

■ Data and methods

Data sources and quality

Cancer registrations

The principal source of data on cancer registrations is the New Zealand Cancer Registry (NZCR), which is administered as part of the National Minimum Data Set (NMDS) by the New Zealand Health Information Service (NZHIS). The NZCR is a population-based register of all primary malignant cancers (other than non-melanotic skin cancers) diagnosed in New Zealand. The registry has a long history, dating back to 1948. Data from 1950 to 1998 were made available for this study in an electronic version by the NZHIS.

The data collected were unit records for individuals, including data on the site, stage and pathology of the cancer, as well as personal information on age, gender, ethnicity and domicile of residence. Throughout the 1980s and early 1990s there was considerable under-reporting to the registry, particularly of melanoma. There were also problems with the correct reporting of cancer histology, though information collected on cancer site, age and gender was considered reliable. With the passing of the Cancer Registry Act 1993 and the Cancer Registry Regulations 1994, which had the effect of making laboratories responsible for cancer registration, the quality and completeness of the database have increased significantly. This is likely to have resulted in artefactual increases in recent incidence that will translate into projections that are upwardly biased, especially for melanoma (which was most frequently under-reported prior to 1994).

Currently the NZCR uses a variety of sources to gather data. The major sources of new registrations include laboratory reports, post-discharge reports from publicly funded hospitals, death certificates and autopsy reports, and discharge reports from private hospitals. Coverage is now almost complete (Jim Fraser, personal communication, September 2002). The NZHIS produces an annual report *Cancer: New Registrations and Deaths*, which summarises some historical trends for the more common cancers, as well as giving rates and numbers of deaths for each of the 3-digit ICD codes for cancer by five year age group in the total population and among Māori.

Cancer mortality

Unit record mortality data for the period 1970 to 1999 were made available for this study by the NZHIS. Relevant data collected in the NMDS and extracted for the purpose of this report include year of registration, ICD code, date of birth, date of death, gender, ethnicity and domicile of residence.

The specified underlying cause of death is based on information from a range of sources, including death certificates, coroners' reports and autopsy reports. Additional information may be sought; for example, from the NZCR, public hospital discharge records and general practitioners. The quality of the mortality data set is considered to be satisfactory (Jim Fraser, personal communication, September 2002).

Mortality data for the years prior to 1970 could have been obtained from the WHO (these data are no longer available in a usable electronic form from the NZHIS). However, the WHO data set was only available with the cancer sites aggregated differently to the classification used in this report, and so could have been used for only a subset of the selected cancer sites. Accordingly, it was decided not to use the mortality data for earlier years in this study. Pre-1970 mortality data are probably of little relevance to post-2000 projections in any case.

Population data

Mid-year population estimates between 1950 and 2001 and projections based on the 1996 Census (series 4, assuming 'medium' mortality, fertility and net migration) for the years 2002 to 2016 were obtained from Statistics New Zealand (SNZ). The population data obtained are finely calibrated by single calendar year, single year of age, and gender.

Population estimates prior to 1990 were prepared by SNZ under the de facto population concept, which includes all New Zealand residents in the country plus all people from overseas. Population estimates from 1991 onwards were prepared under a new, resident population concept, which includes all New Zealand residents (adjusted for census undercount) both in the country as well as those temporarily overseas, but excludes all visitors from overseas. All population projections produced since 1996 are based on this population concept.

While the change from de facto to resident population concepts (including adjustments for census undercount) represents an interruption in the population time series, the net effect was estimated to be small (Statistics New Zealand 1999). For example, the excess of resident population over the corresponding de facto estimates varied little in the first half of the 1990s, at about 2.1% of the resident population. The effect on time series cancer rates is insubstantial (equivalent to discounting rates prior to 1990 by 2%, assuming a uniform effect across all age and gender groups) compared to other sources of uncertainty in the data (eg cancer registry coverage in the earlier years).

Additional data were drawn from the 1996 Census for the ethnic and deprivation analyses, to estimate 'sole Māori' and 'total Māori' (Māori ethnic group) populations as well as NZDep96 quintiles (NZDep96 is a small area measure of deprivation based on the 1996 Census (Salmond et al 1998)). For this report NZDep96 scores have been grouped into quintiles from quintile 1, the least deprived fifth of small areas, to quintile 5, the most deprived fifth).

Both 'sole Māori' and 'total Māori' numerators have been corrected for over- or undercounting, using adjustors derived from the New Zealand Census-Mortality Study for 1996–99¹ (Ajwani et al 2002).

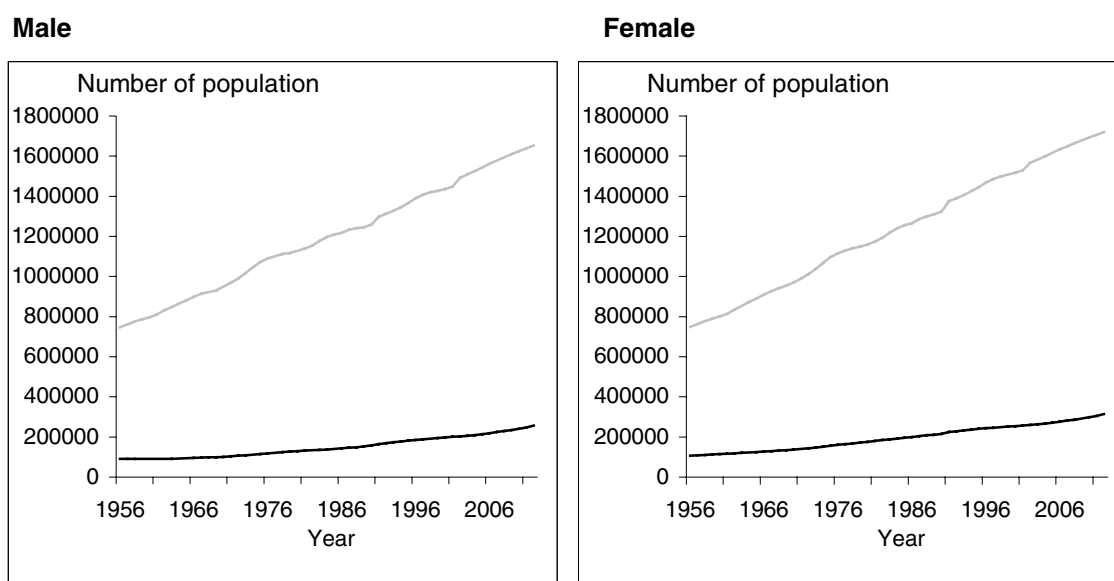
¹ Average adjustors were 0.80 (ie 20% reduction) for 'sole Māori' and 1.07 (ie 7% increment) for 'total Māori'.

Population changes and the demographic driver

As will be shown in each of the chapters in Part II, population growth and structural ageing tend to be the dominant forces driving change in cancer registration and death counts, sometimes overwhelming the effect of changes in cancer risk. A brief overview of demographic changes over the study period is therefore given below.

Figure 3.1 illustrates the steady increase in the adult population (15 years and above) and in the older population (65 years and above) observed and projected over the period 1956 to 2012.

Figure 3.1 Historical and projected New Zealand population, all adults, and older people, 1956 to 2012



The adult population of New Zealand underwent substantial growth over the second half of the 20th century, from around 750,000 males and females in the mid 1950s to 1.4 million males and 1.5 million females in the late 1990s. This steady growth in the adult population is projected to continue through the forecasting period, reaching over 1.6 million males and over 1.7 million females by the early 2010s. Over the entire period (1956 to 2012) the adult population grew and is projected to grow at a steady annual rate of 2.2% among males and 2.3% among females.

The number of older people has also increased sharply, from around 90,000 males and 110,000 females in the mid 1950s to 190,000 and 250,000, respectively, in the late 1990s, and is projected to increase further to 260,000 and 310,000, respectively, by the early 2010s. Over the entire period the older population grew and is projected to grow at an average annual rate of 3.3% among males and 3.5% among females.

To illustrate the ageing effect, the older population as a percentage of the total adult population increased from 12% for males and 14% for females in the mid 1950s to 13%

and 17%, respectively, in the late 1990s, and is projected to increase further to 15% and 18%, respectively, by the early 2010s.

Age standardisation

Age standardisation is a conventional technique used to summarise age specific rates into a summary index, which is not affected by differing population size and age structure, so permitting meaningful comparison between populations and over time.

In this report all age standardised rates refer to rates for adults (15 years and above), calculated using the direct standardisation method. The recently introduced WHO world population (WHO 2000b) is used as the standard population, in preference to the more commonly used Segi population, since it more closely resembles the New Zealand population as at the 2001 Census. A comparison of the WHO world, Segi and New Zealand 2001 Censal populations is provided in Table 3.1.

Table 3.1 WHO world population, Segi's population and the New Zealand 2001 Censal population

Age group (years)	WHO world population		Comparisons		
	Weights	%	Segi's population (%)	New Zealand	
				Total (%)	Māori (%)
15–19	8.47	11	13	9	15
20–24	8.22	11	12	8	13
25–29	7.93	11	12	9	12
30–34	7.61	10	9	10	12
35–39	7.15	10	9	10	12
40–44	6.59	9	9	10	10
45–49	6.04	8	9	9	8
50–54	5.37	7	7	8	6
55–59	4.55	6	6	6	4
60–64	3.72	5	6	5	4
65–69	2.96	4	4	4	2
70–74	2.21	3	3	4	2
75–79	1.52	2	1	3	1
80–84	0.91	1	1	2	0
85+	0.63	1	1	2	0
Total (15+)	73.88	100	100	100	100

Cancer site selection

The selection of cancer sites for this study was guided by a thorough review of similar exercises done locally and overseas (Australian Institute of Health and Welfare 2001a, Black and Stockton 2001, Cox 1995, Department of Health 2000), and after consultation with the study's Expert Advisory Group.

Excluding 'childhood cancer' and 'all adult cancer', a total of 26 specific cancer sites were selected for modelling and forecasting: 22 among males and 24 among females (Table 3.2).

This selection of cancer sites is by no means definitive. Some of these sites include two or more distinct cancer types with different epidemiologies and implications for prevention and treatment. Examples include colorectal cancer, which comprises both colon and rectal cancers; cancer of the oesophagus, which comprises both squamous cancer of the upper and middle third and adenocarcinoma of the lower third; and cervical cancer, which comprises both squamous and adenocarcinomas (only the former being consistently detectable through screening). However, the ability to model different cancer types within a site was limited by the available data. Future editions of this report may be able to differentiate distinct cancer types with greater precision.

While cancer is rare among youth (15–24 years), this age group was included in modelling some sites where cases do occur among young people. To ensure comparability across all cancer sites, all age standardised rates reported for adult cancers were calculated for ages 15 years and above. For those sites where the age range used for the projections was 25 years and above, it was assumed that there were no cases in the 15–24 age group.

Data processing

Unit record data (on individuals) were extracted by specific cancer site according to the ICD9 code, as shown in Table 3.2, and were then aggregated into quinary quinquennia. The population data were aggregated into matching quinary quinquennia.

A number of filtering steps were applied to the aggregated data set to exclude anomalies in the data. These steps are shown in Table 3.3.

Since prostate cancer registrations account for over 10% of all cancer registrations among males, an adjustment had to be made to the male 'all adult cancer' registration model to exclude the 'PSA effect' (see Table 3.3). This was achieved by substituting the projected number of prostate cancer registrations for the observed number for the period 1994–98.

In addition to these data filtration steps, an adjustment was made to the pre-1994 melanoma registration data to compensate for the serious under-reporting of this cancer in earlier years (see Chapter 24). Many other cancers (eg breast cancer) were also affected by under-reporting in earlier years, but to a much lesser extent than was the case for melanoma. Adjustment of the registration data was not considered necessary for these cancers, though their projected incidence may be slightly over-estimated as a result, while the corresponding mortality projections are likely to be more reliable.

Table 3.2 Selected cancer sites, associated ICD9 codes² and relevant age ranges

	Cancer sites	ICD9 code	Age range (years)	Comments
1	All childhood cancer	140–208	0–14	All registrations or deaths among children
2	All adult cancer	140–208	≥15	All registrations or deaths among adults
3	Bladder	188	≥25	
4	Bone and connective tissue	170–171	≥15	
5	Brain	191	≥15	
6	Breast	174	≥25	Female breast cancer only
7	Cervix	180	≥25	Females
8	Colorectal	153–154	≥25	
9	Endometrium (uterus)	182	≥25	Females
10	Gallbladder	156	≥25	
11	Hodgkin's disease	201	≥15	
12	Kidney	189	≥25	
13	Larynx	161	≥25	
14	Leukaemia	204–208	≥15	
15	Lip, mouth and pharynx	140–149	≥25	
16	Liver	155	≥25	
17	Lung	162	≥25	
18	Melanoma	172	≥15	
19	Myeloma	203	≥25	
20	Non-Hodgkin's lymphoma	200–202	≥15	
21	Oesophagus	150	≥25	
22	Ovary	183	≥15	Females
23	Pancreas	157	≥25	
24	Prostate	185	≥25	Males
25	Stomach	151	≥25	
26	Testicular	186	≥15	Males
27	Thyroid	193	≥15	
28	Adult cancer of other sites	140–208	≥15	Not included elsewhere

² Some of the historical data were originally coded to ICD6, 7 or 8. Data for those years have been recoded by NZHIS to ICD9 (the original codes are not included in the electronic copy supplied by NZHIS).

Table 3.3 Data filtration, by cancer site

	Data filtered out	Reasons / comments
1	Cervical cancer stage 0 registrations	Detection of these precancerous and non-invasive lesions has been markedly affected by screening.
2	Gallbladder cancer registrations prior to 1969	There were no registrations in the years to 1963 and unrealistically few cases registered in the five year period 1964–68 (an age standardised rate of less than one-quarter of the subsequent years), which is believed to reflect the coding conventions applied at the time.
3	Laryngeal cancer registrations prior to 1964	Changes in coding practice associated with ICD7 are believed to be responsible for the sudden drop in registrations in the five year period 1959–63, particularly among females (the age standardised rate was around one-third the level of adjacent five year periods).
4	Lip, mouth and pharynx cancer registrations prior to 1959	For reasons not immediately clear, unexpectedly high numbers of registrations were recorded in the five year period 1954–58, particularly among males (at double the age standardised rate of subsequent years).
5	Liver cancer registrations prior to 1969	Clear discontinuity exists between pre- and post-1969 registration data, for reasons that are not clear. The age standardised rate dropped sharply between the five year periods 1964–68 and 1969–73, by one-quarter among males and half among females. After 1969 the incidence rate increased steadily.
6	Prostate cancer registrations between 1994 and 1998	Registrations in the most recent five year period 1994–98 were inflated by widespread use of the PSA test. This ‘PSA effect’ is expected to diminish over the forecasting period, so registrations after 1994 were excluded (and were replaced by the modelled estimates based on data up to 1993).
7	Thyroid cancer registrations prior to 1969	Unrealistically low (or zero) numbers of cases were recorded in the earlier years, but the number increased rapidly after 1968 (eg the age standardised rate among males increased by as much as 15 times between two five year periods).

Modelling and forecasting methods

This section gives a brief description of the methods used to fit statistical models to the cancer incidence and mortality data (for each site and gender), and to obtain projections for each five year period and five year age group from the fitted models. Additional methods and analyses used for certain sites are summarised at the end of the section. A separate detailed account of the modelling and forecasting methods used is also available on request from the Ministry of Health.

Modelling approach

Cancer incidence and mortality rates can be thought of as realisations of three time dimensions: age (at diagnosis or death), period (calendar year of diagnosis or death), and cohort (year of birth). Given a sufficiently long historical time series of cancer incidence or mortality data, statistical models can be constructed to project these rates based on the historical pattern of age, period and cohort effects.

The approach adopted was to fit such models to the available data (typically from 1954 for registrations and 1970 for deaths), using five year age groups and five year periods (quinary quinquennia). The use of quinary quinquennia gives 10 year overlapping cohorts.³ Typically, the models involved observed data for nine five year periods for incidence and six five year periods for mortality, 13 or 15 age groups (from 25–29 years, or 15–19 years, respectively), and 21 or 23 (incidence) and 18 or 20 (mortality) birth cohorts.

The use of quinary quinquennia is common in the literature. The main advantages are that firstly, population projections are usually made available by five year age group, and to use any other age groups would require interpolation, which would introduce another source of unquantifiable error in the data. Secondly, the use of five year age groups and single year periods would give six year overlapping cohorts (ie there would be a larger degree of overlap between cohorts, further blurring the cohort effects).

An added advantage of the quinary quinquennia approach is that it forces a measure of smoothing on the series (more important in rare cancers) and reduces the chance of zero observations in age-period cells (zero counts can pose problems when fitting some models).

A disadvantage of quinary quinquennia is that where a trend is about to change direction (eg a historical rate of increase is about to become a rate of decrease – examples of this are melanoma and lung cancer), the extent of the change may be hidden within the quinquennium, but would be more obvious with annual data.

³ To see this, compare the date of birth of the oldest person in a particular age group in the first year of a particular five year period (this gives the first year of the cohort) with that for the youngest person in the age group in the last year of the five year period (the last year of the cohort) — their dates of birth will be 10 years apart. The cohorts overlap in that the first years of birth in the cohort selected will also be in the preceding cohort, and the last years of birth will be in the succeeding cohort. On average, 75% of individuals in any age-period cell are expected to have been born in the middle non-overlapping five years.

Model identification

The first step was a thorough literature review and expert consultation to identify possible models. The following categories of model were selected for use:

- generalised linear models (so-called age-period-cohort or APC models) (Holford 1991, Clayton and Schifflers 1987a, 1987b, Osmond 1985, Brillinger 1986, Robertson and Boyle 1998)
- non-parametric generalised additive (GAM) models (Wahba 1990, Hastie and Tibshirani 1990, Wood 2001)
- non-linear models (referred to here as Dyba and Hakulinen or DH models) (Dyba 2000⁴)
- versions of APC models with autoregressive prior distributions for the parameters, fitted using Bayesian Markov chain Monte Carlo (MCMC) methods, referred to as BUGS (Bayesian inference Using Gibbs Sampling) and BAMP (Bayesian Age-period-cohort Modelling and Prediction) models (Berzuini and Clayton 1994, Bray 2002, Knorr-Held and Rainer 2001, Spiegelhalter 1998, Schmidt 2002, Spiegelhalter et al 1998).

All of these models are broadly within the framework of the generalised linear model and share the following characteristics: Poisson random errors, the log of the person-years at risk as the offset, and usually a log link function typically with some combination of age, period and cohort fitted as factors.

Some of these models have alternative forms or ‘sub-models’:

- In the *APC* models it is possible to seek the most parsimonious model (in a hierarchy: age effect only; age effect + linear period effect or ‘drift’; age + period effects; age + cohort effects; the full age + period + cohort effects model). However, as the APC model was only one of the models to be fitted, and as APC models require the use of a ‘secondary’ fitting process to obtain projections (see page 26 and also Osmond 1985, Bray 2002), it was decided to fit only the full APC model, which will always give the best – if not the most parsimonious – fit among this class of models.
- In *GAM* models it is possible to fit models with bivariate smoothing of age and period effects, or age and cohort effects, or to fit models with univariate smoothing of age and period, or age and cohort.
- The *DH* model is linear in time (ie assumes that there is a straight-line increase in rate over time) but not in parameters (and so is considered to be ‘non-linear’). This class of model is only used for rates that increase linearly with time (for decreasing

⁴ Dyba’s PhD thesis (freely available from the author) contains a series of articles, some co-authored by Hakulinen, which were written while the non-linear model was being developed.

rates an APC model with age and linear drift would be used).⁵ The DH model could be fitted with either period or cohort as the time variable, but the fitting process was found to fail to converge⁶ more often for the age-cohort model. For our DH models, only the age-period form was used.

- Both BUGS and BAMP were tested. BAMP offers Bayesian models using either first or second differences as autoregressive priors for the time variables (age, period and cohort). The models were fitted using Bayesian MCMC methods.⁷ It is the model using second differences that was recommended to be more appropriate for making forecasts (Knorr-Held and Rainer 2001), so that was the model we used.

Model selection

The models listed above were tested to see how well they fitted the historical data, how similar their projections were, how robust they were when the data were very variable (in the cases of some rare cancers) or when there were zero cases in some age group–period cells, and how the models performed given curvilinear trends over time.

The candidate models were subjected to two types of statistical tests for model selection. Firstly, for a subset of the cancer sites we examined the fit of the model within the observed data and the projections for the next three periods. Secondly, we fitted the models to all but the last three observed periods and then projected values for the last three observed periods. This is known as an *ex post* test or empirical projection.⁸ In this test candidate models were compared based on the sum of their squared standardised residuals.⁹ Some models (not discussed here) were eliminated from consideration during this phase.

During the testing process it became apparent that no one model consistently outperformed all others across all data sets:

- Where the rate was increasing relatively fast, models fitting an exponential rate of increase (APC, BAMP) tended to give a projection that was too high.

⁵ A model that has a *linear* increase/decrease in time will, in the case of a decrease, after a long enough time interval, give negative rates (ie cross the *x*-axis). However, in the case of an increase over time, the model will take a very long time to predict extremely high values, even if the increase is relatively rapid. A model that has an *exponential* (log-linear) increase/decrease in time will, in the case of a decrease, reach a minimum value (horizontal asymptote) that is (just) positive, and then remain constant; but in the case of an increase over time, it will rapidly predict extremely high values. In this sense a linear model can be considered to be more suitable for increasing rates, and a log-linear model to be more suitable for decreasing rates.

⁶ The software used was unable to fit the model, as the iterative process used to determine the values of the parameters in the model did not converge on the parameter values to the specified level of accuracy.

⁷ The number of iterations was 120,000 after burn-in of 20,000, sampling every 10th iteration to aid mixing.

⁸ *Ex post* tests, projecting a full three periods, could only be conducted on registration data. For the mortality data there were a total of six periods, but at least four were needed to fit the model.

⁹ The standardised residuals used were the ones used by Bray (2002), which were extremely close in value to the chi-residuals as defined in Clayton and Schiffers (1987a). Within the observed data, these residuals gave values very close to those of the log-residuals and deviance-residuals (Clayton and Schiffers 1987a). Log- and deviance-residuals have the disadvantage that they can only be calculated when there are no age-period cells with zero cases.

- Where the rate was decreasing, all models other than DH tended to give very similar projections (the DH model was not appropriate in this situation, as it could give negative projections).
- DH models gave very good projections where the rate of increase was approximately linear, but not otherwise. This model was also the most difficult to fit, in that the model sometimes failed to converge.
- GAM models were sensitive to changes in rate over the last two or three periods, and could model a future decrease where no other model did. However, the true trend may or may not show a continuation of such a short-term decrease in the rate, so in some instances the projected decrease in rate may be too rapid.
- BAMP and APC models were relatively unlikely to project a continued decrease after a one- or two-period decrease in the rate, but tended to project a steep increase on the basis of a one- or two-period sharp increase in the rate.
- The BAMP and APC models gave the best fits within the observed range of the data (as measured by the sum of standardised residuals), but the APC models gave the poorest fits for the projections. One possible explanation as to why the APC models were found to project relatively poorly is that the projection assumption (that the period and cohort effects will continue to change as they have over the previous three periods – see next page) is a strong one, and has considerable influence on the projected values.
- In a very few instances, the APC model gave a ‘point of singularity’ where the fitted rate was very high for some age-period combinations, but gave reasonable values elsewhere.

Model averaging

Rather than selecting a single model from those that passed the initial testing phase, a *model averaging* process was used. Our final fitted values were calculated as the mean of the modelled number of cases, for each age-period combination, of the three or four models selected and fitted for each cancer site. Where the rate was increasing linearly or exponentially over time we fitted up to four basic models (using the best-fitting sub-model where there was a choice). We fitted three models (not the DH model) where the rate was decreasing over time, or where the rate was markedly curvilinear over time and the straight line fitted by the DH model was not suitable, or where the DH (or occasionally the APC) model failed to converge.

In further *ex post* tests the model average proved to be conservative, and tended to give ‘better’ fit than any component model. This is because the weaknesses of each of the component models used to calculate the model average are in a sense moderated by the strengths of the others. Where the rates are increasing, the model average was not as high as it would have been had only BAMP or APC models been used, nor as low as it would have been had only the DH model been used. Where the rates were stabilising or decreasing, the model average was not as low as it would have been had only the GAM model been used.

Projection

Once the final set of models for each cancer had been determined, these models were all fitted again using all available data (up to the 1996/97 five year period for most cancers). The projected number of cases in each age-period combination for each model was then determined, and the mean number of cases across the models was calculated as the projected model average value.

For the component BAMP, GAM and DH models, projections are obtained by extending the modelled risk surface to further time periods using the fitted parameters. To project the component APC model, however, an assumption was made that the age effects would remain the same in the future (which is usually accepted to be a reasonable assumption), so that the fitted age parameters could be used for projection. Future period and cohort effect parameters were estimated by fitting a linear regression line to the most recent three (occasionally more) period or cohort fitted parameters, and then extrapolating the next three (or more) (Osmond 1985). This assumes that the recent observed period and cohort trends will continue unchanged into the future, which is a strong assumption and somewhat limits the usefulness of APC models for projection.

Uncertainty estimation

A BAMP 90% credible interval was obtained during the model fitting process.¹⁰ This includes both uncertainty associated with the Bayesian model (or ‘within model’ uncertainty)¹¹ and random fluctuations in the data. These credible intervals are not centred about the average projections, but rather about the BAMP projections. Credible intervals from models fitted using Bayesian methods are generally acknowledged in the literature to be wide, thereby providing a conservative proxy indicator of uncertainty for the projection obtained by the model averaging process.

Ninety percent confidence intervals obtained using bootstrap methods¹² were considered, as were the confidence and prediction intervals proposed for the DH models. However, the variability of the projections between models, the often poor fit of some of the models, the regular failure to converge, and the fact that the selection of models varied from site to site, led to the decision not to use any of the confidence intervals, but to rely instead on the Bayesian equivalent to give an idea of the relative uncertainty of the projections.

¹⁰ The number of iterations was 120,000 after burn-in of 20,000, sampling every 10th iteration to aid mixing.

¹¹ In our model fitting process, which involves several different models, there is also the ‘between model’ uncertainty, in the sense that different models can give markedly different projections. There is not yet an established way to quantify this uncertainty.

¹² Bootstrap methods are those in which the variability in the model (as measured by the residuals) is used to generate new ‘data sets’ based on the fitted model plus a random selection of residual errors, the model is fitted again to each such ‘data set’, and percentiles of the fitted models are used to approximate confidence intervals.

The width of the Bayesian 90% credible intervals varies considerably between cancer sites. On the whole it is true that where:

- the rate is increasing rapidly, the intervals will be very wide (this is because of the exponential nature of the model, where relatively small differences on the log scale become considerably magnified when converted back to the original scale)
- the rate is stable, and any change in the rate over time is moderate, the intervals tend to be moderately wide
- the rate is decreasing steadily, the intervals tend to be narrow (because in all the models the rates are log-linear and approach zero – their theoretical minimum – as they decrease).

Thus the width of the credible interval is in part dependent on the variability in the data, in part on the ‘model uncertainty’, and in part on the actual trend (steepness of increase or decrease), as well as on the number of cases.

The age standardised rates derived from the lowest and highest rates estimated for each age-period combination by any of the component models are also shown on the trend charts in Part II. These lowest and highest estimates provide an indicator of ‘between-model’ (but not ‘within-model’) uncertainty.

Independence of incidence and mortality estimates

Mortality is modelled here independently of incidence rather than as a function of incidence and survival. Hence the forecasts of incidence and mortality for a particular cancer could be inconsistent with each other (declining incidence is unlikely to be associated with increasing mortality, except perhaps for a very short period of time).

Thus once the final projections had been obtained, *mortality:incidence ratios* were calculated to check that the projections were reasonable. It was found that all mortality:incidence ratios derived from the projections were less than 1,¹³ and were consistent with those from the observed data.

Additional methods or analyses for individual cancers

Tobacco related cancers (Chapter 4)

Tobacco was considered to be causally related to the following cancers: lip, mouth and pharynx; oesophagus; stomach; pancreas; larynx; lung; bladder; kidney; colorectal; and cervical (Chao et al 2000, English et al 1995, Thun et al 2000).

Tobacco attributable fractions for lung cancer were calculated by the method of Peto and Lopez (Peto et al 1994). This method recognises that the risk of tobacco related cancer is related to cumulative exposure to tobacco smoke. Current smoking prevalence is a poor proxy for cumulative exposure, which depends on the age at which smoking began,

¹³ However, oesophageal cancer in males and brain cancer in females showed small projected decreases in incidence accompanied by small increases in mortality rates.

duration of smoking, number of cigarettes smoked per day, degree of inhalation, and cigarette characteristics such as tar content and type of filter.

The lung cancer attributable fractions were calculated by comparing the observed age-gender specific lung cancer mortality rates in each historical period with the corresponding lung cancer mortality rates of non-smokers, derived from the American Cancer Society Cancer Prevention Study II (CPS-II) (Chao et al 2000, Thun et al 2000). These attributable fractions were then used together with the relative risks of smoking for lung cancer (for each age-gender group) from the CPS-II study to derive estimates of ‘synthetic’ prevalence of smoking. The ‘synthetic’ prevalence so obtained can be considered to be the equivalent prevalence of lifelong smokers that would have given rise to the calculated attributable fraction.

For future periods the ‘synthetic’ prevalence estimates for males were obtained in the same way, but using the projected lung cancer mortality rates for this gender. For females, convergence to the corresponding male rate by the end of the forecasting period was modelled, reflecting the convergence of observed male and female smoking rates over the past 20 years.

These ‘synthetic’ prevalence rates were then combined with relative risk estimates of smoking for the other cancers listed above to obtain the respective attributable fractions for these other cancers. The relative risk estimates for the other cancers were derived from the CPS-II study.

For each of the selected cancers, tobacco attributable fractions were calculated by combining these relative risk estimates with the ‘synthetic’ prevalence of smoking for each age-gender group in each period (observed and projected). These fractions were then applied to the relevant observed and projected number of cases (registrations and deaths) for each of the cancer sites concerned, and the attributable counts so obtained for each cancer were then summed to calculate estimates for tobacco related cancer as a whole. Non-tobacco related cancer counts were then derived by subtraction from the total number of cases. The same attributable fractions were assumed to apply to incidence as to mortality.

Note that the Peto and Lopez method is based on a number of assumptions, and also relies on the availability of accurate estimates of the relative risk of smoking for the different cancers. Accordingly, the quantification of tobacco related versus non-tobacco related cancer counts and rates reported here should be considered an approximation only.

All adult cancer (Chapter 7)

‘All adult cancer’ was modelled as a site in its own right. These projections were then compared with those obtained by summing the projections for the individual sites.

For male ‘all adult cancer’ incidence, the registration data used for 1994–98 were not the observed data, but modified estimates derived by substituting modelled for observed prostate cancer registration data, as explained below (see ‘Prostate cancer’).

Breast cancer (Chapter 12)

BreastScreen Aotearoa (BSA), the national breast cancer screening programme, is too recent (1998) to have yet impacted on the historical time series. The standard projection therefore may overestimate future breast cancer mortality.

To model this process, the incremental decrease in mortality anticipated to result from the BSA was modelled, given estimates of the coverage of the programme, the effectiveness of screening, and the time lag involved between screening and mortality.

Cervical cancer (Chapter 13)

Two sets of projections were made: the standard set using the full range of observed data available (ie to the 1996/97 five year period), and a second set using only the observed data to the 1986/87 five year period (the National Cervical Screening Programme (NCSP) began at the end of this period).

Comparison of the two sets of projections allows conclusions to be drawn concerning the cohort and period effects operating over the interim (ie late 1980s to mid 1990s).

Melanoma (Chapter 24)

Projections using the observed melanoma registration data show continued rapid rise in the incidence of this important cancer. However, it is known that melanoma was seriously under-reported prior to 1994 (when the Cancer Registry Act 1993 came into effect), so producing an artefactual sharp rise in melanoma incidence post 1994 – a situation not unlike the ‘PSA effect’ (discussed below).

Accordingly, we also modelled this cancer using data adjusted for the pre-1994 undercount. Adjustment was based on the ratio of melanoma registration to mortality rates pre- and post-1994.

Prostate cancer (Chapter 30)

Projections using the full data set for prostate cancer incidence were unreasonably high (using all possible models). The explanation seems to be widespread testing for prostate specific antigen (PSA), which has produced a dramatic increase in registrations of this cancer since the mid 1990s (ie in the last observed period). However, this ‘PSA effect’ is likely to be transient.

Instead, prostate cancer incidence was modelled using observed data only to 1993. Thus it is the underlying trend in prostate cancer incidence that is being modelled, unaffected by the ‘PSA effect’. This means that, at least for the near future, prostate cancer (and hence total male cancer) incidence has been purposely underestimated.

