Burden of Disease Epidemiology, Equity and Cost-Effectiveness (BODE³) Study Protocol

Version 1.0

Burden of Disease Epidemiology, Equity and Cost-Effectiveness Programme

Technical Report: Number 3

Principal Authors:
Tony Blakely
Rachel Foster
Nick Wilson

BODE³ Co-Investigators and contributing authors:
Roy Costilla, Martin Tobias, Diana Sarfati, Des O’Dea, Melissa McLeod, Matthew Soeberg, David Hadorn, Ken Richardson, Andrew Simpson, June Atkinson, Giorgi Kvizhinadze, Nhung Nghiem, Theo Vos, Jan Barendregt, Linda Cobiac

March 2011

A technical report published by the Department of Public Health,
University of Otago, Wellington

Acknowledgements

The Burden of Disease Epidemiology, Equity and Cost-Effectiveness (BODE³) Programme is funded by the Health Research Council of New Zealand.

BODE³ is conducted in collaboration with:

- the University of Queensland Burden of Disease and Cost-effectiveness Centre, more specifically the ACE-Prevention (Australia) Programme, and Deakin University.
- the Ministry of Health, particularly:
  - the Burden of Disease study
  - HealthTracker
  - numerous senior advisors
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC-CBA</td>
<td>Aotearoa Burden of Cancer and Comparative Benefit Assessment study</td>
</tr>
<tr>
<td>ACC</td>
<td>Accident Compensation Corporation</td>
</tr>
<tr>
<td>ACE</td>
<td>Assessing Cost-Effectiveness</td>
</tr>
<tr>
<td>AIHW</td>
<td>Australian Institute of Health and Welfare</td>
</tr>
<tr>
<td>BDS</td>
<td>Burden of Disease study</td>
</tr>
<tr>
<td>BMI</td>
<td>Body Mass Index</td>
</tr>
<tr>
<td>BODE³</td>
<td>Burden of Disease Epidemiology, Equity and Cost-Effectiveness programme</td>
</tr>
<tr>
<td>CHHP</td>
<td>Community Heart Health Programmes</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>CIAG</td>
<td>Cancer Interventions Advisory Group</td>
</tr>
<tr>
<td>CPI</td>
<td>Consumer Price Index</td>
</tr>
<tr>
<td>CRA</td>
<td>Comparative Risk Assessment</td>
</tr>
<tr>
<td>CVD</td>
<td>Cardiovascular disease</td>
</tr>
<tr>
<td>DALY</td>
<td>Disability Adjusted Life Year</td>
</tr>
<tr>
<td>DCIS</td>
<td>Disease Costs and Impacts Study</td>
</tr>
<tr>
<td>DES</td>
<td>Discrete Event Simulation</td>
</tr>
<tr>
<td>DHB</td>
<td>District Health Board</td>
</tr>
<tr>
<td>DRG</td>
<td>Diagnosis-Related Group</td>
</tr>
<tr>
<td>DW</td>
<td>Disability Weights</td>
</tr>
<tr>
<td>GBD</td>
<td>Global Burden of Disease</td>
</tr>
<tr>
<td>GCEA</td>
<td>Generalised Cost-Effectiveness Analysis</td>
</tr>
<tr>
<td>GP</td>
<td>General Practitioner</td>
</tr>
<tr>
<td>GST</td>
<td>Goods and Services Tax</td>
</tr>
<tr>
<td>HOP</td>
<td>Health of Older People</td>
</tr>
<tr>
<td>HRC</td>
<td>Health Research Council</td>
</tr>
<tr>
<td>HTA</td>
<td>Health Technology Assessments</td>
</tr>
<tr>
<td>ICD</td>
<td>International Classification of Diseases</td>
</tr>
<tr>
<td>ICER</td>
<td>Incremental Cost-Effectiveness Ratio</td>
</tr>
<tr>
<td>NGO</td>
<td>Non-Governmental Organisation</td>
</tr>
<tr>
<td>NHB</td>
<td>National Health Board</td>
</tr>
<tr>
<td>NHI</td>
<td>National Health Index</td>
</tr>
<tr>
<td>NHL</td>
<td>Non-Hodgkin’s Lymphoma</td>
</tr>
<tr>
<td>NICE</td>
<td>UK National Institute for Health and Clinical Excellence</td>
</tr>
<tr>
<td>NRT</td>
<td>Nicotine Replacement Therapy</td>
</tr>
<tr>
<td>NZACE</td>
<td>New Zealand Assessment Cost-Effectiveness Prevention Study</td>
</tr>
<tr>
<td>NZCMS</td>
<td>New Zealand Census-Mortality Study</td>
</tr>
<tr>
<td>NZHIS</td>
<td>New Zealand Health Information Services</td>
</tr>
<tr>
<td>PAG</td>
<td>Programme Advisory Group</td>
</tr>
<tr>
<td>PHO</td>
<td>Primary Health Organisation</td>
</tr>
<tr>
<td>PSA</td>
<td>Prostate Specific Antigen, or Probabilistic Sensitivity Analysis [depending on context]</td>
</tr>
<tr>
<td>QALY</td>
<td>Quality Adjusted Life Year</td>
</tr>
<tr>
<td>RR</td>
<td>Rate Ratio</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>SIGN</td>
<td>Scottish Intercollegiate Guidelines Network</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomised Controlled Trial</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
<tr>
<td>YLDs</td>
<td>Years of Life in Disability</td>
</tr>
<tr>
<td>YLLs</td>
<td>Years of Life Lost</td>
</tr>
</tbody>
</table>
### Glossary

| **@RISK** | Risk analysis software from Palisade that works with Microsoft Excel for Windows. Used most ACE-Prevention (Australia) models for all uncertainty analyses. The BODE³ team will use @RISK or **Ersatz** (see entry) for uncertainty analyses in NZACE-Prevention. |
| **Average cost-effectiveness** | The ratio of difference in costs for intervention compared to base-case (e.g., partial null), to the difference in utility (e.g., QALY, DALY averted) between intervention and base-case. The slope of the line from the origin to the interventions x-y coordinate on a cost-effectiveness plane. |
| **Bottom-up costing** | Bottom-up costing means determining the number of units of a particular resource used, then multiplying it by the unit cost for that resource, and then aggregating. (Term used in context of **cost offsets**.) |
| **Cost-effectiveness expansion pathway** | The curve on a cost-effectiveness plane formed by joining the origin to the intervention with the lowest average cost-effectiveness, then the next lowest, and so on. Any intervention above and to the northwest of the curve is “dominated”. |
| **Cost offsets** | These are future health system costs incurred or averted by an intervention that prevents, or reduces severity of, disease in the future or prolongs life that would not have been incurred/averted without the intervention under consideration. (Using Drummond et al’s framework¹, these are S1 (and usually S3) costs.) |
| **Cure time** | (See **statistical cure time**.) |
| **Ersatz** | Ersatz is an uncertainty and bootstrap add-in for Microsoft Excel, developed by Jan Barendregt. The BODE³ Team may use this instead of @RISK (see entry) in future work. See: [http://www.epigear.com/index_files/ersatz.html](http://www.epigear.com/index_files/ersatz.html) |
| **Heterogeneity** | Heterogeneity is where costs, consequences and cost-effectiveness may vary by ‘fixed’ socio-demographics such as sex, age, ethnicity and socioeconomic position, and other ‘fixed’ population characteristics (e.g., cancer stage). It differs from uncertainty, where there is uncertainty in estimates due to one of model structure, input parameter and stochastic uncertainty. Heterogeneity is of direct interest. For example, varying cost-effectiveness by age is critical to inform decision making. Also, and importantly in BODE³, it is likely that ethnic and socioeconomic equity considerations will be largely understood through heterogeneity on these variables. |
### Incremental cost-effectiveness ratio (ICER)

The ratio of difference in costs for intervention A compared to intervention B, to the difference in utility (e.g., QALY, DALY averted) between treatment A and treatment B. The slope of the line from intervention A’s to intervention B’s coordinates on a cost-effectiveness plane (assuming that intervention B’s consequences and costs have been calculated as though intervention A has already been applied).

### Input parameter uncertainty

(Also see entry under “Uncertainty”) One of three forms of uncertainty in economic decision modelling. It arises due to uncertainty in the input parameters to the model. For example, there may be uncertainty about the benefit of a new treatment, often expressed as a confidence interval about an effect estimate such as a rate ratio. Of note, the confidence interval captures random error about the input parameter; often there will be likely residual systematic error (i.e., confounding, selection or information biases that are thought to still be present in systematic reviews, meta-analyses, etc). BODE will on occasion attempt to include combined estimates of random and systematic error about input parameters, using quantitative bias analysis techniques (in addition to usual methods for measuring random error). There are well specified distributional forms for input parameter uncertainty – at least the random error component. For example, log normal for rate ratios, gamma for costs, etc.

### Input parameter uncertainty analysis ("uncertainty analysis" for short)

Input parameter uncertainty effects on total cost, consequence and cost-effectiveness uncertainty will be quantified using Monte Carlo simulations, with each iteration selecting values from the distribution about each input parameter with probability proportional to the density function, using distributional forms that best capture input parameter uncertainty. In NZ-ACE-Prevention, @RISK and Ersatz software will be used. In ABC-CBA, direct programming within Stata or TreeAge will probably be used.

### Intervention pathway

(See Cost-effectiveness expansion pathway)

**Macro-costing**

Macro-costing uses cost estimates for units of input and output that are large relative to the intervention being analysed. For example, macro-costing uses cost estimates for hospital stays or doctor visits rather than for the procedures and professional time expended during these encounters. (Term used in context of direct intervention costs.)

**Micro-costing**

Micro-costing requires the direct enumeration and costing out of every input consumed in the treatment of a particular patient. Micro-costing is laborious to implement. (Term used in context
of direct intervention costs.)

<table>
<thead>
<tr>
<th>Model structure uncertainty</th>
<th>(See entry under Uncertainty.)</th>
</tr>
</thead>
</table>

**Partial Null**

The ‘base-case’ scenario against which proposed intervention costs and health consequences are compared. Often it is simply our best estimate of the current (and projected into future) state of disease incidence and prevalence, preventive programmes, treatment coverage, etc. However, if we are to undertake economic decision modelling about interventions that are currently in place, or responsible for some of the projected future ‘business as usual’ scenario, we need to remove the current/projected effects and cost of these interventions from the base-case scenario. For example, in a comparison of cardiovascular disease prevention and treatment programmes that includes comparing current practice with cholesterol lowering drugs with alternative use of thiazide diuretics, a new ‘null’ base-case scenario needs to be constructed. Usually it is a ‘partial null’ as the costs and health consequences of current/projected interventions that impact only the domain of interest are stripped out of the base-case model – not the costs and consequences of all health system interventions.6

**Probabilistic analysis**

Probabilistic analysis is used in input parameter uncertainty analysis in BODE. To capture the full uncertainty of the input parameter, a distribution around the best estimate value is defined. The uncertainty is then propagated through the model using simulation techniques (see input parameter uncertainty analysis).

**Probabilistic sensitivity analyses**

(See entry under Sensitivity analysis.)

**Sensitivity analysis**

Sensitivity analysis has a variety of meanings. In BODE, we use this term to refer to changes or ‘sensitivity’ of modelling outputs (eg, DALYs averted, costs, ICER) to model structure and inputs. There are two important types of sensitivity analysis:

1. **Sensitivity to model structure assumptions.** Here sensitivity analysis is the testing of scenarios around variations in key design features and structural assumptions of the model (eg, varying the discount rate, adding states for relapse, including or excluding certain costs [although the latter two might be

---

6 Terminology around sensitivity analysis is not uniform across practitioners. The terms and definitions used in this Glossary are influenced by schema in Briggs et al (2006) and that used by Jan Barendregt (course teaching notes).
considered model extensions — but there is a grey zone between what might be called sensitivity analyses and model extensions/variations).

2. Sensitivity to input parameter uncertainty. Here sensitivity analysis is the determination of changes in the DALYs averted, costs or ICER arising from input parameter uncertainty. Results of input parameter uncertainty sensitivity analyses are often shown as a Tornado plot. There are two common ways to quantify this:

a. One-way (input parameter) sensitivity analyses. Determination of sensitivity or change in output variable for a meaningful unit change (e.g., 1 standard deviation) in each input variables, where each sensitivity analysis is for only one variable at a time.

b. Probabilistic sensitivity analyses. Probabilistic analyses must be undertaken first, generating multiple sampled values of input parameters and multiple values of output variables across all iterations. The probabilistic sensitivity analysis is then the correlation between each input parameter (that has uncertainty and sampling from a probability distribution across iterations) and the output parameters. It will differ from one-way sensitivity if there are correlations between parameters in their impact on outputs.

---

**Statistical cure time**

The time post diagnosis when excess mortality has (essentially) reached zero, or the relative survival curve has (essentially) become horizontal. That is, death from the cancer is negligible, and all survivors are assumed to be cured.

**Stochastic uncertainty**

(See entry under Uncertainty.)

**Top-down costing**

Top-down costing is perhaps best thought of in terms of aggregate Vote:Health (and private) outlays, broken down by main expenditure categories or for a specific condition. This approach is appropriate where it is not important to know with precision what the cost drivers are. An advantage of top-down costing is that one may know the total health expenditure (e.g., Vote:Health plus total private health expenditure from survey data), giving a known ‘envelope’ to work within. (Term used in context of cost offsets.)

**Uncertainty**

There is often considerable uncertainty in costs, consequences and cost-effectiveness estimates. We identify three sources of this uncertainty:

1. **Model structure uncertainty** arises due to uncertainty about the assumptions implicit in setting the model structure. For example, the disease model, discount rates, etc. In BODE³, model
structure uncertainty will be assessed by \textit{sensitivity analysis} about given alternatives.

2. \textbf{Input parameter uncertainty} arises due to uncertainty in the true value of parameters input into the model, such as uncertainty about the consequences of a new treatment. Within BODE$^3$, \textit{“input parameter uncertainty analysis”} will be used to address this source of uncertainty.

3. \textbf{Stochastic uncertainty} arises due to random variability in individual experiences (e.g., time to event). We will not usually consider stochastic uncertainty, as the macro-simulation models most commonly used in BODE$^3$ are estimating population-level expectations, with input parameter uncertainty capturing likely variation about mean estimates.

| Uncertainty analysis | Abbreviation for \textit{“input parameter uncertainty analysis”}. |
# Table of Contents

Acknowledgements.............................................................................................................. ii  
Abbreviations ........................................................................................................................ iii  
Glossary ........................................................................................................................................ v  
Table of Contents ......................................................................................................................... x  
List of Tables ............................................................................................................................. xiii  
List of Figures ........................................................................................................................... xiii  
List of Equations ........................................................................................................................... xiii  
List of Required Methodological Developments ........................................................................... xiv  

**Executive Summary** .................................................................................................................... 1  

## 1  Key principles .......................................................... 4  
1.1  General approach .................................................................................................................. 4  
1.2  Overarching Framework of BODE3 .................................................................................... 5  
1.2.1  Allocative efficiency ........................................................................................................ 5  
1.2.2  National decision context ............................................................................................... 5  
1.2.3  Balance of rigour, relevance and process ....................................................................... 5  
1.2.4  Study perspective for health consequences ................................................................... 5  
1.2.5  Study perspective for costing .......................................................................................... 6  
1.2.6  Reference year .................................................................................................................. 8  
1.2.7  Target population ........................................................................................................... 8  
1.2.8  Time horizons ................................................................................................................. 9  
1.2.9  Defining the intervention ............................................................................................... 9  
1.2.10  Defining the comparator .............................................................................................. 10  
1.2.11  Adherence to principles and base-case models stated in the Protocol ......................... 11  
1.3  Study Design ....................................................................................................................... 11  
1.3.1  Models ............................................................................................................................. 11  
1.3.2  Interacting consideration of time horizon for application of intervention, and comparability across interventions .............................................................................................................. 13  
1.3.3  Consequences: general measurement .......................................................................... 16  
1.3.4  Consequences: allowing for background mortality rates ............................................. 17  
1.3.5  Consequences: allowing for background comorbidity ............................................... 19  
1.3.6  Costs: measurement and valuation ............................................................................... 19  
1.3.7  Discounting .................................................................................................................... 20  
1.3.8  Uncertainty analysis ........................................................................................................ 20  
1.3.9  Equity analyses .............................................................................................................. 23  
1.3.10  Effectiveness and consequences of interventions, and classification of the strength of evidence ........................................................................................................................................... 23  
1.3.11  Extrapolating intervention effects over time ................................................................. 27
BODE\textsuperscript{3} Protocol, Version 1.0

1.4 Advisory committees ........................................................................................................ 27
1.5 Other criteria for policy-making, or second-stage filters .................................................. 29

2 Selection of interventions to evaluate ............................................................................. 31
  2.1 ABC-CBA Selection of Interventions to Evaluate ......................................................... 31
    2.1.1 Cancer intervention selection criteria ........................................................................ 31
    2.1.2 Cancer intervention selection process ..................................................................... 32
  2.2 NZACE-Prevention ......................................................................................................... 34
    2.2.1 Risk factor selection ................................................................................................. 34
    2.2.2 Specification of interventions, and groupings of interventions, for NZACE-Prevention analyses .................................................................................................. 36
    2.2.3 Process for Stakeholder Critique ............................................................................ 36
  2.3 Specification of the Intervention ..................................................................................... 37

3 Assessment of health consequences of interventions ....................................................... 38
  3.1 Levels I to III: Literature synthesis ................................................................................. 38
    3.1.1 Methods and steps for undertaking a systematic review ........................................ 38
  3.2 Lower level evidence synthesis ..................................................................................... 40
    3.2.1 Level IV evidence ................................................................................................... 40
    3.2.2 Indirect and parallel evidence .................................................................................. 41
    3.2.3 Expert opinion and consensus ................................................................................ 41

4 Assessment of costs .......................................................................................................... 43
  4.1 Overview of costing methods ......................................................................................... 43
  4.2 HealthTracker ................................................................................................................. 44
  4.3 Criteria for inclusion and exclusion of costs ................................................................. 45
  4.4 Intervention Costs .......................................................................................................... 47
  4.5 Cost Offsets from Disease Averted/Incurred ............................................................... 49
  4.6 Methodological Considerations ...................................................................................... 50
    4.6.1 Assuming the CPI applies to future health care costs ........................................... 50
    4.6.2 Discount rate .......................................................................................................... 51
    4.6.3 Goods and services tax ............................................................................................ 51
    4.6.4 Annuityisation of capital costs ................................................................................. 51
    4.6.5 Marginal vs Average Costing .................................................................................. 52
    4.6.6 Estimating Cost Aggregates .................................................................................... 52
    4.6.7 Data gaps for New Zealand cost data ....................................................................... 52
    4.6.8 Miscellaneous ......................................................................................................... 53

PART II: ABC-CBA METHOD AND PROTOCOL ................................................................... 54

5 Structure and data inputs for baseline models ................................................................. 58
  5.1 Cancer sites ................................................................................................................... 58
5.2 Cancer disease model structure, cure rates and disability weights .................. 60
  5.2.1 Extending the cancer disease models to include stage/sub-type .................. 61
  5.2.2 Duration and disability weights, by state, for each cancer model .................. 62
  5.2.3 Allowing for two or more disability weights ........................................... 64
  5.2.4 Cure times ................................................................................................. 65
  5.2.5 Long-term sequelae of cancer ...................................................................... 65

5.3 Data inputs to baseline models ........................................................................ 66
  5.3.1 Expected background mortality ................................................................. 66
  5.3.2 Cancer incidence ......................................................................................... 66
  5.3.3 Relative survival and excess mortality rate modelling .................................. 68

5.4 Cost off-sets ..................................................................................................... 72

5.5 Baseline model calibration ................................................................................ 72

6 Modelling interventions ....................................................................................... 73
  6.1 Disease progression for 2006 incident cases only .......................................... 73
    6.1.1 Multi-cohort Markov models ................................................................. 73
    6.1.2 Discrete event simulation ........................................................................... 74
  6.2 Incident cancer cases in years beyond 2006 ..................................................... 75
  6.3 Disease progression for future incident cancer cases ....................................... 75
  6.4 Model variations and extensions ...................................................................... 75
  6.5 Intervention model calibration ......................................................................... 75

PART III: NZACE-PREVENTION METHOD ................................................................. 76

7 NZ-ACE: Overview of ACE-Prevention model ...................................................... 76
  7.1 General comments ......................................................................................... 76
  7.2 Particular assumptions implicit in ACE-Prevention (Australia) modelling ......... 76
    7.2.1 Age weights and ‘fair innings’ ................................................................. 76
    7.2.2 Time lags to health benefits ....................................................................... 77

8 NZ-ACE: Modification to New Zealand setting ................................................... 77
  8.1 Modification of epidemiological parameters .................................................... 77
  8.2 Modification to costing ................................................................................... 77
  8.3 Intervention selection ...................................................................................... 77
  8.4 Other modifications .......................................................................................... 78

REFERENCES ........................................................................................................ 79
List of Tables
Table 1: BODE^3 base-case approach to costing .......................................................... 8
Table 3: Preliminary selection of cancer control interventions – the ten non-italicised interventions were prioritised and included as ‘preliminary’ in the HRC application. .......... 33
Table 4: Final prioritised list of the top six major risk factors for evaluation in NZACE-Prevention ......................................................................................................................... 35
Table 5: Cancer groupings used by studies relevant to core ABC-CBA model .................. 59
Table 6: Time spent in each state of the disease model (months), for different scenarios for a cancer with given maximal durations in each state as shown ........................................................................... 61
Table 7: Disability weights (DW) and duration time (T, in years) for the disease model states used in the Burden of Cancer report (ie, excluding consideration of cancer stage/sub-type) 1 ................................................................. 63
Table 8: Sequelae for cancer model from the Australian burden of disease study .......... 66

List of Figures
Figure 1: Components of economic decision modelling in BODE^3 – adapted from Figure 2.1 of Drummond et al (2005) 1. Components shaded in white boxes are routinely in scope in BODE^3, and in half-tone boxes are included either as practicable or as scenarios analyses – see text ................................................................................................................................. 6
Figure 2: Intervention pathway of the most cost-effective interventions for blood pressure- and cholesterol-lowering interventions, including the polypill, compared to current practice. Source: Vos et al (2010) 11 ................................................................................................................................. 11
Figure 3: Schematic example of DALYs averted by annual cohort in which preventive intervention applied (ie, line series), presented by calendar year in which the DALY was actually averted ................................................................................................................ 15
Figure 4: Schematic presentation of input parameters that may have uncertainty. The darker the grey fill, the more important is modelling uncertainty about this parameter. .......... 22
Figure 5: Programme structure and advisory groups ................................................................................................................................. 28
Figure 6: Conceptual cohort approach to modelling in ABC-CBA ................................ 56
Figure 7: General ABC-CBA Markov cancer disease model ........................................... 60
Figure 8: Excess breast cancer (female) mortality rate by sex, age, ethnicity and year (t): observed (−ln [obs RSR t / obs RSR t − 1]) (dotted lines); and predicted from Poisson model (solid lines) (Source 76) .................................................................................................................. 70

List of Equations
Equation 1 ....................................................................................................................... 51
Equation 2 ....................................................................................................................... 66
Equation 3 ....................................................................................................................... 68
Equation 4 ....................................................................................................................... 68
List of Required Methodological Developments

Required methodological development 1: Incorporating New Zealand-specific prevalent YLDs by age, and projecting them out to 2026 by ethnicity and sex .................................................. 19

Required methodological development 2: Development and testing of methods for using HealthTracker data to obtain cost-offsets for BODE³ .......................................................... 20

Required methodological development 3: Determining situations where allowing for parameter uncertainty in the baseline model, as opposed to changes in parameters, is necessary in BODE³ .......................................................... 22

Required methodological development 4: Determining situations when allowing for both systematic and random error in parameter uncertainty is appropriate, and methods to do the same .......................................................................................................................... 22

Required methodological development 5: Finalising the approach to undertaking and presenting equity analyses ............................................................................................................. 23

Required methodological development 6: Assess rapid search methods for possible unitisation in BODE³ ............................................................................................................. 40

Required methodological development 7: Development of approach to expert opinion as a source of information on input parameters and their uncertainty ..................................... 42

Required methodological development 8: Determine when direct costing should and can be undertaken separately by sub-populations .............................................................................. 44

Required methodological development 9: Determine when calculations of cost-offsets should and can be undertaken separately by sub-populations .......................................................... 44

Required methodological development 10: Scope and assess the validity and utility of HealthTracker for costing in BODE³ .................................................................................. 45

Required methodological development 11: Determine BODE³ practice for the costing of overheads ................................................................................................................................. 47

Required methodological development 12: Select best option for generating cost offsets in BODE³ .............................................................................................................................. 50

Required methodological development 13: Determine cost data availability for outpatient and private clinics, especially for cancer patients, through either HealthTracker or other mechanisms .................................................................................................................. 53

Required methodological development 14: Selection of state/sub-types to use for breast, colorectal and cervical cancer, including consultation with New Zealand cancer experts .... 62

Required methodological development 15: Specification of duration and DW for each stage/sub-type in the cancer models, including consultation with New Zealand cancer experts .................................................................................................................. 64

Required methodological development 16: Determine whether cure times should vary by cancer stage/sub-type .................................................................................................................. 65

Required methodological development 17: Develop methods for projecting cancer incidence rates to 2026, including by stage/subtype in some instances, using a mix of empirical data from New Zealand and internationally, and expert opinion ......................................................... 68

Required methodological development 18: Develop full method, and trial with HealthTracker and/or other data, for determining cost-offsets by ABC-CBA cancer model states ................ 72
Required methodological development 19: Develop a model calibration process to apply to baseline ABC-CBA models. ........................................................................................................... 72
Required methodological development 20: Development of criteria for when DES might need to be used.................................................................................................................. 75
Required methodological development 21: Development of process and method for applying DES in ABC-CBA............................................................................................................ 75
Required methodological development 22: Develop a model calibration process to apply to intervention ABC-CBA models.......................................................................................... 75
Executive Summary

This Protocol is for the Burden of Disease Epidemiology, Equity and Cost-Effectiveness programme (BODE³), funded by the Health Research Council of New Zealand (HRC) for the period 2010-15. The aim of the Burden of Disease Epidemiology, Equity & Cost-Effectiveness Programme (BODE3) is:

To build capacity and academic rigour in NZ in the estimation of disease burden, cost-effectiveness and equity impacts of proposed interventions, and undertake a range of such assessments.

The objectives are:

1. To estimate the impact (total & equity-related) and cost-effectiveness of cancer control interventions using Markov time dependent macrosimulation models for 30 types of cancer (Aotearoa Burden of Cancer and Comparative Benefit Assessment study; ABC-CBA).
2. To estimate the impact (total & equity-related) and cost-effectiveness of preventive interventions using multistate lifetables (NZ-Assessing Cost-Effectiveness: Prevention; NZACE-Prevention).
3. To build capacity and academic rigour in disease, intervention, equity, uncertainty and cost-effectiveness modelling.

This report is Version 1.0 of the BODE³ Protocol, published in March 2011. The Protocol will periodically be updated to include ongoing methodological developments, many of which are anticipated and identified as ‘work in progress’ in this version of the Protocol.

The protocol is in three Parts. Part I describes the overall principles that apply to the whole programme. These include:

- A general approach that uses burden of disease methods. This means that epidemiological models across interventions will be comparable, as they share similarly derived epidemiological parameters.
- Emphasis on a health system perspective, in that health consequences (averted disability ability adjusted life years (DALYs)) and costs are health-related and (usually) health system costs.
- A focus on allocative efficiency in economic evaluation.
- The use of probabilistic analysis methods to model input parameter uncertainty (eg, uncertainty around costs and intervention effectiveness).
- A focus on equity analyses, by which we mean the comparative impact (costs, health consequences and cost-effectiveness) of interventions by ethnicity and socioeconomic position (and by sex and age). A key facilitator of such a focus is the rich New Zealand data by ethnicity and socioeconomic position that allows heterogeneity to be specified and modelled.
- The processes that BODE³ will use to:
  o select interventions to model
assess the likely health consequences, and uncertainty, of the selected interventions (e.g., evidence hierarchies, review and expert consensus methodologies)

- assess the likely costs of interventions, including the direct ‘upfront’ costs of the intervention to the health system (and occasionally more widely) and the cost-offsets that may be accrued in the future by preventing death and prolonging life.

- The modelling approaches to be used in BODE³, namely:
  - multistate lifetables in NZACE-Prevention
  - Markov models, and possible discrete event simulation (DES) in ABC-CBA.

- The minimum outputs that BODE³ will generate are:
  - total estimated cost of the intervention, with uncertainty
  - total estimated DALYs averted by the intervention, with uncertainty
  - estimates of cost-effectiveness, with uncertainty
  - a description of the relative magnitude of disease burden (both DALY and (where relevant) comparative risk assessment (CRA) output) that the intervention is addressing, capitalising on the burden of disease study (BDS) underpinnings of BODE³
  - and estimates of ethnic and socioeconomic equity impacts.

Part II provides detail specific to the ABC-CBA methods and models. ABC-CBA gathers together a wide range of New Zealand and international data on cancer incidence, survival, disease models and quality of life in the form of large and heterogeneous (by cancer-site, sex, age, ethnicity, deprivation and sometimes stage/sub-type of cancer) Markov models that represent the business as usual scenario for incident cancers in 2006. This baseline model is extended out into the future for incident cases in years 2007 and beyond. Laid over the states of the model are disability weights and average health system costs for being in that state (the latter being sourced from Ministry of Health “HealthTracker” data if feasible).

Having developed these large and heterogeneous baseline models, cancer control interventions are then modelled by changing key parameters from the baseline model for sub-populations of interest. That is:

- treatment interventions are modelled by changing cancer mortality and disability weight parameters (and possibly time spent in various states), and the occurrence of sequelae
- preventive interventions are modelled by changes in the future incidence rate
- screening and early diagnosis interventions are modelled by changes in the stage or severity distribution at diagnosis, and incidence rates themselves.
- palliative and supportive care interventions are modelled by changes in disability weights.

It is assumed that many (if not most) interventions can be modelled by extracting portions of the baseline Markov cancer model and specifying input parameters (and uncertainty) directly to this model structure. However, it is also envisaged that:
• Disease models will need specific modification for some interventions (eg, the addition for serial relapse and remission stages if the treatment has varying effectiveness dependent on previous history of relapses (ie, adding more ‘memory’ to the Markov model)).
• Due to either computationally over-whelming heterogeneity (ie, just too many states for probabilistic analyses and simulation), or non-linearity in the model (eg, aggressiveness of tumour inversely proportion to time since last ‘clear’ screening test, where that relationship is not efficiently captured by adding further Markov states), discrete event simulation (DES) or other microsimulation modelling approaches may be used in specific instances. An over-riding principle, however, of ensuring comparability of epidemiological and cost parameters with the baseline models will be adhered to.

Part III provides summary detail on the NZACE-Prevention model. The NZACE-Prevention model is largely based on that used in ACE-Prevention in Australia, which has its existing study protocol.\textsuperscript{7} Therefore, Part II of the protocol is brief, focusing on issues such as the adaptation of multi-state lifetables to the New Zealand setting, and incorporating output from the parallel NZ BDS revision. Given the established nature of ACE methodology, specific pieces of additional work have already been undertaken for NZACE-Prevention and are published separately on the BODE\textsuperscript{3} website (www.uow.otago.ac.nz/BODE3-info.html), namely reports detailing options for selecting risk factors, then interventions per se, to model in NZACE-Prevention.\textsuperscript{9,10} Of note is that minor differences will apply to some of the methods NZACE-Prevention compared to ACE-Prevention (Australia) in the domains of what is included in the “health perspective”, and how certain costs are managed (eg, health costs relating to extra life lived as a result of the intervention(s); and set-up costs).
PART I: OVERALL BODE\textsuperscript{3} PRINCIPLES AND PROTOCOL

1 Key principles

1.1 General approach

BODE\textsuperscript{3} is a Health Research Council of New Zealand funded programme, running from 2010 to 2015. The stated aim of BODE\textsuperscript{3} is:

\begin{center}
To build capacity and academic rigour in NZ in the estimation of disease burden, cost-effectiveness and equity impacts of proposed interventions, and undertake a range of such assessments.
\end{center}

The objectives during 2010 to 2015 are:

1. To estimate the impact (total & equity-related) and cost-effectiveness of cancer control interventions using Markov time dependent macrosimulation models for 30 types of cancer (Aotearoa Burden of Cancer and Comparative Benefit Assessment study; ABC-CBA).
2. To estimate the impact (total & equity-related) and cost-effectiveness of preventive interventions using multistate lifetables (NZ-Assessing Cost-Effectiveness: Prevention; NZACE-Prevention).
3. To build capacity and academic rigour in disease, intervention, equity, uncertainty and cost-effectiveness modelling.

Beyond 2015, and perhaps added on during 2010 to 2015, there may be extensions to include other domains (e.g., mental health).

Stated briefly, both ABC-CBA and NZACE-Prevention will calculate the disease impact and cost-effectiveness of interventions by propagating the impacts of the interventions through core (and on occasion additional) Markov and multistate lifetable models, with costing done both external to the models (i.e., intervention costs per se) and internal to the models (i.e., incurred/averted health care system costs, or cost offsets). The models will be created using epidemiological parameters (e.g., incidence of disease and sequelae, disease stage at presentation, survival, mortality and survival rates) from a variety of sources, such as the parallel NZ Burden of Disease study (BDS) being conducted by the Ministry of Health, cancer registry data, etc. The modelling of the effects of interventions will be done by altering at least one of these parameters, or by using the baseline models as a departure point (or ‘inventory’ of epidemiological and costing parameters) to undertake more particular modelling specific to the intervention (e.g., additional states for serial relapse in a Markov model, or discrete event simulation for a cancer screening programme).
1.2 Overarching Framework of BODE3

1.2.1 Allocative efficiency

BODE3 will be evaluating allocative efficiency. Allocative efficiency answers questions of ‘what to do’, ideally across the health sector but more often within domains such as prevention or disease groupings (e.g., cardiovascular disease). For example, should we introduce a new drug for colon cancer or increase palliative care nursing? This also means that when we investigate drug treatments, where feasible analysis will preferentially focus on classes of drugs (e.g., receptor-specific drugs for breast cancer) rather than single drugs (e.g., trastuzumab).

Technical efficiency answer questions of ‘how to do it’. For example, which regimen of a new drug to use for colon cancer. BODE3 will not be addressing technical efficiency in the first instance.

1.2.2 National decision context

BODE3 will be primarily addressing questions about the national allocation of health resources. However, the boundaries with DHB-level down decision-making are inevitably blurred in a health system where much of the decision-making resides at the DHB-level. Key non-academic audiences for BODE3 findings will include researchers, the Ministry of Health, the National Health Board, the National Health Committee, Pharmac, and national organisations (e.g., Medical Colleges, National Heart Foundation).

1.2.3 Balance of rigour, relevance and process

NZACE-Prevention draws from the ACE-Prevention (Australia) project, which emphasises a balance of technical rigour (what matters to academics), relevance (what matters to policy makers), and due process (what matters to stakeholders). BODE3 will follow this balance.

1.2.4 Study perspective for health consequences

Consequences, that is health impacts sometimes more loosely referred to as benefits, will follow a cost-utility framework. Healthy life-years gained (or potentially lost) for a given intervention will be quantified using a hybrid measure of mortality and morbidity, namely disability-adjusted life-years (DALYs) averted. It is important to note that DALYs averted in cost-utility analyses are not the same as DALYs in a BDS (see fuller exposition later). Rather, they are the same as QALYs except that disability weights are used to quantify quality of life lost/gained rather than utilities. Our rationale for using DALYs is that they more directly link to burden of disease study methods and data, and DALYs provide a metric that is comparable across subpopulations and disease states. Regarding the latter, the utility component of a QALY is often disease or survey specific. This may be appropriate if one is just assessing interventions for (say) end-stage renal failure. However, an overarching goal of BODE3 is to make comparisons across multiple interventions, risk factors and disease states, with the greatest comparability possible. Use of DALYs will also assist with
international comparison and comparisons with extensive Australian work (ie, comparing with ACE-Prevention (Australia)).

Consequences beyond directly measurable mortality and morbidity effects to the individual are not included. Thus, issues such as the impact on partner or family/whanau wellbeing from an intervention directly affecting just one person are not quantified. Nor are issues such as reassurance from a true negative screening test (a potentially positive consequence of an intervention). Such ‘extra-DALY’ aspects of the consequences of an intervention, if assessed as substantial, will be flagged in discussion.

1.2.5 Study perspective for costing

Costing is primarily from the health system perspective. However, there are some variations as outlined below.

Drummond et al (2005)² propose a framework for costing (and consequences) that is reproduced below in Figure 1. From the costing perspective, the first distinction is between the direct costs of the intervention (‘C’ costs in figure) and the costs incurred/averted due to flow-on changes in disease incidence and prevalence (‘S’ costs). Both C and S cost can then be disaggregated into four further costs: 1. health sector (including government and private funded, and voluntary); 2. other sectors (eg, road traffic safety; again could be government, private or voluntary); 3. patient and family (eg, co-payments, cost of travelling to clinics); and 4. productivity.

Figure 1: Components of economic decision modelling in BODE³ – adapted from Figure 2.1 of Drummond et al (2005)¹. Components shaded in white boxes are routinely in scope in BODE³, and in half-tone boxes are included either as practicable or as scenarios analyses – see text.
There are many (often subtle) variations in what costs are included. A health funder only perspective would include only C1 and S1 costs, although there may be variation on whether just Government funded activity or Government plus private and voluntary sector activity is included. Furthermore, there are boundary issues to consider between what is the health sector, and what is not. For example, disability support and occupational safety and health.

A health system perspective would normally extend out to include both C1 and C3 and S1 and S3 costs. Within C3 costs, though, issues then arise as to whether one includes just direct patient co-payments and costs (eg, petrol for travel), or one additionally includes a cost for patient/participant time taken up by the intervention (which is an opportunity cost). Including a costing of time raises further issues, for example: how to cost it (average wage, or leisure time cost?); full attribution, or partial (eg, a patient engaged in a treatment may also derive value from reading during that time); and ensuring that the utility component of health consequences is not also quantifying time considerations.

In prevention, the intervention cost is often external to the health sector (ie, C2 costs). For example, a Ministry of Transport may fund breathalysers and traffic safety campaigns. It would seem to be nonsense to exclude these costs that are so obviously and directly part of the intervention – unless perhaps you were a policy-maker with concerns only about the government Vote:Health budget. However, what about the cost to the food industry of improved labelling on products? Should that be included in the analysis? Clearly ‘yes’ if a welfare approach was being taken. But arguably not from a health system point of view.

There will also often be pragmatic limitations on costing too. For example, determining the S3 cost off-sets as well as S1 cost off-sets may be difficult.

In BODE³ we will adopt a base-case model, and then possibly assess variations about this base case scenario.
Table 1: BODE$^3$ base-case approach to costing

<table>
<thead>
<tr>
<th>Costs of intervention</th>
<th>Cost-offsets</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>C1: Health sector</strong>, including:</td>
<td><strong>C1: Health sector</strong>, including:</td>
</tr>
<tr>
<td>• Government costs to Ministry of Health (including disability support and Government-funded primary care), DHBs, ACC.</td>
<td>• Principally Vote:Health (ie, government) costs as captured by HealthTracker and/or DRG costs. But excluding due to pragmatic reasons many of the (likely) smaller C1 costs.</td>
</tr>
<tr>
<td>• Voluntary and NGO costs, such as Cancer Society and Heart Foundation costs running a health education of supportive care programme, and recognising that costs will often either be: reimbursed by subcontract to Government agency; or passed on to (or raised from) patients/public.</td>
<td></td>
</tr>
<tr>
<td><strong>C3: Patient/family</strong>, including:</td>
<td><strong>C3: Patient/family</strong>, including only:</td>
</tr>
<tr>
<td>• Patient co-payments and out-of-pocket costs, such as prescription costs, physiotherapy, etc.</td>
<td>• Other costs as captured by HealthTracker, including average co-payments for primary care. But excluding due to pragmatic reasons many potential C3 costs.</td>
</tr>
<tr>
<td>• Direct travel costs (e.g. petrol). But excluding costing of patient/family time.</td>
<td></td>
</tr>
</tbody>
</table>

In addition to this base-case costing, the most likely alternative scenario will additionally include C2 costs. This would be most likely for preventive interventions that involve the non-health care sector. Whether it is just government-funded or wider C2 costs will be explicitly stated in scenario analyses. (S2 cost offsets also arise. For example, anti-acute psychosis drugs may reduce downstream "Law and Order" costs.)

Regarding S1 (and S3) costs, they will include both related and unrelated health care costs (see Section 4.5). The rationale for not including productivity costs is further discussed in Section 4. The many particularities of costing will be covered further in Section 1.3.6 and 4.

1.2.6 Reference year

The reference year is 2006, the year for which the baseline epidemiological models are constructed. For example, the prevalence of various risk factors is matched as closely as possible to the state of New Zealand in 2006. Likewise real costs. 2006 is also the reference year for the current BDS revision.

1.2.7 Target population

The default target population is the resident New Zealand population alive in 2006 who are potential recipients of the intervention. For treatment interventions in ABC-CBA, this means the newly incident cases in 2006. (See Section 1.3.2 for a discussion of why prevalent cases at 2006 probably do not need to be modelled.) For preventive interventions, this means the population identified for the preventive measure in 2006. However, many interventions will
be targeted at subpopulations, such as an appropriate age group for a specific screening programme.

Note that a closed cohort approach will be used. That is, only the usually resident 2006 population will be included in models. Put another way, people born after 2006 or immigrating to New Zealand are not included. (Also note that we will not be able to allow for emigration out of New Zealand from our closed cohort either.)

Allowing for the examination of heterogeneity across subpopulations in costs and/or consequences, and hence cost-effectiveness, is a critical feature of both ABC-CBA and NZACE-Prevention. Both models disaggregate the population into strata of sex, age, ethnicity, deprivation and such like, allowing highly specific ‘business as usual’ parameterisation (e.g., cancer survival by age by sex by ethnicity by stage group), as well as the ability to simulate interventions through selected subpopulations. Indeed, a major aim of BODE$^3$ is to quantify equity impacts, and this will be achieved by leveraging off these disaggregated strata allowing modelling of consequences and costs separately by subpopulation.

1.2.8 Time horizons

There are two time horizons to consider: the time over which the intervention is applied (and hence costs of the intervention itself accrue); and the time over which both consequences and costs averted/incurred are tracked and summed.

*Intervention time frame.* The time frame for the application of the intervention will follow how the intervention would be applied in real life, with the default position being a time-frame of one year (i.e., the reference year 2006). For instance, a three month nicotine replacement therapy (NRT) could be modelled as applying to eligible members of the 2006 New Zealand population if this is consistent with the interventions specification; it might be, however, that NRT is one component of a more comprehensive five year programme.

There are challenging study design interactions with the intervention time horizon (e.g., comparability of prevalent versus incident disease, time-lags of prevention/screening programmes and attribution of consequences in future years). These issues need to be carefully considered in the final determination of time horizons. Therefore, we defer further discussion of intervention time frames to a more detailed consideration below under Study Design (1.3.2, page 13).

*Follow-up time frame.* The 2006 population will be followed up for consequences and costs (averted and incurred specifically due to the intervention) till death or 100 years of age.

1.2.9 Defining the intervention

In modelling interventions we fully specify all activities (i.e., who does what, to whom, when, how many times, where and how often?). Set up costs and time delays may be modelled, in addition to assuming the intervention is fully implemented and operating at full potential (i.e,
in ‘steady-state’ operation). Whether or not to include set-up costs and phases will depend on how important they are for cost-effectiveness, and will be clearly documented in evaluations.

1.2.10 Defining the comparator

One of two comparators will be used. First, ‘current practice’ or ‘business as usual’. Here, any effect of an intervention is assumed to be over and above the cumulative effect of current interventions, and acting on the baseline of future projected disease incidence, duration and survival parameters. This ‘current practice’ comparator is adequate for interventions that occur in addition to the current array of interventions. It gives rise to incremental cost-effectiveness ratios (ICER).

A problem, however, with current practice as the comparator is that inefficient current practice will make many alternative interventions look very cost-effective. For example, if a very expensive and moderately effective drug is currently being used for osteoarthritis, an incremental analysis that substitutes this drug with a cheaper but equally effective intervention will be cost saving with no change in DALYs. However, if one considered the comparator as no use of any drugs for osteoarthritis, both drugs may be cost ineffective.

An alternative approach, therefore, is to use a ‘do nothing’ or ‘partial null’ comparator, sometimes known as generalised cost-effectiveness analysis (GCEA). This is the preferred option if the policy-maker wants to know what the cost-effectiveness of both current practice and an ‘ideal’ combination of interventions might be had the health system not have evolved the way it has in recent decades. Figure 2 below is a good example. It shows cholesterol and blood pressure lowering medications including a polypill (a low-cost combination of three generic blood-pressure-lowering drugs and a cholesterol-lowering drug (statin) in a single pill) compared with current practice, from the ACE-Prevention (Australia) project. Current practice is shown as costing about $12 billion AUD, averting about 380,000 DALYs, for a cost-effectiveness of about $32,000 AUD per DALY averted (which is cost-effective for a common threshold of about $50,000 per DALY). Importantly, the comparator is a revised model state, whereby the effects of currently used interventions (eg, including more expensive cholesterol lowering drugs) are taken out of the model by revising mortality and incidence rates for CVD as though these existing interventions were not occurring. That is, a ‘partial null’ of no existing interventions. The advantage of this approach is that one can now see the cost-effectiveness of both ‘current practice’ as opposed to what might be alternative packages, namely the comparison of average cost-effectiveness ratios. The graph shows the same DALYs averted with current practice could be achieved with Community Heart Health Programmes (CHHP) and polypills, and actually save about $1 to $2 billion AUD (as opposed to the current $12 billion AUD cost of current practice). That is, the ‘partial null’ approach allows a full evaluation of current practice against alternative practices, potentially identifying major areas for practice change that may free-up resources for use elsewhere.
BODE\textsuperscript{3} Protocol, Version 1.0

Figure 2: Intervention pathway of the most cost-effective interventions for blood pressure- and cholesterol-lowering interventions, including the polypill, compared to current practice. Source: Vos et al (2010)\textsuperscript{11}

Polypill for those at ≥5%, ≥10% or ≥15% five-year cardiovascular disease risk is considered. The polypill is assumed to cost $200 p.a. per person.

BODE\textsuperscript{3} will use a ‘partial null’ comparator where justified, and otherwise a ‘current practice’ comparator. The choice of comparator will be explicitly stated.

1.2.11 Adherence to principles and base-case models stated in the Protocol

Peer reviewers of editors may wish to alter our BODE3 default approaches on a case by case basis. Wherever possible, this will be resisted so as to ensure as much comparability across published outputs. Reviewers and editors will be alerted to the Protocol.

1.3 Study Design

1.3.1 Models

The base-line models in ABC-CBA and NZACE-Prevention are Markov and multistate lifetable models, respectively. That is, macrosimulation models that describe the New Zealand 2006 population now and into the future, assuming ‘business as usual’. Many states in the models are specified for the New Zealand population in 2006, disaggregated by sex, age (single year), ethnicity (Māori, non-Māori initially), and deprivation (three groups: deciles 1-3, 4-7, 8-10). For ABC-CBA, there may be additional states for stage or severity of cancer – especially those cancers suitable for screening and early detection programmes (i.e., cervix, breast and colorectal). Starting distributions of the NZ population in 2006 are derived from the 2006 census, CancerTrends\textsuperscript{13} and cancer projections data\textsuperscript{14,15}, and the NZ-BDS revision
(in progress). Transition probabilities are used to stream the simulated population in each stratum of the initial 2006 population over time, transitioning to various disease states: cured, alive, living with sequelae, and such like. These transition probabilities are taken from the NZ-BDS and survival analyses conducted on cancer registry data.

These base-line models can then be used directly in economic decision modelling, by simply changing key parameters such as the mortality transition probabilities. That is, Markov or multistate lifetable modelling of interventions. We expect that on occasion, however, we may use the base-line models as a ‘data-bank’ of parameters for then specifying more nuanced models to evaluation interventions. For example, discrete event simulation (DES) may be used to evaluate a cancer screening intervention.

There are many key decisions to make in any modelling project that balance parsimony with complexity. Issues include:

1. **Whether to model a single age group, or multiple age groups.**
   Understanding heterogeneity across age groups is an important feature of BODE³, and we have sufficient data to do so. Therefore, our default approach is to model all age groups considered eligible for the intervention, using the 2006 population distribution of age (ie, the population alive in 2006 is used for multi-cohort modelling). This will often be achieved by ‘multiple cohort Markov modelling’.

2. **Whether to include prevalent cases of disease in 2006, or just incident cases of disease in 2006 that are prevalent thereafter.**
   For ABC-CBA at least, our default position will be to model incident cancer cases in 2006 only. Treatments for people in the second year or later post-diagnosis will be captured as the population ages. This will mean that due to discounting the absolute costs and consequences will be less than if prevalent cases in 2006 were modelled, but the relativities (and hence the ICER) will be unaffected. This issue, and its interaction with other model design assumptions, is considered further in Section 1.3.2 below.

3. **Whether to include incident cases in future years (ie, 2007 and beyond) as the 2006 population ages.**
   This depends on the analysis question. For treatment, supportive care and palliation it makes little difference to relativities, and thus parsimony dictates just using incident cases in 2006. For preventive interventions, however, one obviously needs to model future changes in incidence. The approach in both ABC-CBA and NZACE-Prevention is to model changes in future incidence among the population alive in 2006. Again, this issue is considered further in Section 1.3.2 below.

All base-line or business as usual models (ie, before any intervention is modelled) require projections into the future. For example, projected cancer incidence and survival for the remainder of the 2006 population’s life. As a general principle, where there is sufficient data (eg, cancer incidence projections¹⁴⁻¹⁵) or strong theoretical expectation (eg, steady
improvements in ovarian cancer survival across multiple countries; phasing of the tobacco epidemic associated with delayed phasing of disease rates) we will model changes in such epidemiological parameters for 20 years out to 2026, then assume they are constant thereafter. Specific details or projections will be described later in this protocol, and in subsequent reports and papers.

1.3.2 Interacting consideration of time horizon for application of intervention, and comparability across interventions

There are two key issues to consider:

- Treatment, supportive care and palliative interventions for prevalent and incident cases of disease
- Preventive and screening and early detection interventions and attribution of future consequences.

1.3.2.1 Treatment, supportive care and palliative interventions for prevalent and incident cases of disease

Should one model incident cases just in 2006, or incident cases in all subsequent years? So long as the discount rates are the same for consequences (ie, DALYs) and costs, and treatment costs (real dollars, non-discounted) and treatment effectiveness do not change much in the future, the ICER for the 2006 and future incident cohorts will be the same. An exception to this generalisation is if the ICER varies markedly by age, as future incident cases will arise from a modified age distribution (especially if it is only the 2006 population that is modelled out into the future). However, we will explore this by analyses directly of heterogeneity (ie, examining ICER by age group).

For a treatment applied to the 2006 population, does it matter whether one includes incident, prevalent or both incident and prevalent cases within the population during that year? Quite often, ‘yes’. If the disease in question has a reasonably long duration (ie, years to decades) such that average age differs between incident and prevalent cases, and is progressive such that disease severity increases with duration, and treatment effectiveness (or cost) varies with age and or severity, then the ICER among incident and prevalent cases may vary (often substantially). The exact amount of variation is context specific. Sometimes the cost-effectiveness may be better among prevalent cases; for example, more severe cases of CVD may be more likely to benefit in terms of deaths prevented than less severe CVD. Sometimes the cost-effectiveness may be worse among prevalent cases; for example, a new treatment for osteoarthritis may generate less health gain among older people with more severe disease.

This difference in ICER between prevalent and incident cases can be thought of as heterogeneity of the ICER across subpopulations. Both ABC-CBA and NZACE-Prevention models are finely disaggregated by sex, age, ethnicity (and sometimes deprivation and stage). The models are less prone to produce different ICERs for incident and prevalent cases if results are determined by subpopulation. Treatments are more likely to be modelled in ABC-CBA than NZACE-Prevention. Cancer is often not a long duration (ie, many decades)
chronic and progressive disease, although exceptions such as leukaemia, breast and prostate cancer definitely exist. Otherwise, cancer is often a rapidly progressive disease (be it to death or cure). Thus, this incident versus prevalent disease ICER issue is not too concerning within BODE3.

Perhaps more important in ABC-CBA are two more subtle questions:

- Do we just model cancers incident in 2006, and assume that results apply also to prevalent cancers in 2006?
- For interventions that often occur in later years of the disease course (e.g., palliative care, treatment of relapse), do we just model the costs and consequences of these interventions in out-years (e.g., 2011) as some of the cases incident in 2006 become eligible?

These two questions are actually different sides of the same coin, and thus have the same answer. Most cancers have curative treatment applied in the first year or so; therefore, there will not often be the need to consider prevalent cases. Considering supportive or palliative care, costs and consequences will occur together at some number of years post diagnosis. Given the same discount rate for costs and consequences, the ICER should be much the same for (say) palliative care applied to those needing it in 2006 (i.e., arising from many previous annual cohorts of incident cases) compared to that arising in future years among the 2006 incident cases.

Thus, for the issues considered above it appears that we are justified in modelling only incident cases in 2006 for treatment, supportive care and palliative interventions.

1.3.2.2 Preventive and screening interventions and attribution of future consequences

The second issue identified at the outset of this section was the time horizon over which to apply a prevention or screening intervention. There are two scenarios to consider. First, if the intervention only needs to occur in one year (such as changing a legislation), then one theoretically only needs to model the direct costs of the intervention as incurred in base-year 2006 and sum the DALYs averted and costs incurred/averted thereafter (allowing for time-lags and discounting). (That said, such single year interventions also have monitoring, enforcement or other costs in out years.)

Second, some preventive interventions need to be applied continuously to maintain the effect. For example, a future 30% reduction in cancer X mortality requiring screening every 3 years, or a sustained healthy eating campaign that is required to maintain dietary changes. We consider this now.

Figure 3 schematically shows DALYs averted by some preventive campaign applied sequentially in subsequent years to the same (aging) population (series), shown by subsequent calendar year in which the DALYs were actually averted (x-axis). The future stream of DALYs averted is assumed to be a log normal function of time for each single calendar year in which the intervention was applied. For example, and shown in the graph as the 2006 series, one can plot the future stream of DALYs averted caused by the intervention.
in 2006 only. Figure 3 shows such individual-year-DALY-attributions for every tenth year only to avoid clutter. The total DALYs averted in each calendar year post-2006 is also shown (ie, the ‘all yrs’ series), which is the sum across all annual applications of the intervention. Note that no allowance has been made for attrition (death mainly) of the starting population, which would mean that the all years line in the ‘no discounting’ sub-figure would actually peak and decline well before 2056, and the peak in the 3% discounting sub-figure would occur earlier than 2020.

**Figure 3**: Schematic example of DALYs averted by annual cohort in which preventive intervention applied (ie, line series), presented by calendar year in which the DALY was actually averted

There are three options for calculating the ICER in the above scenario. First, one could assume the intervention was at steady state and applied to 2006 only. The DALYs averted...
would be that of the 2006 line only in Figure 3. When it is possible to calculate the stream of DALYs averted in the future due to the intervention effect in one year only, this is viable. The ICER will be the same as that modelled for the programme applied over many years (so long as the discount rate of costs and consequences are the same, and intervention effectiveness and cost structures are similar in the future). However, it is often not possible to estimate the DALY’s averted from just one year of the preventive interventions, when the actual intervention or screening programme needs to run continuously. For example, reductions in cancer mortality are usually estimated for a long-running programme as a whole, not allocated per annum of the programme and strung out over future years.

The second option is to simply model the intervention as applying indefinitely from 2006 (or until all members of the simulation population have died), and calculate all of the (discounted) costs and consequences over this same intervention time horizon. This modelling indefinitely into the future will be our default intervention time frame option for preventive programmes.

The third option is to model the intervention as running for, say, 10 years (as in WHO-CHOICE \(^6\)) to 20 years, allowing for specification of an aggregate time lag to full programme effectiveness and decline after 10-20 years. However, this is still computationally problematic as one needs to model the decrease in effectiveness after stopping the intervention.

Regardless, the ICER will theoretically be much the same for the above three options with a closed cohort, unless future disease incidence (due to other secular trends) is projected to change markedly, or future (real) cost structures and programme effectiveness are projected to change markedly. Any such differences in ICER between approaches will reduce further with discounting and attrition of the base-year cohort. The choice of intervention time frame is driven more by pragmatic and empiric considerations.

1.3.3 Consequences: general measurement\(^b\)

We measure the size of the health gain associated with each intervention in ‘health-adjusted life years’ where we value the loss of health due to non-fatal health states with the appropriate disability weight(s) used to estimate disability-adjusted life years (DALYs) in burden of disease studies. When we present our results we equate these health-adjusted life years gained to DALYs averted by the intervention. However, there are important differences between DALYs calculated in burden of disease studies and DALYs averted used in disease and economic modelling, such as BODE\(^3\). First, in a burden of disease study, the health status of a population is estimated in a particular year. It is, therefore, a cross-sectional measure. Economic evaluation methods always have a time dimension. Health gain is calculated as the difference in mortality and morbidity outcomes between a comparator and the intervention option over a defined period of time (the ‘time horizon’).

\(^b\) This section is taken from the ACE Prevention protocol, with minor amendments only.
Second, in burden of disease studies the DALY is constructed as a health gap measure, i.e., an ideal is set (everyone ought to live into old age free of disease) and contrasted with the current health status of a population. Thus, years of life lost, the mortality component of the DALY, are calculated as the difference between age at death and a standard life expectancy at that age for each death. In the economic analyses of BODE³, we do not use the standard life table to give a value to loss of healthy life. Instead, we keep track of a target population over time and count the health-adjusted years of life lived in intervention and comparator scenarios assuming realistic mortality risks as people age (i.e., the population’s own lifetables – not an external ‘ideal’ standard; the population’s own lifetable concept will be extended to sex by ethnic (and occasionally by deprivation) specific lifetables, and is described in more detail below in Section 1.3.4).

This counting of health-adjusted years of life also includes an adjustment for expected levels of disability by age and sex for conditions not immediately affected by the intervention of interest. In other words, extra years of life gained from an intervention are counted as less than full years taking into account the probability that the person will suffer from osteoarthritis, dementia, hip fracture or any other condition as they age. That is, we allow for co-morbidities. Operationally, this involves assuming an average disability weight (DW) by sex and year of age, using averages from an appropriate BDS (which will be the New Zealand BDS for 2006 in due course). We allow for expected background disabilities in order to measure realistic health gains, rather than hypothetical health gains assessed against perfect health. Actual DWs used in BODE³ by sex, age are described in more detail in Section 1.3.5 below.

### 1.3.4 Consequences: allowing for background mortality rates

Background mortality rates for lifetables are used for two main purposes in BODE³. First, they can be transformed into transition probabilities to ‘death from other causes’ in Markov, multistate lifetable and perhaps DES models (having first subtracted away the mortality rate(s) for the disease(s) addressed by the intervention). Second, for modelled deaths, the lifetables are then used to estimate years of life lost.

Two questions arise at this point:

1. Should we allow for varying mortality rates not only by sex and age, but also by ethnicity and deprivation?
2. Should we allow for varying mortality rates into the future?

Our answer to both questions is ‘yes’, as to not allow for these variations is equivalent to assuming no variation in mortality by ethnicity and deprivation, and no variation over time in mortality rates. Both assumptions are clearly false.

Regarding ethnic and deprivation-specific lifetables, and hence mortality rates, for base-year 2006, we undertook the following process:
Official Māori and non-Māori complete lifetables for 2006 were sourced from Statistics New Zealand.

Abridged (i.e., <1, 1-4, 5-9,..., 85+ year increments only) lifetables for Māori and non-Māori, each disaggregated by quintile of NZDep 2006, were sourced from the Ministry of Health.

The mortality rates from the abridged lifetables for each quintile of deprivation (within ethnic and sex groups, and age groups) were then submitted to a simple linear regression with quintiles coded as 0.1, 0.3, 0.5, 0.7 and 0.9. Thus, the intercept was the expected mortality rate for the least deprived (i.e., 0th percentile) for any sex by age by ethnic groups, and the slope was expected increase in mortality from the least to most deprived (i.e., 100th percentile).

The slope and intercepts from these models were then used to estimate the mortality rate at the mid-point of the three deprivation groupings used in BODE3 (i.e., deciles 1-3, 4-7 and 8-10), and thence to estimate mortality rate ratios for these three groups. (Rate ratios above the age 85 were assumed to linearly approach 1.0.)

These rate ratios were then applied to the Māori and non-Māori lifetables to generate 12 complete (i.e., single year of age) lifetables for 2006. That is, sex (2) by ethnic group (2) by deprivation (3) lifetables.

Further details on this method, and the actual lifetables, can be obtained from the BODE3 website [www.uow.otago.ac.nz/BODE3-info.html](http://www.uow.otago.ac.nz/BODE3-info.html).

Regarding annual percentage changes into the future for mortality rates, we used two sources of information. First, Statistics New Zealand has used cohort lifetables and age period cohort modelling to estimate future annual percentage changes in mortality (averaged over ages). The 2007 estimates for medium scenario projections term projections hovered around 2% per annum reductions in mortality rates for the four combinations of sex by ethnicity. Second, Blakely et al (2010) noted linear trends in the long-run trends in life expectancy for Māori and non-Māori since about 1900 to 2006.21 (Trends vary decade by decade (e.g., flat in the 1960s and 1970s for non-Māori, and likewise in the 180s and 1990s for Māori), but over the long-run the trends in both ethnic groups are remarkably linear.) Assuming such a linear trend continues till 2026, annual percentage changes in mortality rates of about 1.5% and 2.5% for non-Māori and Māori are required. Simply averaging these two sources of estimates gives a 1.75% annual reduction for non-Māori, and 2.25% for Māori. These average estimates have been applied to the above lifetables, giving calendar-year specific lifetables to 2026, then assumed constant.

(Elsewhere, we have calculated ethnicity by income by smoking status lifetables.22 Smoking-specific expected mortality rates may be required for some simulation models in the future in BODE3, but we will not routinely include smoking strata in the baseline model.)
1.3.5 Consequences: allowing for background comorbidity

As stated above in Section 1.3.3, it is unrealistic to assume that a year of life saved at age 85 is a year of life in perfect health, or more specifically a year of life with full utility or no disability. National burden of disease studies can be used to estimate the average DW by single year of age (by sex), or what has been termed the ‘prevalent YLD’ in the ACE-Prevention (Australia) project. They are typically close to 0.0 through most years, but increase to 0.2 to 0.4 at older ages (eg, 75 to 95 years of age) as comorbidities tend to increase. These average disability weights by sex and age are then used to estimate healthy life years gained, in that expected years of life gained (eg, due to prevention of a CVD death) are adjusted for disability.

At the time of writing this version of the Protocol, the New Zealand BDS has not been completed. Thus, Australian prevalent YLDs will initially be assumed to apply to New Zealand. Once the New Zealand BDS is completed, we will incorporate New Zealand-specific prevalent YLDs by sex and ethnic group – but probably not by deprivation. Second, just as life expectancy is increasing, so too is healthy life expectancy. Thus, it seems reasonable to assume that this trend will continue into the future (till 2026 at least), meaning that prevalent YLDs need to be reduced at each year of age by some percentage into the future (just as mortality rates above are reduced annually). The 1.75% and 2.25% options above used for life expectancy increases seem a sensible starting point, but further work is required.

Required methodological development 1: Incorporating New Zealand-specific prevalent YLDs by age, and projecting them out to 2026 by ethnicity and sex.

1.3.6 Costs: measurement and valuation

The assessment of costs will be briefly overviewed here, and considered in detail in Section 4. The costs (more strictly the opportunity costs) are to be measured in ‘resource cost’ terms; ie, the economic or ‘dollar’ costs of the resources involved, approximating market prices. The costing perspective has already been outlined in Section 1.2.5.

In brief, the costs of the intervention will be costed directly, using activity costing of the event pathway as the default position. Macro- and micro-costing methods may also be used as appropriate.

If feasible, a top-down approach will be used to estimate downstream costs averted and incurred (cost offsets). This will be attempted using HealthTracker data from the Ministry of Health. These costs will include those averted by reduced incidence of disease and sequelae, as well as future costs incurred by improved survival.

---

"Top-down costing apportions total expenditure to the desired measure of output (eg, per person in a health state). Bottom-up, or micro costing, identifies and costs the resources used by a specific individual."
Required methodological development 2: Development and testing of methods for using HealthTracker data to obtain cost-offsets for BODE³.

The choice of costing method in each model will be determined by what data are available, what is most appropriate for the specific research question, and what can be feasibly achieved with the resources and time available.

1.3.7 Discounting
Discounting is standard practice for economic evaluation to incorporate time preferences for current and future costs and benefits. In BODE³, discounting will be applied to both costs and benefits (DALYs). The default discount rate will be 3% per annum.

Use of a discount rate of 3% p.a. is in line with ACE-Prevention (Australia). This is also the rate recommended by a consensus panel of health economists in the USA for cost-effectiveness analysis. Use of this 3% figure will optimise comparisons with existing international work. Nevertheless, to assist with NZ Government and Pharmac policy-making, we will include other discount rates in sensitivity analyses (ie, definitely 0%, but also at times 3.5%, 5% and 7%); Pharmac applies a discount rate of 3.5% p.a. In some cases it may be appropriate to consider a threshold analysis (eg, to determine at what discount rate an intervention becomes no longer cost-saving or cost-effective).

1.3.8 Uncertainty analysis

There is usually considerable uncertainty in the outputs of economic decision modelling. There are different frameworks for considering uncertainty. They may be thought of at three levels:

- Model structure uncertainty
- Input parameter uncertainty
- Stochastic uncertainty.

Model structure uncertainty is where the structure of the model may not be a good approximation of reality. This is difficult to incorporate explicitly into analyses, although it is still important and must be considered in interpretation. It may also be possible to consider in model structure uncertainty analyses (sometimes referred to as "sensitivity analyses" for short), where the model structure is altered for a different set of assumptions. For example, additional remission and relapse stages might be added to the model, to facilitate increased ‘memory’ in modelling as to previous disease progression. (Some practitioners might refer to not call this sensitivity analyses, but ‘model elaboration’ or some such term.) Other model assumptions that may be assessed include the intervention time frame (eg, where it is inappropriate to model just 2006 as steady-state), duration of effect of an intervention, etc.

---

Many of terms used in this section are also defined in the Glossary.
Stochastic uncertainty, or what is sometimes referred to as first-order uncertainty or random variability, is simply the variability in responsiveness and cost between different individuals. Modelling of stochastic uncertainty is necessary in microsimulation models, but not macrosimulation models. BODE³ mainly uses macrosimulation models (microsimulation models may occasionally be used for, say, screening in cancers), and will therefore be using averages (and uncertainty) for parameters in the model.

Thus, input parameter uncertainty is the major focus in BODE³.

We will use Monte Carlo simulation to model input parameter uncertainty. This involves thousands of iterations of the model calculations. Each iteration involves a random draw of a value from the probability distribution specified for each input parameter about which there is uncertainty. For example, we may specify a normal distribution about the natural logarithm of a rate ratio (ln(RR)) for change in a mortality rate. Each draw has a maximum probability of being close to the best or central estimate of the ln(RR), but it might also be one, two or more standard deviations above or below the central estimate with a probability proportional to that under the normal curve. Having run thousands of iterations, we then have thousands of (paired) values of both the cost and DALYs, giving an uncertainty range for the cost-effectiveness ratio (or the net benefit, etc). A 95% uncertainty interval would be bounded by the 2.5 and 97.5 percentile values across all iterations.

It is then possible to determine which of those input parameters with uncertainty confer the most total uncertainty in costs, consequence (ie, DALYs averted) and cost-effectiveness estimates. (This as a form of sensitivity analysis, in particular, probabilistic sensitivity analysis, in that we determine how sensitive output parameters are to each of the input parameters.) Those input parameters (with their associated uncertainty) that determine the greatest amount of output parameter uncertainty can then be prioritised for further attention and research. Tornado plots are often used for this type of analysis.

Note that heterogeneity is a separate issue, and relates to expectations of costs and consequences that systematically vary by variables such as sex, ethnicity, age, and so on, and sometimes by variables we do not understand but we assume exist (eg, tumour progression by some genotype). The BODE³ models are highly disaggregated, and hence allow modelling of cost-effectiveness for many (heterogeneous) populations. But parameter uncertainty will still exist within sub-populations, requiring uncertainty analyses by sub-population.

There are many input parameters for the modelling about which there is likely to be uncertainty. Figure 4 depicts generic types of input parameters. For assessing total parameter uncertainty, it is the uncertainty about the difference between the baseline (or partial null) to intervention scenario that matters most. For example, the difference in disease mortality between no intervention and intervention A. Or the difference in downstream costs between no intervention and intervention A. In this instance, one may not need to allow for uncertainty in the baseline or partial null parameters (ie, they are treated as certain or ‘fixed’), and only allow for uncertainty in the intervention scenario.
(which is equivalent to just allowing for uncertainty in the difference between the partial null and intervention).

**Figure 4:** Schematic presentation of input parameters that may have uncertainty. The darker the grey fill, the more important is modelling uncertainty about this parameter.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Baseline, or partial null</th>
<th>Intervention</th>
<th>Difference partial null to intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transition probabilities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incidence †</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Background mortality †</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Disease mortality ‡</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Quality of life</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disability weight</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Costs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention (eg, C1 and C3)</td>
<td>[ ]</td>
<td>[ ]</td>
<td>n.a.</td>
</tr>
<tr>
<td>Off-sets (eg, S1 and S3)</td>
<td>[ ]</td>
<td>[ ]</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

† This includes incidence into different stages, and between different states.
‡ Background mortality is that from diseases other than the diseases directly under consideration. Disease mortality is that from the disease directly under consideration.

However, there may be instances where it is important to model uncertainty in both the partial null and intervention states. For example, the difference in mortality rates between the partial null and intervention scenario may vary in proportion to the partial null mortality rate (which is also prone to uncertainty). That is, the expected value (and distribution) of the parameter estimate for the intervention is correlated with the partial null mortality rate (which is uncertain). However, it is unlikely that this situation will both arise and be of a magnitude sufficient enough to worry about on many occasions. For example, in specifying a partial null one is essentially calculating back with the same intervention assumptions (eg, effectiveness, adherence, coverage) as for the interventions that will then be evaluated. Thus, it is the same uncertainty that is being used to strip the model back to the partial null, as it is for then modelling these interventions. Therefore, we do not anticipate having to often explicitly model parameter uncertainty for both partial null and intervention states – just the intervention parameters themselves.

**Required methodological development 3:** Determining situations where allowing for parameter uncertainty in the baseline model, as opposed to changes in parameters, is necessary in BODE³.

In health economics, analysts usually use confidence intervals from other research studies or meta-analyses to specify probability distributions about input parameters. This focus on only random variation may be due to the heavy reliance on randomised trial evidence where, theoretically, systematic error is less problematic than in observational studies. Epidemiologists on the other hand are increasingly focusing on systematic error (ie, selection, information and confounding biases), and using methods (eg, quantitative bias analyses, including probabilistic sensitivity analysis methods) to address residual systematic
A methodological objective of this BODE3 programme will be to bring these two approaches together, and more explicitly model both random and systematic error.

**Required methodological development 4: Determining situations when allowing for both systematic and random error in parameter uncertainty is appropriate, and methods to do the same.**

### 1.3.9 Equity analyses

Our main focus will be Māori vs non-Māori comparisons, but differences by level of deprivation, gender and age may also be analysed. An aim will be to ascertain the relative advantages and disadvantages of various methods for equity incorporation, both for academic and policy communities, ranging from simpler to more complex methods (as ordered below):

- Separate modelling by social group, eg, Māori-specific patient navigator programme. If the ‘targeted’ programme is cost-effective, then it is both equity promoting and cost-effective on a total utilitarian perspective. However, there will often be equity-efficiency tradeoffs, hence the options below.
- Presenting DALYs-avoided separately by social group. This can be presented as a ratio to the ‘standard’ total DALY from a BDS for that social group (as a marker of ‘need’), to determine whether DALYs avoided are accruing more (or less) to those with the greatest disease burden. (At the time of writing this version of the Protocol, it seems that this option may be the most useful. It draws on the rich heterogeneity intrinsic in BODE3 models, and the strengths of being linked to a BDS.)
- We will trial measures of cost expressed per unit change in absolute difference in DALYs avoided. For example, if the total population cost was $1 million, but for Māori 0.1 DALY per capita (age-standardised) was averted compared to 0.08 for non-Māori, then this equity-change ratio is $1m/ per 0.02 change in DALY difference between Māori and non-Māori.
- Although not fully developed, equity-weighted benefit measures (either QALYs (quality adjusted life years) or DALYs) are the most commonly anticipated method of quantitatively incorporating equity. However, empirical analyses are rare – probably due to a lack of high-quality data, which is where BODE3 has an advantage. We will make a contribution to this literature by examining what happens to cost per DALY rankings if DALYs avoided for (say) Māori/young are weighted by 10% to 50% more than non-Māori/old.

In sum, our equity work will respond to the emerging call internationally for explicit equity methods that inform, rather than obfuscate. We will test the utility of such methods with policy-makers.

**Required methodological development 5: Finalising the approach to undertaking and presenting equity analyses.**

### 1.3.10 Effectiveness and consequences of interventions, and classification of the strength of evidence

The determination of effectiveness of interventions, and hence impact on consequences in the modelling, is of critical importance. Exact methods of information synthesis (eg,
literature reviews, meta-analysis, expert consensus) and parameterisation for modelling will be detailed later in this protocol (Section 3). Here, three general issues are considered:

- How is the intervention’s impact on DALYs averted in the model going to be parameterised, and what ‘link’ models are required to do this?
- General approach to synthesising evidence.
- Classification of strength of evidence, and implications for uncertainty analysis, sensitivity analysis and second stage filters.

1.3.10.1 How is effectiveness going to be parameterised?

Cost-utility analysis captures health consequences of interventions through changes in the QALYs or DALYs. Changes in the DALYs are driven by changes in life-years lost and years of life lived with disability. These in turn are driven by changes in transition probabilities due to changes in death, survival and cure rates, disease incidence, remission, and so on, produced by the intervention.

However, the majority of the research evidence on effectiveness of interventions does not use these types of variables as the outcome. For example, much evidence of preventive programme effectiveness uses changes in risk factors, physiological markers or other intermediary outcomes. This requires ‘link models’ to convert research evidence into parameter changes for the economic decision models of BODE\(^3\). The comparative risk assessment (CRA) models used in BDS are one example, whereby changes in risk factors are ‘converted’ to changes in incidence rates.

Such link models will often be needed, and may be intervention specific. Hence, the specification of the intervention, and how it will be parameterised in the model, is a critical first step to informing what research findings are going to be synthesised (and hence the nature of the literature search strategy).

1.3.10.2 General approach to synthesising evidence

A high quality systematic review and meta-analysis may take one analyst a full year. We do not have such resources to apply to each key input variable for modelling. Rather, a parsimonious and efficient approach will be required, guided by principles such as:

- making use of the growing number of international clearing houses for quality-appraised systematic reviews
- utilising evidence synthesis from ACE-Prevention (Australia), the UK National Institute for Health and Clinical Excellence (NICE) and elsewhere
- using other published systematic reviews that meet quality appraisal criteria.

Where there is an appropriate good-quality systematic review available, we may update it, and where necessary tailor it to the NZ context. We will perform meta-analysis, as needed, to determine single point estimates (with variance) for the intervention’s effectiveness. On occasion, we may also use quantitative bias analysis methods to allow for likely residual systematic error in point estimates (ie, likely residual measurement error, confounding or selection bias).
Where there is not a suitable systematic review available, we will conduct an appropriate review de novo.

Given that many of the interventions that will be assessed in BODE$^3$ will not have randomised control trial evidence, there will not always be a strong evidence base (see following section). We may need to seek and quantify expert opinion. In such cases, there will be greater parameter uncertainty about the effect size of an intervention.

1.3.10.3 Strength of evidence, and implications for evaluation

Table 2 below presents the classification of strength of evidence used in ACE-Prevention (Australia), with some minor modifications for application to BODE$^3$ (consideration of confidence intervals rather than p values; mention of quantitative bias analysis). We will use this classification to structure our approach to evidence synthesis. It provides a useful overview of what constitutes evidence, and what constitutes better evidence.

As has been described in the context of the SIGN (Scottish Intercollegiate Guidelines Network) grading system$^{39}$, grading aims to differentiate between results based on strong evidence and those based on weak evidence. This does not reflect the importance of the finding, but is a measure of the accuracy of the estimates. For the policy-maker, the evidence grading will form part of the ‘picture’ in understanding the likelihood that the predicted outcome will be achieved if the recommendation is implemented. Of note, and as found in the ACE-Prevention (Australia) report,$^{11}$ it is often the interventions with level IV, indirect or parallel evidence that have the greatest health impacts and best cost-effectiveness.
### Table 2: Approach to classification of strength of evidence in BODE³. Source: Vos et al (2010) ¹¹

<table>
<thead>
<tr>
<th>Conventional approach based on epidemiological study design</th>
<th>Additional categories to be utilised in BODE³</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evidence* from level I-III study designs</td>
<td>Evidence from level IV studies, indirect or parallel evidence and/or from epidemiological modelling using a mixture of study designs</td>
</tr>
</tbody>
</table>

#### A. 'Sufficient evidence of effectiveness’

*Effectiveness is demonstrated by sufficient evidence from well-designed research that the effect:*

- is unlikely to be due to chance (e.g. 95% CI excludes the null); and
- is unlikely to be due to bias, e.g. evidence from:
  - a level I study design;
  - several good quality level II studies; or
  - several high quality level III-1 or III-2 studies (from which effects of bias and confounding can be reasonably excluded on the basis of the design and analysis, and/or quantitative bias analysis.)

#### B. 'Likely to be effective’

*Effectiveness results are based on:*

- sound theoretical rationale and programme logic; and
- level IV studies, indirect† or parallel‡ evidence for outcomes; or
- epidemiological modelling to the desired outcome using a mix of evidence types or levels.

The effect is unlikely to be due to chance (the final uncertainty interval does not include zero, and there is no evidence of systematic bias in the supporting studies and/or quantitative bias analysis suggests bias an unlikely explanation.)

#### C. 'Limited evidence of effectiveness’

*Effectiveness is demonstrated by limited evidence from studies of varying quality that:*

- the effect is probably not due to chance (e.g. 90% CI excludes the null); but
- bias, while not certainly an explanation for the effect, cannot be excluded as a possible explanation (e.g., evidence from:
  - one level II study of uncertain or indifferent quality;
  - one level III-1 or III-2 study of high quality;
  - several level III-1 or III-2 studies of insufficiently high quality to rule out bias as a possible explanation; or
  - a sizeable number of level III-3 studies of good quality and consistent in suggesting an effect.)

#### D. 'May be effective’

*Effectiveness results are based on:*

- sound theoretical rationale and programme logic; or
- level IV studies, indirect† or parallel‡ evidence for outcomes; or
- epidemiological modelling to the desired outcome using a mix of evidence types or levels.

The effect is probably not due to chance. But bias, while not certainly an explanation for the effect, cannot be excluded as a possible explanation.

Would benefit from further research and/or pilot studies before implementation.

#### E. 'Inconclusive evidence of effectiveness’

*Inadequate evidence due to insufficient or inadequate quality research. No position could be reached on the presence or absence of an effect of the intervention (e.g. no evidence from level I or level II studies, and level III studies are available but they are few and of poor quality.)

#### F. 'No evidence of effectiveness’

No position could be reached on the likely credentials of this intervention. Further research may be warranted.

---

*Evidence classifications based on those of the Australian National Health and Medical Research Council.⁴⁰

I - Evidence obtained from a systematic review of all relevant randomised controlled trials.

II – Evidence obtained from at least one properly designed randomised controlled trial.

III-1 – Evidence obtained from well-designed pseudo-randomised controlled trials (alternate allocation or some other method).

III-2 – Evidence obtained from comparative studies with concurrent controls and allocation not randomised: cohort studies; case-control studies; or interrupted time series with a control group.

III-3 – Evidence obtained from comparative studies without concurrent controls: historical control studies, two or more single-arm studies, or interrupted time series without a parallel control group.

IV – Evidence obtained from case series, either post-test, or pre-test/post-test outcomes.

†Information that strongly suggest that the evidence exists (e.g. a high and continued investment in food advertising is indirect evidence that there is positive (but proprietary) evidence that food advertisement increases sales of those products).

‡Evidence of intervention effectiveness for another public health issue using similar strategies (e.g. the role of social marketing, regulation or behavioural change initiatives in tobacco control, sun exposure, speeding).
**1.3.11 Extrapolating intervention effects over time**

Trials often only measure outcomes over a limited time period, while epidemiological and economic decision models need estimates of the true impact on disease outcomes and costs well into the future. This is a major problem in economic decision modelling. For example, does an obesity prevention programme in childhood alter people’s future BMI trend or track, or does any effect completely disappear in 10 years? Does a new treatment producing 30% lower mortality in the first year result in 30% lower mortality continuously into the future, or only in the first year?

One option is to limit the modelling to the duration of the trial or observational data, but this does not adequately reflect reality (eg, health improvements are likely to be underestimated). The alternative is to make assumptions about the impact beyond the duration of the available evidence (ie, to assume either a continued impact over time, a lessening of the impact over a period beyond the known impact time from underlying studies, or the abrupt disappearance of the impact).

The choice of whether to model future attenuation (or not) of treatment effectiveness will depend on:

- the intervention in question;
- discussions with technical experts; and
- the most plausible way of modelling.

Often, however, there is no clear choice and the solution we adopt is to present results as discrete scenarios using different choices as a sensitivity analysis. For instance, ACE-Prevention (Australia) assumed a best estimate of an annual decay of the impact of GP-mediated physical activity interventions of 50% - but then varied this between 0% and 100% in sensitivity analyses.

**1.4 Advisory committees**

Final autonomy about what BODE³ evaluates, and how BODE³ evaluates it, rests with the Director, co-Directors and named investigators of BODE³. However, BODE³ also has strong collaborative links with the Ministry of Health and the sector, and a range of advisory groups (Figure 5). The Programme Advisory Group (PAG) will meet annually to:

- Review and monitor strategic direction
- Advise on sector interface
- Provide overall advice on process
- Provide specific advice on 2nd stage filters.

The membership of the PAG will include representatives from: Ministry of Health— including National Health Board (NHB); University of Otago; Pharmac; and the National Health Committee when restructured and functioning. At the inaugural meeting of PAG in January 2011, members included:
At this inaugural meeting, advice was received on general issues, second stage filters, and intervention selection criteria. Content of this version of the Protocol reflects that advice.

Figure 5: Programme structure and advisory groups

At the time of writing this version of the Protocol, PAG is acting as the Prevention Intervention Advisory Group.

The Cancer Interventions Advisory Group membership is likely to include:

- National Clinical Director Cancer Control (John Childs)
- National Programme Manager, Cancer (Deborah Woodley)
- Clinical Director Regional Cancer Network (Richard Sullivan)
- Cancer Society (either Dalton Kelly (CEO) or Jan Pearson (Manager))
- Medical oncologist (Andrew Simpson)
- As required based on evolving intervention mix:
  - Mix of primary care, secondary care, palliation, population health, and equity perspectives
  - More specific clinical expertise.
1.5 Other criteria for policy-making, or second-stage filters

The main focus of BODE$^3$ will be technical analysis (epidemiological and economic decision modelling, and quantification (where possible) of uncertainty and equity) and building capacity. The outputs will be academically relevant, but of course we also hope relevant and useful for policy-making in New Zealand (and perhaps elsewhere). Some institutions have formalised the role of cost-effectiveness analysis in decision making (eg, NHS through NICE in the UK; Pharmac in New Zealand). But no institution uses only cost-effectiveness analysis in decision making, and nor should it. There are other considerations, for example: equity (eg, socioeconomic and ethnic, those in poor health compared to those in good health), age (eg, the ‘fair innings’ argument$^{41}$), total cost of the intervention, total burden of disease being addressed, total health impact, capacity of sector to implement intervention, political will, rule of rescue, and societal values.

In New Zealand these other considerations have been formalised by at least three influential institutions: the Ministry of Health, the National Health Committee and Pharmac. The Ministry of Health’s suggested criteria for prioritisation on the “Best Use of Available Resources” (2005) include three high-level principles: effectiveness, equity and value for money. Concepts of whānau ora from He Korowai Oranga$^{42}$ are also adopted, to ensure that Māori health is taken into account. For example, “Whānau ora means considering effectiveness, value for money and equity for Māori from a Māori perspective. It also recognises that prioritisation processes should enable Māori to participate in and contribute to strategies for Māori health improvement, and foster the development of Māori capacity to participate in the health and disability sector.”$^{43}$

Other considerations identified by the Ministry of Health included:

- the acceptability of the proposal, including the degree of acceptability to, and participation by, Māori, other population groups and other stakeholders
- the ethical dimensions of the proposal
- the impact on the sector
- the ability to manage potential risks
- other legislative requirements.

Pharmac$^{44}$ and the National Health Committee (NHC)$^{45}$ also have criteria, however they are similar to, precede, and form the basis of the above Ministry of Health criteria.

The ACE-Prevention (Australia) project developed ‘second stage filters’ that were applied to their main analyses, and have been included in both academic and policy outputs.$^{11}$ The core filters used in all ACE studies are:

- capacity of the intervention to reduce inequity
- acceptability to stakeholders
- feasibility of implementation
• strength of the evidence base.\textsuperscript{a}

The ACE-Prevention (Australia) Project Steering Committee specified additional filters:
• sustainability; and
• potential for other consequences (side effects).

The Indigenous Steering Committee in ACE-Prevention (Australia) also specified two additional filters:
• cultural security; and
• community health gain.

It will be routine in BODE\textsuperscript{3} for ethnic inequalities (and probably socioeconomic inequalities), to be explicitly and quantitatively included in analyses. Other equity considerations (e.g., age, gender, severity of illness) are beyond the scope of ‘routine’ BODE\textsuperscript{3} output in the first instance.

Beyond BODE\textsuperscript{3}’s focus on equity analyses, it is a moot issue how much the research should explicitly consider all these decision criteria. One could argue that it is not the role of researchers, but that of policy institutions and elected representatives. Alternatively, one could argue that as part of dissemination, translation and maximising the impact of research, researchers should actively engage in providing information for all decision criteria.

Following consultation with the BODE\textsuperscript{3} PAG, at the minimum BODE\textsuperscript{3} outputs will provide:
• total estimated cost of the intervention, with uncertainty
• total estimated DALYs averted by the intervention, with uncertainty
• estimates of cost-effectiveness, with uncertainty
• a description of the relative magnitude of disease burden (both DALY and (where relevant) comparative risk assessment (CRA) output) that the intervention is addressing, capitalising on the burden of disease study (BDS) underpinnings of BODE\textsuperscript{3}
• and estimates of ethnic and socioeconomic equity impacts.

\textsuperscript{a} The strength of the evidence base should be reflected in the uncertainty estimates about cost-effectiveness. Nevertheless, a separate itemisation of strength of evidence is often useful in its own right.
2 Selection of interventions to evaluate

The selection of interventions to evaluate is critical. For example, electing very similar interventions for a narrow range of diseases would not achieve the goals of BODE\(^3\) of comparing interventions across the health services and disease spectrum. There are perhaps three over-riding principles to consider:

1. Relevance. Selected interventions should inform decision-making in the next two to five years. This does not mean ‘only select interventions that are likely to be directly considered by policy-makers in the near future’; there will also be a need for good comparators or benchmarks, and the need for academic innovation and agenda-setting.

2. Academic leadership. Academia has a position in society that allows it to lead policy thinking, and propose interventions that may be currently beyond what is considered viable by policy makers and society at large.

3. Academic rigour. The work must be academically rigorous. With respect to selection, this means selecting interventions that can be reliably specified and plausibly parameterised in terms of best estimates and uncertainty. To be clear, this does not mean excluding interventions that don’t have (say) randomised trial evidence.

These three over-riding principles were reinforced by the BODE\(^3\) PAG. Other general considerations included:

- Costs of interventions may rapidly change. Therefore, it may be warranted to revisit evaluations at times in the future.
- Developing infrastructure that allows rapid evaluations is highly desirable (and indeed is a motivation for BODE\(^3\)).
- Initially selecting some evaluations that can be rapidly undertaken as ‘pilot’ cases, perhaps focused around a risk factor or disease cluster.
- Retaining the focus on generality. For example, considering the viability of evaluations of classes of treatments rather than specific treatments per se.

The remainder of this Section considers specific ABC-CBA and NZACE-Prevention criteria and processes for selecting interventions to evaluate, and a brief conclusion on issues to consider when ‘specifying’ the intervention.

2.1 ABC-CBA Selection of Interventions to Evaluate

2.1.1 Cancer intervention selection criteria

Individual interventions for ABC-CBA will first be selected on the basis of three key criteria:

1. Interventions which are likely to have a substantial impact on cost and/or cancer burden, and for which the evidence on programme effectiveness is likely to be in categories A, B, C or D shown in Table 2 (page 26).
   - For interventions not already implemented, weighting will be given to those that are likely to be cost-effective or cost saving
   - For interventions already implemented (i.e., the policy question is potential disinvestment), weighting will be given to those interventions that are unlikely to
be cost-effective compared to standard thresholds or likely to be dominated by (new) alternative programmes addressing the same problem.

2. Interventions that do not meet the criteria above, but for other reasons (e.g., public pressure, political lobbying) are highly likely to require active decision making in the next 2-5 years, and where economic decision modelling is likely to be useful and influential in that decision making process.

3. Interventions satisfying either criteria 1 or 2 will be further weighted for inclusion if they are likely to have substantial equity relevance, where:
   - Dimensions of social group equity – in order of priority – are ethnic (Māori:non-Māori mostly), socioeconomic and regional.
   - Equity relevance includes:
     - Interventions designed for particular social groups (e.g., Māori patient navigators)
     - Total population interventions that address a cancer that varies in incidence, survival or mortality between social groups
     - Total population interventions that are likely to have varying cost-effectiveness between social groups.

The final set of cancer interventions will include:
   - a range of interventions across the cancer spectrum from prevention through to palliation, and across a range of cancer sites;
   - some interventions that can be treated as validation case studies (e.g., drugs previously assessed by Pharmac, prevention interventions also assessed within NZACE-Prevention prevention)
   - some interventions for cancers that specifically address Māori / non-Māori health inequalities
   - for the first round of analyses, clusters of interventions for the same cancer (to facilitate assessment (and validation) of methods during the early stages)

2.1.2 Cancer intervention selection process

The selection process will be primarily undertaken by the BODE³ researchers and the Cancer Interventions Advisory Group (CIAG; Figure 5 page 28). Additional input for interventions to be included will be sought from PAG and others as appropriate.

Members of the Cancer Interventions Advisory Group (CIAG) will be asked to individually identify interventions that meet the above criteria. The actual process of selecting and finalising the interventions will be a series of iterations backwards and forwards between the BODE³ researchers and the CIAG:

1. Initial horizon scanning and preliminary selection. [Completed in 2009/10 as part of the HRC submission. See Table 3 below.]
2. **BODE³ researchers.** Summary, assessment and analysis of the preliminary interventions in Table 3 below against the above criteria. Tighter specification (and options) of the actual interventions.¹ Possible addition of other interventions.

3. **CIAG.** A meeting of the CIAG will be held to discuss the above assessment and analysis. Particular input for Criteria 2 above (ie, likely policy relevant in next 2-5 years). Second round of scanning and suggestion of potential interventions (eg, based on recent clinical conference updates). Revised and prioritised list agreed.

4. **BODE³ researchers.** Final assessment and analysis of intervention selections, and recommendations to CIAG. Consult more widely on this final proposal with PAG and others (as agreed by CIAG).

5. **CIAG.** Minor modifications only. Approval of list. Discuss and agree to role of CIAG as potential consultative body (eg, technical advice on parameters for some interventions) during actual analysis, interpretation and dissemination.

A second round of this selection process above will be undertaken to select interventions for modelling in years 3 to 5 of the programme (ie, the above applies only to the first round of selection).

Table 3: Preliminary selection of cancer control interventions – the ten non-italicised interventions were prioritised and included as ‘preliminary’ in the HRC application.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Intervention</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevention</td>
<td>1. Doubling number of calls to Quitline (for Māori and non-Māori analysed separately)</td>
<td>Cross-reference with NZACE-Prevention. Includes Māori-focused mass media campaign.</td>
</tr>
<tr>
<td></td>
<td>2. Reducing tobacco imports by 10 % per yr for 10 yrs → &lt;2% smoking prevalence.</td>
<td>Equates to assessment of total tobacco-cancer burden. Societal costing necessary.</td>
</tr>
<tr>
<td></td>
<td>3. Ensuring target 80% of smokers receive ABC (Ask, Brief Advice, Cessation).</td>
<td>Good trial of a systems intervention. Will be a challenge to cost.</td>
</tr>
<tr>
<td></td>
<td>4. Low-dose aspirin of cancer prevention</td>
<td>Possible to compare with NZACE-Prevention</td>
</tr>
<tr>
<td>Screening/early detection</td>
<td>5. Colorectal cancer screening programme.</td>
<td>Validation with results from Ministry, and possibly MoDCONZ project (led by Sarfati).</td>
</tr>
<tr>
<td></td>
<td>6. Instigation of prostate cancer screening using PSA testing.</td>
<td>Useful test of ABC-CBA ability to triage interventions. Shifts in disability weights (or utilities) will be important.</td>
</tr>
<tr>
<td>Diagnosis and Treatment</td>
<td>7. 12 mth Trastuzumab (Herceptin) for early breast cancer.</td>
<td>Lots of data, and validation with Pharmac and Ministry.</td>
</tr>
</tbody>
</table>

¹ Specification should include: the target population; specific technologies used; the type of personnel delivering the service or treatment; the site of delivery; whether the service is bundled or piggy-backed with other services; and the timing of the intervention.
9. Ensuring all rectal cancer patients receive surgery in specialist unit/team. | Little scrutiny of effectiveness of surgery compared to drug therapy.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Intervention</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.</td>
<td>As above, but for lung cancer patients.</td>
<td>Ditto.</td>
</tr>
<tr>
<td>Support and rehabilitation</td>
<td>11. Improved adult survivorship care, as per late effects programme used for adolescents.</td>
<td>System intervention. Under consideration.</td>
</tr>
<tr>
<td>Palliation</td>
<td>12. Radiation for prostatic cancer bone pain, including fractionation options.</td>
<td>RCTs emerging on radiotherapy.</td>
</tr>
<tr>
<td>13. Community care (e.g. Liverpool Pathways of Care model)</td>
<td>Currently being considered. Some evidence from UK.</td>
<td></td>
</tr>
<tr>
<td>Pan-Spectrum</td>
<td>14. Scaling up of patient navigators.</td>
<td>Challenging, but current. Trial scenarios that estimate the improvement in DALYs that would be necessary to justify the cost, rather than CEA per se.</td>
</tr>
<tr>
<td>15. Enhanced patient management support and/or IT support (e.g. Map of Medicine)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 2.2 NZACE-Prevention

The selection of interventions in NZACE-Prevention differs from that for ABC-CBA in the following ways. First, domains of priority risk factors were selected. This follows the modelling approach in ACE-Prevention (Australia), which uses CRA approaches, making it most sensible and pragmatic to evaluate interventions by domains. Second, a list of potential preventive interventions were collated for further stakeholder assessment. Unlike ABC-CBA, we did not rely on a specific advisory group for three reasons: the BODE³ researchers have substantial expertise in preventive interventions; the BODE³ PAG also has substantial experience with preventive interventions, and we drew on this advisory group as required rather than a specific prevention advisory group; and the previous experience of ACE-Prevention (Australia) researchers (e.g., Vos, Barendregt, Cobiac) included in the NZACE-Prevention team gives a stronger departure point.

#### 2.2.1 Risk factor selection

The process of selection of risk factor topics for NZACE-Prevention is described in detail in an online Background Paper on Risk Factors (see Report under Publications at [http://www.uow.otago.ac.nz/BODE3-info.html](http://www.uow.otago.ac.nz/BODE3-info.html)).

To be considered, a risk factor had to be in the top 15 for causing lost DALYs for high-income countries (from recent WHO work on DALYs). The following steps were applied to further select and prioritise the various risk factors:

- The risk factor had to be amendable to at least one preventive intervention for which there was a good evidence-base for effectiveness and likely cost-effectiveness.
- The risk factor had to contribute to health inequalities in the New Zealand setting in terms of the gap between Māori and non-Māori.
The risk factor was given less priority if study of the effectiveness and cost-effectiveness of preventive interventions would be particularly demanding because of the need for complex new burden of disease data.

The final prioritised risk factor list is shown in Table 4 below. For further details see the separate report on the BODE³ website (http://www.uow.otago.ac.nz/BODE3-info.html).

Table 4: Final prioritised list of the top six major risk factors for evaluation in NZACE-Prevention

<table>
<thead>
<tr>
<th>Risk factor (prioritised order)</th>
<th>Rationale and comment</th>
<th>Relevance to ongoing research around cancer control (ABC-CBA)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Highest priority</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tobacco use</td>
<td>A major cause of disease burden and especially of inequalities in the NZ setting(^6)</td>
<td>High overlap given the number of tobacco-related cancers</td>
</tr>
<tr>
<td>High blood pressure</td>
<td>A more important cause of lost DALYs than cholesterol, contributes to inequalities, and many effective interventions are available.</td>
<td>Small overlap: salt intake is a joint risk factor for high blood pressure and stomach cancer</td>
</tr>
<tr>
<td>High cholesterol</td>
<td>This risk factor was upgraded in priority because interventions appear more promising than for most other risk factors in this list (and there is some overlap with the blood pressure interventions if an absolute risk approach is adopted eg, for considering a polypill intervention).</td>
<td>Nil</td>
</tr>
<tr>
<td><strong>Medium priority</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol use</td>
<td>This risk factor is important but is complex to study as there are over 200 ICD-10 three-digit disease codes in which alcohol is part of a component cause. Intervention analyses therefore should follow the completion of the NZ Burden of Disease Study revision.</td>
<td>High overlap given the number of alcohol-related cancers</td>
</tr>
<tr>
<td>Overweight and obesity</td>
<td>An important risk factor, but there is uncertainty around the persistence of intervention effects.</td>
<td>Some overlap given that obesity is a risk factor for some cancers</td>
</tr>
<tr>
<td>Physical inactivity</td>
<td>An important risk factor but the possible impact on health inequalities is indirect and there are uncertainties around the persistence of intervention effects (especially for interventions applied to children).</td>
<td>Some overlap given that physical inactivity and obesity are risk factor for some cancers</td>
</tr>
<tr>
<td><strong>Lower priority</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low fruit and vegetable intake</td>
<td>This risk factor is ranked relatively low as past work may have over-estimated the benefits of its reduction given the findings in a recent and very large cohort study.</td>
<td>Modest overlap given a potential role for low fruit and vegetable intake as a risk factor for certain cancers</td>
</tr>
<tr>
<td>High blood glucose</td>
<td>This risk factor is of relatively lower priority given that interventions addressing blood glucose directly are not particularly cost-effective. Also this</td>
<td>Nil</td>
</tr>
</tbody>
</table>

\(^6\) We note however, that ACE-Prevention (Australia) work on tobacco control intervention modelling is unlikely to be completed before mid-2012.
<table>
<thead>
<tr>
<th>Risk factor (prioritised order)</th>
<th>Rationale and comment</th>
<th>Relevance to ongoing research around cancer control (ABC-CBA)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>risk factor will be partly addressed by considering other risk factors eg. “physical inactivity”, “overweight and obesity” (see above) and possibly vegetable intake.</td>
<td></td>
</tr>
</tbody>
</table>

### 2.2.2 Specification of interventions, and groupings of interventions, for NZACE-Prevention analyses

The process for selecting potential interventions for NZACE-Prevention modelling work is detailed in the *ACE-Prevention Interventions Report*. To summarise, the list of potential interventions was based on considerations of likely effectiveness, likely cost-effectiveness and likely potential to reduce health inequalities. In particular they arose from the following:

- Details from published ACE-Prevention (Australia) interventions (especially the September 2010 Report). Relevant aspects of ACE-Prevention (Australia) work have also been published in the journal literature in such topic areas as: alcohol use, overweight and obesity (particularly for children/adolescents), skin cancer, pre-diabetes and physical inactivity.
- Knowledge of the ACE-Prevention (Australia) Team members (particularly in the area of tobacco control) and research around innovative edges of tobacco control.
- Knowledge of the ACE-Prevention (Australia) Team members in pharmacological interventions and pharmacoeconomics.
- Frontier scanning for innovative interventions in cardiovascular disease prevention, particularly from Scandinavian countries (Google Scholar and Medline searches).
- Consideration of the literature around the co-benefits to health from climate change interventions (eg, reference) which are likely to benefit cost-effectiveness from a societal perspective.
- Within-team review of the draft lists of possible interventions.

### 2.2.3 Process for Stakeholder Critique

Stakeholders (including PAG) were asked to critique about 20 specified interventions within each domain of the three highest priority risk factors: tobacco, high blood pressure and high cholesterol. Stakeholders were particularly asked to focus on the *relevance* of introducing the intervention in the NZ setting as the key criterion for stakeholder consideration. In particular, stakeholders were asked to:

- Tick the **top five most relevant interventions** (next 5 years in NZ), for each of the three domains of risk factors.
- Cross out the **five least relevant interventions** (next 5 years), within each of the three domains.
- Add **critical comment** if they thought there were problems with the argument for considering a particular intervention.
• Provide new intervention ideas.

Stakeholders found the task challenging. Feedback included a suggestion to merge some of the interventions across risk factor domains, consistent with many of the risk factors addressing CVD and an absolute risk approach to CVD. At the time of writing Version 1.0 of the Protocol, work on revising the selected preventive interventions is still proceeding. An initial focus on salt reduction interventions has been selected for a range of reasons, including being a suitable domain in which to train staff in the use of ACE multistate lifetables.

2.3 Specification of the Intervention

In order to determine health consequences and costs of an intervention, the intervention needs to be clearly specified. Key steps include:

1. Retrieving the relevant published studies (including those relevant to effectiveness, cost and cost-effectiveness; section 3).
2. Summarising the basic activities for the intervention and key design issues in each of the key papers. (Note that there is a strong link here with the event pathway specification needed in costing (Section 4.4)).
3. Specifying that version of the intervention where ‘best evidence’ exists or those which are most appropriate in the New Zealand health service context.

There may be many versions of an intervention evaluated in the literature. It is beyond the scope of BODE³ to perform extensive evaluations on different versions of a single intervention. Thus, the ‘most promising’ or ‘prototypical’ version of the intervention needs to be identified from the outset – namely that version with the best evidence base and which is the most appropriate from a policy standpoint. The weight put on these two factors can be somewhat flexible, but the intervention must be modelled to fit into the current New Zealand health service structure. An example might be where the intervention in question uses pharmacist prescribers to provide a service, but in New Zealand only prescribing by GPs would be acceptable. In this instance, we might model our intervention with GPs, but still use the efficacy data from the trial, along with documenting support for the assumption that GPs are as efficacious as pharmacist prescribers for the task.
3 Assessment of health consequences of interventions

Effect sizes of interventions investigated in BODE³ will be derived from existing literature whenever possible (evidence levels I to III of the Australian National Health and Medical Research Council [40]). However, many interventions will not have been subjected to rigorous evaluation, especially non-clinical interventions. When needed, BODE³ will consider less robust levels of evidence, such as case series (level IV studies), indirect or parallel evidence and/or epidemiological modelling (see Table 2 page 26). We may at times need to involve expert opinion if existing data are not sufficient. Importantly, BODE³ will specify the uncertainty surrounding such inputs, and additionally identify lack of good evidence as an issue in a brief second-stage filter stage.

This Section provides a brief outline of the methods that will be used (and need further developing) in BODE³ for assessing health consequences of interventions, in three parts:

1. Higher quality or levels I to III evidence
2. Indirect and parallel evidence
3. Expert opinion.

3.1 Levels I to III: Literature synthesis

The first choice will be to use existing systematic reviews (including meta-analyses and health technology assessments [HTAs]) where an appropriate, high-quality review exists. This approach is largely pragmatic. Systematic reviews are highly time and resource intensive. This would not be the best use of the project’s resources when organisations such as the Cochrane Collaboration, the Campbell Collaboration and other HTA bodies have the resources to produce systematic reviews to very high standards with many internal quality assurance and peer review procedures and methods aimed at reducing bias.

Use of existing systematic reviews is in line with the protocol of ACE-Prevention (Australia). Further, Pharmac recognises high-quality meta-analyses and systematic reviews of RCTs as being the highest level of evidence.

Appropriate high-quality systematic reviews will be updated using a time-limited search strategy. Where there is no existing appropriate systematic review, a de novo systematic review will be conducted following the literature search and selection procedures outlined below.

3.1.1 Methods and steps for undertaking a systematic review

A systematic review aims to address a specific research question by evaluating all evidence that fits pre-specified eligibility criteria. Bias is minimized by using explicit, systematic methods. To meet good standards for systematic review, BODE³ will whenever possible
adhere to the following principles when compiling de novo systematic reviews, and when updating existing systematic reviews:

- The research question will be clearly defined in advance.
- Literature will be identified according to an explicit search strategy
- Literature will be selected according to predefined inclusion/exclusion criteria
- Literature will be evaluated for quality using consistent methodological standards.
- Data will be collected systematically from the included literature.

The ideal is to adhere to the predetermined protocol (in terms of the search strategy, the inclusion/exclusion criteria, and the critical appraisal tools) but this may not always be possible or appropriate. Any deviations from the protocol will be documented. It is acknowledged that there will be instances where the above strategy will not provide sufficient evidence to derive an effect size, for instance, where evidence for an intervention is only in grey literature. In such cases, an alternative appropriate strategy will be explicitly defined, and may include expert opinion (see Section 3.2).

The steps to be followed to produce, or update, a systematic review are briefly listed below:

- Precise description of the research question
- Creation and recording of search strategies, including language and time limits etc.
- Searching of systematic review/HTA websites and literature databases (and other sources as needed) for relevant literature:
  - existing systematic reviews
  - literature to update existing systematic review, or create a de novo systematic review
- Expansion of the search by snowballing (identification of other useful articles from reference lists of retrieved articles)
- Application of inclusion and exclusion criteria to identify relevant papers for possible inclusion in the review
- If sufficient evidence is available, application of the hierarchy shown in Table 2 to prioritise best evidence
- Critical appraisal of identified literature
- Systematic collection of data.

This can be conceptualised as an iterative process to which a common research question is applied. The first iteration will be to identify existing systematic reviews from HTA/systematic review sites and literature-indexing databases. If a systematic review is available, the inclusion/exclusion criteria will be applied to determine that it is relevant, and it will be subjected to critical appraisal (if the systematic review is not from a quality-assured source). Subsequent iterations will identify relevant literature to update the systematic review, or to conduct a de novo systematic review if there is no appropriate existing systematic review.
Critical appraisal is a key aspect of this process for several reasons: (i) critical appraisal will determine whether any existing systematic reviews are of sufficient quality to be used to contribute to determination of the effect size; reviews that do not meet quality criteria will be disregarded; (ii) all relevant literature identified for possible inclusion in writing or updating systematic reviews will be subjected to critical appraisal to determine their quality and reliability, and the strength of evidence that they provide (as per Table 2); where sufficient data exist, papers of lower quality or providing less strong evidence may be disregarded; (iii) it is also expected that the process of critical appraisal will help identify areas where there is uncertainty around parameters (e.g., large variation in reported values, differing of opinions between experts, differential response between patient subgroups etc), and additionally may identify issues that should be raised in the second-filter stage.

3.1.1.1 Possible development of rapid search methods

With the advent of extremely quick and comprehensive internet search engines, and the availability of most research through the web, it may be possible to develop and use novel rapid search methods. For example, using Google Scholar.

Required methodological development 6: Assess rapid search methods for possible unitisation in BODE3.

3.2 Lower level evidence synthesis

The existence of a body of higher-quality evidence is not a prerequisite for an intervention to be evaluated in BODE3. Rather, interventions are also selected for their policy and equity relevance (see Section 0). The exact methods of assessing lower level evidence will vary depending on the context and intervention. This Section is therefore more about the principles than the exact method.

3.2.1 Level IV evidence

Level IV evidence is that obtained from case series describing post-test, or pre-test/post-test outcomes. This information will be identified by the literature searches described earlier in this Section. Where stronger evidence exists, level IV evidence will be given less weight or disregarded. However, where no stronger evidence is available, level IV evidence may contribute to the criteria for an intervention being considered to be “likely to be effective” or “may be effective” (see Table 2). These terms reflect the lesser certainty around data from level IV findings.

Case series that provide level IV evidence involve a single group of people exposed to the intervention of interest. The individuals within the group may have different characteristics with regards to demographics, disease severity and prognosis etc. In a post-test report, only outcomes after the intervention is given are recorded, so no comparisons can be made. With a pre-test/post-test design, outcomes are compared before and after the intervention is introduced. Such studies are not well controlled, and may be more susceptible to bias (particularly confounding and selection biases) and chance findings. Thus, level IV evidence
must be considered in the context of other supporting factors such as theoretical rationale and logic.

3.2.2 Indirect and parallel evidence
Along with level IV evidence, indirect and parallel evidence may contribute to the evidence for effectiveness of an intervention (see Table 2). Indirect evidence is information that strongly suggests the intervention is likely to be effective, but intervention-specific evaluations have not been done, or at least not in the public arena. For instance, the food industry’s willingness to invest in gaining the Health Foundation Food Tick provides indirect evidence that this strategy increases sales of their products. Parallel evidence is provided when effectiveness is established for an intervention that uses similar strategies. For example, the proven effect of tobacco tax and alcohol tax in reducing harm to health provides parallel evidence that an unhealthy food tax may be effective.

The quantification (including uncertainty) of interventions with only indirect or parallel evidence will be context specific. The rationale and workings will be clearly documented.

3.2.3 Expert opinion and consensus
Given the nature of the interventions being investigated in BODE³, it is anticipated that for some models expert opinion will need to be drawn upon. Expert opinion is a legitimate source of information for modelling parameters when there is not sufficient data from other valid sources.

Guidelines have recommended various methods including Delphi, modified Delphi and nominal group techniques. As has also been emphasised by others, the methods used in BODE³ will not force a consensus because we want to capture the variability in opinions and investigate (through uncertainty analysis) the impact that this diversity has on the model outcomes. The simplest way to look at this is to average the parameter estimates from the individual experts, and use the lowest and highest estimates to represent the variance for uncertainty analysis. It would also be possible to use the estimated extremes to test ‘best’ and ‘worst’ case scenarios in sensitivity analyses. However, these approaches provide little information on the distribution of the uncertainty surrounding the estimate. It may be possible to determine a statistical distribution to the range of values provided (e.g., using bootstrap methods), and test uncertainty e.g., with probabilistic sensitivity analysis.

There is no single consensus on the optimal number of experts from whom opinions are elicited, but existing literature tends to suggest that 8–20 subjects is reasonable. A larger number than this is likely to incur extra expense and time with no significant change to the findings.

There is also no consensus on the exact criteria that should be applied in selecting experts, and is to a large extent a matter of judgement by the principal investigators. The subjects being knowledgeable in the topic area is essential, but may not be sufficient in itself. It has been suggested that subjects should also be respected opinion leaders, be from diverse
practice settings and geographical areas, and ideally should include representatives of the key policy-makers and stakeholders in the topic area.\textsuperscript{70,73} For BODE\textsuperscript{3}, diversity of experts is considered a key requirement, with focus on incorporating different stakeholder viewpoints, eg, clinicians, Ministry of Health, DHB, and Māori Health advocacy.

The methods used to identify experts and methods used to elicit their opinions must be documented clearly.

\textit{Required methodological development 7: Development of approach to expert opinion as a source of information on input parameters and their uncertainty.}
4 Assessment of costs

In Section 1.2.5 we stated that the study perspective is (principally) that of the health system, involving C1, C3, S1 and S3 costs as per Drummond’s framework. This Section provides more detail on the approach to costing that will be used in BODE³, in both NZACE-Prevention and ABC-CBA.

4.1 Overview of costing methods

- BODE³ will treat intervention costs (C1, C2 and C3) and downstream cost offsets (S1 and S3) separately.
- Intervention costs will primarily be estimated by standard activity costing methods using event pathways and patient flowcharts, but other macro- and micro-costing methods will be used as required.
- Cost offsets will primarily be estimated by a top-down approach. This requires a source of total or average costs for each disease state by sex, age, and, if possible, ethnicity and deprivation (and possibly cancer stage or severity). Our options and preferred approach are still under investigation (see section 4.5).

BODE³ will measure costs in economic terms, i.e., the market costs of the resources consumed, as a substitute for full and proper measurement of opportunity costs. This approach involves 3 basic steps:

- Identifying what costs are to be included
- Measuring the resources consumed (or saved) with and without the intervention
- Valuing these resources.

Costs are to be measured in ‘real’ dollars. That is dollars in costs and prices at a specified date. Costs will be valued at the ‘market value’ of the resources involved. Therefore, if ‘purchase cost’ data that is more readily available for actual costing deviates from the market value (e.g., because of subsidies on GP consultations or drugs) the cost should be adjusted to equate more closely to the ‘market value’ (e.g., by adding the subsidy amount to the purchase cost). Where there is no market to provide market prices, and the cost is important for overall estimates of cost-effectiveness (e.g., > 10% of net cost), sensitivity analyses about the market value of the item will be undertaken.

Importantly, the intervention will be modelled as part of the current New Zealand health system. Even though the project reference year is 2006, from a costing perspective, this simply means that all prices are expressed in 2006 dollars. The organisation of health services however must reflect current practice. For instance, if a drug has been recently listed on the Pharmaceutical Schedule, but was not listed in 2006, then this drug should be priced as listed on the Schedule with prices deflated back to 2006. Where subsidies on doctors’ visit and prescription fees, for example, have changed since 2006, current subsidy levels will be applied.
The direct costs of the intervention will be calculated by standard activity costing methods based on creating event pathways and patient flowcharts, as was done by ACE-Prevention (Australia) (section 4.4). A top-down method of costing will be used for valuing cost offsets (section 4.5).

Ideally, direct costing of the intervention will be disaggregated by sex, age, ethnicity, and deprivation and disease state when costing the intervention. However, this may either be challenging to implement (e.g., inadequate data) or conceptually erroneous (e.g., where an intervention is developed for the whole population, and cannot be divided across individuals).

Ideally, cost offsets will also be disaggregated by population heterogeneity. This seems a feasible goal, assuming analyses and output from HealthTracker (see below) meets our expectations. If HealthTracker is not sufficiently robust for costing by (say) ethnicity and deprivation, estimates by just sex and age may be preferable. Furthermore, sex by age cost-offsets only may be preferred for equity reasons. For example, HealthTracker data may find lower cost offsets for Māori with a given disease than should be the case, if utilisation relative to need is lower among Māori.

Required methodological development 8: Determine when direct costing should and can be undertaken separately by sub-populations.

Required methodological development 9: Determine when calculations of cost-offsets should and can be undertaken separately by sub-populations.

4.2 HealthTracker

HealthTracker is a New Zealand data tool that has become available relatively recently. The unique patient identifier (NHI – the National Health Identifier) now has coverage of at least 98% of the New Zealand population. HealthTracker links, by means of the NHI, health-care events occurring to any individual. The data are validated against PHO registers every 3 months. Reasonable data are available from July 2006, but are better from 2007 onwards. It includes only costs to the Government, and most is claiming data where the Ministry of Health pays for the service. The database contains information on those with an NHI number who have had any contact with health services in the last year. It is estimated that data are missing for about 37,000 people.

Costing data included in HealthTracker includes:

- hospital costs paid by the Ministry of DHBs (case mix cost weights)
- outpatient costs (contracted purchase units)
- GP visits (average capitation cost only, using enrolled capitation costs and funding formula (that includes ethnicity, NZDep, sex and age))
- general medical subsidy for visits to GPs outside of enrolled PHO
- client contract processing system that pays for Disability Support, and Health of Older People (HOP)
• emergency department triage level contracted purchase unit cost for event (ie, costs do differ by seriousness of presentation)
• community pharmacy, and more recently hospital pharmacy costs (excluding non-subsidised medications)
• lab tests funded by Vote:Health.

HealthTracker can track health-care events ‘cross-sectionally’, within a given year or other time-period, or ‘longitudinally’ linking to NHI numbers for earlier years in which NHI coverage was reasonably complete. Alternatively, for cancers, links can be made to the year in which the patient was diagnosed with cancer and was added to the Cancer Register. Note that Cancer Register data are generally robust from about 1996 onwards (after registration procedures were reformed in 1994).

HealthTracker is potentially a useful tool for improving cost data coverage, but what advantages it offers over other cost sources is still to be determined. Scoping the full potential of HealthTracker, and determining how it can usefully contribute to epidemiological and economic modelling, is an important aspect of BODE³.

Required methodological development 10: Scope and assess the validity and utility of HealthTracker for costing in BODE³.

(If HealthTracker proves to be insufficient for determining cost-offset data, then Australian data will be used as a starting point, with some adjustment to the New Zealand context.)

4.3 Criteria for inclusion and exclusion of costs

As previously stated in Section 1.2.5, the perspective is primarily that of the health system, including costs borne by the health sector and patients/families for both publicly funded (Vote:Health) and privately funded healthcare. The perspective may be broadened in specific cases where excluding other costs would substantially misrepresent the value of the intervention (see Section 1.2.5). The perspective we adopt is close to that used by Pharmac,⁴⁴ and the ACE-Prevention (Australia) programme.⁴⁹

Costs to patients for private healthcare and out-of-pocket payments for visits to health professionals, pharmaceuticals and other miscellaneous expenses will be included (where practicable) for both intervention costs and cost offsets. However, out-of-pocket expenditure on over-the-counter medications and alternative providers are excluded. An exception to this is where over-the-counter medication is a core component of the intervention (eg, widespread prophylactic use of low-dose aspirin purchased over-the-counter).

Direct costs of patient travel (eg, petrol) but not time are in scope.

Unpaid caregiver costs for time spent caring for the patient will not routinely be included due to both being outside of the default study perspective and logistical constraints.
Both related and unrelated health care costs are included in cost-offsets (S1 and S3; see Section 4.5 below).

Productivity costs for either the patient or unpaid caregiver will not be included. The primary reason for excluding productivity costs is that they are outside of the study costing perspective. Furthermore, there are several problems with including productivity costs. One is that there is disagreement among experts about whether there is an element of ‘double-counting’ in constructing such estimates (if, for instance, income effects are taken into account in ‘health-related quality-of-life’ measurements). Also, there is uncertainty about the best way of valuing any ‘lost contribution’. If a person is unable to work or prematurely dies due to illness, the actual production loss for society from sickness is likely to be much smaller than the estimated value of potential production lost because a person’s work may be covered by others or made up by the sick person on his/her return to work. If the worker needs to be replaced, a previously unemployed person may fill the vacancy. In our view, this strengthens our default position of not including productivity costs.

Costs associated with non-adherence will be included. The non-adherence rate is important to the incremental cost-effectiveness ratio because the participants who don’t adhere to the intervention would be expected to incur some costs but receive little or no health benefit. Information needs to be sought on the likely subsequent health care costs of non-adherers. In the absence of such information, it will be assumed that the non-adherers incur part of the intervention costs, receive no benefit and have the same subsequent health care costs as those currently not receiving the evidence-based intervention.

Other non-economic costs, such as those associated with pain and suffering (SO CALLED ‘intangible costs’) are to some extent captured in the DALY and will not be separately included.

All set up and ongoing running costs (over and above current practice) of the intervention will be included in direct costing, from the point in time of a decision being made to implement the intervention by Government. This means that:

- Research and development by universities and private industry are not included (although they may be factored in the purchase price of, say, a drug).
- Initial training of staff to administer the new intervention is included, and ongoing training of staff is included if it is not absorbed back into ‘routine’ continuing professional development and staff training. (There is some inescapable arbitrariness here, and thus clear documentation will be required.)
- Additional equipment or other resources needed to implement the intervention will be included.
- The costs of developing and enforcing regulations (say) will be included if the activity post-dates a Government decision in principle to implement a legislative or regulatory change.
As a principle, incremental changes in overhead costs resulting from the intervention will be included. Note that DRG costs already include overheads. For other direct costs, it will be necessary to determine if overheads are already factored into the cost, or not. If not, an overhead rate will need to be applied. One way to allow for overheads is to add a percentage amount to salaries. The ACE-Prevention (Australia) protocol uses a 60% overhead for new staff, and a 30% overhead for existing staff – the latter including just salary on-costs (eg, superannuation, ACC, etc), not other overheads, on the premise that there is less overhead opportunity cost for existing staff.

Required methodological development 11: Determine BODE³ practice for the costing of overheads.

Even when a health sector perspective is taken, the revenue gained from a tax intervention (eg, higher taxes on alcohol, tobacco or unhealthy food) will not be included in the costs because it is a transfer payment. This is even so for “dedicated” taxes that are specifically returned to the health sector. This is because it is assumed that the overall funding of the health sector will be adjusted down to balance any new dedicated taxes. Likewise, income support payments such as Sickness or Invalids' benefits are not included, as they are 'transfer payments'. Finally, we do not intend to include so-called ‘dead-weight costs’ of any tax increases required to fund an intervention.

4.4 Intervention Costs

Costs of interventions will be costed directly. The default position is to use activity costing, with event pathways and patient flowcharts. Macro- and micro-costing will also be used as appropriate, depending on the individual intervention. For instance, macro-costing is appropriate for discrete events such as doctor’s visits and hospital stays. Micro-costing may be used when sufficient detail is available, such as individual patient data from a clinical trial.

The specific costs included will vary widely by intervention; for instance, some cancer interventions may require hospitalisation, while other preventive interventions may primarily involve media campaigns. Following standard costing methodologies (including the approach of ACE-Prevention (Australia)⁴⁹), pathways will be constructed to describe the major components of the intervention, and the activities that occur within each component. The detailed specification of the activities for the average client provides the foundation for identifying data needs and to decide how data will be organised and collected. To avoid bias, the process for identifying costs should be identical for the intervention and comparator. Well-defined steps for both the intervention and comparator will be constructed as follows:

- an ‘event pathway’: “who does what, to whom, when, where and how often?”
- when required, a ‘patient flowchart’ that describes how we get from the target population to those who actually participate in the activities
- when required, a ‘practitioner recruitment flowchart’ that describes how we recruit and train the practitioners who will provide the intervention
Through these pathways, we will identify what events (activities) happen, the probability of each event occurring for an “average patient”, and how many times each event occurs. Through this process, the resources consumed are identified, and quantity consumed is multiplied by the relevant unit cost. The unit of resource may be either a specific item/service or a macro-costing item such as a doctor’s visit or hospital admission, depending on which is more practical and appropriate to measure.

Costs can be described as falling within the following 5 broad expenditure categories.

- salary and wages
- capital (land, buildings and equipment)
- consumables
- overheads
- other.

Costs within each of these categories above will be included, as appropriate, in costing the interventions. The type of direct costs will vary widely by intervention, but are likely to include costs arising within some or all of the following areas:

- Programme-based interventions:
  - Programme development, marketing and implementation
  - Programme administration, space and utilities
  - Identification and enrolment of eligible participants (including screening)
  - Recruitment and, when appropriate, training of providers
  - Key intervention elements (such as advice, consultations, care, change in legislation or regulations)
  - Ongoing monitoring of programme performance
  - Downstream effects
  - Assessment of patient outcomes.

- Treatment-based interventions:
  - Diagnostic tests and procedures
  - Key intervention elements (such as drug treatment, surgical or other procedures, patient services)
  - Patient care within hospital-based services (eg, outpatient consultations, inpatient care), and/or in the community (GP, nurse and other healthcare professional consultations)
  - Monitoring and treatment of both short- and long-term adverse effects
  - Routine systems to improve adherence
  - Monitoring of treatment effectiveness and patient response.

The individual costs within these categories to be included will be defined separately for each intervention.
A cost may be excluded on protocol grounds (not relevant to the perspective) or practicality grounds (if it cannot be easily attained) when it is considered very unlikely that it will contribute significantly to the total cost, and its omission will not substantially bias the results (ie, it is not of importance at the margin). For each intervention studied, identified costs that have been excluded will be specified and justified. The same criteria will be applied to both the intervention and its comparator(s) to avoid bias.

4.5 Cost Offsets from Disease Averted/Incurred

As stated in Section 1.2.5, BODE³ will include both related and unrelated health system costs in costs averted/incurred (ie, S1 costs as per Figure 1, page 6).

Unrelated costs include costs from future admissions and treatment of diseases not targeted by the intervention, but arising due to prolongation of life. This has been a contentious issue, although there seems to be a growing theoretical consensus to include both related and unrelated costs. Moreover, it is often empirically difficult to determine what is related and unrelated. For example, presumably only some of future heart disease treatment costs can be attributed to a smoking intervention, as heart disease is due to multiple risk factors.

Van Baal et al (2011) provide a useful framework that can be applied to BODE³. They define types of cost as follows:

\[
\begin{align*}
\text{oc} &= \text{direct costs of the intervention (C1 costs as in Figure 1)} \\
\text{ac} &= \text{average health care costs, by sex and year of age} \\
\text{dc} &= \text{average health care costs in last year of life, by sex and year of age} \\
\text{sc} &= \text{survivor average health care costs, by sex and year of age (ie, for all but last year of life, and having excluded costs incurred in last year of life)}
\end{align*}
\]

The most commonly used ICER in preventive interventions for intervention y compared to intervention x, and as used in ACE-Prevention (Australia), is:

\[
ICER1 = \frac{\text{oc}_y + \text{ac}_y - \text{oc}_x + \text{ac}_x}{\text{DALYS averted } y - \text{DALYS averted } x}
\]

An improvement on ICER1 that allows for costs in last year of life is:

\[
ICER2 = \frac{\text{oc}_y + \text{sc}_y + \text{dc}(y) - \text{oc}_x + \text{sc}_x + \text{dc}(x)}{\text{DALYS averted } y - \text{DALYS averted } x}
\]

This could be calculated for NZACE-Prevention using HealthTracker data (as per below). For ABC-CBA, it could also be calculated both during and beyond cancer diagnosis states. If as in NZACE-Prevention the remaining years of life lived is just an expected value, one would need to assume that the last year of this incurred the increased death costs. Such a method will not be as accurate as a full microsimulation model that tracks people to death through a Markov model, but it should be sufficient.

Regarding the data to calculate these component costs, BODE³ will use a top-down method to estimate downstream costs incurred or averted due to the effect of an intervention on
future disease incidence, duration, severity, sequelae and survival. Broadly, we will follow the methods of ACE-Prevention (Australia), who created a grid of treatment costs for disease states (by age and gender) derived from AIHW Disease Costs and Impacts Study (DCIS) data. The disease cost per prevalent case for each year was calculated by inputting the DCIS data for the year as the numerator, and the number of prevalent cases for that year from the Australian BoD study as the denominator. This was applied to each year lived by any individual with the disease of interest. In some instances where prevalence was hard to define (eg, the prevalence of a wrist fracture) and costs were (almost) exclusively clustered around an incident event, ACE-Prevention (Australia) would apply a cost per incident case and apply this as a one-off cost in their models.

Regarding New Zealand, we do not have an equivalent of the AIHW top-down costing data. However, the existence of routine health data in New Zealand with over 70% of Vote:Health attributed across all citizens (ie, HealthTracker) creates an exciting opportunity for further development of a top-down method. We need to establish whether HealthTracker can provide similar costs by disease states as AIHW DCIS. The final method of generating New Zealand costs is still pending. Options include:

1. Simple tabular extractions of average annual cost per person, by state (ie, sex by age [and possibly also by ethnicity by deprivation]), from HealthTracker
2. If 1 above is unstable for some states/strata, use regression methods to ‘smooth’ the cost data.
3. Incorporate the AIHW DCIS data as a ‘prior’, and use Bayesian regression methods to estimate cost-offsets.
4. If all else fails, simply default to Purchasing Power Parity- and inflation-adjusted AIHW data – perhaps also with adjustments for known major variations in health costing between Australia and New Zealand (eg, pharmaceuticals).

Required methodological development 12: Select best option for generating cost offsets in BODE³.

### 4.6 Methodological Considerations

#### 4.6.1 Assuming the CPI applies to future health care costs

The reference year for cost values is 2006, in line with the burden of disease data being for the 2006 population. Cost beyond 2006 will need to be CPI adjusted to 2006 real costs. Historically, the health inflator has increased at a greater rate than the CPI in New Zealand. However, with recent strategies to improve health sector productivity, it is not clear whether this trend will continue into the future; our proposed method, therefore, is to just use the CPI for all costs (ie, health sector and non-health sector).

Note that there are some exceptions to the above assumption of the CPI applying to future costs that need to be captured if possible. For example, Pharmac may bind drug companies into extended-period fixed-price contracts (eg, 5 years), meaning the price of these drugs in
real terms will actually decrease over that time. Also account should be taken of drug and other technological prices falling precipitously when they come off patent.

4.6.2 Discount rate
A discounting rate of 3% p.a. will be the default for both costs and benefits (see section 1.3.7).

4.6.3 Goods and services tax
Goods and Services Tax (GST) is a ‘transfer payment’ rather than a payment for some good or service. It will therefore be excluded from all costs used in the cost-effectiveness analysis.

4.6.4 Annuitisation of capital costs
In economic evaluation, we do not use the accounting concepts of depreciation. Capital costs can be either annuitised or not. The results of the analysis are not affected. Annuitisation will be applied in BODE\(^3\) only when, and if, required. Of note, ACE-Prevention (Australia) annuitised the costs of constructing a new fluoridation plant, but did not apply annuitisation in most models.

The standard annuitisation formula is \(^1\):

\[
E = \frac{K}{(1 - (1 + r)^{-n}) / r}
\]

Equation 1

The equivalent annual cost (E) = capital (K) divided by the annuity factor: \((1-(1+r)^{-n})/r\), where:

- \(E\) = annual cost
- \(K\) = capital
- \(r\) = interest rate (ie, discount rate)
- \(n\) = number of years over which capital depreciated

In economic analyses, the discount rate is the ‘interest rate’ to be used. The number of years is the expected working lifetime of the asset (or the number of years over which the capital is to be valued). This formula assumes that the asset is new and has no resale value.\(^h\)

The alternative to annuitisation is to fully cost the capital in the year of purchase. The expected lifetime of the capital item determines when a replacement will need to be purchased. The capital cost is discounted if the purchase occurs in the future. This method may be preferred from an intuitive point of view because it can be more clearly seen when

\(^h\) The formula when the asset is not new or has a resale value is given in Drummond et al. (2005), pg 75.
funds have been required for capital, and given the discount rate is also the interest rate in the annuitisation formula no mathematical difference results.

4.6.5 Marginal vs Average Costing

Average costs will be used for stand-alone, mutually exclusive programmes. Marginal costs will be used as appropriate for scaling up or down of interventions (when there is no substantive change in fixed costs), and for interventions that occur in series. For example, where the eligibility criteria are changed and 100 extra people are eligible in addition to an initial 1,000 people, the costs to be used should be ‘marginal costs’ for these 100 people not the average cost (of either the initial 1,000 or total 1,100). The marginal cost excludes any ‘fixed cost’ component in the market price of a good or service. However, this principle should not be applied too inflexibly. For instance, drug prices often include a substantial R&D component. The ‘marginal cost’ principle implies this component should be excluded. In practice it makes better sense to include this fixed cost, as a cost which in any case is recaptured in the market price and still has to be met from health sector budgets.

4.6.6 Estimating Cost Aggregates

This is analogous to Drummond’s ‘Case-mix group’ category$^1$, aspects of which may be incorporated (as appropriate) when costing interventions (or as a fallback position for calculating cost offsets). In particular to New Zealand, DRG (diagnosis-related group) costs per hospital stay can be calculated for a given disease, then aggregated over all hospital stays in the selected disease category. These estimates of hospital costs could then be supplemented by estimates of other healthcare cost components, eg, GP consultation unit cost multiplied by average number of consultations, drugs by unit cost by quantity, laboratory tests by code by unit fee, home-care and palliative care by hour or by week by average cost per hour or week. (Appendix 5 to Pharmac’s Prescription for pharmacoeconomic analysis$^67$ provides a useful list of unit costs for most of these items, in June 2006 dollars.) Note that some further modelling and data analysis (for instance computing population averages) will be needed in some instances to identify which events are caused by the specified disease, and which are not. For example, not all GP consultations by a person with cancer will be a consequence of the cancer. HealthTracker could be useful in this regard.

In general, cost aggregate estimates would be attempted for selected age-gender categories, and also by ethnicity and socio-economic status as appropriate. The quality of such estimates will not necessarily be high. However appropriate smoothing techniques could be applied, interpolating or extrapolating to fill gaps in the available data.

4.6.7 Data gaps for New Zealand cost data

The main cost data gap in New Zealand is for procedures provided to outpatients, specialist consultations, and procedures and hospital stays at private clinics (publicly or privately funded). This will impact upon ABC-CBA in particular (eg, oncology outpatient consultations, biopsies, chemo- or radio-therapy). The non-private costs should in theory be available to
the Ministry of Health, since payments to District Health Boards will be based on a count of specific procedures multiplied, usually, by a standard fee. And it should be contained within HealthTracker.

**Required methodological development 13: Determine cost data availability for outpatient and private clinics, especially for cancer patients, through either HealthTracker or other mechanisms.**

### 4.6.8 Miscellaneous

There are also a number of other miscellaneous issues that have not been substantively addressed at the time of writing Version 1.0 of the Protocol, namely:

- How will ACC costs be included?
- How are publicly funded stays in private hospitals funded?
- How to cost private hospital/specialist care?
- Identifying and costing hospital pharmacy drugs (that are currently included in DRG costs).
PART II: ABC-CBA METHOD AND PROTOCOL

ABC-CBA will be a platform of data and parameters on cancer incidence and survival, projected into the future, by cancer sites, sex, age, ethnicity and deprivation (and sometimes stage). This platform of data will be captured in the form of Markov models, with (for example) survival rates transformed into time-dependent transition probabilities of death from the cancer in question. Costs and disability weights are laid over the Markov model states, so that each state (usually of a month’s duration) has an attributed cost to the health sector and a disability (or more generally utility) weight assigned. This platform will then be used for evaluations of proposed interventions, where the interventions are parameterised in terms of cost of the intervention, and changes in incidence, survival, disability weight, stage distribution and occurrence of sequelae. This intervention modelling may either use the already existing Markov model and Monte Carlo simulation, or use the existing platform of data to specify parameters for discrete event simulation (DES).

Initial ABC-CBA model development occurred as part of a Burden of Cancer study, conducted by Blakely, Costilla and Tobias during 2009-10 at the Ministry of Health. This study estimated the burden of cancers, using DALYs and for cancers incident in the year 2006 – not the burden of cancer cross-sectionally present in 2006. This focus on incident cancers was deliberate, both due to the good cancer registry data (for incidence and thence survival) and to allow an easier migration to economic decision modelling (ie, ABC-CBA as described in this protocol). The burden of cancer estimates used an external model lifetable; for the economic decision modelling in ABC-CBA we will be using each population’s own lifetable (eg, Māori males) so that we evaluate that population’s potential DALY change due to an intervention.

There are both similarities and differences between the ABC-CBA and NZACE-Prevention models. The similarities dominate, and include: the same ‘Key Principles’ and assumptions as outlined in Section 1 of this Protocol; same datasets and methods for costing; default Markov models, with strong reliance on lifetables; the use of DALYs (although the only real difference from QALYs is the use of disability weights rather than utilities); and an emphasis on a common epidemiological model structure (and inherent cost-offset model) that is utilised by all intervention modelling. The differences with NZACE-Prevention largely arise in the core models, and include:

- ABC-CBA is a model of incident cancers in 2006 (and subsequent years for prevention and screening models; Figure 6) and their future stream of mortality (and hence YLLs) and morbidity (and hence YLDs). ABC-CBA is not a model of both incident and prevalent cancer in 2006.
- ABC-CBA has a greater reliance on (cancer) disease models with multiple time-dependent states, as opposed to simpler prevalent states in a multi-state lifetable.
- ABC-CBA makes much use of cancer registry data (incidence and survival estimates), and the University of Otago’s CancerTrends study (for estimates of how both cancer incidence and survival varies by ethnicity and deprivation, and over time from 1981 to 2004). NZACE-Prevention makes much use of input parameters arising from a
national revision of the BDS being conducted by the Ministry of Health in parallel to BODE³.

The overall conceptual approach of the ABC-CBA models is outlined in Figure 6, sub-figures a through d:

a. **Baseline** model. The first task is to determine the baseline or ‘business as usual’ number and/or rates of:
   - incident cases (by cancer site, sex, age, ethnicity, deprivation, and occasionally stage or other category of disease severity at presentation (‘sub-type’)) in 2006
   - incident cases by calendar year after 2006 (arising from the closed cohort of the New Zealand population alive in 2006)
   - survival and disease progression among the 2006 incident cases, and all other future calendar year cohorts of incident cases.

b. **Treatment** intervention models, that consider just 2006 incident cancer cases. Cancer interventions are modelled by changes in survival, disability weights and sequelae after diagnosis.

c. **Prevention** intervention models, that allow changes in future cancer incidence (including distribution or rates by stage or sub-type), and thus can model cancer prevention and screening and early diagnosis interventions. (It is assumed that there is no impact on ‘business as usual’ disease progression post diagnosis.)

d. Both **Treatment and Prevention** models that merge both b. and c. That is, cancer control interventions that impact of both incidence and disease progression. (An example might a screening programme accompanied by improved treatment services.)

This conceptual approach will be repeatedly used in this Part II of the Protocol.

The majority of this Part II of the protocol will assume that Markov modelling is used – both in the creation of the baseline models (Figure 6a), and intervention modelling (b, c, and d in Figure 6). However, we may use micro-simulation or discrete event simulation (DES) in some instances for modelling interventions (i.e., Figure 6b, and perhaps d), and some sections are included below on possible approaches.

It is useful to think of ‘a’ in Figure 6 (baseline models) as a large databank of input parameters for subsequent ‘real’ modelling of the interventions (models b, c and d). This applies regardless of whether Markov or other modelling of the interventions is used. This separation is also useful when one considers the enormous heterogeneity in the baseline models; we will often only model through a few ‘types of people’ in intervention modelling (e.g., 50-59 year old with regional stage cancer). That is, the baseline model will often provide a databank of parameters, from which only some are selected to go forward to intervention modelling.
Figure 6: Conceptual cohort approach to modelling in ABC-CBA

a. Baseline model

b. Intervention modelling of disease progression for 2006 incident cases only

c. Intervention modelling of future incident cases only

d. Intervention modelling of future incident cases and disease progression

Years post diagnosis

Calendar year of incidence

Years post diagnosis

Calendar year of incidence

Years post diagnosis

Calendar year of incidence

Years post diagnosis

Calendar year of incidence
The remainder of Part II of this Protocol provides more detailed information on the assumptions, logic and development of the models necessary to achieve the modelling objectives of ABC-CBA.

- Section 0 describes the data inputs to the baseline or business as usual model. This includes both epidemiological inputs, and cost off-set inputs.
- Section 6 describes the approach to modelling interventions, most notably the parameters that need to be varied to model a cancer control intervention.

The ABC-CBA model prototypes are implemented initially in Excel, then programmed in Stata. TreeAge will be used for validation, and perhaps some intervention analyses. This protocol does not include details of the computer programs and code. Rather, this protocol serves as the higher-level specification.
5  Structure and data inputs for baseline models

5.1  Cancer sites

Table 5 shows the cancer site groupings and ICD10 codes used. In the majority of cases cancer site definitions were the same across the sources that were used to build the core ABC-CBA model. Some exceptions, and the ABC-CBA definitions, are:

- a narrower definition of brain, myeloma and ovarian cancer, to be consistent with survival analyses
- colon and rectal cancers were combined
- non-melanoma skin cancers were excluded.

Likewise, ethnic differences in incidence data were sourced from the CancerTrends study, which used the groupings shown in Table 5. When inconsistencies in site groupings occurred, the ethnic variations for the closest matching analysis in CancerTrends were used, or analyses rerun with different groupings if possible.

Existing Ministry of Health survival estimates were not available for gallbladder, bone and connective tissue cancers, ‘other adult’ cancers and childhood cancers. For these four sites, specific data was taken from the Cancer Registry and mortality files to allow excess mortality rate modelling; otherwise, the data sets already in existence for previous relative survival analyses 77 were used for modified excess mortality rate modelling.
<table>
<thead>
<tr>
<th>Cancer</th>
<th>ABC-CBA</th>
<th>Ministry of Health cancer trends and projections&lt;sup&gt;14,15&lt;/sup&gt;</th>
<th>CancerTrends (UOW, HRC-funded)&lt;sup&gt;13&lt;/sup&gt;</th>
<th>NZHIS Survival&lt;sup&gt;17&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICD10</td>
<td>ICD9</td>
<td>ICD10</td>
<td>ICD10</td>
</tr>
<tr>
<td>All childhood</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(&lt; 15)</td>
<td>C00–C096</td>
<td>140–208 (&lt;15)</td>
<td>C00–C96</td>
<td>140–208 (&lt;15)</td>
</tr>
<tr>
<td>All adult</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(≥ 15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bladder</td>
<td>C67</td>
<td>188</td>
<td>C67</td>
<td>188</td>
</tr>
<tr>
<td>Bone and connective</td>
<td>C40–41</td>
<td>170–171</td>
<td>C40–41</td>
<td>170–171</td>
</tr>
<tr>
<td>Brain</td>
<td>C71</td>
<td>191</td>
<td>C70–72</td>
<td>191</td>
</tr>
<tr>
<td>Breast (female)</td>
<td>C50</td>
<td>174</td>
<td>C50</td>
<td>174</td>
</tr>
<tr>
<td>Cervix</td>
<td>C53</td>
<td>180</td>
<td>C53</td>
<td>180</td>
</tr>
<tr>
<td>Uterus</td>
<td>C54–55</td>
<td>182</td>
<td>C54–55</td>
<td>182</td>
</tr>
<tr>
<td>Colon</td>
<td>C18–21</td>
<td>153–154</td>
<td>C18</td>
<td>153–154</td>
</tr>
<tr>
<td>Rectum, sigmoid, anus&lt;sup&gt;3&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>C19–21</td>
<td></td>
</tr>
<tr>
<td>Gallbladder</td>
<td>C23–24</td>
<td>156</td>
<td>C23–24</td>
<td>156</td>
</tr>
<tr>
<td>Hodgkin’s disease</td>
<td>C81</td>
<td>201</td>
<td>C81</td>
<td>201</td>
</tr>
<tr>
<td>Kidney and other urinary</td>
<td>C64–66, C68</td>
<td>189</td>
<td>C64–66, C68</td>
<td>189</td>
</tr>
<tr>
<td>Larynx</td>
<td>C32</td>
<td>161</td>
<td>C32</td>
<td>161</td>
</tr>
<tr>
<td>Lip, mouth</td>
<td>C00–14</td>
<td>140–149</td>
<td>C00–14</td>
<td>140–149</td>
</tr>
<tr>
<td>Pharynx</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leukaemia</td>
<td>C91–95</td>
<td>204–208</td>
<td>C91–95</td>
<td>204–208</td>
</tr>
<tr>
<td>Liver</td>
<td>C22</td>
<td>155</td>
<td>C22</td>
<td>155</td>
</tr>
<tr>
<td>Lung, trachea, bronchus</td>
<td>C33–34</td>
<td>162</td>
<td>C33–34</td>
<td>162</td>
</tr>
<tr>
<td>Melanoma</td>
<td>C43</td>
<td>172</td>
<td>C43</td>
<td>172</td>
</tr>
<tr>
<td>Myeloma</td>
<td>C90</td>
<td>203</td>
<td>C88, C90</td>
<td>203</td>
</tr>
<tr>
<td>NHL</td>
<td>C82–85, C96</td>
<td>200, 202</td>
<td>C82–85, C96</td>
<td>200, 202</td>
</tr>
<tr>
<td>Non-melanoma skin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oesophagus</td>
<td>C15</td>
<td>150</td>
<td>C15</td>
<td>150</td>
</tr>
<tr>
<td>Ovary</td>
<td>C56</td>
<td>183</td>
<td>C56, C57.0–57.4</td>
<td>183</td>
</tr>
<tr>
<td>Pancreas</td>
<td>C25</td>
<td>157</td>
<td>C25</td>
<td>157</td>
</tr>
<tr>
<td>Prostate</td>
<td>C61</td>
<td>185</td>
<td>C61</td>
<td>185</td>
</tr>
<tr>
<td>Stomach</td>
<td>C16</td>
<td>151</td>
<td>C16</td>
<td>151</td>
</tr>
<tr>
<td>Testis</td>
<td>C62</td>
<td>186</td>
<td>C62</td>
<td>186</td>
</tr>
<tr>
<td>Thyroid</td>
<td>C73</td>
<td>193</td>
<td>C73</td>
<td>193</td>
</tr>
</tbody>
</table>

---

<sup>1</sup> CancerTrends input is the first diagnosed malignant cancer after each census date – which is much the same as any cancer.

<sup>2</sup> C21 (anal cancers) were not included in the HRC-funded CancerTrends grouping.

<sup>3</sup> ICD code for anus is C21.

<sup>4</sup> Regression analyses for a combined C82–85 and C96 group have been conducted on CancerTrends data, and are used in ABC-CBA.
5.2 Cancer disease model structure, cure rates and disability weights

For each cancer, a model of disease progression is needed to estimate the years of life lost due to disability (and to allow for subsequent modelling of interventions that impact on morbidity rather than mortality). The general ABC-CBA disease model is shown below in Figure 7. Sex-, age-, ethnicity-, deprivation- (and occasionally stage/sub-type-specific) inputs for incidence and survival are specified for each cancer, using a common model structure (states, duration in each state, disability weight for each state, and sequelae). Each state has a one month duration. Also, for initial ABC-CBA modelling we set $T_D$ (duration in disseminated state) to zero for all cancers, and incorporate this state with the pre-terminal state. The disseminated state is conceptually identified both to be consistent with other burden of disease models and to allow for flexibility in future scenario modelling.

Figure 7: General ABC-CBA Markov cancer disease model

- **Susceptible / Population**
- **Diagnosis & Treatment**
  - Duration ($T_{DT}$) 2-14 mth
- **Remission**
  - Duration variable: $T_R = T_C - (T_{DT} + T_D + T_{PT} + T_T)$
- **Disseminated (or irradically treated)**
  - Duration ($T_D$)
- **Pre-terminal**
  - Duration ($T_{PT}$) 3-18 mth
- **Terminal**
  - Duration ($T_T$) 1 mth
- **Death from other causes**
- **Death from cancer**

$T_C$ = total cancer duration; $T_{DT}$ = time in diagnosis and treatment state; $T_R$ = time in remission state; $T_D$ = time in disseminated state; $T_{PT}$ = time in pre-terminal state; $T_T$ = time in terminal state.
Note that cancer subjects can move to the ‘death from other causes’ state from any disease state within the model, based on the background mortality risks by sex, age, ethnicity and deprivation from population life tables. There is also a parallel chain of states for those people who have permanent sequelae (eg, leg amputated due to bone cancer), but this is not shown in Figure 1.

The maximum durations in each state (excluding remission) are given in Table 7 (page 63). They are maximums in that if a person dies relatively soon after diagnosis, they will not live enough months post diagnosis to traverse all possible states. Accordingly, the duration in the terminal state takes priority over that in the pre-terminal state, which in turn takes priority over that in diagnosis and treatment. The duration in the remission state is simply the residual of the duration in all other states, subtracted from the total cancer duration time ($T_c$). Table 6 gives examples of the time in different states for four scenarios. For example, a person dying 6 months

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Total cancer duration: $T_c$</th>
<th>Time in state (months)</th>
<th>Remission: $T_R$ (duration = residual)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person dying of cancer at the end of year 3</td>
<td>36</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>Person dying of cancer at the end of month 6</td>
<td>6</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Survivor</td>
<td>48</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Person dying of a cause other than cancer at the end of year 3</td>
<td>36</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Note that this has implications for how YLDs are calculated in ABC-CBA; they are tallied up ‘backwards’ once people have died from the cancer, died from other causes, or survived. That is, we do not actually model transition probabilities between these disease states (eg, pre-terminal to terminal), but rather model transition probabilities to ‘death from cancer’ or ‘death from other causes’ at any stage from diagnosis to the cure time. For those transitions to ‘death from cancer’, we then back-calculate time and DWs in the various disease states. Likewise for ‘death from other causes’.

5.2.1 Extending the cancer disease models to include stage/sub-type

The cancer disease models need to be respecified and recalculated by stage (eg, localised, regional spread, or metastasised; TNM classification) or subtype (eg, size of tumour, if this is more relevant than stage) for colorectal, cervical and breast cancers, at least initially. Many cancers have a stage/sub-type template to follow from the Australian BDS. However, work is
required (and has not been undertaken at the time of this version of the Protocol) to develop New Zealand-specific cancer disease models by stage/sub-type, including:

- a desktop analysis and proposal of initial stage or sub-type disease models (with duration and DW specified), taking into account:
  - existing international disease models
  - cancer stage/sub-type data availability in New Zealand (e.g., the pros and cons of using SEER stage versus tumour size versus TNM for breast cancer)
  - recently developed cancer pathways of care in New Zealand (and how prognosis and treatment best match what staging or sub-typing)
  - nationally and internationally available data on incidence and survival by stage or sub-type
- consultation with clinicians and relevant cancer control experts (e.g., National Cancer Registry staff) on the initially proposed disease models, and subsequent refinement.

**Required methodological development 14: Selection of state/sub-types to use for breast, colorectal and cervical cancer, including consultation with New Zealand cancer experts.**

### 5.2.2 Duration and disability weights, by state, for each cancer model

Table 7 below lists the duration and disability weights (DWs) for each state and each cancer site included in the ABC-CBA baseline model, and as used in the *Burden of Cancer* report. These durations and DWs in Table 7 are taken or adapted from the Australian burden of disease 2003 study, which in turn largely uses the Dutch disability weights. Adaptations made for the *Burden of Cancer* report are described elsewhere (www.uow.otago.ac.nz/abc-cba-info.html, pages 11-17). Most notably, they include weighted averages of the DWs across stage or subtype distributions, to allow analyses aggregated by stage (or sub-type).

The duration and DWs for each stage/sub-type need specification. The Australian BDS provides an initial template, although the exact structure of the ABC-CBA model will depend on ‘Required methodological development 14’ above. Of note, the overall or average DW for a cancer site will be the average of the stage/sub-type DWs, weighted by the proportion or incidence rate for cancer patients in each stage/sub-type. Thus, the average DW for a given cancer site in ABC-CBA may vary modestly from that shown in Table 7, especially in the future if the stage/sub-type distribution is forecast to change.

Updated disability weights (DWs) from the GBD are projected to be made available in May 2011. These will be incorporated. Detailed consultation with cancer experts on DWs by stage/sub-type, etc (i.e., ‘Required methodological development 15’ below), will most sensibly occur once the updated DWs are available.
Table 7: Disability weights (DW) and duration time (T, in years) for the disease model states used in the Burden of Cancer report (ie, excluding consideration of cancer stage/sub-type) †

<table>
<thead>
<tr>
<th>Cancer site</th>
<th>Statistical cure time (years)</th>
<th>Diagnosis and treatment</th>
<th>Remission</th>
<th>Pre-terminal (including disseminated cancer)</th>
<th>Terminal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>T_DT DW †</td>
<td>T_R DW †</td>
<td>T_PT DW †</td>
<td>T_T DW †</td>
</tr>
<tr>
<td>All childhood</td>
<td>5</td>
<td>0.67 0.66</td>
<td>Residual 0.20</td>
<td>0.50 0.75</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Bladder</td>
<td>10</td>
<td>0.17 0.27</td>
<td>Residual 0.18</td>
<td>0.92 0.64</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Bone and connective</td>
<td>10</td>
<td>0.50 0.41</td>
<td>Residual 0.30</td>
<td>0.92 0.75</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Brain</td>
<td>5 (&lt; 55 years); 10 (≥ 55 years)</td>
<td>0.25 0.68</td>
<td>Residual 0.18</td>
<td>0.67* 0.75</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Breast (female)</td>
<td>20</td>
<td>0.33 0.29</td>
<td>Residual 0.26</td>
<td>0.92 0.79</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Cervix</td>
<td>5</td>
<td>0.25 0.43</td>
<td>Residual 0.20</td>
<td>0.42 0.75</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Colorectal</td>
<td>8</td>
<td>0.75 0.43</td>
<td>Residual 0.25</td>
<td>0.25 0.83</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Gallbladder</td>
<td>7</td>
<td>0.17 0.43</td>
<td>Residual 0.20</td>
<td>0.92 0.73</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Hodgkin’s disease</td>
<td>10</td>
<td>0.33 0.66</td>
<td>Residual 0.19</td>
<td>0.42 0.75</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Kidney and other urinary</td>
<td>10</td>
<td>0.17 0.27</td>
<td>Residual 0.18</td>
<td>0.92 0.64</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Larynx</td>
<td>10</td>
<td>0.25 0.56</td>
<td>Residual 0.37</td>
<td>0.67 0.90</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Leukaemia, &lt; 45 years</td>
<td>10</td>
<td>1.17 0.55</td>
<td>Residual 0.19</td>
<td>0.25 0.75</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Leukaemia, ≥ 45 years</td>
<td>10</td>
<td>0.50 0.55</td>
<td>Residual 0.19</td>
<td>0.25 0.75</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Lip, mouth, pharynx</td>
<td>10</td>
<td>0.25 0.56</td>
<td>Residual 0.37</td>
<td>0.67 0.90</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Liver</td>
<td>7</td>
<td>0.17 0.43</td>
<td>Residual 0.20</td>
<td>0.92 0.73</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Lung, trachea, bronchus</td>
<td>6</td>
<td>0.42 0.70</td>
<td>Residual 0.47</td>
<td>0.42 0.83</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Melanoma</td>
<td>6</td>
<td>0.17 0.22</td>
<td>Residual 0.19</td>
<td>0.25 0.81</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Myeloma</td>
<td>20</td>
<td>0.75 0.19</td>
<td>Residual 0.19</td>
<td>0.42 0.75</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>NHL</td>
<td>20</td>
<td>0.33 0.66</td>
<td>Residual 0.19</td>
<td>0.42 0.75</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Oesophagus</td>
<td>6</td>
<td>0.17 0.56</td>
<td>Residual 0.37</td>
<td>0.92 0.90</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Ovary</td>
<td>10</td>
<td>0.25 0.43</td>
<td>Residual 0.20</td>
<td>0.42 0.75</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Pancreas</td>
<td>5</td>
<td>0.17 0.43</td>
<td>Residual 0.20</td>
<td>0.92 0.73</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Pleura, thymus, heart</td>
<td>5</td>
<td>0.25 0.35</td>
<td>Residual 0.30</td>
<td>0.67* 0.75</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Prostate</td>
<td>20</td>
<td>0.17 0.27</td>
<td>Residual 0.20</td>
<td>1.50 0.64</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Stomach</td>
<td>6</td>
<td>0.50 0.53</td>
<td>Residual 0.38</td>
<td>0.92 0.73</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Testis</td>
<td>3</td>
<td>0.25 0.27</td>
<td>Residual 0.18</td>
<td>0.75 0.64</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Thyroid</td>
<td>5</td>
<td>0.17 0.27</td>
<td>Residual 0.18</td>
<td>0.75 0.64</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Uterus</td>
<td>6</td>
<td>0.25 0.43</td>
<td>Residual 0.20</td>
<td>0.42 0.75</td>
<td>0.08 0.93</td>
</tr>
<tr>
<td>Other adult cancer**</td>
<td>10</td>
<td>0.35 0.44</td>
<td>Residual 0.24</td>
<td>0.66 0.75</td>
<td>0.08 0.93</td>
</tr>
</tbody>
</table>

* The Australian Burden of Disease study Excel spreadsheets state one-year duration on the flow diagram and 0.67 years in text notes. We have elected to follow the text notes.

** The duration and DWs for ‘other adult cancer’ are simply averages of the specified adult cancer sites.

† Note that as disease models are developed by stage or sub-type the ‘fixed’ DWs will be those at this most disaggregated level. The ‘average’ DW for all stages or sub-types combined may no longer exactly equal that.
reported in this Table. For example, is stage distribution in the future varies, then the weighted (by stage distribution) average of stage-specific DWs (fixed) will vary.

‡ DWs and duration by cancer stage/sub-type will be presented in future versions of this Protocol.

Another issue is that of the DW in the remission state for cancers with a long statistical cure time. For example, breast cancer deaths can occur many years after diagnosis, and even at 20 years post-diagnosis the relative survival is still lower than that expected based on population mortality rates. For those women who survive to the nominated statistical cure time (ie, 20 years), and even those who relapse or die of breast cancer (say) 10 to 20 years after diagnosis, it seems inappropriate to assume that the loss of quality of life and hence the DW is constant throughout the (up to) 20 years. Rather, it seems more realistic to assume that the DW reduces with each subsequent year of disease-free survival. (Other burden of disease studies circumvent this issue by assuming five years’ duration for all YLD calculations, even if the statistical cure time is in excess of five years. However, we wanted a state-based model for future scenario modelling of interventions that may occur any time until ‘cure’ was pronounced.) In the Burden of Cancer report we assumed that the remission disability weight reduces by 20% per annum from the first year onwards. We need to consult more generally on this default assumption, and also undertake sensitivity analyses for this 20% (or whatever is decided) per annum reduction in the DW.

Required methodological development 15: Specification of duration and DW for each stage/sub-type in the cancer models, including consultation with New Zealand cancer experts.

5.2.3 Allowing for two or more disability weights

A person with a particular cancer may also have other diseases or co-morbidities, with their own attendant impact on quality of life. Considering individuals for now, this creates challenges in assigning a combined DW to someone with two or more diseases and hence two or more DWs. Additionally, everyone will be assigned an average background DW for average comorbidities (see Section 1.3.5, page 19). For example, if someone has another disease with a DW of 0.40, and their current cancer state has a DW of 0.30, what is the combined DW for this individual? A value of 0.70 (ie, 0.40 + 0.30) is one possibility, but two problems arise with this system:

- For someone with multiple diseases and a simple sum of DWs greater than 1.0, do we assign them a DW > 1.0 (ie, a state worse than death)?
- Thinking in terms of capacity to benefit in intervention modelling, do we assume that the reduction in DW is just that for the disease under question, or some fraction of the individual’s remaining quality of life (ie, 1 – DW)?

The usual way of accommodating multiple DWs is to let the total disability weight \( = 1 - \prod_i (1-\text{DW}_i) \). In the case above, this would be \( 1 - (1-0.3) \times (1-0.4) = 0.58 \). That is, the ‘independent’ DW from cancer added another \( 0.30 \times 0.60 = 0.18 \) to the already existing 0.40 DW. It is this marginal change in the DW that needs to be captured by simulation modelling.
5.2.4 Cure times
Statistical cure times are also presented in Table 7. Determining the time of statistical cure is critical for a state-based model like ABC-CBA. (An alternative is to assume asymptotic reduction in excess mortality due to cancer, and use parametric functions (e.g., Weibull curves) to estimate survivorship. For flexibility of modelling we have elected to use a state-based model that requires specifying statistical cure times. However, if we do model some interventions with DES, then a parametric specification of survival would be desirable – see later sections of this Protocol.) The methods used to determine statistical cure times are specified in detail in the *Burden of Cancer* report 76 (available at [www.uow.otago.ac.nz/abc-cba-info.html](http://www.uow.otago.ac.nz/abc-cba-info.html), pages 17 to 22). Briefly, the process required inspection of relative survival curves from New Zealand and Sweden, and determining the time after diagnosis at which the excess mortality from the cancer under question was trivial.

When stage or sub-type disease models are developed, consideration will need to be given to the possibility of cure time varying by stage or sub-type.

*Required methodological development 16: Determine whether cure times should vary by cancer stage/sub-type.*

5.2.5 Long-term sequelae of cancer
Standard burden of disease methods also estimate health loss due to disease sequelae (e.g., a disability weight for amputation is assigned for the person’s remaining life). The GBD study included long-term sequelae for colorectal cancer, breast cancer, female reproductive cancers and male genitourinary cancers. In addition, the recent Australian burden of disease study included removal of one eye for eye cancer, removal of the larynx for larynx cancer, amputation for bone cancer and long-term brain injury for brain cancer. These sequelae and their associated severity weights, as used in the Australian burden of disease study, are listed in Table 8. Disability weights for the sequelae were derived from the GBD and Dutch BD studies.

For ABC-CBA, such sequelae present another challenge in terms of extra strata that, when combined with sociodemographic and cancer subgroup data, make the total number of states cumbersome. They were modelled as parallel chains to the normal disease model, with the only difference being the presence of sequelae. Sequelae were assumed to commence at the end of year 1 of the disease model. Consistent with routine burden of disease methods, we did not just combine sequelae and state DWs additively (sum of the sequelae and state DWs). Rather, we combined them multiplicatively as described above in Section 0.

YLDs are allowed to accrue for people with sequelae beyond the statistical cure time for the expected remaining life expectancy at the point at which statistical cure occurs.
Table 8: Sequelae for cancer model from the Australian burden of disease study

<table>
<thead>
<tr>
<th>Site – sequelae</th>
<th>Proportion of survivors with sequelae (%)</th>
<th>Severity weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colorectal cancer – stoma</td>
<td>0.09</td>
<td>0.21</td>
</tr>
<tr>
<td>Bone and connective tissue cancer – amputation</td>
<td>0.08</td>
<td>0.30</td>
</tr>
<tr>
<td>Breast cancer – mastectomy</td>
<td>0.51</td>
<td>0.09</td>
</tr>
<tr>
<td>Female reproductive cancer – infertility</td>
<td>Cervix: 0.46</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>Uterus: 1.00</td>
<td>(ages under 40 only)</td>
</tr>
<tr>
<td></td>
<td>Ovary: 0.64</td>
<td></td>
</tr>
<tr>
<td>Male genitourinary cancer – impotence and incontinence</td>
<td>Prostate: 0.53</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>Bladder: 0.12</td>
<td></td>
</tr>
<tr>
<td>Brain cancer – long-term brain injury</td>
<td>0.05</td>
<td>0.35</td>
</tr>
<tr>
<td>Eye cancer – removal of an eye</td>
<td>0.45</td>
<td>0.30</td>
</tr>
<tr>
<td>Larynx cancer – removal of the larynx</td>
<td>0.35</td>
<td>0.04</td>
</tr>
</tbody>
</table>

5.3 Data inputs to baseline models
Section 5.2 above describes the baseline model structure. This section describes the data inputs necessary for the baseline model, i.e., that shown in Figure 6a.

5.3.1 Expected background mortality
See Section 1.3.4 (page 17) for the generation of background mortality rates. For the purposes of the Markov models, the background mortality rates need converting to transition probabilities. This will be done by first subtracting the excess mortality rate for the cancer in question (Section 5.3.3 below, and using ‘business as usual’ cancer excess mortality rate rather than intervention simulated rate), then the using exponential link formula:

\[1 - s = 1 - \exp(-r_t T)\]

Equation 2

Where:
- \(s\) = survival proportion (and therefore 1-\(s\) = mortality risk)
- \(r_t\) = rate at time \(t\) (in our case, length of time after diagnosis)
- \(T\) = duration of time included in state (in our case 0.0833, or one month).

Note that it is important to keep track of time since diagnosis as the excess mortality rate (\(r_t\)) decreases with time since diagnosis.

5.3.2 Cancer incidence
The base year is 2006, requiring estimated 2006 incidence rates by cancer site, sex, age, ethnicity and deprivation (and for some cancers stage or sub-type). This is a demanding task, but one for which we have reasonably good data in New Zealand.
Our general approach is to merge three sources of data:

- Ministry of Health estimated projections by sex and age
- CancerTrends estimated differences in incidence by ethnicity and deprivation
- Stage or sub-type distribution from Cancer Registry, international studies, or expert opinion.

5.3.2.1 Ministry of Health projections for 2006 by sex and age

The Ministry of Health has undertaken analysis on trends and projections of cancer incidence in New Zealand.\(^14\)\(^15\) We use these estimates, rather than actual rates, for two reasons:

- They are smoothed, and therefore not prone to random variation in 2006. (For example, just by chance there may have been more or less than expected cases in certain strata in 2006.)
- We also need to include future cancer incidence rates in the ABC-CBA model, which will come from these Ministry of Health projections. Therefore, it is consistent to also use the Ministry of Health ‘projections’ for 2006.

These incidence projection models are only by sex and age, due to a lack of ethnic and socioeconomic data back to 1950 – the data that was used to drive the projection models.

5.3.2.2 Incorporating ethnic and deprivation variation in 2006 incidence estimates, and beyond

In the previous Burden of Cancer report\(^16\), we allowed for ethnic variation in cancer incidence rates by mathematically solving what the disaggregated 2006 projected rates would be by ethnicity, using published ethnic incidence rate ratios from CancerTrends.\(^13\) We will use a similar approach to generate cancer incidence rates by both ethnicity and deprivation for ABC-CBA, but apply a more generalised framework (utilising new regression analyses on CancerTrends data) that will also be applicable to 2007 and beyond projections (see below).

The goal of the regression modelling is to produce rate ratios for the joint distribution of ethnicity (Māori, non-Māori only) and deprivation (deciles 1-3, 4-7, and 8-10), by sex and age group. These rate ratios can then be combined with the above sex by age incidence projections, and the (projected) population distribution by sex and age, to generate cancer incidence rates by any combination of sex by age by ethnicity by deprivation (and by year for 2007 and beyond). Parsimony is important - it is the baseline model ‘only’ we are building (not the intervention effects per se). We will follow a similar approach to that used for NZCMS analyses to build sub-population lifetables.\(^22\)

Methods also need to be developed to generate rates by stage/subtype, where applicable.
Required methodological development 17: Develop methods for projecting cancer incidence rates to 2026, including by stage/subtype in some instances, using a mix of empirical data from New Zealand and internationally, and expert opinion.

5.3.3 Relative survival and excess mortality rate modelling

Background or expected mortality rates provide the transition probabilities for ‘other causes of death’. Relative survival and excess mortality rate modelling provides the transition probabilities for ‘deaths from cancer’.

5.3.3.1 Background

Relative survival methods have been applied to New Zealand Cancer Registry data by both the Ministry of Health and university researchers. There are also extensive cross-sectional estimates of cancer-specific survival using Kaplan–Meier and Cox proportional hazards modelling for the total population, and for Māori and non-Māori. There is currently no extended times series on relative cancer survival by ethnicity or socioeconomic position, although estimates are forthcoming from CancerTrends.

Relative survival is the difference between observed and expected (in the absence of the disease in question – in this case cancer) survival. More precisely, we can define it as follows:

\[ \text{Relative survival ratio (RSR)} = \frac{\text{observed survival}}{\text{expected survival}} \]

The observed survival is that up to t years post-diagnosis, regardless of the cause of death, and the expected survival is that determined for a comparable cohort based on population life tables – in our case, 2006 period life tables.

Relative survival and excess mortality are mirror image concepts. The excess mortality rate is simply the difference between the mortality rate of the cancer populations and the background population mortality rates. Excess mortality rates are hazards, and are expressed per unit of time. Relative survival ratios are survival proportions; however, the use of the term relative survival ‘rates’ is, unfortunately, longstanding and embedded.

Using Equation 2 (page 66), it is straightforward to move backwards and forwards between rates and proportions, and parameterise Markov model transition probabilities for excess mortality rates.

Excess mortality rate modelling was used extensively in the Burden of Cancer report, and will be used extensible in ABC-CBA. Assuming a Poisson distribution of excess deaths due to cancer, the excess mortality model can be written as:

\[ \ln(u_j - d_j) = \ln(y_j) + x \beta \]

where:

\[ u_j = \text{expected number of all deaths (} d_j \text{; using the period method for 2002–06) for observation } j \text{ (eg, a cross-classified covariate pattern – see below)} \]
\[ d_{ij} = \text{expected number of deaths for observation } j, \text{ due to causes other than the cancer of interest and estimated from general population mortality rates (ie, for the baseline model, as described in Section 5.3.1)} \]

\[ y_{ij} = \text{person time for observation } j \text{ (ie, offset)} \]

\[ x = \text{vector of variables that predict excess mortality: ethnicity (binary variable for Māori:non-Māori), sex, age group (five-level categorial variable: } 15\text{–}44, 45\text{–}54, 55\text{–}64, 65\text{–}74, 75+), \text{ years (t) post-registration (10-level), and interactions between the latter two to allow for commonly observed higher initial excess mortality for older people early in follow-up (65–74, first year; 75+, first year; 65–74, second year; 75+, second year; all other age-by-year combinations as reference).} \]

Observed excess mortality rates, or relative survival, are often unstable – particularly for Māori. Figure 8 shows the observed and modelled excess mortality rates for breast cancer from the Burden of Cancer report.\(^7^6\) The predicted rates are ‘smoothed’ by virtue of the Poisson model and its covariate structure. (Note that the same relative shape of the excess mortality rate curve by time post-diagnosis (across sex by age by ethnic strata) results from time being a main effect in the Poisson model, except among those aged 75 years and over in the first year of follow-up, where the excess mortality curve starts from a higher position. This latter exception is because of the interaction terms of age and time included routinely in the models, due to excess mortality often being particularly high in the first one or two years of follow-up for older cancer patients.)
Figure 8: Excess breast cancer (female) mortality rate by sex, age, ethnicity and year (t): observed (− \ln [obs RSR t / obs RSR t − 1]) (dotted lines); and predicted from Poisson model (solid lines) (Source 76)

Excess mortality rates for base-year 2006

Ethnic-specific excess mortality rates for base-year 2006, and hence ‘cancer death’ transition probabilities, for cancer sites non-disaggregated by deprivation and stage/sub-type will be sourced directly from the Burden of Cancer study.

Regarding disaggregation by deprivation, differences in cancer survival are not as marked by deprivation in New Zealand as they are by ethnicity. Nevertheless, survival tends to be
worse among more deprived populations. After adjusting for stage or extent of disease, deprivation gradients in five year cancer survival are smaller again, ranging from 0.07 greater among the most deprived (brain) to 0.09 greater among the least deprived (kidney) for solid tumours with greater than 60% of cases having SEER stage data. However, whilst there was a general trend to worse stage-adjusted survival among deprived people, none of the differences were statistically significant (and measurement error of stage may have reduced the ability to completely adjust for stage).

Therefore, we will assume that there are no differences in relative survival or excess mortality by deprivation in New Zealand, having already adjusted for sex, age, ethnicity and stage/sub-type. (This has the added advantage of simplifying calculation of input parameters.)

Regarding disaggregation by stage/sub-type, this will first require the specification of cancer stages/sub-types to be used in ABC-CBA (Section 5.2.1), and then a judicious combination of the following methods:

- Where appropriate New Zealand data exists (e.g., SEER stage is selected for a given cancer, and at least 60% of cancer registrants have SEER stage data), undertake excess mortality rate modelling to directly generate estimates of excess mortality rate by stage/sub-type.
- Where appropriate New Zealand data does not exist, use comparable international data on excess mortality rate differences by stage/sub-type.
- Where no appropriate data exists, use expert opinion to determine likely differences.

Note that regardless of which of the above methods are used, the weighted average (where stage/sub-type proportionate distribution are the weights) of stage/sub-type excess mortality rates must equal the non-stage/sub-type disaggregated excess mortality rates by strata of sex, age and ethnicity. (Deprivation can be put aside here, for reasons described above.)

### 5.3.3.3 Excess mortality rates for years 2007 onwards

The default assumption will be that stratum-specific excess mortality rates are unchanged into the future.

Specification of future change (presumably always decreasing) in stratum specific excess mortality rates will only occur when at least one of the following applies:

- New Zealand data suggests changing relative survival or excess mortality rates over time (e.g., Ministry analyses and forthcoming CancerTrends analyses), and expert opinion suggests changes will occur into the future.
- International data suggests changing relative survival or excess mortality, and such trends are likely to apply into the future and in New Zealand.
• Expert opinion (e.g., with respect to new treatments) suggests that future changes are highly likely.

Specification of future differential rates of change by stage/sub-type may be warranted (e.g., if new treatments only apply to certain stage of sub-type cancers).

Specification of future differential rates of change by ethnicity or socioeconomic position will only occur when there is such evidence from forthcoming CancerTrends analyses and/or very strong theoretical or expert opinion of such Differentiality.

The actual specification and calculation of such future trends will be documented in future versions of this Protocol.

Note that it will also be important to differentiate ‘business as usual’ changes in survival, and those changes due to interventions we wish to simulate; the latter should not be included in the baseline models.

5.4 Cost off-sets

Cost off-set values will be derived from Ministry of HealthTracker data, with possible augmentation from Australian AIHW data. (See Section 4.)

There is insufficient data or knowledge to estimate how cost-offsets will vary in the future. Therefore, base-year 2006 estimates will be assumed to apply to states in 2007 and beyond.

The actual cost-offset values will be ‘attached’ to each state in the same way that DWs are.

Further work is necessary to disaggregate cost off-sets by strata of time (likely to be higher immediately after diagnosis), and stage/sub-type. These methods will be described in future versions of this Protocol.

Required methodological development 18: Develop full method, and trial with HealthTracker and/or other data, for determining cost-offsets by ABC-CBA cancer model states.

5.5 Baseline model calibration

The Burden of Cancer study included a basic calibration of the number of modelled deaths compared to deaths occurring in recent calendar years. However, it was a crude calibration only, as there were many reasons why the number of deaths among incident cancer cases in 2006 might vary from the number of deaths expected in a given calendar year.

A process of model calibration needs to be developed for ABC-CBA baseline models, and intervention simulation models.

Required methodological development 19: Develop a model calibration process to apply to baseline ABC-CBA models.
6 Modelling interventions

This section describes the actual simulation process, corresponding to subfigures b to d of Figure 6 (page 56). At a simple conceptual level, one just uses Monte Carlo simulation to simulate the intervention 1000’s of times, with each iteration involving a separate set of samples from the uncertainty distribution about each input parameter. Operationally, however, the core ABC-CBA has many strata or much heterogeneity. Modelling each iteration for all possible sub-populations will not only be time consuming, but also inefficient and (often) unnecessary. For example, it will often not be necessary to model both young and old – and it may only be necessary to model age in five year (rather than single year) increments. There are also mathematical and operational methods to improve efficiency, eg, the use of matrices for transition probabilities, etc. Finally, there may be occasions where either/both for validity or efficiency that we need to use a micro-simulation model (eg, discrete event simulation), utilising the core ABC-CBA model more as a databank of input parameters. This section will describe the modelling approach(es) that we will use.

6.1 Disease progression for 2006 incident cases only

This section applies to treatment, supportive care and palliation interventions that occur at or after diagnosis (Figure 6b).

6.1.1 Multi-cohort Markov models

Using the intervention specification (see Sections 0and 3 of this Protocol), the target population is identified (eg, age range, cancer site(s), and stage(s)/sub-types) and the relevant cohorts extracted. If results by sub-population are required for equity analyses, then disaggregated cohorts are required (eg, by ethnicity by deprivation).

For example, assume we are interested in simulating a treatment for both sexes with stage II of cancer X, aged between 40 and 79. And we are interested in possible equity impacts by ethnicity and deprivation. Assume also that we have aggregated single age-groups into five year age-bands. This means we extract 2 sex × 8 age groups × 2 ethnic groups × 3 deprivation groups = 96 cohorts to model simultaneously.

If the intervention under consideration, or any of the comparator interventions under consideration, might be considered as contributing to changes and/or parameters of the business as usual scenario (ie, the baseline model in Figure 6a), then a modified ‘partial null’ baseline model for the cancer (stage/sub-type) in question will need to be specified.

Next, the intervention is specified in terms of input parameters and their attendant uncertainty distributions. Extending the above example, this might be:

- A rate ratio of 0.80, with 95% probability distribution specified on a log scale of -0.223 +/- 0.2 to give a 95% interval for the RR ranging from 0.655 to 0.977.
• A risk ratio of 0.90 to apply to the disability weights (as the treatment not only saves lives, but also improves quality of life), with some similarly specified probability distribution.

• And a central estimate and probability distribution for the direct costs of the intervention. Note this direct cost of the intervention may be split into two components:
  o A variable component that varies with number of survivors and duration of survivorship (e.g., monthly treatment costs), and therefore needs to be modelled through the Markov disease model
  o A fixed cost component that does not vary with change in number of survivors and duration of survival, and therefore only enters calculations at the point of tallying up net cost and the ICER.

Next, a decision is required as to whether the sampled values at each iteration for each input parameter are the same across all cohorts, or sampled independently. The former is the default option. Possible attenuation of effect (e.g., diminishing relative impact on excess mortality following cessation of treatment) may also need parameterisation. Any correlations between input parameters also need consideration, and possible parameterisation.

If the intervention under consideration is one of multiple mutually exclusive interventions, then a pathway series of analyses will also need to be conducted with revision of the ‘baseline’ model at each step to allow for the interventions that have gone before.

Regarding the baseline model input (e.g., background mortality rates, baseline excess mortality rates, etc), the default assumption is that they are fixed and therefore do not require uncertainty intervals being specified. However, where there is notable uncertainty about the baseline model, or the health gains and or net costs would vary notably depending on the baseline model parameters, we may need to additionally model uncertainty of these baseline parameters. Such instances where uncertainty analyses about baseline model input parameters are required are expected to be uncommon.

Further details about specific models will be documented with that analysis.

It is expected that actual Markov modelling will be conducted in Stata, although TreeAge may on occasion be used.

**6.1.2 Discrete event simulation**

It is envisaged that discrete event simulation (DES) may on occasion be warranted or highly desirable. For example, where it is too complex to fully specify the disease model due to many strata, and the need for memory of previous states to be built into the model. Criteria as to when DES is used in ABC-CBA, and how, will be presented in a future version of this Protocol.
6.2 Incident cancer cases in years beyond 2006
This section applies to prevention, screening and early diagnosis interventions (Figure 6c).

The range of cancer sites affected needs to be identified, and the age range and sub-population groups, as per above. Likewise, whether the intervention (or any of set of comparator interventions) requires the specification of a partial null or revised baseline model.

The number of future incident cancers – including shifts in stage/sub-type distribution to approximate a screening or early diagnosis intervention – needs to be parameterised in terms of changes in the future incidence rates (including differentially by stage/sub-type).

In addition to the general issues identified above, particular attention must be given to lag times, then attenuation effects, for any impact of the intervention of future incidence rates. This will require careful consideration of the intervention time frame as well (eg, is the intervention applied for just one year, or 20 years?).

6.3 Disease progression for future incident cancer cases
This section applies to prevention, screening and early diagnosis interventions, where there is also reason to suspect that future treatment, supportive care or palliation will also change (Figure 6d). Further details, if necessary, on this modelling option will be provided in future versions of this Protocol.

6.4 Model variations and extensions
The default disease models described above are designed to accommodate most interventions. However, it is likely that variations to the model structure will be required for some interventions. For example, treatments for cancer patients with multiple remissions will require serial remission states to be added to the model. Such variations and extensions will be intervention specific, and may be presented as examples in future versions of the Protocol.

6.5 Intervention model calibration
A process of model calibration needs to be developed for ABC-CBA intervention simulation models.

Required methodological development 22: Develop a model calibration process to apply to intervention ABC-CBA models.
PART III: NZACE-PREVENTION METHOD

7 NZ-ACE: Overview of ACE-Prevention model

The ACE-Prevention (Australia) model is well established with its own protocol, peer reviewed published methodological approach, numerous outputs with modelling-specific descriptions of methods, and existing Excel spreadsheet models and programmes.

7.1 General comments

ACE-Prevention (Australia) is fairly conservative in some of its assumptions. For example, some ACE-Prevention (Australia) interventions conservatively assumed no benefit to younger people (e.g., the salt reduction interventions do not assume any benefit (ever) for those initially aged <30 years). This is a reasonable simplifying step – but may generate quite a gap in the DALY averted benefit. One of the overall principles of BODE will be to input best estimates, not conservative estimates, wherever possible and where the impact of so doing is more than negligible.

For simplicity, disability weights (and associated DALYs) are not attributed to risk factor states, e.g., having high blood pressure / hypertension, high cholesterol, etc. This is despite considerable evidence that getting such diagnoses do often contribute to morbidity (associated with “disease labelling” but also directly e.g., headaches from hypertension).

7.2 Particular assumptions implicit in ACE-Prevention (Australia) modelling

ACE-Prevention (Australia) models follow the principles and assumptions outlined in Section 1 of this Protocol. Here we just list some particular aspects of the existing or preceding ACE-Prevention (Australia) models that illustrate some of these assumptions, focusing on those that we may vary slightly (or conduct sensitivity analyses on) in NZACE-Prevention.

7.2.1 Age weights and ‘fair innings’

The original DALYs used in the GBD study incorporate age weights, reflecting an assumption of maximal social value of life at around 20. This assumption proved highly controversial, and age weights were subsequently dropped in future iterations of the GBD and country BDS. They have also not been included in economic evaluations using DALYs.

However, just as ethnicity is an equity concern, so too is age. Williams advances a fair innings argument that society values health gains for the young more than the old. One option in NZACE-Prevention will be to trial using the original DALY age weights. That said, the nature of preventive interventions having time lags, the lifetable modelling that already assigns more potential health gain to younger people, and discounting, all mean that age weights may not make much relative difference to the final cost-effectiveness ranking.
7.2.2 Time lags to health benefits

It is possible to set time lags from intervention establishment to actual change in disease incidence. The ACE-Prevention (Australia) salt analyses allowed for some time lag between intervention (i.e., year 2003) and changes in stroke and IHD incidence as a sensitivity analysis. The physical activity analyses allowed for no time lags (e.g., colorectal cancer incidence changed from year 2003). The assumption of ‘running at steady state’ can be taken to mean that no time-lags to disease prevention apply (i.e., has been running at steady state for decades).

In NZACE-Prevention we will model both with and without time lags, the former being the cost-effectiveness for at intervention actually starting in 2006, and the latter the cost-effectiveness for an intervention in steady state in the future. Similarly we will also consider phase-in periods for some interventions e.g., mandated step-wise reductions in national-level of tobacco for sale and food-grade salt for sale.

8 NZ-ACE: Modification to New Zealand setting

8.1 Modification of epidemiological parameters

The model structure for NZACE-Prevention will be the same as that for ACE-Prevention (Australia) – but with the addition of additional strata for ethnicity and deprivation.

Many input parameters to the baseline model will be changed, for example:

- Population socio-demographics will be those for New Zealand in 2006, and subsequent projections.
- Epidemiological parameters will be sourced from the New Zealand BDS
- Prevalent YLDs will be sourced from the New Zealand BDS.

8.2 Modification to costing

As described above for BODE\(^3\) generally, and ABC-CBA specifically, direct costing will be conducted specifically for BODE\(^3\) simulated interventions, and cost off-sets will (hopefully) be determined from HealthTracker data with AIHW data as a possible prior or default option.

8.3 Intervention selection

A two stage process was used to select interventions to model in New Zealand. First, salt reduction strategies were selected to test out implementation processes in the New Zealand context (since the ACE-Prevention (Australia) salt model was well developed and relatively simple). Second, priority risk factors were selected, and included tobacco, cholesterol and hypertension in the first wave of analyses. Third, a process of selecting exact interventions from these risk factor domains was undertaken. (See Section 2.2, page 34.)
8.4 Other modifications

Various minor differences will apply to some of the methods NZACE-Prevention compared to ACE-Prevention (Australia) in the domains of:

- What is included in the “health perspective” with the routine exclusion of industry costs (though these may be included in scenarios that move towards a societal perspective).
- The routine inclusion of health costs relating to extra life lived as a result of the intervention/s.
- Possibly a slightly different approach to set-up costs (see Section 1.2.9).
- A greater focus on discussing aspects of health inequalities (particularly for Māori vs non-Māori, but also for gender and age).

Further details on these issues will be dealt with in topic specific technical reports on the BODE^3 website, eg, on the methods used in the salt reduction modelling work.
REFERENCES


40. NHMRC additional levels of evidence and grades for recommendations for developers of guidelines - Stage 2 Pilot. Canberra (ACT): National Health and Medical Research Council, 2009.


