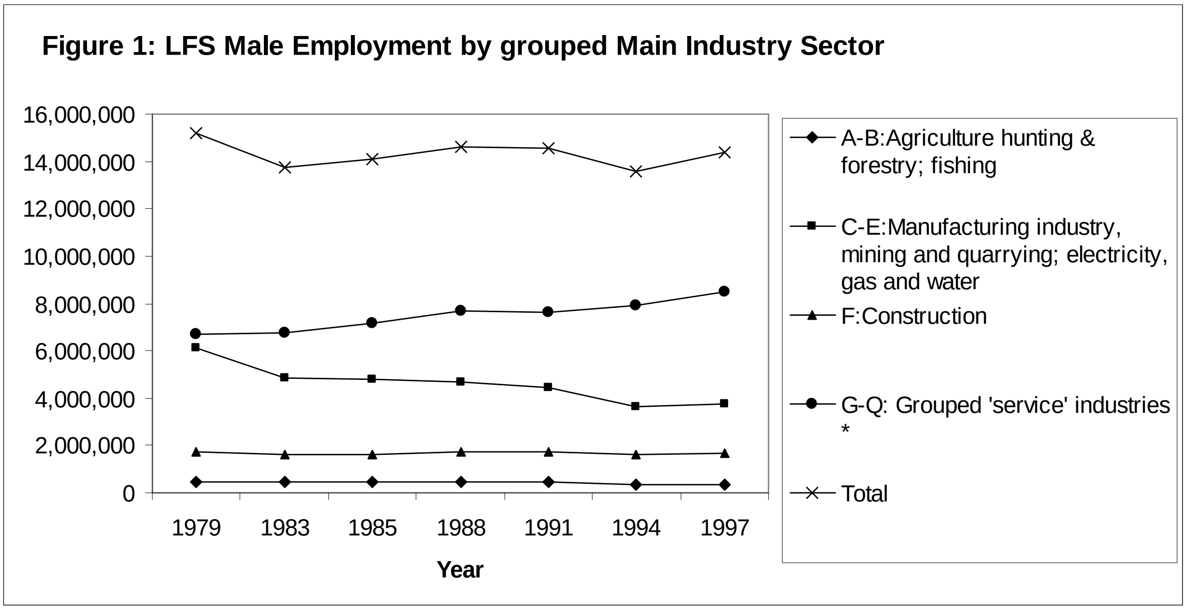
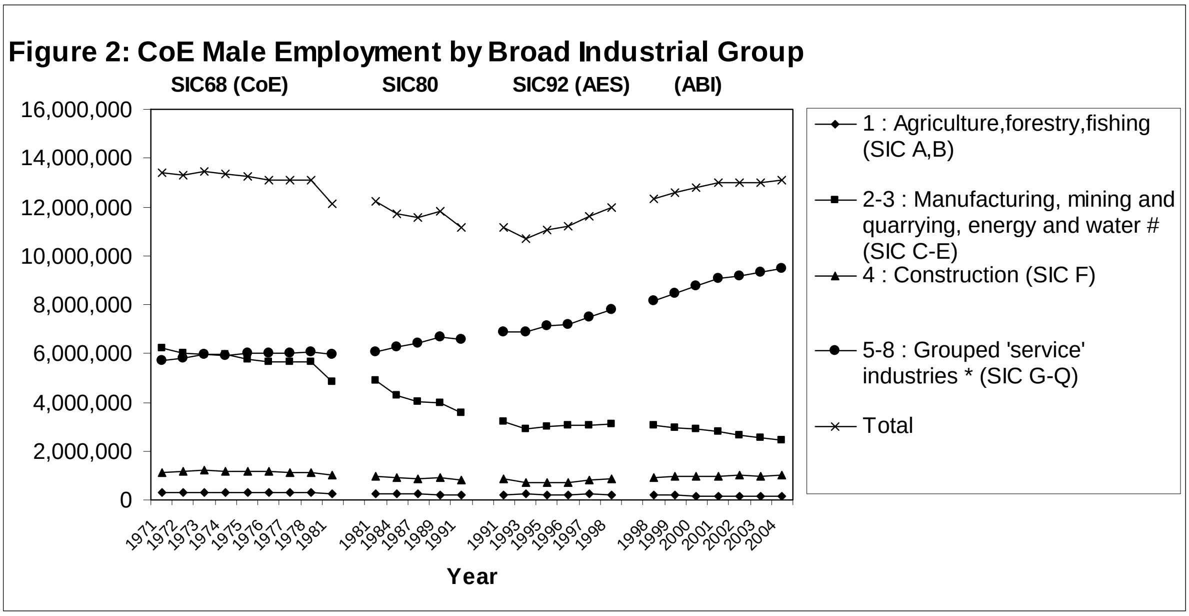


Figure A5.2



SIC68: 02.Mining and quarrying, 2.Gas, electricity and water, 3.Manufacturing industries.
SIC80: 2.Energy and water supply, 3.Manufacturing industries.
* SIC68: 5.Distributive trades, 6.Transport/communication, banking, finance, 7.Public administration and defence, 8.Miscellaneous services.

SIC80:	5.Distribution, hotels/catering, repairs, 6.Transport/communication, banking, finance, 7.Public administration and defence, 8.Other service industries

The CAREX data were for 1990-93. The risk exposure period for solid tumour cancers was 1956 to 1995, with current cancers taken as relating on average to exposures in the middle of this period in the early 1970s. We therefore needed to adjust the CAREX data to take account of these changing employment levels between the early 1970s and early 1990s. For the shorter latency neoplasms the risk exposure period was 1986 to 2005, peaking in the early 1990s and therefore requiring no adjustment.

To decide on the factors to use to backdate CAREX data for the solid tumour cancers, ratios of numbers employed for the broad industry categories that were consistent across the classifications were used. These ratios along with the data supporting the calculations are given in Table A5.1, and the equation from which the ratios were obtained can be summarised as:

$$f_l = N_{\sum SIC}(\text{REP mid year}) / N_{\sum SIC}(\text{CAREX year}) \quad (5.2)$$

where f_l = CAREX adjustment factor

$N_{\sum SIC}$ = CoE or LFS numbers employed in the relevant SIC or other classification codes, in the middle year of the REP (REP mid year) or 1991 (for CAREX year)[9].

Although not available back to 1973 (the mid-REP year chosen for the adjustment), the LFS was the better source of data to use for factors applying to the main industry groups A-B (agriculture, forestry and fishing), F (construction) and for the total. All these factors were close enough to one (implying no significant change in numbers between the early 1970s and early 1990s) and were therefore ignored, at least for men. However the increase in paid employment amongst women has resulted in rising employment in all these sectors. CAREX numbers were not separately available for women, but where assumptions were made for women’s exposure based on the CAREX data, the adjustment factors in Table A5.2 were applied to deflate the numbers of women who would have been exposed in the early 1970s. For the grouped service industries, Figures A5.1 and A5.2 indicate rising numbers in these sectors across the period. The CoE adjustment factor for 1973 to 1991 was nearly equal to the LFS factor for 1979 to 1991, and so the LFS factor was used. For women the rise in employment was again significant; and the LFS factor has been used for the sake of consistency. The most important factor was for the combined manufacturing, mining and quarrying and energy and water industry sectors, where there has been a substantial fall in the numbers employed, most of which has occurred since 1978 (Figure A5.2); this indicated that the LFS factor was the preferred choice as it took account of the self-employed. The factor was 1.4 for men, and 1.5 for women.

The factors used are given in Table A5.2.

Then the appropriate factor is applied to the point estimate, by industry/occupation, to obtain an adjusted point estimate as:

$$n_{0(adj)} = f_l * n_0 \quad (5.3)$$

It was not necessary to adjust CAREX data for short latency cancers, or to adjust data from other sources available for a wider range of years.

9 Similar factors derived for future burden estimates use LFS data entirely and a 1990-93 average for (CAREX year).

Table A5.1 Calculation of CAREX adjustment factors

	CoE Employee counts (excludes self-employed)						LFS Employees + self-employed **			
Broad Industrial Group	**SIC68**		**SIC80**	**For SIC68 to SIC80 conversion (on 1981 overlap data)**	**CAREX adjustment factor 1991 to 1973**	**CAREX adjustment factor 1991 to 1978 (for comparison with LFS adjustment)**			**CAREX adjustment factor 1991 to 1979**	**LFS main industry group**
Male	**1973**	**1978**	**1991**				**1979**	**1991**		
1: Agriculture, forestry, fishing	305,062	281,059	211,398	0.92	1.57	1.45	437,013	457,595	0.96	A-B:Agriculture hunting & forestry, fishing
2-3: Manufacturing, mining and quarrying, energy and water #	5,984,066	5,633,601	3,554,973	0.99	1.71	1.61	6,096,181	4,468,216	1.36	C-E:Mining and quarrying, electricity, gas and water, manufacturing industry
4 Construction	1,243,966	1,114,207	833,213	1.03	1.45	1.30	1,705,582	1,725,143	0.99	F:Construction
5-8: Grouped 'service' ndustries *	5,942,661	6,046,512	6,570,033	0.99	0.92	0.93	6,690,244	7,600,644	0.88	G-Q: Grouped 'service' industries
Total	**13,476,203**	**13,076,046**	**11,169,618**	**0.99**	**1.22**	**1.18**	**15,165,159**	**14,565,591**	**1.04**	Total
Female										
1: Agriculture, forestry, fishing	113,i889	91,355	78,678	0.93	1.55	1.25	93,080	124,162	0.75	A-B:Agriculture hunting & forestry, fishing
2-3: Manufacturing, mining and quarrying, energy and water #	2,375,996	2,169,121	1,443,922	0.98	1.68	1.53	2,392,842	1,625,075	1.47	C-E:Mining and quarrying, electricity, gas and water, manufacturing industry
4: Construction	93,901	106,703	137,377	1.03	0.67	0.76	129,657	193,275	0.67	F:Construction
5-8: Grouped 'service' industries *	6,119,483	6,782,548	8,739,414	0.99	0.71	0.78	6,943,567	8,549,368	0.81	G-Q: Grouped 'service' industries
Total	**8,703,718**	**9,150,297**	**10,399,390**	**0.99**	**0.85**	**0.89**	**9,747,605**	**11,453,404**	**0.85**	Total

SIC68: 02. Mining and quarrying, 2.Gas, electricity and water, 3.Manufacturing industries.
SIC80: 2.Energy and water supply, 3.Manufacturing industries
* SIC68: 5.Distributive trades, 6.Transport/communication, banking, finance, 7.Public administration and defence, 8.Miscellaneous services
SIC80: 5.Distribution, hotels/catering, repairs, 6.Transport/communication, banking, finance, 7.Public administration and defence, 8.Other service industries

** includes government schemes and family workers

Table A5.2 Factors used for adjusting CAREX exposed numbers to the risk exposure period (solid tumour cancers)

	Main Industry Sector		**CAREX adjustment factor**
Men	A,B	Agriculture, hunting and forestry; fishing	1
	C,D,E	Mining and quarrying, electricity, gas and water; manufacturing industry	1.4
	F	Construction	1
	G-Q	Service industries	0.9
		Total	1
Women	A,B	Agriculture, hunting and forestry; fishing	0.75
	C,D,E	Mining and quarrying, electricity, gas and water; manufacturing industry	1.5
	F	Construction	0.67
	G-Q	Service industries	0.8
		Total	0.9

A 5.3 Estimation of the numbers ever exposed [$N_{e(REP)}$]

In order to estimate the *numbers ever exposed in the target risk exposure period*, the point estimate obtained from CAREX or national employment data was adjusted to take into account staff turnover during the REP. Staff turnover was estimated from the LFS 'length of service of employees' data (see section A5.5). The equation used to calculate numbers ever exposed in the target REP (using life table estimates of the proportions surviving to the target year 2005) is as follows:

$$N_{e(REP)} = \sum_{i=a}^{i=b} \{l_{(adj15)i} * n_0/(R-15)\}$$

$$+ \sum_{k=0}^{k=(age(u)-age(l))} \sum_{j=c+k}^{j=d+k} \{l_{(adj15)j} * n_0 * TO /(age(u)-age(l)+1)\} \quad (5.4)$$

where

n_0 = point estimate of numbers employed at a midpoint of the REP [= $n_{0(adj)}$ for the CAREX data from equation (5.3) or N_e otherwise where $Pr_{wk}(E)$=1]

TO = staff turnover per year (see section A5.5)

R = retirement age

$l_{(adj15)i}$ = the proportion of survivors to exact age i of those alive at age 15 who are assumed to be subject throughout their lives to the mortality rates experienced in the three year period to which the relevant Interim Life Table relates[10], (1980-82 for the solid tumour REP, 1994-96 for the short latency REP). The proportions are sex specific.

a to b = age range achieved by the original cohort members by the target year

c to d = age range achieved by the turnover recruited cohort members by the target year

age(u) and age(l) = upper and lower recruitment age limits

Assumptions underlying this estimator were (1) there was an even distribution of ages across the cohort in its first year, and (2) recruitment into the cohort was confined to a specific age range, defined by the user.

[10] Government Actuary's Department, Interim Life Tables, at http://www.gad.gov.uk/Life_Tables/Interim_Life_Tables.htm. $l_{(adj15)i}$ is calculated as lx(age i)/ lx (age 15)

It is unlikely to be realistic to assume that recruitment into the cohort was only at age 15, the assumption that would give the largest estimates of surviving exposed numbers. An alternative assumption, that recruitment into the cohort is evenly spread across ages from 15 to retirement age, results in greatly reduced estimates of surviving exposed numbers, and is also probably unrealistic. For this study, it was assumed that recruitment to the cohort was in the age range 15-24 only. The impact this has on the estimates will be examined in a sensitivity analysis to follow in conjunction with the development of 'credibility intervals'. For this study therefore:

age(u) and age(l) = 24 and 15
a to b = age 65 to 100 for the solid tumour REP; age 35 to 84 (M) or 79 (F) for the short latency REP
c to d = age 25 to 64 for the solid tumour REP; age 15 to 34 for the short latency REP

Retirement age for this study was also taken as
R = 65 for men, 60 for women

A description of how equation (5.4) was derived from the simple turnover equation:

$N_{e(REP)} = n_0 + \{n_0 * TO * t\}$
where n_0 = numbers employed in the exposed job/industry at a mid-point in the REP,
TO = employment turnover rate,
t = length of risk exposure period in years.

is given in section A5.7.

The equation can be represented as a single factor acting as a multiplier for n_0, calculated by setting n_0 to one in the above equation, so that the factor varies only with TO (see Section A5.5). Separate values for TO and therefore this multiplier were used for men and for women, as their employment turnover rates differed.

A5.4 Estimation of the proportion of the population exposed

The numbers ever exposed in the target REP from equation (5.4) were divided by an estimate of *numbers ever of working age during the REP* (calculated from population estimates by age cohort in the target year[11]) to give the proportion of the population who were employed during the risk exposure period, by industry or occupation.

Numbers ever of working age during the REP were determined as follows [12]:

$$N_{p(REP)} = \sum_{age(j=f)}^{age(j=g)} (N_{age(j)}) \text{ for the target year (2005)} \qquad (5.5)$$

where $N_{p(REP)}$ = Population ever of working age during the REP
$N_{age(j)}$ = population in age group j in the target year
f = youngest working age group (15-24) + (TY-UL)[13]
g = oldest working age group e.g. (60-64) + (TY-LL), or age group (90+) if this is reached first[14]

[11] Mid-2005 Population Estimates: Great Britain; estimated resident population by single year of age and sex; Table 2. http://www.statistics.gov.uk/statbase/ssdataset.asp?vlnk=9083&More=Y
[12] This is the GB population of working age and over, up to the maximum age that someone retiring during the REP could have reached by the last year of the REP (1995 for solid tumours), moved forward in age from then to the target year (2005).
[13] i.e. (25-34) for the solid tumour REP, (15-24) for the short latency REP
[14] i.e. (90+) for the solid tumour REP, (80-84) for the short latency REP for males, (75-79) for females

LL = first year of REP
TY = target year (2005)
UL = last year of REP

Again, this value becomes a constant across cancer/exposure pairings when based on the standard risk exposure periods. For the solid tumour REP, the numbers were 19.2 million men and 20.9 million women. For the short latency cancers REP, the numbers were 20.9 million men and 19.8 million women.

Then $Pr(E) = N_{e(REP)} / N_{p(REP)}$ (5.6)

where $Pr(E)$ = proportion in the population exposed

A5.5 Employment Turnover Calculation (TO)

A5.5.1 Data available from the Labour Force Survey

The LFS collects data for the length of time with current employer in eight durations of employment categories from <3 months to >20 years. These data are available by occupation codes and by SIC main industry sector four digit codes and by year; for this study 1983, 1984, 1991, 1998, 2000 and 2003 were available.

The years chosen for the calculation of TO need to be as close as possible to the mid-point of the risk exposure period for the cancer. For the risk exposure period for solid tumours (1956 to 1995), 1984 was the earliest available year with SIC coding consistent through to 2003. As the LFS is a sample survey, estimates based on those codes, which represented small numbers of workers, would have been unstable. Consequently, an estimate for broad industry sectors has the advantage both of being stable and providing consistency across multiple exposures. The grouped main industry sectors used for the CAREX adjustment factors were also used for the turnover estimates, so that the partitioning of the numbers exposed was the same for both factors.

A5.5.2 Alternative ways to calculate employment turnover (TO)

For the cancer burden study, we are interested in estimating numbers ever exposed in a defined period, the risk exposure period. Therefore the purpose of calculating employment turnover is to estimate numbers entering exposed employment during the period. Leavers during this period have been exposed and therefore need to be retained in the estimate. We also need to exclude those employed for less than one year, as many epidemiological studies (and therefore our RRs) are based on workers employed for at least one year; as those possibly exposed for very short periods would not normally be considered at risk..

There are two quite different approaches: (i) to base employment turnover on average length of employment, or (ii) to base it on estimates of numbers starting and/or leaving work in a defined period. The two methods are equivalent when applied to a full dataset of starters and leavers. For the cancer burden study we have chosen to use the second approach and in particular to base the estimation on new starters rather than job separation. Estimating TO directly for ‘new starters’ in years within the REPs will give the best approximation of the target ‘ever exposed’ that the study requires. Estimates using an alternative i.e., the inverse of average length of employment method, based on LFS data, exclude the contribution of leavers’ person-years-in-work, and so underestimate the turnover (assuming that leavers have shorter average lengths of employment than those continuing in employment).

A5.5.3 Modification of the new starters (accessions) rate

Employment turnover can be estimated as new starters (accessions) rate, i.e. starters in the past year as a proportion of the average number employed. Gregg and Wadsworth (2002) estimate from LFS and

GHS data that the average share of new jobs was approximately 18% between 1975 and 1995. To adjust for exclusion of short-term employees (to match the studies from which RRs are taken) an adaptation of the new starters rate is to take starters in the past year, who are expected to remain employed for at least one year, as a proportion of all those expected to be employed for at least 1 year. Short-term labour turnover, in this case employment for less than one year, is excluded. It is estimated as the difference between numbers employed for less than one year and the numbers employed for 1-2 years, as a proportion of those employed less than one year (Gregg and Wadsworth, 1995). For the LFS data:

$$\text{Short-term turnover (STTO)} = (\Sigma_{j=0\text{-}3mths\ to\ 6\text{-}12mths}\, n_j - n_{1\text{-}2yrs}) / \Sigma_{j=0\text{-}3mths\ to\ 6\text{-}12mths}\, n_j \quad (5.7)$$

Where n_j = LFS numbers in length of employment category j

Short-term turnover is averaged over three LFS data years (1984, 1991 and 1998), as the estimates for 1984 alone were high (51% for men), and thought not to be representative. As estimates based on new starters are not robust to changes across the economic cycle and in expanding and contracting industry sectors, an average of new starter ratios is also taken over the available LFS years as above. Therefore, as 'continuers' are estimated as (1- short-term turnover) as a proportion of those employed less than one year, TO is estimated as:

$$\text{TO (modified new starters)} = (1\text{-STTO})*(\Sigma_{j=0\text{-}3mths\ to\ 6\text{-}12mths}\, n_j) / (n_{1\text{-}2yrs} + \Sigma_{j=1\text{-}2\ yrs\ to\ 20+\ yrs}\, n_j) \quad (5.8)$$

This can be simplified (if no adjustment is made for proportion changing employment but not industry, see Section A5.5.5) to

$$\text{TO} = n_{1\text{-}2yrs} / (n_{1\text{-}2yrs} + \Sigma_{j=1\text{-}2\ yrs\ to\ 20+\ yrs}\, n_j) \quad (5.9)$$

A5.5.4 Using TO to estimate numbers ever exposed

In all cases, annual employment turnover (TO) is applied to a point estimate of exposed numbers via the standard turnover equation (Equation (5.4)), in order to estimate numbers ever exposed. The point estimate is generally obtained from CAREX, or as numbers employed in a single representative year from LFS or Census of Employment data. In theory the numbers include those employed for less than one year on notional entry to the REP. However the numbers can also be taken to represent exposed 'posts' in that year through which a series of employees will pass during the risk exposure period. Therefore to make an adjustment to exclude the less than one year employed from these point estimates would lead to a serious underestimate of numbers ever exposed for at least one year. This adjustment has been confined to the TO element.

The estimated annual employment turnover (TO) short-term turnover (STTO) and turnover factors are in Table A5.3. The turnover factors were calculated using a modification of the Global Burden of Disease formula in Nelson *et al.* (2005)[15], and are given for comparison with other authors' estimates of employment turnover.

A5.5.5 Adjustment for changed employer but not changed exposure with the modified new starter method

'Length of time with current employer' does not correctly represent the 'length of exposed employment' that is required. Although it is possible that workers reporting any length of employment with one employer may have changed occupation in the previous year within that employment, it is more likely that a change in employment may not have resulted in a change in exposure, if the

[15] Turnover factor $= N_{e(REP)}/n_0$
where $N_{e(REP)}$ = numbers ever exposed during the REP, from equation (5.4)
n_0 = point estimate of numbers exposed at a midpoint in the REP

employment remained within the same industry. As a result length of employment as a proxy for length of exposure will generally underestimate this, and the TO we need will be inflated. For the modified new starters method used here we are interested only in the effect this will have on workers employed in their current job for less than one year but who are continuing in employment in the same industry for at least a year. If p% of these report no change in occupation in the past year in spite of the change of employer, these should be excluded from the TO estimate. Table A5.4 shows estimates of the effect this has on the modified new starters TO, for an estimate of p=6% for men, p=7% for women, obtained from 1984 LFS data (OPCS, 1984) on whether respondents reporting less than one year in their current job also reported no change to the industry in which they were employed one year ago. The complement of these percentages is applied to the estimate of those employed for less than one year in equation (5.8)[16].

A5.5.6 Other authors' estimates of TO

The estimates from the modified new starters method are also similar to other authors estimates.

HSE (2002) estimated a 10-fold lifetime (over 40 years) turnover factor (implying a 25% annual rate), for demolition and asbestos-removal workers (consistent with data on individuals having statutory asbestos medicals as asbestos removal workers over the previous 14 years). 5-fold (12.5%) and 2.5-fold (6%) turnover factors were determined for maintenance and other building work respectively. These were based on the LFS, which provides estimates of time with current employer, and also on whether the respondent's occupation had changed over the last year.

A factor of four has been used by NIOSH on a cross-sectional exposure survey from early 1981, to estimate the number of workers ever exposed to lung carcinogens as of 1997, in order to estimate AFs in 1997 (communication from Kyle Steenland[17]). The factor of four came from assuming a 10% turnover per year, along with a 40 year period of interest (1957-1997); a given age distribution at baseline in 1957; an average duration of exposure of 10 years (log normal distribution); and a mortality loss of 20% over the 40 year period. This scenario assumed an average latency of approximately 20 years. This estimate has been used in the Global Burden of Disease methodology (Nelson *et al,* 2005).

Dreyer *et al.* (1997) used a turnover assumption that 85% of blue collar production workers were engaged in the same trade for at least one year during the 15 year period 1970-84, implying a yearly turnover of the workforce of 15%. They also assumed 1970-84 was the REP for the cancer pattern in the year 2000.

Sorahan et al (1989) estimate an average duration of exposed employment of 10.5 years for 36,700 male workers in the UK rubber industry first employed between 1946 and 1960 and with job histories covering the period 1976 to 1985. This gives a turnover estimate of 9.5% using the alternative average length of employment method.

Data on occupational history for 2,300 controls in a population based case-control study of lung cancer in Great Britain (Peto *et al,* 2008) indicated annual employment turnover of 9.3% for men (estimated as the inverse of average length of employment in an industry defined at 3-digit SIC code level), and a turnover factor of 3.6.

---

[16] Applying a p% reduction to those employed for less than one year in equations (5.7) and (5.8) cancels out in equation (5.8) (exactly when short-term labour turnover is estimated from the same data source as TO as a whole). However, those employed for 1-2 years were also subject to the STTO rate the previous year, so are also correctly subject to the p% reduction applying to the previous year. In practice then the p% reduction is applied to $n_{1\text{-}2yrs}$ in equation (5.7) and therefore (5.9).

[17] GB Burden of Occupational Cancer: Summary Report of Workshop Held on the 22nd and 23rd November 2004 in Manchester: HE/05/03

Table A5.3 Employment turnover factors used in the calculation of numbers ever exposed, using a modified 'new starters' method, based on LFS data

Main Industry Sector	Method: Continuing ((1- short-term TO)% of) starters within past year as proportion of all expected to be employed at least 1 year. Based on LFS data averaged over 1984/91/98.		
	Annual TO	TO factor over 40 years	Short-term (<1 year) labour turnover
	MEN		
Agriculture, hunting and forestry; fishing	7%	3	0.56
Mining and quarrying, electricity, gas and water; manufacturing industry	9%	3	0.40
Construction	12%	4	0.46
Service industries	11%	4	0.44
Total	**10%**	**4**	**0.44**
	WOMEN		
Agriculture, hunting and forestry; fishing	10%	4	0.44
Mining and quarrying, electricity, gas and water; manufacturing industry	14%	5	0.44
Construction	15%	6	0.37
Service industries	15%	6	0.46
Total	**14%**	**6**	**0.45**

Table A5.4 Effect of assuming a proportion of new starters are changing employer but not occupation

Main Industry Sector	Annual TO % (modified new starters method)	Adjusted for p% changing employer but not occupation
MEN		***p=6%***
Agriculture, hunting and forestry; fishing	7%	7%
Mining and quarrying, electricity, gas and water; manufacturing industry	9%	8%
Construction	12%	11%
Service industries	11%	10%
Total	**10%**	**10%**
WOMEN		***p=7%***
Agriculture, hunting and forestry; fishing	10%	10%
Mining and quarrying, electricity, gas and water; manufacturing industry	14%	13%
Construction	15%	14%
Service industries	15%	13%
Total	**14%**	**13%**

A5.6 Partitioning the exposed population

A partition of the estimate of the 'ever exposed' population is generally made by level of exposure, into 'higher' and 'lower' exposed groups generally defined by (i) industry/ occupation. It may also be needed for (ii) sex and (iii) time period.

A5.6.1 Partition by industry/occupation

Data are not available on the levels of exposure in all industry sectors for all the carcinogens considered, nor the numbers exposed at these levels. The CAREX industry sectors were thus allocated to 'higher' or 'lower' exposure categories assuming distributions of exposure and risk that corresponded broadly to those of the studies from which the risk estimates were selected. The initial allocations were based on the judgment of an experienced human exposure scientist; each assessment was then independently peer-reviewed and if necessary, a consensus assessment agreed. The 'higher' and 'lower' counts were also split between the grouped main industry sectors, so that separate CAREX adjustment and turnover factors were applied before the proportion exposed was calculated for the group as a whole. See also Appendix 3.

A5.6.2 CAREX partition by sex

CAREX data are not available separately for men and women. They also do not include white collar workers in the industry groups covered. To split the CAREX numbers between men and women, where there were thought to be women exposed, the male to female proportions were obtained as follows.

1) If the risk estimate studies included very few women and it is thought that this also applied to the UK then the CAREX figures were assumed to be all men.
2) If the risk estimate studies included quite a large proportion of women then the same proportion as that in the studies was assumed for CAREX if the UK population was thought to be similar to the study population.
3) Alternatively the AES or LFS was used to obtain the ratio of males to females for all those employed in the exposure scenario and then, using knowledge of the jobs that women do in the industry, a judgement was made about what proportion of the women are exposed to levels of concern i.e. not just background, and the CAREX figures were apportioned accordingly. For example if the ratio of men to women was 50:50 overall in an industry group from the AES/LFS but it is thought that 50% of the women will have had little or no exposure the ratio of M:F in the CAREX data is 75:25.
4) In practice an occupation by industry cross tabulation by sex of numbers over age 16 employed or self-employed in Great Britain has been used, from the 1991 Census. From these data proportions were estimated, using mainly standardised sets of the major occupation groups, for the broad industrial groups that were used for the CAREX adjustment and turnover factors described above. The estimated male percentages are given in Table A5.5. Separate percentages are available for all workers in the broad industrial groups and for appropriate combinations of these groups. The most useful are for blue collar workers, defined as those in Standard Occupational Classification major groups 5, 8 and 9 which cover skilled trades, shop floor and transport operatives; for white collar workers, SOC major groups 1, 2 and 4, managerial, professional, administrative and secretarial; and for associate professional, technical, personal and customer service occupations, SOC major groups 3, 6 and 7. In practice the percentage most appropriate to the particular exposure scenario has been used, with flexibility to use percentages for single rather than grouped SOC major groups if exposed numbers were clearly concentrated in these occupations. Transport operatives (SOC group 8) in the service industries (industry sectors G-Q) were an example of the appropriate use of a single SOC group for exposure to diesel engine exhaust.

Table A5.5 Numbers and percent of men and women in employment by industry and occupation

SIC Main Industry Sectors:	Agriculture, hunting and forestry; fishing			Mining and quarrying, electricity, gas and water; manufacturing industry			Construction			Service industries		
	A-B			**C-E**			**F**			**G-Q**		
	Male	Female	Male %	Male	Female	Male %	Male	Female	Male %	Male	Female	Male %
Standard Occupational Classification Major Groups:												
1 Managers and administrators	17,628	2,962	86%	55,941	14,129	80%	14,711	2,396		163,799	99,599	62%
2 Professional occupations	163	44		24,721	2,099		4,618	164		94,340	76,298	55%
3 Associate professional and technical occupations	77	65		22,497	7,154		4,558	398		75,062	93,838	44%
4 Clerical and secretarial occupations	125	1,195		20,182	42,506	32%	1,801	9,818		65,259	235,196	22%
5 Craft and related occupations	5,793	1,114		124,302	28,768	81%	96,292	745	99%	75,548	5,907	93%
6 Personal and protective service occupations	78	170		2,576	1,785		314	90		75,410	131,163	37%
7 Sales occupations	107	176		10,979	5,759		1,262	730		46,927	101,490	32%
8 Plant and machine operatives	708	515		95,967	41,146	70%	15,649	284		75,067	10,680	88%
9 Other occupations	11,052	3,243	77%	17,681	7,172	71%	17,331	873		52,174	90,917	36%
Occupation not stated or inadequately described	205	220		1,635	574		944	82		3,801	2,684	
TOTAL PERSONS	**35,936**	**9,704**	**79%**	**376,481**	**151,092**	**71%**	**157,480**	**15,580**	**91%**	**727,387**	**847,772**	**46%**
"White collar" Managerial, professional, administrative, secretarial: SOC groups 1,2,4	17,916	4,201	81%	100,844	58,734	63%	21,130	12,378	63%	323,398	411,093	44%
Associate professional, technical, personal and customer service occupations: SOC groups 3,6,7	262	411	39%	36,052	14,698	71%	6,134	1,218	83%	197,399	326,491	38%
"Blue collar" Skilled trades, shop floor and transport operatives: SOC groups 5,8,9	17,553	4,872	78%	237,950	77,086	76%	129,272	1,902	99%	202,789	107,504	65%

Residents aged 16 and over - employees and self-employed, 1991 Census 10% sample
Male % is given for single SOC major group only if it represents >20% of any one industry class in broad industry group

A5.6.3 Partition by time period

If it was known that the exposure level had changed sufficiently during the target risk exposure period to move from 'higher' to 'lower' exposure, the REP needed to be split between these periods to calculate Pr(E). This was done by partitioning the estimate for the numbers 'ever exposed' in the REP in the industry/occupation-defined group ($N_{e(REP)}$) between the 'higher' and 'lower' exposure periods ($N_{e(REP)\,1}$ and $N_{e(REP)\,2}$) according to the age ranges achieved by the target year by the members of the cohorts in each successive period. Note that members of the original cohort were only counted once, in the first period. Examples are given below for the solid tumour REP (1956-1995) split between the periods 1956-62 and 1963-95.

From (5.4)

$$N_{e(REP)} = \sum_{i=a}^{i=b} \{l_{(adj15)i} * n_0/(R-15)\} + \sum_{k=0}^{k=(age(u)-age(1))} \sum_{j=c+k}^{j=d+k} \{l_{(adj15)j} * n_0 * TO/(age(u)-age(l)+1)\}$$

Where

a to b = age range achieved by the original cohort members by the target year (age 65 to 100 for the solid tumour REP; age 35 to 84 (M) or 79 (F) for the short latency REP)

c to d = age range achieved by the turnover recruited cohort members by the target year (age 25 to 64 for the solid tumour REP; age 15 to 34 for the short latency REP)

For period 1956-62 a modification is:

$$N_{e(REP)\,1} = \sum_{i=a}^{i=b} \{l_{(adj15)i} * n_0/(R-15)\} + \sum_{k=0}^{k=(age(u)-age(1))} \sum_{j=c1+k}^{j=d1+k} \{l_{(adj15)j} * n_0 * TO/(age(u)-age(l)+1)\} \qquad (5.10)$$

Where

a to b = age 65 to 100

c1 to d1 = age 57 to 64

For 1963-95 the equation is

$$N_{e(REP)\,2} = \sum_{k=0}^{k=(age(u)-age(1)+1)} \sum_{j=c2+k}^{j=d2+k} \{l_{(adj15)j} * n_0 * TO/(age(u)-age(l)+1)\} \qquad (5.11)$$

where c2 to d2 = age 25 to 56

Note that the denominator (i.e., the numbers ever of working age during the REP) remains the same as calculated for the whole period (1956-1995 for the solid tumours) and was not split between the two periods. The denominator also remained the same if an exposure was deemed to have ceased before the end of the REP for the cancer, although the numerator (i.e., the numbers ever exposed) was based on a shorter time period (e.g. 1956-1962).

A5.7 Derivation of the turnover equation (5.4)

The equation is derived from the simple turnover equation:

$N_{e(REP)} = n_0 + \{n_0 * TO * t\}$

where n_0 = numbers employed in the exposed job/industry at a mid-point in the REP,
TO = employment turnover rate,
t = length of risk exposure period in years.

To make adjustment for loss from the cohort from deaths occurring before the target year of 2005, n_0 is replaced in this equation by an estimate of only those surviving to 2005, estimated using GB life table data[18]. For the original cohort recruited at the beginning of the REP, the estimate is obtained as the sum of the proportions of survivors in GB at all the ages that would have been represented in the cohort at the outset. As an example, ages from 15 to retirement age (R) in 1956 for the solid tumour standard REP, who would have been aged from 65 upwards by 2005 ($\Sigma_{i=a}^{i=b} l_{(adj15)i}$, a=65, b=100), were divided by the number of ages originally represented (R-15). An even distribution of ages across the original cohort is assumed. Retirement age is taken as 65 for men, 60 for women. This proportion is then applied to the original cohort size (n0) to get an estimate of the numbers from this original cohort who would still be alive in 2005. For the part of the cohort recruited through employment turnover, t in the simple turnover equation is replaced by the sum of the proportions of survivors in the age range achieved by the turnover recruited cohort members by 2005 (i.e. all those aged 25 to 64 for the solid tumour REP, $\Sigma_{j=c}^{j=d} l_{(adj15)j}$, c=25, d=64). This proportion is then applied to the original cohort size n_0, multiplied by the percent turnover (TO). Summing the proportions over this age range is equivalent to multiplying by the number of years in the REP (t). However, it would not be realistic to assume that all newly recruited members of the cohort enter at age 15, as this summation of survivor proportions would assume. Therefore to allow recruitment across the age range u to l, a further summation is introduced into the turnover part of the equation so that the range of ages is counted for the survivor proportions, divided by the length of the age range (giving an 'average' survivor proportion for the age range). An entry age range of 15 to 24 was used in this study, giving an age range length of 10 years. This seemed more realistic than assuming entry at all ages from 15 to retirement, but more conservative than allowing entry at age 15 only.

The proportions of survivors to age i (the l_i) were obtained from GB life tables, adjusting to include only those still alive at age 15 by taking ($l_{(adj15)i}$) = l_i / l_{15}. The l_i are the proportions of survivors per live birth at age i, assuming sex-specific death rates current in 1980-82 (the earliest tables available, for the solid tumour REP).

[18] Government Actuary's Department, Interim Life Tables, at http://www.gad.gov.uk/Life_Tables/Interim_Life_Tables.htm

APPENDIX 6: CALCULATING THE ATTRIBUTABLE FRACTION

A6.1 Population-based studies – Miettinen's method

For population-based studies an estimate of the proportion of the population exposed, or of the proportion of cases exposed, was obtained directly from the chosen study using the exposure distribution for the controls and the cases respectively. The cases had to represent a random sample of all the cases in the population. Using a population-based study has the advantage that Miettinen's estimator of the attributable fraction can be used. This is unbiased when adjusted relative risk estimates are being used as is often the case in more recent studies (Benichou, 2001).

Miettinen's formula for the attributable fraction (AF) is:

(6.1) $AF = Pr(E|D) * (RR – 1) / RR$
where $Pr(E|D)$ = proportion of cases exposed

This formula requires an estimate of the proportion of the cases exposed, which was not available other than from the study data.

If the alternative Levin's formula (see below) is used with adjusted RR estimates, the results will be biased (Greenland, 1984; Rockhill et al, 1998) but the direction and degree of this bias can be determined, and will not normally be significant in comparison with other potential sources of bias in the AF calculation.

AFs calculated using Miettinen's formula can also be combined across 'sufficient cause' sets of exposures (see Section A6.3), where interaction and / or confounding between exposures is thought to be a problem. These sets also can be partitioned without introducing bias into the resulting AF (Bruzzi *et al,* 1985).

In general, as multiple industries, exposure levels and time periods are involved in estimating an exposure AF, a more complex form of this equation was used. To sum the exposure levels (h) for a single exposure (h = 1 to H, with H usually equal to just 2 levels 'higher' and 'lower' for this study) they must have a common reference level. The equation used for Meittinen's estimator was

$$AF = \Sigma_{h=1 \text{ to } H} Pr(E_h|D) (RR_h - 1) / RR_h = Pr(E|D) – \Sigma_{h=1 \text{ to } H} Pr(E_h|D) / RR_h \quad (6.2)$$

A6.2 Industry-based studies – Levin's method

In the case where the estimate of relative risk was obtained from an industry based study, Levin's formula was used to calculate AF. In its simplest form

$$AF = Pr(E) * (RR – 1) / \{1 + Pr(E) * (RR – 1)\} \quad (6.3)$$
where $Pr(E)$ = proportion of the population exposed

The form of the equation used for multiple exposure levels for Levin's estimator was

$$AF = \{\Sigma_{h=1 \text{ to } H} Pr(E_h)(RR_h - 1)\}/\{1+\Sigma_{h=1 \text{ to } H} Pr(E_h)(RR_h - 1)\} \quad (6.4)$$

This equation, and (6.2) above, can be adapted for estimating AFs across multiple industries, time periods and exposures. For these, a common denominator must be used that incorporates all the Pr(E) and RR estimates which contribute to the final exposure AF. Using Levin's method for example all separate exposure level AF estimates have the common denominator $\{1+\Sigma_{h=1 \text{ to } H} Pr(E_h)(RR_h - 1)\}$ for exposure levels h. Additional summations for industry and time period are similarly incorporated in the denominator and the numerator according to the degree of disaggregation required.

A6.3 Combining AF across exposures

A number of strategies were adopted to address the problem of combining multiple exposures, whether they were known to overlap in the working population or were disjoint. AFs cannot in general be summed directly because (1) summing AFs for overlapping exposures may result in an overall AF exceeding 100%, and (2) summing disjoint exposures also introduces upward bias. The source of this bias is investigated and described in full elsewhere (Hutchings, 2010c). Disjoint exposure AFs can only be summed if the risk structure is additive.

Where there were multiple occupational exposures responsible for cancer at a single site, and it could be assumed that where the exposures overlapped in the working population, the exposures were independent and their joint effect on initiating or promoting cancer was multiplicative, i.e. $RR_{(\text{exp 1 and exp2})} = RR_{(\text{exp1})} * RR_{(\text{exp2})}$, the AFs for each exposure in the overlapping set were combined into an overall AF for the set by taking the complement of the product of complements:

$$AF_{overall} = 1 - \Pi_k(1\text{-}AF_k) \text{ for the k exposures in the set.} \qquad (6.5)$$

This is also an appropriate equation to use for combining disjoint exposure AFs, as it introduces less bias into the results compared with, for example, direct summing of the AFs (Hutchings, 2010c).

However, in some instances it is not appropriate to assume independence of overlapping exposures. An alternative approach adopted was to partition sections of the workforce known to be subject to multiple exposures between each exposure, allocating occupation/industry codes to the 'highest risk' or dominant exposures. Equation 6.5 was then used to combine the now disjoint exposure AFs, as direct summing would still introduce more upward bias. There were various strategies to achieve this and examples of their use are given below for lung cancer.

1. Exposure scenarios were excluded entirely, if they wholly overlapped with another dominant exposure. Exposures in boxes with dotted lines in Figure 2 (Section 3.2), were in general excluded for this reason, and did not have an AF calculated separately.

2. Where the exposed populations only partially overlapped, exposed numbers were partitioned between the overlapping exposures. This strategy was adopted for the exposures circled in hashed lines in Figure 2 in Section 3.2 for lung cancer. Although these clearly overlapped in the working population with other exposures (as indicated by the joining lines in this figure), AFs were calculated for a restricted group of workers only, which (as far as possible) did not include workers in the populations for other exposures with which they overlapped. Tin miners, steel foundry workers and mineral oils (printers only) were treated in this way. The exposed populations (iron and steel basic industries as defined in the CAREX data, and printers as narrowly defined by occupation from the LFS, and tin miners from an established UK cohort) were excluded from the other exposures with which they overlapped.

3. It is also possible to identify sets of exposures that can be designated as a 'sufficient cause' set i.e. a set of exposures that may overlap and not be independent of one another, but when taken as a set do not overlap or interact with other exposures or sets of exposures (Hoffmann *et al,* 2006). However, this requires a single estimate of RR and Pr(E) or Pr(E|D) for the whole set of

exposures. This may be available from a population-based study, but it will be very difficult in practice to separate out the sufficient cause classes and therefore the contributions of the individual exposures within the set. As the relative contributions of the different occupational exposures were needed to prioritise interventions, separate AFs were required in this study for each cancer/exposure pairing. These may have to be obtained from an alternative ‘best study’ to the one chosen for the ‘sufficient cause’ set, so leading to inconsistency between the individual and combined AFs. Although sets of overlapping exposures were identified for each occupational cancer (e.g. Figure 2 in Section 3.2), in practice independence and multiplicativity were assumed in general and equation (6.5) always used to combine exposures within these sets. The exposures circled in bold in Figure 2 (Section 3.2) were treated in this way (termed ‘AF product’ sets). The AF for the set as a whole could be summed with other non-mutually overlapping ‘AF product’ sets or individual exposure AFs if could be assumed that the risks were additive but in general sets were combined using equation (6.5) to minimise the upward bias.

A6.4 Estimating attributable numbers

Attributable numbers (AN), either as deaths or registrations were estimated using the following equation:

$$AN = AF * \Sigma_j D_j \tag{6.6}$$

where D_j is the number of cancer deaths or registrations for age class j for the target year.

For solid tumours, only cancers in ages 25+ were counted as these could have been initiated during the REPs being considered. For haematopoietic neoplasms ages 15 to twenty years beyond retirement age only were included.

APPENDIX 7: ESTIMATION FOR LUNG CANCER AND ASBESTOS AND RADON, AND FOR ANY CANCERS ASSOCIATED WITH IONISING RADIATION

A7.1 Lung cancer due to exposure to asbestos

An alternative method of calculating an attributable fraction for lung cancer due to asbestos exposure was to estimate the number of excess deaths for lung cancer and for mesothelioma in job categories where there was known to have been exposure to asbestos. This was based on the incidence of mesothelioma amongst workers in these job categories. The ratio of the excess lung cancers to mesotheliomas was then taken as an indication of the numbers of lung cancers that were attributable to the exposure. The AF was then derived from this. The method was applicable only to asbestos exposure, as mesothelioma is considered uniquely to be caused by exposure to asbestos. This method had the advantage that it took account of the current impact that past levels of exposure to asbestos are known to be having on the incidence of cancer. There is a direct link to mesothelioma deaths, which are climbing rapidly, whereas lung cancer in general is declining due to the fall in the number of smokers. The link does, however, depend on the assumption that lung cancer has a similar pattern of latency to mesothelioma. Numbers of mesotheliomas in GB were obtained from the British Mesothelioma Register.

The ratio of excess lung cancers to mesotheliomas was estimated as follows. In populations with heavy asbestos exposures there have typically been at least as many (but sometimes up to ten times as many) excess lung cancers as there have been mesotheliomas. The ratio depends on a range of factors - the most important of which are the types of asbestos (the highest ratios have been linked to chrysotile exposure), the level of exposure, age at exposure and smoking. Two lines of argument suggest that the ratio of asbestos related lung cancers to mesotheliomas in the British population as a whole is towards the lower end of the range of 1-10 estimated from the epidemiological studies. Firstly a study of lung cancer mortality in relation to indices of asbestos exposure and smoking habits in the west of Scotland suggested a ratio of around two asbestos lung cancers per mesothelioma for a region known to be associated with fairly high asbestos exposures. Secondly, analyses of mesothelioma deaths in GB by occupation and geographical area has suggested that substantial numbers of deaths may have arisen in workers other than those that were most heavily exposed[19].

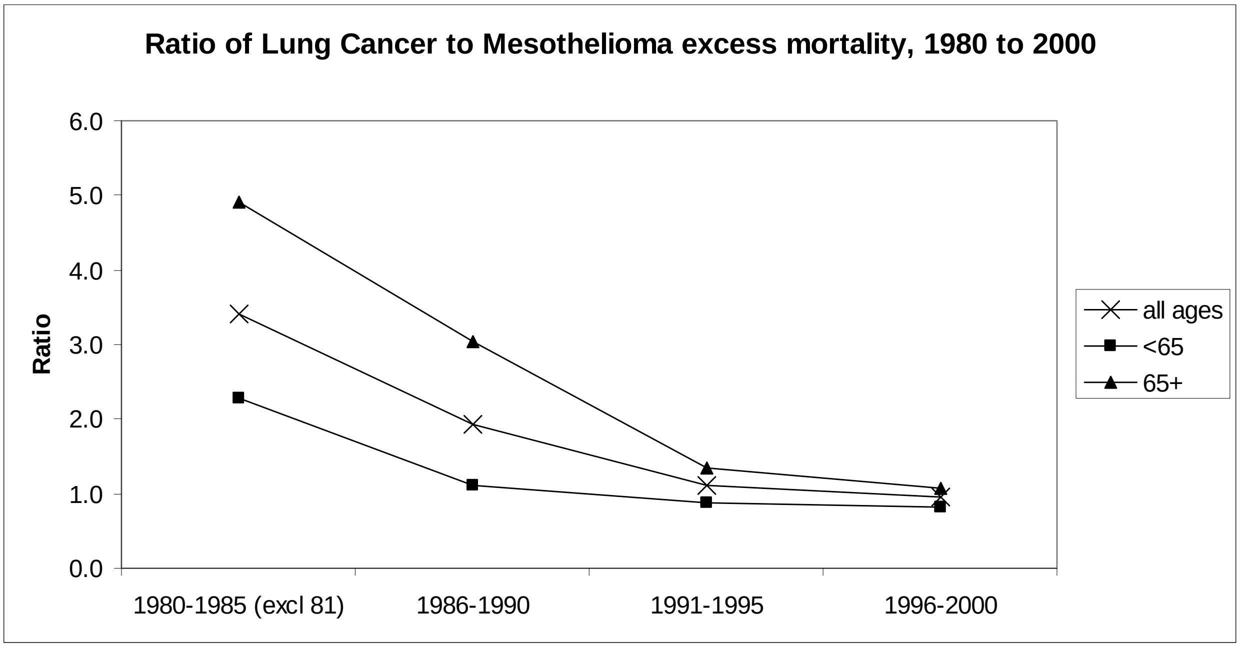
More recent evidence has suggested that this ratio may be at the bottom end of the range of 1-2. Asbestos is a more potent cause of mesothelioma than lung cancer. Smoking is also thought to interact with asbestos exposure in the causation of lung cancer. Thus going forward in time the ratio of lung cancers to mesotheliomas is likely to fall, because the mesotheliomas will increasingly be generated by low exposure levels of asbestos that are less likely to cause lung cancer and because smoking levels have fallen since the 1960s (factors that, together, mean fewer lung cancers per mesothelioma). A recent analysis of lung cancer mortality for the whole of GB (by occupational group in relation to indices of asbestos exposure and smoking habits) suggested that the ratio of asbestos related lung cancer to mesothelioma deaths was between two-thirds and one (Darnton *et al,* 2006).

To examine the trend in the ratio estimated by Darnton et al for the period 1980-2000, occupations with exposure to asbestos were identified from their excess mortality from mesothelioma and were estimated using the proportional mortality ratio (PMR) data obtained from an analysis of death certificate data held on the HSE's Mesothelioma Register. Observed minus expected deaths (estimated as the average across all occupations) for mesothelioma and for lung cancer were calculated for each occupation, and summed

[19] http://www.hse.gov.uk/statistics/causdis/lungcan.htm

across occupations that had a mesothelioma PMR with a 95% confidence interval excluding 100, to give an estimate of asbestos-attributable 'excess' deaths, for a particular time period (1980-2000, in five-year blocks, for this analysis). A suitable ratio of 'excess' lung cancer to mesothelioma deaths from this analysis was then applied to the count of mesothelioma deaths in the target year, to estimate attributable numbers. A ratio of 1:1 was chosen, as the ratio had clearly declined over the twenty year period as mesothelioma numbers have risen, although it appears to have levelled off at this ratio from 1990 onwards (Figure A7.1). The AF was derived as the attributable numbers of lung cancers divided by total numbers of lung cancers in the target year.

Figure A7.1



In order to allocate the estimate of total attributable lung cancers between asbestos exposed industries, estimates of relative risk for 'higher' and 'lower' level exposed workers (taken from an appropriate 'best study' – see the Lung Cancer Technical Report) and of the proportions ever exposed by industry (based on CAREX data) were used to estimate attributable numbers for each industry as:

$$AN_j = \{Pr(E_{jh})(RR_h - 1)\}/\{\Sigma_{h=1 \text{ to } H}\ \Sigma_{j=1 \text{ to } J} Pr(E_{hj})(RR_{hj} - 1)\}*AN_{tot} \qquad (7.1)$$

for exposure level h of H levels and industry j of J industries

A7.2 Mesothelioma due to exposure to asbestos

The AF for mesothelioma was derived directly from several UK mesothelioma studies that suggest between 96% and 98% of male mesothelioma cases are due to occupational or paraoccupational (e.g. exposure from living near an asbestos factory or handling clothes contaminated due to occupational exposure) exposure (Yates *et al*, 1997; Howel *et al,* 1997; Rake *et al,* 2009). Combining the results from Rake *et al* (2009) with those from two studies in which results were reported separately for females (Goldberg *et al*, 2006; Spirtas *et al,* 1994) gave estimates of 75%-90% for females. A full description is in the Mesothelioma Technical Report. When allocating attributable mesotheliomas between industries, the fraction attributed to paraoccupational exposure was excluded.

A7.3 Lung cancer due to exposure to radon

The AF for radon was derived from an NRPB estimate of the numbers of lung cancer deaths attributable to domestic exposure to radon. Full details are in the Lung Cancer Technical Report.

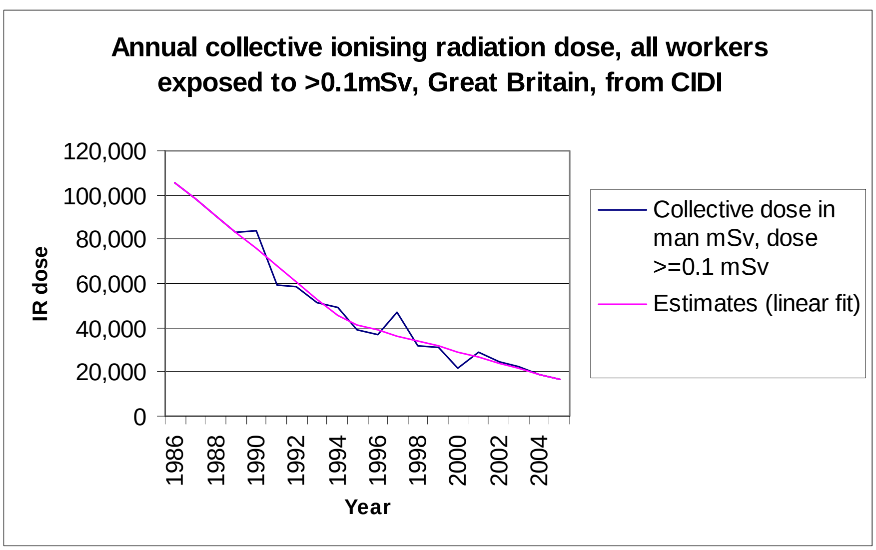
A7.4 Burden of cancer due to occupational exposure to Ionising Radiation, using the UNSCEAR model. Relative risks and numbers exposed for leukaemia, lung, thyroid, bone, liver and all solid tumours.

A7.4.1 Relative risks and estimated lifetime dose

Airline cockpit crews are occupationally exposed to ionising radiation (IR) of cosmic origin. Radiation workers in the nuclear industry and medical and laboratory staff are the other principal group exposed. The relative risks for occupational exposure to ionising radiation were obtained from UNSCEAR, 2006, using models of excess relative risk (ERR) per unit of radiation dose, estimated as RR=1+ERR. Details of the model used are described in the appendix for each cancer site.

Dose was assumed to be an individual's cumulative dose received over the risk exposure period (REP) for each cancer (1986-2005 for leukaemia, 1956-1995 for the solid tumours). For workers exposed to ionising radiation, lifetime doses were estimated using data from the Central Index of Dose Information (CIDI). Data on collective doses for the years 1990 to 2004 were used to estimate total collective dose for the REPs, by linear extrapolation forward and back from the year 1994 to obtain estimates for 1986-89 and 2005 respectively for the 1986-2005 REP (Figure A7.2), and by assuming a constant 1990 rate prior to 1990 for the 1956-1995 REP. The estimated REP collective dose was then divided by an estimate of the numbers ever exposed to ionising radiation during the REP. These estimates, (Table A7.1) were obtained by multiplying the CIDI point estimates of IR exposed workers (see section A7.4.2 by the employment turnover factors in Table A7.2 and by the number of years in the REP (20 for leukaemia, 40 for the solid tumours).

Figure A7.2



Aircrew are not covered by the CIDI data. An estimate of lifetime dose was obtained from Langner et al. (2004). In a large cohort of airline pilots from 7 European countries employed from the earliest days of air transport (1921, Finland to 1965, Italy) up to between 1994 and 1997, the mean total lifetime radiation dose per pilot for all pilots in the cohort was 15.3 mSv, (median 10.7 mSv, maximum 78.5 mSv). The annual mean dose rate of all active pilots was 2.96 μSv per block hour flying time, for an average of 7,031 block hours. Pilots in the cohort were employed for an average 14.6 years (Langner et al, 2004). The lifetime dose estimate of 15.3 mSv per worker is used to estimate ERR for aircrew.

ERR(a) was estimated for ages (a) that could be attained by workers in 2005 who had been exposed during the REP between the ages of 15 and 65 (an even distribution of ages from 15 to 65 in the exposed cohorts was assumed). ERR (all ages) was then obtained as the average across these ages. Then RR = 1+ERR.

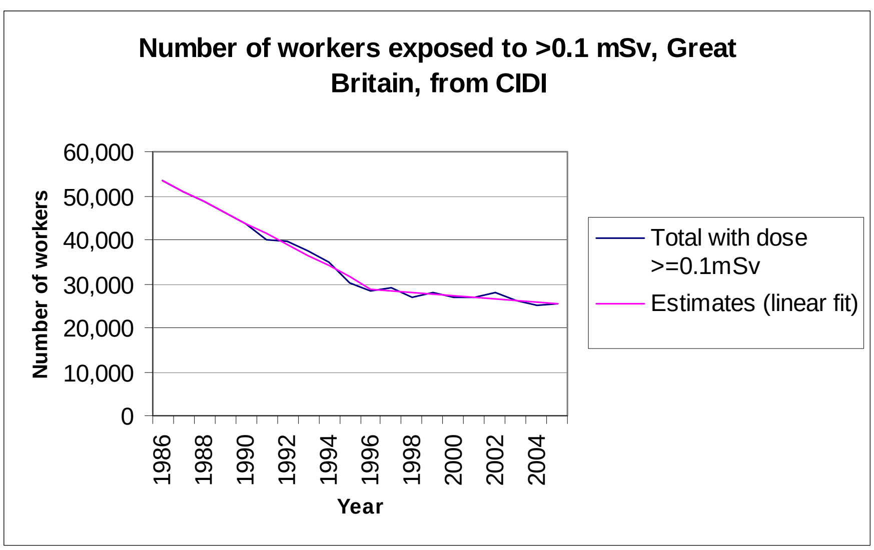
Standard errors were not available for the estimates from the UNSCEAR data at the time of publication.

The estimated RRs are given in Table A7.2.

A7.4.2 Numbers exposed

Data from the HSE's Central Index of Dose Information (CIDI) (HSE, 1998) indicates that there were 43,805 people exposed above 0.1mSv in GB in 1990, and 30,289 in 1995. The data exclude aircrew. A breakdown by occupation is given in Table 7.1. Estimated numbers exposed over 0.1mSv are split between men and women in proportion to the proportion of men (93%) with recorded doses between 1997 and 2004. Estimates of numbers of aircraft flight deck officers and male travel and flight attendants estimated from the LFS for 1979 and 1991, are also given in Table 10.1. CIDI data from 1990 and LFS data from 1979 are used as a best available point estimate for numbers exposed in the 'solid tumour' REP, 1956-1995, and CIDI data from 1995 and LFS data from 1991 are used as a best available point estimate for numbers exposed in the 'short latency' REP for leukaemia, 1986-2005. The number of workers exposed to >= 0.1 mSv radiation by year, from CIDI data, is shown in Figure A7.3.

Figure A7.3



For female air stewardesses, complete information on the numbers employed since 1958 was available from the British Airways Stewards and Stewardesses Union (for women only). The 1979 classifications 67.3+69.1 'Travel stewards and attendants', and 2000 class 630 'travel and flight attendants' used by the LFS appear to include non-air travel attendants, as the numbers drop sharply with re-classification as 'air travel assistants' from 2003 for women (Figure A7.4). For men the fall is also evident but less influential on overall numbers. In 2003 the number of women stewardesses employed by BA (11,479) was 48% of the LFS 'air travel assistants' total (23,890), and 55% of the CAA 'cabin attendants' total (20,761) (Table A7.3). Thus, doubling the BA numbers of new starters during the REP gives an appropriate estimate of stewardesses 'ever employed' in the period (13,902 in 1956-95, and 22,980 in 1986-2005). These 'ever exposed' numbers for air stewardesses are given in Table A7.3, and are used in the estimation of AF for this part of the exposed population (bypassing the usual turnover equation estimate).

Figure A7.4 Numbers of female flight personnel (cabin attendants) according to LFS, CAA and BA stewards and stewardesses union, 1979 to 2005. Numbers for CAA refer to personnel employed by UK-based airline companies but need not be resident in the UK

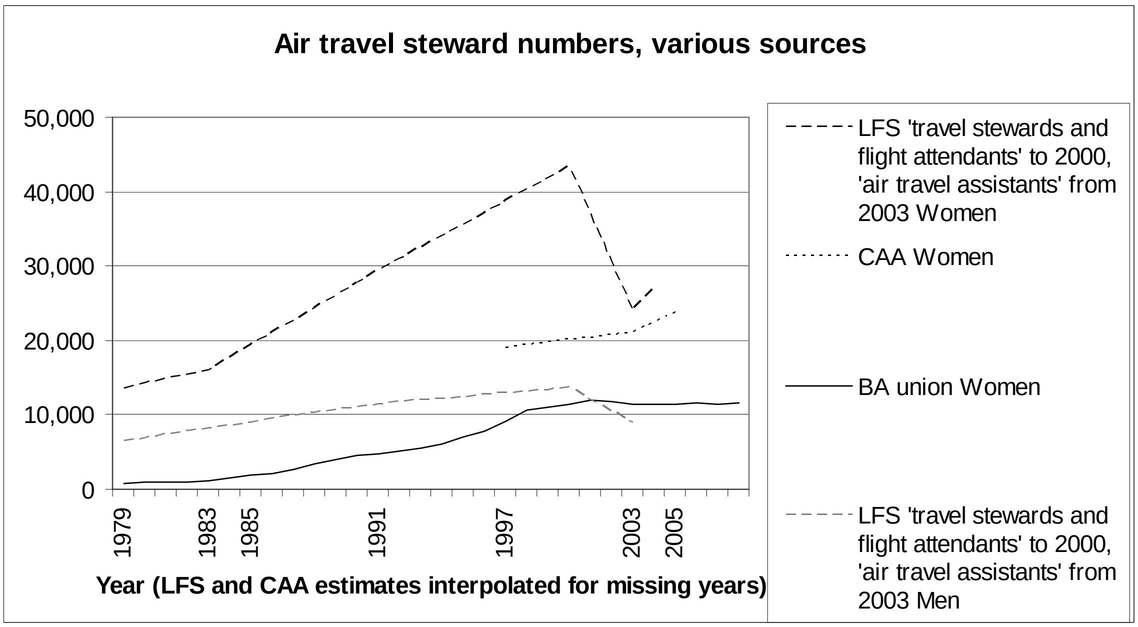


Table A7.1 Numbers of workers exposed to >0.1mSv ionising radiation in GB in 1990, from CIDI (1998), LFS 1979 and 1991, and average lifetime dose.

	Industry/occupation	Numbers exposed >0.1 mSv				Average length of employment (years) (1)		Numbers ever exposed at >0.1mSv in REP (2)			Total collective dose (mSv)	Lifetime dose /worker (mSv)		Annual dose (7)
	REP 1956-95	**M**	**F**	**Total**	**%male (3)**	**M**	**F**	**M**	**F**	**Total**				
	CIDI 1990													
C-E	Nuclear Power	13414	1010	14424	93%									
C-E	Nuclear Fuel Fabrication/ Reprocessing	7376	555	7931	93%									
C-E	General Industry	7489	564	8053	93%									
C-E	Industrial Radiography	2614	197	2811	93%									
C-E	Non-coal Mining	264	20	284	93%									
C-E	Radiation Protection	2407	181	2588	93%									
C-E	Waste Treatment	1202	90	1292	93%									
C-E	Nuclear Industry Misc.	683	51	734	93%									
C-E	Other	4275	322	4597	93%									
	Sub-total	**39724**	**2990**	**42714**		**11**	**7**	**182,730**	**19,734**	**202,464**				
G-Q	Medical/Dental	408	31	439	93%									
G-Q	Transport	179	13	192	93%									
G-Q	Academic	428	32	460	93%									
	Sub-total	**1015**	**76**	**1091**		**9**	**7**	**5,479**	**535**	**6,014**				
	CIDI Total >0.1 mSv	**40739**	**3066**	**43805**				**188,209**	**20,268**	**208,478**	**3,190,100**	**15.3**	**(4)**	**1.9**
	LFS 1979													
G-Q	Aircraft Flight Deck Officers	6915	-	6915										
G-Q	Supervisors of Travel Stewards and Attendants	258												
G-Q	Travel Stewards and Attendants	6248												
	BA stewardesses union data, 1956-95													

	Industry/occupation	Numbers exposed >0.1 mSv				Average length of employment (years) (1)		Numbers ever exposed at >0.1mSv in REP (2)			Total collective dose (mSv)	Lifetime dose /worker (mSv)		Annual dose (7)
	Air stewardesses								13,902					
	Aircrew Total	*13421*				9	7					**15.3**	(5)	2 - 6
	REP 1986-2005	***M***	***F***	***Total***	***%male (3)***	***M***	***F***	***M***	***F***	***Total***				
	CIDI 1995													
C-E	Nuclear Power	8575	645	9220	93%									
C-E	Nuclear Fuel Fabrication/ Reprocessing	6258	471	6729	93%									
C-E	General Industry	4185	315	4500	93%									
C-E	Industrial Radiography	1549	117	1666	93%									
C-E	Non-coal Mining	244	18	262	93%									
C-E	Radiation Protection	1618	122	1740	93%									
C-E	Waste Treatment	931	70	1001	93%									
C-E	Nuclear Industry Misc.	690	52	742	93%									
C-E	Other	3718	280	3998	93%									
	Sub-total	**27768**	**2090**	**29858**		**11**	**7**	**77,750**	**7,942**	**85,692**				
G-Q	Medical/Dental	130	10	140	93%									
G-Q	Transport	89	7	96	93%									
G-Q	Academic	181	14	195	93%									
	Sub-total	**401**	**30**	**431**		**9**	**7**	**1,283**	**121**	**1,403**				
	CIDI Total >0.1 mSv	**28169**	**2120**	**30289**				**79,033**	**8,063**	**87,096**	**998,549**	**11.5**	**(6)**	**1.4**
	LFS 1991													
G-Q	Aircraft Flight Deck Officers	14630	766	15396										
G-Q	Travel and flight attendants	11495												
	BA stewardesses union data, 1986-2005													
	Air stewardesses								22,980					
G-Q	**Aircrew Total**	**26125**				**9**	**7**					**15.3**	**(5)**	**2 - 6**

Notes

(1) Estimated as the inverse of the relevant employment turnover factor in Table 3.

(2) Numbers ever exposed estimated as [1991 numbers exposed at >0.1mSv * (1+TO*20)] for REP = 1986-2005,
[1979 numbers exposed at >0.1mSv * (1+TO*40)] for REP = 1956-1995, TO = annual turnover rate from Table 3.
For air stewardesses, numbers ever employed estimated as double the sum of BA union new starters during the REP.
(3) Estimated from CIDI data for 1997-2004
(4) Estimated as [CIDI total collective dose 1956-1995] divided by [numbers ever exposed 1956-1995].
CIDI collective dose for years 1956-1989 assumed = 1990 dose.
(5) From Langner, Blettner et al, 2004
(6) Estimated as [CIDI total collective dose 1986-2005] divided by [numbers ever exposed 1986-2005].
CIDI doses for 1986-1989 and 2005 estimated from linear fits to the 1990-95 and 1996-2004 data respectively see Figure 1).
(7) For CIDI data taken as estimated collective dose 1986-2005 divided by estimated average annual number exposed.

Table A7.2 Estimated relative risks for cancers associated with occupational exposure to ionising radiation

		Mortality				Incidence				
Cancer site		**Leukaemia**		**Solid tumours**		**Thyroid**	**Bone**	**Lung**		**Liver**
ICD10		C91-C95		C00-C80, C97		C73	C40-C41	C33-C34		C22
Model (UNSCEAR table)		Lin.-quad. (46)		Linear-quadratic (45)		Linear (57)	Quadratic (52)	Linear (51)		Linear (50)
		M+F		**M**	**F**	**M+F**	**M+F**	**M**	**F**	**M+F**
Using UNSCEAR models:										
RR (CIDI exposed):	11.5 *	1.03	15.3 **	1.005	1.010	1.03	1.09	1.005	1.021	1.01
RR (aircrew):	15.3 *	1.04	15.3 **	1.005	1.010	1.03	1.09	1.005	1.021	1.01
Comments		*Very similar estimate with linear component only*		*Same estimate from linear-only component of lin.-quad. model, but higher estimates with alternative linear ERR models.*			*No linear model available*			

** Dose in mSv, REP 1986-2005*
*** Dose in mSv, REP 1956-1995*

Table A7.3 Numbers of stewardesses employed by British Airways since records began in 1958

Year	BA New starters (women)	BA Leavers (women)	Numbers employed (women)	LFS, Air travel stewards (women)	CAA, Cabin attendants (women)
1958	1	0	1		
1959	0	0	1		
1960	1	0	2		
1961	0	0	2		
1962	1	0	3		
1963	1	0	4		
1964	2	0	6		
1965	7	0	13		
1966	9	0	22		
1967	7	0	29		
1968	18	0	47		
1969	28	0	75		
1970	39	0	114		
1971	30	0	144		
1972	57	0	201		
1973	98	0	299		
1974	140	0	439		
1975	71	0	510		
1976	71	0	581		
1977	89	0	670		
1978	111	0	781		
1979	47	0	828	13237 *	
1980	40	0	868		
1981	27	0	895		
1982	40	0	935		
1983	128	0	1063	15706 *	
1984	396	0	1459		
1985	401	0	1860		
1986	193	0	2053		
1987	703	0	2756		
1988	643	0	3399		
1989	570	0	3969		
1990	666	0	4635		
1991	52	0	4687	29440 **	
1992	507	0	5194		
1993	312	0	5506		
1994	536	0	6042		
1995	909	0	6951		
1996	848	0	7799		

Year	BA New starters (women)	BA Leavers (women)	Numbers employed (women)	LFS, Air travel stewards (women)	CAA, Cabin attendants (women)
1997	1390	0	9189		18765
1998	1421	0	10610		
1999	452	0	11062		
2000	431	3	11490	43336 **	
2001	579	4	12065		
2002	90	333	11822		
2003	205	548	11479	23890 ***	20761
2004	476	539	11416	26978 ***	
2005	563	504	11475		23189
2006	753	601	11627		
2007	326	626	11327		
2008	626	330	11623		
Estimation period	***Sum of new starters***	***Doubled for all air stewards employed***		***BA as % of LFS, 2003***	***BA as % of CAA, 2003***
REP 1956-1995	6951	13,902		48%	55%
REP 1986-2005	11546	22,980			

CAA Civil Aviation Authority. Numbers for CAA refer to personnel employed by UK-based airline companies but need not be resident in the UK

* 69.1 ‘Travel stewards and attendants’

** 630 ‘Travel and flight attendants’

*** 6214 ‘Air travel assistants’

A7.4.3 Models and parameters used to estimate excess relative risks, from UNSCEAR 2006

Leukaemia

Model used to derive RR:
ERR(a) = $[\alpha \,.\, D + \beta \,.\, D^2] \,.\, \exp[\kappa_1 \,.\, \ln(a)]$, from UNSCEAR 2006, Table 46, mortality model:

Generalized ERR model, linear–quadratic dose response	
$$h_0(a,e,c,s)\cdot\left[1+(\alpha\cdot D+\beta\cdot D^2)\cdot\exp[\kappa_1\cdot\ln[a]]\right]$$	
α =	864.552 Sv^{-1}
β/α =	1.180 92 Sv^{-1}
κ_1 =	-1.647

where D = mean lifetime dose /worker (Sv)
a = attained age; e = age at exposure; s = sex; c = city
h_0 = underlying death/incidence rate

From Table D3: Model Deviance = 2,136.589, df = 31,412
ERR is obtained as average ERR(a), averaged over a = 15-84 (short latency REP 1986-2005).

Solid tumours

Model used to derive RR:
ERR(a) = $[\alpha \,.\, D + \beta \,.\, D^2] \,.\, \exp[\kappa_1 \,.1_{s=female} + \kappa_2 \ln[a-e] + \kappa_3 \ln([a]]$, from UNSCEAR 2006, Table 45, mortality model:

Generalized ERR model (adjustment for attained age, years since exposure), linear–quadratic dose response	
$$h_0(a,e,c,s)\cdot\left[1+(\alpha\cdot D+\beta\cdot D^2)\cdot\exp[\kappa_1\cdot 1_{s=female}+\kappa_2\cdot\ln[a-e]+\kappa_3\cdot\ln[a]]\right]$$	
α =	408.285 Sv^{-1}
β/α =	0.292 23 Sv^{-1}
κ_1 =	0.663
κ_2 =	0.987 1
κ_3 =	-2.636

e = 20 years for all ages, i.e. variable values for a-e.
From Table D1: Model Deviance = 13,376.676, df = 31,394
ERR is obtained as average ERR(a), averaged over a = 25-100 (long latency REP 1956-1995).
Used mortality rather than incidence estimate as the incidence estimate was for sites other than the 11 separately estimated.

Thyroid cancer
Model used to derive RR:
ERR(a) = α . D . exp[κ_1 ln[e] + κ_2 ln([a]], from UNSCEAR 2006, Table 57, incidence model:

Generalized ERR model, linear dose response	
$h_0(a,e,c,s)\cdot\left[1+\alpha\cdot D\cdot\exp[\kappa_1\cdot\ln[e]+\kappa_2\cdot\ln[a]]\right]$	
α =	3.80452×10^4 Sv^{-1}
κ_1 =	-0.4405
κ_2 =	-2.197

e = 20 years
From Table D15: Model Deviance = 2,890.965, df = 42,697
ERR is obtained as average ERR(a), averaged over a = 25-100 (long latency REP 1956-1995).

Bone cancer
Model used to derive RR:
ERR(a) =β . D^2. exp[κ_1 ln([a]], from UNSCEAR 2006, Table 52, incidence model:

Generalized ERR model, quadratic dose response	
$h_0(a,e,c,s)\cdot\left[1+\beta\cdot D^2\cdot\exp[\kappa_1\cdot\ln[a]]\right]$	
β =	$6.903\ 79 \times 10^7$ Sv^{-2}
κ_1 =	-4.472

From Table D10: Model Deviance = 236.222, df = 42,703
ERR is obtained as average ERR(a), averaged over a = 25-100 (long latency REP 1956-1995).

Liver cancer
Model used to derive RR:
ERR(a) = α . D , from UNSCEAR 2006, Table 50, incidence model:

Generalized ERR model, linear dose response	
$h_0(a,e,c,s)\cdot\left[1+\alpha\cdot D\right]$	
α =	$3.951\ 06 \times 10^{-1}$ Sv^{-1}

From Table D8: Model Deviance = 5,370.978, df = 42,690
ERR is obtained as average ERR(a), averaged over a = 25-100 (long latency REP 1956-1995).

Lung cancer
Model used to derive RR:
ERR(a) = α . D . exp[κ_1 . $1_{s=female}$], from UNSCEAR 2006, Table 51, incidence model:

Generalized ERR model, linear dose response	
$$h_0(a,e,c,s)\cdot\left[1+\alpha\cdot D\cdot\exp[\kappa_1\cdot 1_{s=female}]\right]$$	
α =	$3.182\ 24 \times 10^{-1}$ Sv^{-1}
κ_1 =	1.480 8

From Table D9: Model Deviance = 6,181.503, df = 42,695
ERR is obtained as average ERR(a), averaged over a = 25-100 (long latency REP 1956-1995).

APPENDIX 8: RANDOM ERROR CONFIDENCE INTERVALS

For each AF, a confidence interval that takes into account random error only was calculated. The equations for the various calculation methods and data sources are given below and were based upon methods published by Steenland and Armstrong (2006). However, it was acknowledged that the effect of bias needed to be accounted for. A full sensitivity analysis and description of the methodology for estimating confidence limits that incorporate bias, known as 'credibility limits', and examples of these will be the subject of a forthcoming report.

Confidence interval for random error only

Methods used to derive a confidence interval for the AF that take account only of this type of error are well established. The simplest method is to use the upper and lower confidence interval bounds published with the relative risk estimate in the AF calculation to obtain a 'naïve' confidence interval for the AF (Greenland, 2004). This assumes no random variability in the estimate of the proportion of the population exposed [Pr(E)], which may be acceptable where this estimate is based on national census data. This was the method used for the majority of the current AF CI calculations.

Where the proportion of cases exposed [Pr(E|D)] was derived from the same case-control study source as the estimate for RR, and AF was calculated using Miettinen's equation, a formula to estimate the variance and confidence limits is given by Steenland and Armstrong (2006), and Greenland (1987). The equation is:

$$\mathrm{Var}[\ln(1-AF)] = [(AF)^2/(1-AF)^2][\{(\mathrm{var}\ \ln(RR))/(RR-1)^2\} + \{2/n_{ec}(RR-1)\} + (n_{nec}/n_{ec}n_t)] \qquad (8.1)$$

where

n_{ec} = number of exposed cases
n_{nec} = number of non-exposed cases
n_t = total number of cases

Equation (8.1) is used to estimate the upper and lower confidence limits of the AF (ULAF and LLAF); these are 1-exp[ln(1-AF) +/- 1.96(√var(ln(1-AF)))].

Where the estimate of Pr(E) for the AF was obtained from an independent (survey) source and Levin's formula was used, the following equation, from Steenland and Armstrong (2006) and Greenland (2004) can be used:

$$\mathrm{St.dev.}(\ln(1-AF)) = [(O*RR*\mathrm{st.dev.}RR)^2 + (O/T)\{(RR-1)^2 + (RR*\mathrm{st.dev.}(RR)\}^2]^{1/2} *[(1-AF)/(1+O)] \qquad (8.2)$$

where T = number in survey
O = Pr(E) /(1-Pr(E)).

This incorporates the variance of Pr(E) taken from an ancillary survey in the target population, and is used to obtain 95% Wald limits, which are converted back to the original AF scale. However, in the case of CAREX estimates of the numbers exposed, the use of this equation was not appropriate, as the CAREX estimates were not known to be based on a survey of a target population for which T was known. Therefore as noted above 'naïve' limits were used, based on the random variability of RR alone when Levin's formula and CAREX exposed numbers were used to estimate AF. Where LFS data were used, the sample survey size of about 55,000 households resulted in a near zero second term in equation 8.2, so ‘naïve’ limits were also appropriate in this case.

Methods used in this project

Monte Carlo methods can be used (Greenland, 2004), for combined AFs for a single exposure based on more than one RR estimate, or for a combined AF across more than one exposure for a single cancer, based on the use of Greenland’s ‘naïve’ confidence intervals for single relative risk AF estimates. Each RR is assumed to be distributed log-normally. 100,000 random standard normal (0,1) variates are generated for each RR (the RR_k) in the estimate. From these, 100,000 values with mean ln RR_k and its SD, transformed back using the exponential function, are obtained for each RR_k. 100,000 AF_is are then calculated using these log-normally distributed RR_{ki}s applied in the equation that has been used to calculate the AF. The results are ordered, and the upper and lower 95% confidence limits are given by the values of AF_i at the 97.5^{th} and 2.5^{th} percentiles of the distribution of the AF_is. Where a negative lower limit results from a lower CI limit for the RR of less than 1, 0 is assumed for the lower limit of the AF.

When results for all cancers have been obtained, a straightforward extension of this method, using 100,000 random values (the same or a new set) for each RR represented in the combined estimate, can be used to estimate a confidence interval for the overall ‘all cancers’ AF.

An alternative method has been used for situations in which the AF has been estimated using a method that does not depend on an estimate of RR with a suitable confidence interval (e.g. for lung cancer due to exposure to asbestos and radon). In this case if there is an estimated CI for the AF, the assumption is made that the AF is approximately log-normally distributed (this has been tested empirically on a range of AFs for individual cancer-exposure pairings and combined exposure AFs). The above procedure is then followed to generate 100,000 random values for the AF as for the RR.

APPENDIX 9: CLASSIFICATION OF INDUSTRIES AND OCCUPATIONS USED IN THIS STUDY

Description of Industry/Occupation in source coding	Main industry sector code	Source	Description for cancer burden study
Agriculture and hunting	A-B	CAREX	Farming
Fishing	A-B	CAREX	Farming
Agricultural machinery drivers and operators	A-B	LFS1979	Farming
Agricultural machinery foremen	A-B	LFS1979	Farming
All other in farming and related	A-B	LFS1979	Farming
Farm foremen	A-B	LFS1979	Farming
Farm workers	A-B	LFS1979	Farming
Farmers, horticulturalists, farm managers	A-B	LFS1979	Farming
Other foremen in farming and related	A-B	LFS1979	Farming
Agriculture and horticulture	A-B	LFS1991	Farming
Farm owners and managers, horticulturalists	A-B	LFS1991	Farming
Other managers farming, horticulture, forestry and fishing	A-B	LFS1991	Farming
Other occupations in farming and related	A-B	LFS1991	Farming
Forestry and logging	A-B	CAREX	Forestry
Forestry foremen	A-B	LFS1979	Forestry
Forestry workers	A-B	LFS1979-1991	Forestry
Forestry	A-B	LFS1991	Forestry
Foremen of gardeners and groundsmen	A-B	LFS1979	Horticulture
Gardeners and groundsmen	A-B	LFS1979	Horticulture
Horticultural foremen	A-B	LFS1979	Horticulture
Horticultural workers	A-B	LFS1979	Horticulture
Gardeners, groundsmen, groundswomen	A-B	LFS1991	Horticulture
Horticultural trades	A-B	LFS1991	Horticulture
Beverage industries	C-E	CAREX	Beverage industries
Crude petroleum and natural gas production	C-E	CAREX	Crude petroleum and natural gas production
Electricity, gas and steam	C-E	CAREX	Electricity, gas and steam
Nuclear fuel fabrication/ reprocessing	C-E	CIDI	Electricity, gas and steam
Nuclear industry misc.	C-E	CIDI	Electricity, gas and steam
Nuclear power	C-E	CIDI	Electricity, gas and steam

Description of Industry/Occupation in source coding	**Main industry sector code**	**Source**	**Description for cancer burden study**
Other	C-E	CIDI	Electricity, gas and steam
Radiation protection	C-E	CIDI	Electricity, gas and steam
Food manufacturing	C-E	CAREX	Food manufacturing
Grain milling	C-E	LFS1979-1991	Food manufacturing
General industry	C-E	CIDI	General industry
Iron and steel basic industries	C-E	CAREX	Iron and steel basic industries
Iron and steel industry	C-E	CoE 1981, LFS 1991	Iron and steel basic industries
Foremen of labourers and unskilled workers in coke ovens and gas works	C-E	LFS1979	Iron and steel basic industries
Foremen of workers in foundries	C-E	LFS1979	Iron and steel basic industries
Labourers and other unskilled workers in coke ovens and gas works	C-E	LFS1979	Iron and steel basic industries
Labourers and other unskilled workers in foundries in engineering	C-E	LFS1979	Iron and steel basic industries
Fettlers dressers	C-E	LFS1979	Iron and steel basic industries
Shot blasters	C-E	LFS1979	Iron and steel basic industries
Steel tubes	C-E	CoE 1981, LFS 1991	Iron and steel basic industries
Steel wire and wire products	C-E	CoE 1981, LFS 1991	Iron and steel basic industries
Insulated wires and cables	C-E	LFS1991	Iron and steel basic industries
Other drawing cold rolling forming of steel	C-E	CoE 1981, LFS 1991	Iron and steel basic industries
Manufacture of electrical machinery, apparatus, appliances and supplies	C-E	CAREX	Manufacture of electrical machinery, apparatus, appliances and supplies
Manufacture of fabricated metal products, except machinery and equipment	C-E	CAREX	Manufacture of fabricated metal products, except machinery and equipment
Metal polishers	C-E	LFS1979	Manufacture of fabricated metal products, except machinery and equipment
Foremen of metal polishers	C-E	LFS1979	Manufacture of fabricated metal products, except machinery and equipment
Manufacture of footwear	C-E	CAREX	Manufacture of footwear
Footwear	C-E	CoE1971	Manufacture of footwear
Manufacture of furniture and fixture, except primary of metal	C-E	CAREX	Manufacture of furniture and fixture, except primary of metal
Manufacture of glass and glass products	C-E	CAREX	Manufacture of glass and glass products

Description of Industry/Occupation in source coding	Main industry sector code	Source	Description for cancer burden study
Flat glass manufacture	C-E	CoE 1981, LFS 1991	Manufacture of glass and glass products
Glass containers	C-E	CoE 1981, LFS 1991	Manufacture of glass and glass products
Other glass products	C-E	CoE 1981, LFS 1991	Manufacture of glass and glass products
Manufacture of industrial chemicals	C-E	CAREX	Manufacture of industrial chemicals
Organic chemical manufacture	C-E	CoE 1981, LFS 1991	Manufacture of industrial chemicals
Formulated pesticides	C-E	CoE 1981, LFS 1991	Manufacture of industrial chemicals
Manufacture of instruments, photographic and optical goods	C-E	CAREX	Manufacture of instruments, photographic and optical goods
Foremen of precision instrument makers and repairers	C-E	LFS1979	Manufacture of instruments, photographic and optical goods
Foremen of toolmakers tool fitters markers-out	C-E	LFS1979	Manufacture of instruments, photographic and optical goods
Manufacture of leather and products of leather or of its substitutes	C-E	CAREX	Manufacture of leather and products of leather or of its substitutes
Leather (tanning and dressing)	C-E	CoE1971	Manufacture of leather and products of leather or of its substitutes
Leather goods	C-E	CoE1971	Manufacture of leather and products of leather or of its substitutes
Manufacture of machinery except electrical	C-E	CAREX	Manufacture of machinery except electrical
Office machinery mechanics	C-E	LFS1979	Manufacture of machinery except electrical
Industrial radiography	C-E	CIDI	Manufacture of machinery except electrical
Manufacture of miscellaneous products of petroleum and coal	C-E	CAREX	Manufacture of miscellaneous products of petroleum and coal
Manufacture of other chemical products	C-E	CAREX	Manufacture of other chemical products
Manufacture of other non-metallic mineral products	C-E	CAREX	Manufacture of other non-metallic mineral products
Cement, lime, and plaster manufacture	C-E	CoE 1981, LFS 1991	Manufacture of other non-metallic mineral products
Refractory goods	C-E	CoE 1981, LFS 1991	Manufacture of other non-metallic mineral products
Manufacture of paper and paper products	C-E	CAREX	Manufacture of paper and paper products

Description of Industry/Occupation in source coding	**Main industry sector code**	**Source**	**Description for cancer burden study**
Pulp manufacture	C-E	CoE 1981, LFS 1991	Manufacture of paper and paper products
Manufacture of plastic products not elsewhere classified	C-E	CAREX	Manufacture of plastic products not elsewhere classified
Manufacture of pottery, china and earthenware	C-E	CAREX	Manufacture of pottery, china and earthenware
Ceramic goods	C-E	CoE 1981, LFS 1991	Manufacture of pottery, china and earthenware
Manufacture of rubber products	C-E	CAREX	Manufacture of rubber products
Foremen inspectors viewers examiners of rubber goods	C-E	LFS1979	Manufacture of rubber products
Foremen of rubber	C-E	LFS1979	Manufacture of rubber products
Foremen of rubber process workers moulding machine operators tyre builders	C-E	LFS1979	Manufacture of rubber products
Inspectors viewers examiners of rubber goods	C-E	LFS1979	Manufacture of rubber products
Rubber	C-E	LFS1979	Manufacture of rubber products
Rubber process workers moulding machine operators tyre builders	C-E	LFS1979	Manufacture of rubber products
Manufacture of textiles	C-E	CAREX	Manufacture of textiles
Other textile industries	C-E	CoE1971	Manufacture of textiles
Textile finishing	C-E	CoE1971	Manufacture of textiles
Manufacture of transport equipment	C-E	CAREX	Manufacture of transport equipment
Foremen of maintenance fitters (aircraft engines)	C-E	LFS1979	Manufacture of transport equipment
Maintenance fitters (aircraft engines)	C-E	LFS1979	Manufacture of transport equipment
Manufacture of wearing apparel, except footwear	C-E	CAREX	Manufacture of wearing apparel, except footwear
Manufacture of wood and wood and cork products, except furniture	C-E	CAREX	Manufacture of wood and wood and cork products, except furniture
Wood sawmill planning impregnation	C-E	CoE 1981, LFS 1991	Manufacture of wood and wood and cork products, except furniture
Metal ore mining	C-E	CAREX	Mining
Non-coal mining	C-E	CIDI	Mining
Other mining	C-E	CAREX	Mining
Cornish tin miners cohort	C-E	HSE	Mining
Non-ferrous metal basic industries	C-E	CAREX	Non-ferrous metal basic industries
Clydach nickel carbonyl factory	C-E	LFS1979	Non-ferrous metal basic industries
Copper, brass and other copper alloys	C-E	CoE 1981, LFS 1991	Non-ferrous metal basic industries

Description of Industry/Occupation in source coding	**Main industry sector code**	**Source**	**Description for cancer burden study**
Other non-ferrous metals and their alloys	C-E	CoE 1981, LFS 1991	Non-ferrous metal basic industries
Aluminium and aluminium alloys	C-E	CoE 1981, LFS 1991	Non-ferrous metal basic industries
Other manufacturing industries	C-E	CAREX	Other manufacturing industries
Petroleum refineries	C-E	CAREX	Petroleum refineries
Mineral oil refining	C-E	LFS1979	Petroleum refineries
Printing, publishing and allied industries	C-E	CAREX	Printing, publishing and allied industries
Foremen of printers (so described)	C-E	LFS1979	Printing, publishing and allied industries
Foremen of printing machine minders and assistants	C-E	LFS1979	Printing, publishing and allied industries
Printers (so described)	C-E	LFS1979	Printing, publishing and allied industries
Printing machine minders and assistants	C-E	LFS1979	Printing, publishing and allied industries
Services allied to transport	G-Q	CAREX	Services allied to transport
Tobacco manufacture	C-E	CAREX	Tobacco manufacture
Water works and supply	C-E	CAREX	Water works and supply
Coach painters	C-E	LFS1979	Painters, not construction
Foremen of coach painters (so described)	C-E	LFS1979	Painters, not construction
Foremen of other spray painters	C-E	LFS1979	Painters, not construction
Other spray painters	C-E	LFS1979	Painters, not construction
Painting assembling and related occupations, nec	C-E	LFS1979	Painters, not construction
Foremen of machine tool operators	C-E	LFS1979	Metal workers
Foremen of machine tool setter operators	C-E	LFS1979	Metal workers
Foremen of metal working production fitters and fitter/machinists	C-E	LFS1979	Metal workers
Foremen of other centre lathe turners	C-E	LFS1979	Metal workers
Foremen of press and machine tool setters	C-E	LFS1979	Metal workers
Foremen of press stamping and automatic machine operators	C-E	LFS1979	Metal workers
Machine tool operators	C-E	LFS1979	Metal workers
Machine tool setter operators	C-E	LFS1979	Metal workers
Metal working production fitters and fitter/machinists	C-E	LFS1979	Metal workers
Other centre lathe turners	C-E	LFS1979	Metal workers

Description of Industry/Occupation in source coding	Main industry sector code	Source	Description for cancer burden study
Precision instrument makers and repairers	C-E	LFS1979	Metal workers
Press and machine tool setters	C-E	LFS1979	Metal workers
Press stamping and automatic machine operators	C-E	LFS1979	Metal workers
Toolmakers tool fitters markers-out	C-E	LFS1979	Metal workers
Labourers and other unskilled workers in engineering and allied t	C-E	LFS1979	Metal workers
Foremen of painters and decorators	F	LFS1979	Painters and decorators (Construction)
Painters and decorators	F	LFS1979	Painters and decorators (Construction)
Foremen of paviors kerb layers	F	LFS1979	Roofers, road surfacers, Roadmen, Paviors (Construction)
Foremen of road surfacers concreters	F	LFS1979	Roofers, road surfacers, Roadmen, Paviors (Construction)
Foremen of roadmen	F	LFS1979	Roofers, road surfacers, Roadmen, Paviors (Construction)
Foremen of roofers glaziers	F	LFS1979	Roofers, road surfacers, Roadmen, Paviors (Construction)
Paviors kerb layers	F	LFS1979	Roofers, road surfacers, Roadmen, Paviors (Construction)
Road surfacers concreters	F	LFS1979	Roofers, road surfacers, Roadmen, Paviors (Construction)
Roadmen	F	LFS1979	Roofers, road surfacers, Roadmen, Paviors (Construction)
Roofers glaziers	F	LFS1979	Roofers, road surfacers, Roadmen, Paviors (Construction)
Construction	F	CAREX	Construction
Shift work	G-Q	LFS1979 and 1992	Shift work
Air transport	G-Q	CAREX	Air transport
Business, professional and other organisation	G-Q	CAREX	Business, professional and other organisation
Communication	G-Q	CAREX	Communication
Education services	G-Q	CAREX	Education services
Financing, insurance, real estate and business services	G-Q	CAREX	Financing, insurance, real estate and business services
Flight personnel	G-Q	BASSA cohort	Flight personnel
Travel stewards and attendants	G-Q	LFS1979	Flight personnel
Travel and flight attendants	G-Q	LFS1991	Flight personnel
Supervisors of travel stewards and attendants	G-Q	LFS1979	Flight personnel
Aircraft flight deck officers	G-Q	LFS 1979	Flight personnel

Description of Industry/Occupation in source coding	Main industry sector code	Source	Description for cancer burden study
Land transport	G-Q	CAREX	Land transport
Transport	G-Q	CIDI	Land transport
Medical, dental, other health and veterinary services	G-Q	CAREX	Medical, dental, other health and veterinary services
Medical/dental	G-Q	CIDI	Medical, dental, other health and veterinary services
Personal and household services	G-Q	CAREX	Personal and household services
Watch and chronometer makers and repairers	G-Q C-E	LFS1979	Personal and household services
Dry cleaning, job dyeing etc	G-Q	LFS1991	Personal and household services
Dyestuffs and pigments	G-Q	LFS1991	Personal and household services
Chimney sweep	G-Q	LFS1979	Personal and household services
Hairdressers and barbers managers	G-Q	LFS1979	Personal and household services
Hairdressers and barbers	G-Q	LFS1979-1991	Personal and household services
Foremen of motor mechanics auto engineers	G-Q	LFS1979	Personal and household services
Motor mechanics auto engineers	G-Q	LFS1979	Personal and household services
Public administration and defence	G-Q	CAREX	Public administration and defence
Recreational and cultural services	G-Q	CAREX	Recreational and cultural services
Research and scientific institutes	G-Q	CAREX	Research and scientific institutes
Academic	G-Q	CIDI	Research and scientific institutes
Waste treatment	G-Q	CIDI	Sanitary and similar services
Sanitary and similar services	G-Q	CAREX	Sanitary and similar services
Water transport	G-Q	CAREX	Water transport
Welfare institutions	G-Q	CAREX	Welfare institutions
Wholesale and retail trade and restaurants and hotels	G-Q	CAREX	Wholesale and retail trade and restaurants and hotels
Scrap dealers general dealers rag and bone merchants	G-Q	LFS1979	Wholesale and retail trade and restaurants and hotels
Scrap dealers and metal merchants	G-Q	LFS1991	Wholesale and retail trade and restaurants and hotels
Foremen of welders	C-E	LFS1979	Welders
Welders	C-E	LFS1979	Welders

Abbreviations used

CAREX	CARcinogens Exposure database
LFS	Labour Force Survey
CoE	Census of Employment
CIDI	Central Index of Dose Information
BASSA	British Airways Stewards and Stewardesses Union

GLOSSARY OF TERMS

Attributable fraction (AF)	The proportion of cases that would not have occurred in the absence of occupational exposure.
Attributable numbers (AN)	Deaths, or cancer registrations, that would not have occurred in the absence of occupational exposure.
Best source /study	This is the study chosen as the source of a relative risk estimate for one or more exposure pairings.
Cancer/exposure pairing	This is a single carcinogenic occupational exposure and a site at which cancer is induced. It is the basic unit for which AF is calculated.
Causality	IARC criteria for causality include (1) a strong association (e.g. a large relative risk), (2) risk increasing with the exposure, and (3) the demonstration of a decline in risk after cessation, or reduction in exposure.
Confounder	A confounding factor, also known as an adjustment factor, is a factor which can have the effect of biasing the estimate of relative risk (RR) if it is correlated in the study population with the exposure of interest, whether it is in the causal pathway for the disease or not. Examples are age, and smoking status in the case of lung cancer.
Dominant exposure	Where exposures overlap in the working population, the exposed workers are allocated to the exposure deemed to have the highest risk, that is the dominant exposure, for calculation of the attributable fraction.
"Strong" exposure group	Carcinogens established by IARC as Group 1 and with 'strong' evidence of carcinogenicity in humans, such as a large relative risk, for the relevant cancer site (after Siemiatycki et al (2004))
"Suggestive" exposure group	Carcinogens in IARC Groups 1 or 2A and with 'suggestive' evidence of carcinogenicity in humans for the relevant cancer site (after Siemiatycki *et al.* (2004)), plus other suspected carcinogens for which the evidence of carcinogenicity in humans for the relevant cancer site is not yet established.
Exposure map	A diagram setting out links and therefore overlaps between exposures in the working population. Linked exposures may also not act independently of one another in causing disease.
Exposure scenario	The industry process, job or occupation in which a worker is exposed to a carcinogenic substance.
Exposure set	A group of exposures that overlap in the working population, and are not independent of one another in causing disease, so that their

contribution towards attributable fraction cannot be estimated separately.

External data sources	These are sources for an estimate of the proportion of the population exposed that do not come from the study providing the source for the relative risk estimate; either CAREX, national employment data from the Census of Employment (CoE) (and its successors the Annual Employment Survey (AES) and the Annual Business Inquiry (ABI)), or the Labour Force Survey (LFS).
Industry-based study	An epidemiological study whose study members are drawn from a single industry or occupational group. These are usually cohort studies, or case-control studies drawn from within an industry-based cohort.
Occupational exposure	There are two ways of defining exposure, by substance or by occupation. Substances are the IARC agents, groups of agents and mixtures identified as carcinogens. Occupations cover exposure scenarios where the causal agent is unknown.
Overlapping exposures	This is where the distribution of occupational exposures in the population overlaps, so that some workers will be subject to multiple exposures in a single workplace, or in a series of jobs over a working lifetime.
Population-based study	An epidemiological study drawing all study members from a national or large regional population, or from hospital attendees. Census based cohorts as well as hospital or cancer registry or death certificate based case-control studies are included under this heading.
Risk exposure period	The period during which exposure to a causal factor is able to result in a cancer appearing at a certain time. It can also be called a Relevant Latency Window.
Source population	The population in which the study was done from which RR estimates for calculating AF have been taken.
Target population	The population of Great Britain
Target year	The year for which a “current burden” estimate is made. It is normally the most recent year for which mortality or cancer registration data is available.

BIBLIOGRAPHY

Andersen A, Barlow L, Engeland A, Kjaerheim K, Lynge E, Pukkala E (1999) Work-related cancer in the Nordic countries. *Scand J Work Environ Health* **25 (Suppl 2)**: 1–116

BASSA (British Airlines Stewards and Stewardesses Assoc) (2008) Personal communication.

Benichou J (2001) A review of adjusted estimators of attributable risk. *Statistical Methods in Medical Research* **10**: 195-216

Blanc PD, Becker CE. (1992) Occupational health planning in Israel: a US perspective. *Isr J Med Sci* **28**:500–8

Bouchardy C, Schuler G, Minder C, Hotz P, Bousquet A, Levi F, Fisch T, Torhorst J, Raymond L (2002) Cancer risk by occupation and socioeconomic group among men – a study by the Association of Swiss Cancer Registries. *Scand. J Work Environ Health* **28, suppl 1**: 1-88

Breslow NE, Day NE (1989) Statistical Methods in Cancer Research, Volume II: The Design and Analysis of Cohort Studies. International Agency for Research on Cancer, Lyon, France. Page 67

Bruzzi P, Green SB, Byar DP, Brinton LA, Schairer C (1985). Estimating the population attributable risk for multiple risk factors using case-control data. *American J Epidemiology* **122:** 904-914

Carpenter L, Roman E (1999) Cancer and Occupation in women: identifying associations using routinely collected national data. *Env. Health Perspectives* **107**: 299-303

Darnton AJ, McElvenny DM, Hodgson JT (2006) Estimating the number of asbestos-related lung cancer deaths in Great Britain from 1980 to 2000. *Ann. Occup. Hyg* **50**: 29-38

DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. *Controlled Clinical Trials* **7**: 177-188

Deschamps F, Barouh M, Deslee G, Prevost A, Munck J (2006) Estimates of work-related cancers in workers exposed to carcinogens. *Occupational Medicine* **56**: 204–209

Doll R, Peto R (1981) The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *J Natl Cancer Inst* **66(6)**:1191-1308

Dreyer L, Andersen A, Pukkala E (1997) Occupation. In 'Avoidable cancers in the Nordic countries' APMIS **Suppl. 76**: 105, 68-79

Driscoll T, Mannetje A, Dryson E, et al. (2004) *The Burden of Occupational Disease and Injury in New Zealand: Technical Report*. Wellington: National Occupational Health and Safety Advisory Committee

Goldberg M, Imbernon E, Rolland P, Gilg Soit Ilg A, Saves M, de Quillacq A, Frenay C, Chammings S, Arveux P, Boutin C, Launoy G, Pairon J, Astoul P, Galateau-Salle F, Brochard P (2006) The French national mesothelioma surveillance program. *Occup Environ Med* **63:** 390-395

Greenland S (1984) Bias in methods for deriving standardized morbidity ratio and attributable fraction estimates. *Stat Med* **3**:131-41

Greenland S (1987) Variance estimators for attributable fraction estimates consistent in both large strata and sparse data *Stat Med* **6**: 701-708

Greenland S (1998) Meta-analysis. In: Rothman, K.J. & Greenland, S., eds, *Modern Epidemiology,* Philadelphia, Lippincott Williams & Wilkins, pp. 643–673

Greenland S 2004. Interval estimation by simulation as an alternative to and extension of confidence intervals. *International J Epidemiology* **33**: 1389-1397.

Gregg P, Wadsworth J (1995) A short history of labour turnover, job tenure, and job security, 1975-93. *Oxford Review of Economic Policy* **11**: 73-90

Gregg P, Wadsworth J (2002) Job tenure in Britain, 1975-2000. Is a job for life or just for Christmas? *Oxford Bulletin of Economics and Statistics* **64**: 111-134

Gustavsson P, Jakobsson R, Nyberg F, Pershagen G, Jarup L, Scheele P (2000) Occupational exposure and lung cancer risk: A population-based case-referent study in Sweden. *American Journal Epidemiology* **152**: 32-40

Hodgson JT, Darnton A (2000) The quantitative risks of mesothelioma and lung cancer in relation to asbestos exposure *Ann. Occ. Hygiene* **44**: 565-601

Hoffmann K, Heidemann C, Weikert C, Schulze MB, Boeing H (2006) Estimating the proportion of disease due to classes of sufficient causes. *Am J Epidemiol* **163**: 76–83

Howel D, Arblaster L, Swinburne L, Schweiger E, Renvoiz, E, Hatton P (1997) Routes of asbestos exposure and the development of mesothelioma in an English region. *Occup Environ Med* **54:** 403-409

HSE (1998) Occupational exposure to ionising radiation 1990-1996, analysis of doses reported to the Health and Safety Executive's Central Index of Dose Information.

HSE (2002) Amendment to the Control of Asbestos at Work Regulations 1987 and ACoP: Regulatory Impact Assessment (available on the web at http://www.hse.gov.uk/ria/chemical/ria-176.pdf). Health and Safety Executive. P31 para A63.

Hutchings S (2010a) Estimation of the Burden of Cancer due to Occupation in Great Britain, Disability Adjusted Life Years, Technical Report. HSE website (In Preparation)

Hutchings S (2010b) Estimation of the Burden of Cancer due to Occupation in Great Britain, Future Burden Methodology, Technical Report. HSE website (In Preparation)

Hutchings S (2010c) Estimation of the Burden of Cancer in Great Britain due to Occupation, Bias and Credibility Intervals, Technical Report. HSE website (In Preparation)

Kogevinas M, Kauppinen T, Boffetta P, Saracci R (1998) *Estimation of the Burden of Occupational Cancer in Europe. Final Report to the European Commission of a project funded by the Programme "Europe Against Cancer"*. Barcelona: IMIM

Landrigan PL, Schechter CB, Lipton JM, Fahs MC, Schwartz J (2002) Environmental pollutants and disease in American children: estimates of morbidity, mortality and costs for lead poisoning, asthma, cancer, and developmental disabilities. *Environmental Health Perspectives* **110**: 721-728

Langner IM, Blettner M, Gundestrup et al. (2004) Cosmic radiation and cancer mortality among airline pilots: results from a European cohort study (ESCAPE). *Radiat Environ Biophys* **42(4)**: 247-256

Leigh J, Markowitz S, Fahs M (1997) Occupational injury and illness in the United States. Estimates of costs, morbidity and mortality: *Arch. Internal Med* **157**: 1557-1568

Levin M (1953) The occurrence of lung cancer in man, *Acta Unio Internationalis Contra Cancrum* **9**:531-541

Markowitz SB, Fischer E, Fahs MC (1989) Occupational disease in New York state: a comprehensive examination: *Am J Ind Med* **16**: 417–35.

Mathers CD, Vos ET, Stevenson CE, Begg SJ (2000) The Australian burden of disease study: measuring the loss of health from diseases, injuries and risk factors. *Med J Aust* **172**:592–596

Miettinen O (1974) Proportion of disease caused or prevented by a given exposure, trait or Intervention, *Am J Epidem* **99**:325-32

Morabia A, Markowitz S, Garibaldi K, Wynder EL (1992) Lung cancer and occupation: results of a multicentre case– control study. *Br J Ind Med* **49**: 721–727

Morrell S, Kerr C, Driscoll T, Taylor R, Salkeld G, Corbett S (1998) Best estimate of the magnitude of mortality due to occupational exposure to hazardous substances. *Occup Environ Medicine* **55**:634-641

Nelson DI, Concha-Barrientos M, Driscoll T, Steenland K, Fingerhut M, Punnett L, Pruss-Ustun A, Leigh J, Corvalan C (2005) The global burden of selected occupational diseases and injury risks: methodology and summary. *American J Ind. Med* **48**: 400-418

Nurminen M, Karjalainen A (2001) Epidemiologic estimate of the proportion of fatalities related to occupational factors in Finland. *Scand J Work Environ Health* **27**:161–213

ONS, 2005. Mortality statistics. Series DH2 no. 30 Table 2.2

ONS, 2006. Cancer statistics, Registrations. Series MB1 no. 34 Table 1

OPCS (20004) Social Survey Division, *Labour Force Survey, 1984* [computer file]. *3rd Edition.* Colchester, Essex: UK Data Archive [distributor], August 2004. SN: 2143

Pannett B, Kauppinen T, Toikkanen J, Pedersen J, Young R, Kogevinas M (1998) Occupational Exposure to Carcinogens in Great Britain in 1990-93: Preliminary Results. In *Carex: International Information System on Occupational Exposure to Carcinogens.* Finnish Institute of Occupational Health, Helsinki

Peto *et al.* The major role of amosite asbestos in the British mesothelioma epidemic (Submitted for publication)

Rake C, Gilham C, Hatch J, Darnton A, Hodgson J, Peto J (2009) Occupational, domestic and environmental mesothelioma risks in the British population: a case-control study. *B J Cancer* **100**: 1175-1183

Rockhill B, Newman B, Weinberg C (1998) Use and misuse of population attributable fractions. *Am J Public Health* **88**:15-9.

Rothman KJ, Greenland S (1998) Modern Epidemiology. 2nd edition, Lippincott, Williams and Wilkins, Philadelphia, USA, page 660.

Rushton L, Hutchings S, Brown TB (2008) The burden of cancer at work; first steps to prevention: *Occup Environ Med* **65**: 789-800

Siemiatycki J, Richardson L, Straif K, Latreille B, Lakhani R, Campbell S, Rousseau M-C, Boffetta P (2004) Listing Occupational Carcinogens. *Environmental Health Perspectives* **112**: 1447 – 1460

Sorahan T, Parkes HG, Veys CA, Waterhouse JA, Straughan JK, Nutt A (1989) Mortality in the British rubber industry 1946-85. *British Journal of Industrial Medicine* **46 (1)**: 1-10

Spirtas R, Heineman EF, Bernstein L, Beebe G, Keehn R, Stark A, Harlow B, Benichou J (1994) Malignant mesothelioma: attributable risk of asbestos exposure. *Occup Environ Med* **51:**804-811

Steenland K, Burnett C, Lalich N, Ward E, Hurrell J. (2003) Dying for work: the magnitude of US mortality from selected causes of death associated with occupation. *Am J Ind Med* **43**:461– 482

Steenland K, Armstrong B (2006) An overview of methods for calculating the burden of disease due to specific risk factors. *Epidemiology* **17 (5)**: 512-519

Steenland K, Loomis D, Chy C, Simonsen N (1996). Review of occupational lung carcinogens. *American Journal of Industrial Medicine* **29**: 474-490

United Nations Scientific Committee on the Effects of Atomic Radiation (2006) Effects of ionising radiation: UNSCEAR 2006, Volume 1, Report to the General Assembly, Annex A.

Van Loon AJM, Kant IJ, Swaen GMH, Goldbohm RA, Kremer AM, Van den Brandt PA (1997) Occupational exposure to carcinogens and risk of lung cancer: results from The Netherlands cohort study. *Occupational Environmental Medicine* **54:** 817-824

Vineis P, Simonato L (1991) Proportion of lung and bladder cancers in males resulting from occupation: a systematic approach. *Arch Environ Health* **46**:6 –15

Vineis P, Thomas T, Hayes RB, Blot WJ, Mason TJ, Pickle LW, Correa P, Fontham ETH, Schoenberg J (1988) Proportion of lung cancers in males, due to occupation, in different areas of the USA. *Int. J. Cancer* **42**: 851-856

Yates D, Corrin B, Stidolph P, Browne K (1997) Malignant mesothelioma in south east England: clinico pathological experience of 272 cases. *Thorax* **52:** 507-512

Health and Safety Executive

The burden of occupational cancer in Great Britain

Methodology

This project aimed to estimate the current burden of cancer for Great Britain due to occupational exposure to carcinogenic agents/exposure circumstances classified by the International Agency for Research on Cancer as definite (Group 1) or probable (Group 2A) human carcinogens.

The measure of the burden of cancer used was the attributable fraction (AF) ie the proportion of cases that would not have occurred in the absence of exposure. For each cancer/exposure pair, risk estimates, adjusted for confounders, were obtained from key industry-based studies, meta-analyses or reviews. The period of exposure relevant to the development of the cancer in the target year 2005 was defined as the risk exposure period (REP) and was assumed to be 1956-1995 for solid tumours and 1986-2005 for haematopoietic neoplasms. National data were used to derive the proportion of the population ever exposed to each carcinogenic agent or occupation in the REP ie the ratio of the numbers ever exposed within GB over the total number of people ever employed. Adjustment was made for staff turnover, life expectancy and industry-employment trends over the REP. Exposed numbers were obtained for high and low exposures matched to appropriate risk estimates from the literature. The AFs were applied to national cancer deaths and registrations to give occupation attributable cancer numbers.

This report and the work it describes were funded by the Health and Safety Executive (HSE). Its contents, including any opinions and/or conclusions expressed, are those of the authors alone and do not necessarily reflect HSE policy.

RR927

www.hse.gov.uk