Paper

Modelling predictions of cancer deaths in Northern Ireland

D French,[1] D Catney,[2] AT Gavin [2]

ABSTRACT

Background: An ageing population has service planners concerned about future levels of disease which are age dependent. Predictions of mortality for colorectal, lung and breast cancers, which account for 30% of cancer cases and 40% of cancers deaths, were calculated for 2010 and 2015, based on trends in death rates and the predicted change in the demography of the Northern Ireland population.

Methods: The US National Cancer Institute's "Joinpoint" program was used to check for structural breaks in the time series of cancer death rates from 1984 to 2004. The prediction models applied to the data allowed variations in trends across age groups to be taken into account. A linear model was used for increasing or constant trends and a log linear model was used where the trend was decreasing. The models assume the number of deaths in each stratum, defined by age-sex and time-period, is Poisson distributed, with the average value determined by a log or linear function.

Results: Recent trends in rates of cancers studied were downwards except for female lung. Predictions include decreased colorectal cancer deaths in females and lung cancer deaths in males. In females, lung cancer deaths are predicted to more than double by the year 2015 (473 deaths), based on the 1984 level. Colorectal death rates in males are predicted to drop, but the number of deaths will increase by more than 10%, due to demographic change. Numbers of breast cancer deaths are likely to rise slightly, despite falling age standardised death rates, due to an ageing population.

Conclusions: This work has provided estimates of early future trends, useful to service planners, and highlights the need for tobacco control, to reduce numbers of lung cancer deaths in females. The recently announced control of environmental tobacco legislation is one welcome development which should reduce lung cancer mortality in Northern Ireland.

INTRODUCTION

In Northern Ireland cancers of the colorectum, lung and breast account for 30% of all cancer cases and about 40% of cancers deaths.[1] The diagnosis, treatment and care of patients with these diseases use significant resources, so any change in numbers is particularly relevant to service planners. Projections can also identify need for health promotion initiatives to reduce risk in a population.

The predicted changes to the population structure, of increases in the proportion of older people,[2] will result in a rise in the incidence of cancer, a disease more common in older people. Disease risk factors are the other major driver for change.

In Northern Ireland, accurate cancer incidence data is available only since 1993; death information is available for a longer period. For lung cancer, the greatest cause of cancer mortality, numbers of deaths are close to incidence levels, due to poor survival rates. Deaths therefore form the basis of this analysis, which aims to identify for planners, future demands on services and highlight areas for disease prevention.

METHODOLOGY

Annual age specific death rates for the period 1984-2004 were calculated using official mortality figures from the General Register Office for Northern Ireland [3] and mid-year population estimates, provided by the Northern Ireland Statistics and Research Agency. The age groups used were 0-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84 and 85+.

Linear and log-linear regression models have been found to be the most practical means by which relatively short-term future patterns of cancer mortality can be estimated.[4] The use of classical age-period-cohort modelling techniques may well improve the "fit" of the model (albeit at the expense of extra degrees of freedom), but the random variation associated with parameter estimates can lead to erratic projections.[5]

In this study, the identification of the most recent trends was assisted through inflexion point regression analysis of the data, using the US National Cancer Institute's (NCI) "Joinpoint" program (Version 3.0),[6] a Windows-based statistical package used to analyse modelled data where several trend-lines are connected together at "joinpoints". The software takes trend

[1] Project Support Analysis Branch, Annexe 2, Castle Buildings, Stormont, Belfast BT4 3UD.

[2] Northern Ireland Cancer Registry, Mulhouse Building, Grosvenor Road, Belfast BT12 6BJ.

Correspondence to Dr Gavin.

Email: a.gavin@qub.ac.uk

data (in this case age-standardised cancer mortality rates) and fits the simplest inflexion model that the data allows. The programme starts with the minimum number of joinpoints (e.g. 0 joinpoints, which is a straight line) and tests whether more joinpoints are statistically significant and must be added to the model, thus enabling the user to test whether an apparent change in trend is statistically significant.[7]

Existing incidence-based modelling tools developed by Dyba and Hakulinen from the Finnish Cancer Registry,[8,9] were used to predict cancer mortality, on the principle that the underlying theoretical assumptions apply equally to incidence and mortality data. These assumptions are:

1. Future cancer death trends can be modelled by extrapolating a historic trend.

2. The lengths of the data time-series permit the estimation of models that take account of age-sex group specific trends.

3. The numbers of deaths in each age-sex-timeperiod stratum are Poisson distributed.

4. Where historic trends in standardized deaths are decreasing, a log–linear model is appropriate to estimate the average rate, whereas a linear model is used for increasing or constant trends to avoid explosive growth.

Let $M_{it} = \frac{c_{it}}{n_{it}}$ be the mortality rate for age group i in period t, given n_{it}, the number of persons and c_{it} the number of deaths. Let α_i, β_i be the model parameters for age group i.

A description of the models applied in various scenarios is given below (*Table I*).

Model 4 is an alternative to model 3 and may be used where model 3 gives large confidence intervals for predictions or unrealistic trends in some age groups. It proves useful for the case, common in practice, where larger baseline death rates give larger trends over time and this functional form explicitly keeps the ratio between baseline rate and trend constant across age groups.

The 95% predicted confidence intervals generated in the output analysis are used to account not only for the uncertainty in the historical data but also for the randomness in the numbers of deaths themselves. Consequently the "95% prediction intervals" reported here may seem quite large.

Estimates of the parameters for each of the models were calculated using the STATA 8.0 Statistical package for Windows.[10]

The Pearson goodness-of-fit test statistic was used to test the adequacy of the model and is given by:

$$\chi^2 = \sum_{i=1}^{N}\sum_{t=1}^{T} \frac{M_{it} - \hat{M}_{it}}{\sqrt{\hat{M}_{it}}}$$

where N is the number of age-groups and T the length of the time series. This statistic indicates the fit of the data to a Poisson distribution.

The 'deviance' is defined to be twice the difference between the maximum achievable log likelihood and the log likelihood at the maximum likelihood estimates of the regression parameters. It forms a useful basis for choosing between models when two models give a reasonable fit. A log-likelihood ratio test of the form $D = D_a - D_b$, where D_a, D_b are the deviances of fit of models a and b respectively, is used to choose between models. D follows a chi-square distribution, with degrees of freedom equal to the difference in degrees of freedom between models a and b, when both models describe the data well.

As the number of parameters increases, the model will inevitably improve. To enable selection of the most parsimonious model, the Akaike Information Criterion (AIC), which penalises the addition of extra parameters was used. Generally, where the log-likelihood test indicated either of two models was acceptable, the more parsimonious model was chosen.

Once the best model was chosen, 2004-based population projections based on assumptions regarding fertility, mortality and net migration, for N. Ireland[2] were applied to predict cancer deaths by gender and site (error range in brackets).

RESULTS

Joinpoint analysis of mortality rates (1984-2004) revealed significant downward trends in: colorectal cancer in males and females; lung cancer in males and breast cancer in females (*Fig 1*). A significant upward trend was demonstrated for female lung cancer. No change in trend was detected over the period (1984-2004) for any site, i.e. a joinpoint model with zero joins (JP_0) was shown to best represent each time series.

TABLE I

Outline of Models

Data Trend	Decreasing	Decreasing	Increasing	Increasing	
Significant variation in trend across age groups	No	Yes	Yes	Yes	
Most appropriate model	$\ln(M_{it}) = \alpha_i + \beta_i t$ Model 1	$\ln(M_{it}) = \alpha_i + \beta_i t$ Model 2	$M_{it} = \alpha_i + \beta_i t$ Model 3	$M_{it} = \alpha_i + \beta_i t$ where $\frac{\alpha_i}{\beta_i} = \frac{\alpha_j}{\beta_j}$ Model 4	for any two age-groups i, j

Fig 1A. Trends in Age-Standardised Mortality Rates* (3-year moving averages)
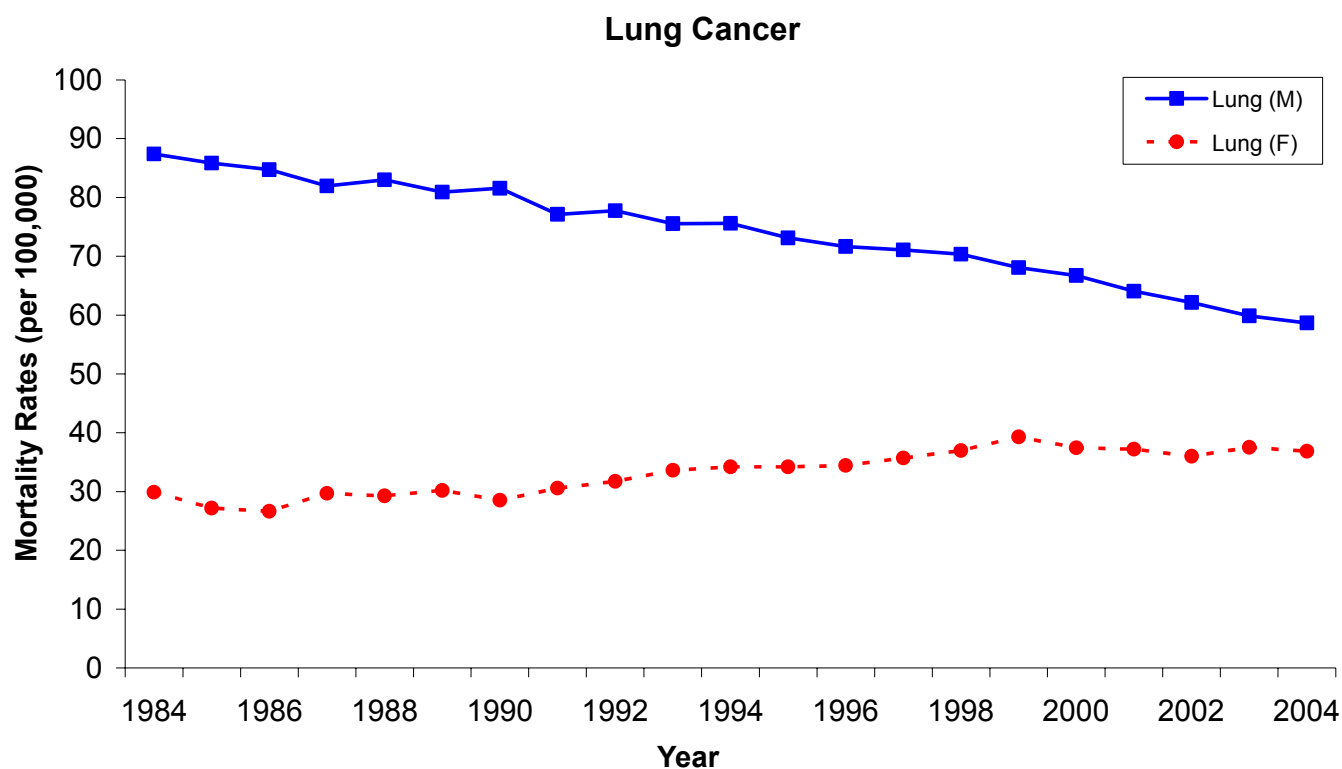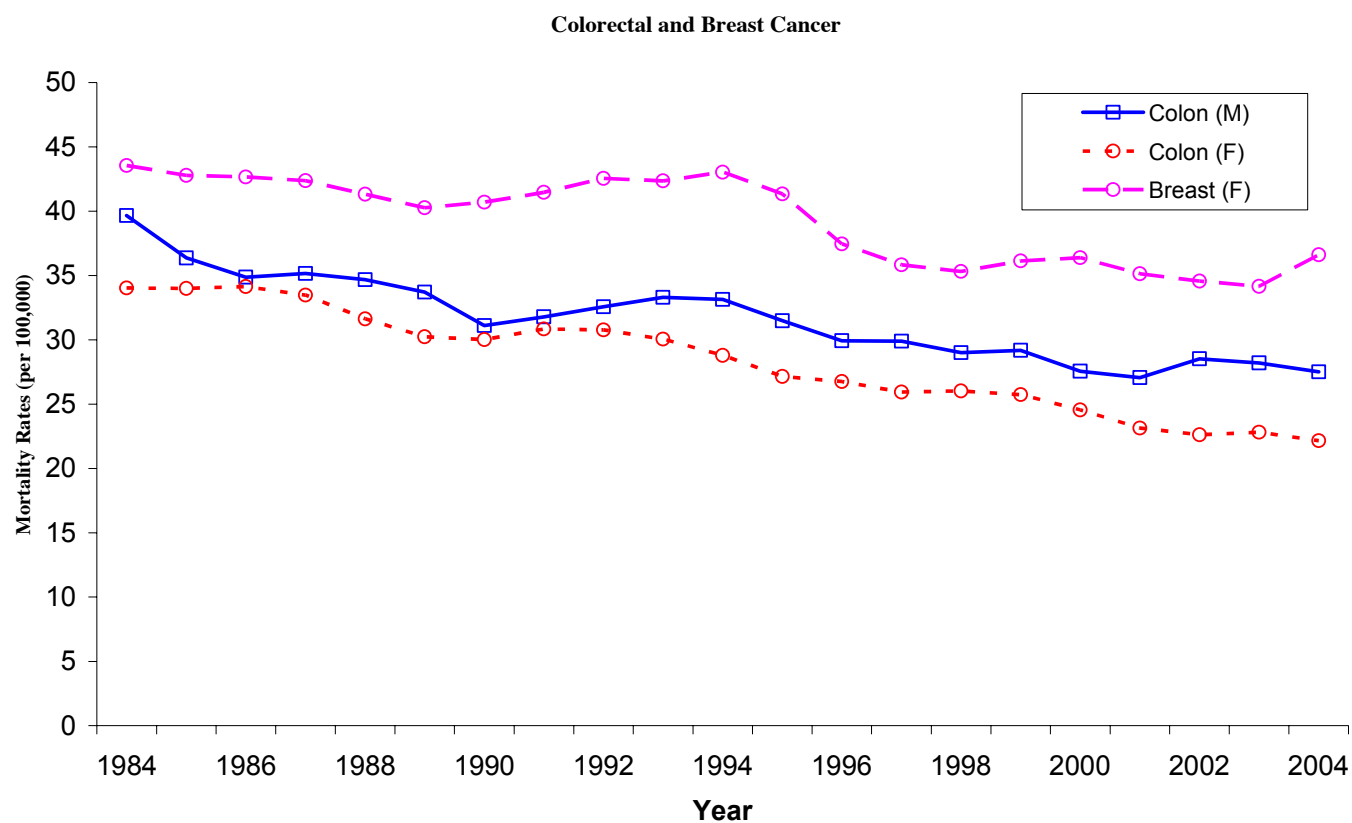*rates standardised to N. Ireland population projection 2005

Lung Cancer

Fig 1B.

Colorectal and Breast Cancer

TABLE II

Diagnostic tests by model and cancer site

Site	Model	Pearsons's χ^2	d.f.	Prob
Colorectal (male)	Log - single trend (1)	201	199	0.448
Lung (Male)	Log - multiple trend (2)	192	190	0.438
Colorectal (Female)	Log - single trend (1)	193	199	0.602
Lung (Female)	Linear - multiple trend (4)	265	190	0.000
Breast	Log - multiple trend (2)	191	190	0.464

Variations in mortality rate trends across age-groups were also observed for some sites. Since the female lung cancer mortality series (1984-2004) was the only one to demonstrate a significant increasing trend, further scrutiny of age-specific rates for this site was of particular interest, to determine which age-groups were driving this increase (*see Fig 2*).

Joinpoint analysis of age-specific mortality rates for female lung cancer revealed no significant trend among females aged 0-64yrs (combined). Significant upward trends were observed within each of the other age-groups, but most notably among females aged 75-79 years [single continuous increase detected (1987-2004)] and 85+years [increases detected (1984-1988), (1992-2001) & (2001-2004)].

According to the principles outlined in the methods, the appropriate model was selected for each site. Table II details the models selected. All models provide a reasonable fit to the data, except for female lung cancer, where the model selected is very poor.

Colorectal cancer death rates are predicted to fall in males and females, but numbers in males may rise due to demographic pressures (*see Table III*). This reflects the more marked historic decreasing trend in females, counterbalancing the pressure due to an ageing population.

Lung cancer deaths fell in males but increased in females between 1984 and 2004. The prediction follows on the trend for males so that despite demographic change, deaths from lung cancer in males do not rise. In females, however death rates (2015) are predicted to more than double, based on the 1985 level, due to a rising rate of disease and an ageing population.

Breast cancer death rates fell between 1984 and 2004, with a more rapid decline noticed after 1997. The model predicts slightly increasing deaths from breast cancer – despite decreasing age standardized death rates. This is driven by the ageing population.

DISCUSSION

Trends based on incidence are preferable to those based on deaths, as death trends are affected by changes in survival as well as changes in incidence. Also, deaths data are less rigorously checked compared to cancer incidence data. The accuracy of trend predictions however depends on the length of the time series examined and hence deaths are preferred in this instance.

Predictions based on historical data do not take into account future improvements in treatment or changes in the way health services are organised, but only extrapolate mortality patterns from a range of observations. These predictions also assume that current trends will remain unchanged in the future. As the predictions calculated here are only for up to ten years, such effects are likely to be minimal. However, the recent fall in breast cancer mortality, with documented significant survival improvements in the population studied[1] may not have been given enough weight in the model used. Deaths from breast cancer are falling rapidly due to earlier diagnosis and advances in treatment and the predicted increase in deaths may not materialise.

Generally, the models provide a very good fit to the data. However, the model of female lung cancer mortality trend is inadequate. Both male and female lung cancer predictions could be improved upon by the inclusion of information on smoking, as about 90% of lung cancers are directly attributed to tobacco, with very low levels of background disease.[11]

The biggest determinant of the level of cancer in the short and long term is the pressure of an ageing population, as cancer is a disease more common in older people. While age standardised rates may decrease, due to public health initiatives, the absolute number of cases diagnosed and requiring treatment is likely to rise. Service planners must take this into consideration now, to ensure adequate service provision in the future. As tobacco is causally implicated in around a third of all cancers,[11] significant efforts made now, in tobacco control, could counterbalance the pressures due to demographic change and halt the predicted rise in cancers.

Trends in lung cancer mortality can be affected by smoking patterns 20-30 years previously, but the cessation of smoking, even well into middle age, can avoid most of the subsequent risk of lung cancer; stopping before middle age avoids more than 90% of the risk attributable to tobacco.[11] If current smokers can succeed in giving up the habit, mortality from lung cancer in the near future could be substantially reduced.[11] The recently announced environmental tobacco control legislation for N. Ireland is a welcome initiative which should reduce lung cancer mortality.

It is planned to introduce colorectal cancer screening in N. Ireland within the next ten years. This should further reduce deaths from colorectal cancer below that predicted by the model. Service planners should note that disease levels will

TABLE III

Numbers [1] (n) and rates [2] (r) (per 100,000) of cancer deaths by gender and site (error range in brackets)

	Year	Males				Females					
		Colorectal		Lung		Colorectal		Lung		Breast	
		N	r	N	r	n	r	N	r	n	r
A C T U A L	1985	222	36	572	89	237	33	181	24	307	44
	1990	214	32	524	78	228	30	242	31	295	40
	1995	236	33	507	71	219	28	272	34	327	42
	2000	205	27	494	65	221	27	346	42	286	35
P R E D I C T E D	2010	237	25	525	55	187	20	404	43	313	33
		(201, 273)	(21, 29)	(472, 578)	(50, 61)	(156, 218)	(16, 23)	(356, 453)	(38, 48)	(271, 354)	(29, 37)
	2015	251	23	554	51	183	18	473	46	328	32
		(210, 291)	(19, 27)	(493, 615)	(45, 56)	(150, 215)	(14, 21)	(415, 530)	(40, 52)	(281, 376)	(27, 36)

[1] predicted deaths calculated by applying predicted age-specific rates to predicted populations

[2] rates standardised to N. Ireland population projection 2005

Fig 2. Trends in Age-Specific Mortality Rates for Female Lung Cancer (3-year moving averages)
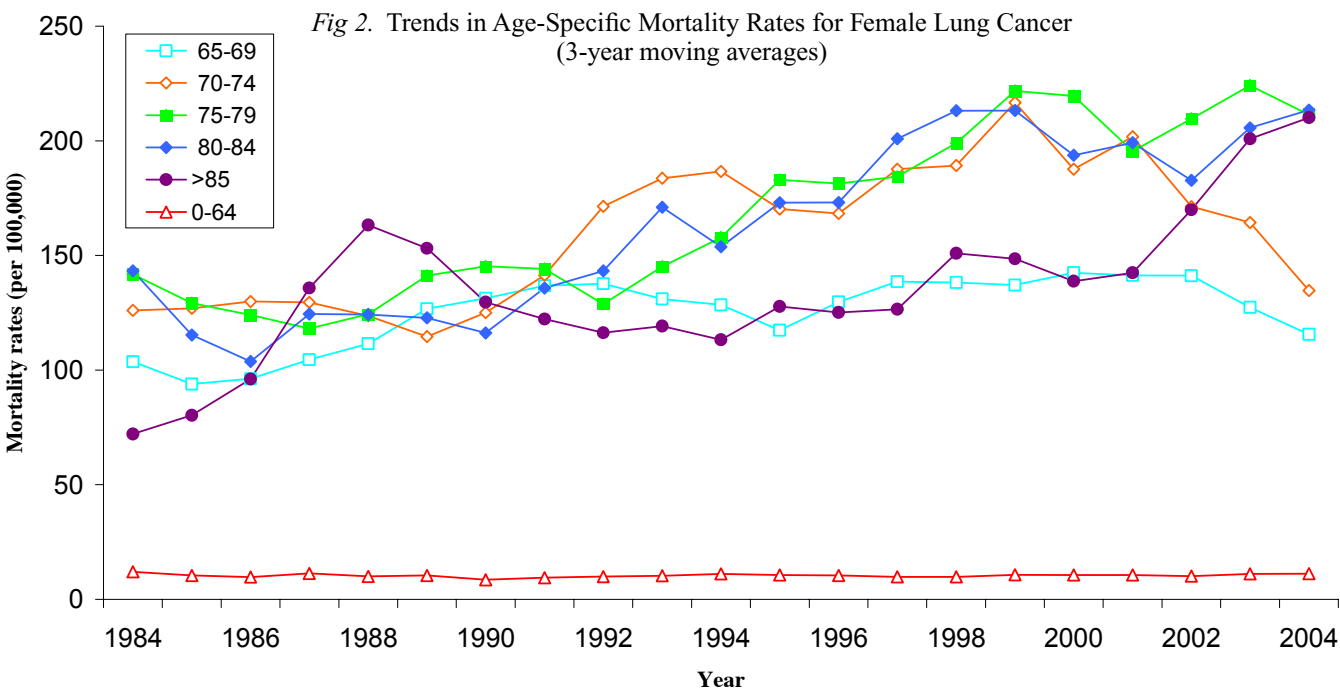
increase due to enhanced detection during the early prevalence round of screening.

Exercise has been shown to reduce colorectal cancer,[12] while obesity in postmenopausal females increases the risk of breast cancer.[13] While changes in these lifestyle factors are unlikely to affect deaths in the short timescale of this study, they will have serious implications in the longer term. Initiatives to increase exercise, increase fruit/vegetable consumption and tackle obesity need to be put in place now.

Possible advances in treatment within the next ten years have not been taken into account. Also specialisation of cancer services which has been shown to reduce mortality [14-16] is ongoing in N. Ireland, since the publication of the Campbell Report.[17]

This work has several limitations but it has allowed service planners to explore the issues based on predictions in a developing and costly area of care. It has also highlighted the need to target tobacco control, especially among females.

The authors have no conflict of interest.

REFERENCE

1. Fitzpatrick D, Gavin A, Middleton R, Catney D. Cancer in Northern Ireland 1993-2001: a comprehensive report. Belfast: Northern Ireland Cancer Registry; 2004.

2. Government's Actuary Department. Population projections by the Government Actuary. United Kingdom. Available from: http://www.gad.gov.uk/Populationl2004/england/weng04mort.xls

3. Great Britain, Department of Health. The Annual Report of the Registrar General. 2003 Available from: http://www.nisra.gov.uk/registrar/rg2003index. Accessed March 2006.

4. Hakulinen T. The future cancer burden as a study subject. *Acta Oncol* 1996; **35(6):** 665-70.

5. Dyba T, Hakulinen, T. Comparison of different approaches to incidence prediction based on simple interpolation techniques. *Stat Med* 2000; **19(13):** 1741-52.

6. U.S. Institute of Health. National Cancer Institute. Cancer control and population sciences. Statistical research and applications branch. Joinpoint regression program: latest release: Version 3.0 (April 2005). Available from: http://srab.cancer.gov/joinpoint. Accessed March 2006.

7. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. *Stat Med* 2000; **19(3);** 335-51.

8. Hakulinen T, Dyba T. Precision of incidence predictions based on Poisson distributed observations, *Stat Med* 1994;**13(15):** 1513-23.

9. Dyba T, Hakulinen T, Paivarinta A. A simple non-linear model in incidence prediction. *Stat Med* 1997;**16(20):** 2297-309.

10. Dyba T, Hakulinen T. European Network of Cancer Registries: Finnish Cancer Registry. Stata (Version 8.0) macros for short-term predictions. Available from: http://www.encr.com.fr/stata-macros1.htm. Accessed March 2006.

11. Cancer Research UK. Lifestyle and cancer risk in the UK population: Cancerstats report. Available from: http://info.cancerresearchuk.org/cancerstats/causes/lifestyle/. Accessed March 2006.

12. Colditz GA, Cannuscio CC, Frazier AI. Physical activity and reduced risk of colon cancer: implications for prevention. *Cancer Causes Control* 1997; **8(4):** 649-67.

13. Vairiio H, Bianchini F, editors. Weight Control and Physical Activity. IARC Handbooks of Cancer Prevention: Vol. 6. Lyon: International Association for Research on Cancer; 2002.

14. Stiller CA, Centralised treatment, entry to trials and survival. *Br J Cancer* 1994; **70(2):** 352-62.

15. Sainsbury R, Haward B, Rider L, Johnston C, Round C. Influence of clinicians workload and patterns of treatment on survival from breast cancer. *Lancet* 1995; **345(8950):** 1265-70.

16. Gillis CR, Hole DJ. Survival outcome of care by specialist surgeons in breast cancer: a study of 3786 patients in the West of Scotland. *BMJ* 1996; **312(7024):** 145-8.

17. Campbell H, Chair of the Working Group. Cancer Services: investing for the future. Belfast: Department of Health and Social Services; 1996.