British Journal of Cancer (2011) 105, S2 – S5

www.bjcancer.com

1.

The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010

Introduction

DM Parkin[*,1]

[1]*Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, UK*

The overall objective of the study is to estimate the percentage of cancers (excluding non-melanoma skin cancer) in the UK in 2010 that were the result of exposure to 14 major lifestyle, dietary and environmental risk factors: tobacco, alcohol, four elements of diet (consumption of meat, fruit and vegetables, fibre and salt), overweight, lack of physical exercise, occupation, infections, radiation (ionising and solar), use of hormones and reproductive history (breast feeding). The number of new cases attributable to suboptimal exposure levels in the past, relative to a theoretical optimum exposure distribution, is evaluated. For most of the exposures, the attributable fraction was calculated based on the distribution of exposure prevalence (around 2000), the difference from the theoretical optimum (by age group and sex) and the relative risk per unit difference. For tobacco smoking, the method developed by Peto *et al* (1992) was used, which relies on the ratio between observed incidence of lung cancer in smokers and that in non-smokers, to calibrate the risk. This article outlines the structure of the supplement – a section for each of the 14 exposures, followed by a Summary chapter, which considers the relative contributions of each factor to the total number of cancers diagnosed in the UK in 2010 that were, in theory, avoidable.

British Journal of Cancer (2011) **105,** S2 – S5; doi:10.1038/bjc.2011.474 www.bjcancer.com

Keywords: cancer; environment; lifestyle; risk factors; UK

The purpose of this study is to estimate the fraction (or percentage) of cancers occurring in the UK in 2010 that were the result of exposure to common and, for the most part, modifiable lifestyle and environmental exposures. A total of 14 major modifiable lifestyle, dietary and environmental metabolic risks are considered (Table 1).

The analyses in the chapters that follow estimate the number of cancer cases diagnosed in the UK in 2010 that were due to such exposures in the past (or that would have been prevented if risk factor exposures had been at some hypothetical alternative optimal distribution from those actually present). The proportion (or percentage) of such avoidable cancers is known as the population-attributable fraction (PAF), which provides a quantification of the total effects of a risk factor (direct, as well as mediated through other factors).

The inputs to each analysis are as follows:

(1) The aetiological effect of risk factor exposures on cancer-specific risk.
(2) The population distribution of risk factor exposure in the past
(3) An alternative exposure distribution.
(4) The projected total number of cancer cases (by type) in the UK population in 2010.

*Correspondence: Professor DM Parkin; E-mail: d.m.parkin@qmul.ac.uk

SELECTION OF RISK FACTORS

Among dietary, lifestyle and environmental factors, those that fulfilled the following criteria were selected:

(i) There was sufficient evidence on the presence and magnitude of likely causal associations with cancer risk from high-quality epidemiological studies.
(ii) Data on risk factor exposure were available from nationally representative surveys.
(iii) There were achievable alternative exposure levels that would modify the risk.

Several other risk factors were considered but were not included because the evidence on causal effects was less convincing, or because their effects on national cancer incidence were likely to have been small and estimates of relevant past exposures difficult to obtain. This is discussed further below.

SOURCES OF DATA

(1) The risks of exposure (aetiological effect sizes) were taken from published systematic reviews and meta-analyses of epidemiological studies.
(2) Risk factor exposure distributions were obtained from nationally representative health examination and interview surveys. Data on prevalence of risk factors from epidemiological studies (cohort or case – control) were not used, as such

studies will almost never provide information relevant to the general population of the UK.

(3) The number of cancer cases in 2010 (by cancer type, sex and 5-year age group) was projected using UK incidence rates for the 15-year period from 1993 to 2007. For such a short-term projection (3 years), most established methods will provide very similar results. For all but two cancers (breast and prostate) the R-based software, 'Nordpred' (Møller *et al*, 2002), was used to project incidence rates from 2008 to 2012, on the basis of the incidence rates from 1993 to 2007, aggregated into three 5-year time periods. National population projections (2008 based) for the UK by sex, 5-year age group and year, from 2008 to 2012, were obtained from the population projections of the Office for National Statistics (Office of National Statistics (ONS), 2009). The estimate for 2010 was taken as the average annual number of cases projected for the period 2008–2012. For cancers of the prostate and female breast, a different approach was used, because recent rates have been modified to a great extent by the increased use of PSA testing and extensions to the breast cancer screening programme. An age–period cohort model based on observations for single years was fitted, but incidence rates from age groups and time periods that were assumed to have been affected by the introduction of screening were not used in the model building (Mistry *et al*, 2011).

Table 2 compares the numbers of cases diagnosed in 2007 with the projected numbers for 2010.

Table 1 Exposures considered, and theoretical optimum exposure level

Exposure	Optimum exposure level
Tobacco smoke	Nil
Alcohol consumption	Nil
Diet	
1 Deficit in intake of fruit and vegetables	$\geqslant$5 servings (400 g) per day
2 Red and preserved meat	Nil
3 Deficit in intake of dietary fibre	$\geqslant$23 g per day
4 Excess intake of salt	$\leqslant$6 g per day
Overweight and obesity	BMI $\leqslant$25 kg m^{-2}
Physical exercise	$\geqslant$30 min 5 times per week
Exogenous hormones	Nil
Infections	Nil
Radiation – ionising	Nil
Radiation – solar (UV)	As in 1903 birth cohort
Occupational exposures	Nil
Reproduction: breast feeding	Minimum of 6 months

AETIOLOGICAL EFFECTS OF RISK FACTORS ON DISEASE-SPECIFIC INCIDENCE

The relative risk (RR) per unit of exposure or for each exposure category (for risks measured in categories) was obtained for cancers with probable or convincing causal associations with each risk factor. The studies used for aetiological effect sizes were observational studies (prospective cohort studies whenever possible) that estimated the effects relative to baseline exposure. The RRs used in the analyses represent the best evidence for the impact of risk factor exposure on cancer risk in the UK population, based on the current causes and determinants of the population distribution of exposure. Relative risks adjusted for major potential confounders were used to estimate the causal components of risk factor–disease associations. With respect to diet, for example, the relative risks for specific components – for example, meat – have generally been adjusted for intake of other components with which they may be confounded, as well as for total energy intake. However, if there is also a correlation between exposure and risk of a specific cancer, due to correlations of exposure with other risks or other unobserved factors, the above equations may result in under- (when there is positive correlation) or over-estimation (negative correlation) of the true PAF when used with adjusted RRs (Bruzzi *et al*, 1985).

The cancers that occur in a particular year, related to specific risk factors, are presumably related to cumulative exposures to the factor concerned over a period of many years. For tobacco smoking, for example, the risk of lung cancer relates to the

Table 2 Numbers of cancers diagnosed in the UK in 2007 (20 most common sites) and estimates for 2010

	Males			Females		
Cancer site	**2007**	**2010 (estimate)**	**Change (%)**	**2007**	**2010 (estimate)**	**Change (%)**
Breast (female)	—	—	—	45 695	48 385	6
Lung	22 355	22 273	0	17 118	18 132	6
Colorectal cancer	21 014	22 127	5	17 594	17 787	1
Prostate	36 101	40 750	13	—	—	—
Non-Hodgkin lymphoma	5881	6297	7	5036	5305	5
Malignant melanoma	4975	6095	23	5697	6822	20
Bladder	7284	6713	−8	2807	2572	−8
Kidney	5165	5697	10	3063	3365	10
Oesophagus	5226	5713	9	2740	2819	3
Stomach	4988	4467	−10	2796	2577	−8
Pancreas	3748	4084	9	3936	4280	9
Uterus (corpus and unspecified)	—	—	—	7536	8195	9
Leukaemias	4069	4639	14	2932	3201	9
Ovary	—	—	—	6719	6820	2
Oral cavity and pharynx	4083	4571	12	2136	2359	10
Brain and CNS	2663	2799	5	2013	1902	−6
Multiple myeloma	2223	2506	13	1817	1994	10
Liver	2152	2270	5	1255	1298	3
Cervix uteri	—	—	—	2828	2691	−5
Mesothelioma[a]	1977	2077	5	424	462	9
All[b]	149 356	158 667	6	148 635	155 584	5

[a]Number of cases estimated from the UK population (2010) and rates in England in 2008. [b]Excluding non-melanoma skin cancer.

cumulative exposure to tobacco smoke (duration and dose), including the time since quitting in ex-smokers. Similarly, the total lifetime exposure to ionising radiation for individuals in each age group in 2010 was estimated on the basis of known or estimated levels of exposure in the past. Such detailed quantification of risk is not available for most exposures, and, even if it was, it would be impossible to partition the 2010 UK population according to the appropriate categories of past exposure. Therefore, for several exposures, an arbitrary latent period was included, which is the average interval between 'exposure' and the appropriate increase in risk of the cancers concerned. The most appropriate period was deemed to be the mean interval between measurement of exposure and cancer outcome in the prospective studies that were used as the source of data on relative risks. For most exposures, this was around 10 years, and thus the effects on cancers occurring in 2010 of suboptimal levels of exposure in 2000 were examined. When there was evidence about the duration between exposure and change in risk (for example, for exposure to radiation, or exogenous and endogenous sex hormones), the appropriate interval was used to select the year for which exposure data were obtained. The method used for estimating the attributable fraction of the most important exposure – tobacco smoking – does not require estimation on the basis of past exposure, and so no such assumptions are needed (although, in fact, the latency between exposure to cigarette smoking and lung cancer risk (at least) is well documented).

Many calculations of PAFs are based on *current* levels of exposure to risk factors; for example, the work of the Global Burden of Disease/Comparative Risk Assessment Group (Ezzati *et al*, 2002; Danaei *et al*, 2005) or the World Cancer Research Fund (WCRF/AICR, 2009). Although this simplifies the business of obtaining data on prevalence of the different exposures, the effect being imputed must relate to cancers that will be caused by these exposures at some variable, and undefined, period in the future.

To measure the effects of non-optimal levels of exposure, one must define, for each exposure, an optimal exposure distribution, sometimes referred to as the theoretical-minimum-risk exposure distribution (TMRED), against which the excess risk due to actual exposure is evaluated. The optimal exposure may be zero for risk factors for which zero exposure is imaginable, and results in minimum risk (e.g., no tobacco smoking, alcohol drinking or consumption of red meat). For some exposures (e.g., BMI, solar radiation, salt consumption), zero exposure is physiologically impossible. For these risks, we used optimal exposure levels corresponding to accepted recommendations for the UK population, or, for UV radiation, corresponding to those observed in a population with an attainable low level of exposure (Table 1). The 'optimum' exposure levels for factors with protective effects (physical activity, and dietary fruit and vegetable and fibre intake) were selected as the intake and activity levels recommended for the UK population (Table 1). Strictly speaking, these baselines should be called 'recommended levels', as benefits may continue to accrue at higher (for preventive exposures) or lower (for carcinogenic exposures) levels, but the terminology of 'optimum' is retained for consistency. The optimum exposure levels (TMREDs) should obviously be identical in calculations for the effect of the same exposure on different cancers.

The fraction of cancer cases considered to be attributable to a given exposure is based on estimating the effect of bringing all those individuals at suboptimal levels to the exact level of the optimum baseline, without changing (improving) the exposure (and risk) of those individuals who already exceed it. This approach is a conservative one. In other studies, for example, that of the WCRF (2009), attributable fractions are based on the estimated effect of moving all those in suboptimal exposure categories to the most favourable one (in which the *mean* exposure is considerably higher than the optimum baseline).

The analyses use data on the fraction of the UK population at different levels of exposure, and estimates of the risk associated with each, relative to the optimum exposure. The PAF is given by the following equation:

$$\frac{(p_1\times \mathrm{ERR}_1)+(p_2\times \mathrm{ERR}_2)+(p_3\times \mathrm{ERR}_3)\ \ldots\ +(p_n\times \mathrm{ERR}_n)}{1+[(p_1\times \mathrm{ERR}_1)+(p_2\times \mathrm{ERR}_2)+(p_3\times \mathrm{ERR}_3)\ \ldots\ +(p_n\times \mathrm{ERR}_n)]}$$

where p_x is the proportion of the population in exposure level x and ERR_x the excess relative risk (relative risk−1) at exposure level x.

The calculation is carried out separately by sex and age group (the choice of which depended on availability of exposure data).

The method of estimation of PAF follows the same principle for the different exposures, although some variations to the formula above are necessary depending on the type of exposure and the availability of pertinent data; they are presented in detail in each chapter. For tobacco smoking, the method developed by Peto *et al* (1992) was used, which relies on the ratio between observed incidence of lung cancer in smokers and that in non-smokers, to calibrate the risk.

Because the current (2010) cancer risk is, for most of the factors considered, related to past exposures that occur only in adulthood (age 15+), or for which data are available only for adults, PAFs can be calculated only for ages $\geqslant$25, when the latency between exposure and outcome is 10 years. Even where a fraction of cases occurring at ages $<$25 are related to childhood exposure, the effect of ignoring these on the estimate of the total PAF (at all ages) will be very small, owing to the rarity of cancer in the age group of 15–24 years.

A separate section is devoted to each lifestyle/environmental factor, for which the number of cases of different cancers attributable to suboptimal levels exposure is estimated. This is expressed also as a percentage of the observed number of cases in 2010. The total number of cancer cases (all sites) attributable to each risk factor was obtained by summing the numbers at the individual sites. Cases of different cancers attributable to a single risk factor are additive because each cancer case is assigned to a single ICD category.

In a summary chapter, the estimates for the 14 different exposures are listed together, and the numbers of cancer cases caused by all of them functioning individually, or in combination, are estimated.

See acknowledgements on page Si.

Conflict of interest

The author declares no conflict of interest.

REFERENCES

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

Danaei G, Vander Hoorn S, Lopez AD, Murray CJ, Ezzati M (2005) Causes of cancer in the world: comparative risk assessment of nine behavioural and environmental risk factors. *Lancet* **366:** 1784–1793

Ezzati M, Lopez AD, Rodgers A, Vander Hoorn S, Murray CJ (2002) Selected major risk factors and global and regional burden of disease. *Lancet* **360:** 1347–1360

Mistry M, Parkin DM, Ahmad AS, Sasieni P (2011) Cancer incidence in the United Kingdom: projections to the year 2030. *Br J Cancer* **105:** 1795–1803

Møller B, Fekjaer H, Hakulinen T, Tryggvadottir L, Storm HH, Talback M, Haldorsen T (2002) Prediction of cancer incidence in the Nordic countries up to the year 2020. *Eur J Cancer Prev* **11**(Suppl 1)**:** S1 – S96

Office of National Statistics (ONS) (2009) 2008-based National population projections. http://www.statistics.gov.uk/downloads/theme_population/NPP2008/NatPopProj2008.pdf

Peto R, Lopez AD, Boreham J, Thun M (1992) Mortality from tobacco in developed countries: indirect estimation from national vital statistics. *Lancet* **339:** 1268 – 1278

World Cancer Research Fund (WCRF)/American Institute for Cancer Research (AICR) (2009) *Policy and Action for Cancer Prevention. Food, Nutrition and Physical Activity: a Global Perspective.* AICR: Washington, DC