

# Assessing Cost-Effectiveness in the Prevention of Non-Communicable Disease (ACE–Prevention) Project 2005–09

## Economic Evaluation Protocol (As per September 2007)

### Chief Investigators:

- A/Professor Theo Vos (Epidemiologist & Leader, University of Queensland team)
- A/Professor Rob Carter (Health Economist & Leader, Deakin University team)
- A/Professor Chris Doran (Health Economist, UQ)
- Professor Ian Anderson (Medical Anthropologist/Sociologist University of Melbourne)
- Professor Alan Lopez (Medical Statistics, UQ)
- Professor Andrew Wilson (Health Policy, QLD Health)

### Research Team:

#### The University of Queensland:

- Theo Vos (Epidemiologist & Leader, University of Queensland team)
- Chris Doran (Health Economist, now at UNSW)
- Jan Barendregt (Epidemiologist and Modeller)
- Linda Cobiac (Associate Lecturer)
- Megan Forster (Associate Lecturer)
- Lennert Veerman (Research Fellow)
- Melanie Bertram (PhD Student until July 2008)
- Joshua Byrnes (PhD Student from Jan 2007, now at UNSW)
- Vittal Mogasale (PhD student from August 2007)
- Angela Wallace (PhD student from July 2008)
- Daphni Chao (MHEc student)
- Kathryn Arnett (MHEc student)

#### The Deakin University/ University of Melbourne:

- Rob Carter (Health Economist & Leader, Deakin University team)
- Cathy Mihalopoulos (Senior Research Fellow and Project Coordinator for Deakin Uni until 2007)
- Sandra Younie (Senior Research Fellow and Project Coordinator for Deakin Uni from 2007)
- Liliana Buffone (Senior Research Fellow joined in 2007)
- Sophy Shih (Research Fellow)
- Michael Otim (PhD Student)
- Kuisiang Tan (PhD Student)
- Katherine Ong (PhD Student)

## Table of contents

1	Introduction to the economic protocol .....	3
2	Overview of key issues in economic evaluation .....	6
3	Measurement of costs .....	26
4	Measurement of health benefits .....	41
5	Cost-effectiveness ratios .....	54
6	Sensitivity and uncertainty analyses .....	57
7	Second stage filter criteria.....	66
8	Discussion and presentation of study results .....	71
9	References .....	73

# 1 Introduction to the economic protocol

The Assessing Cost-Effectiveness in Prevention Project (ACE-Prevention) is a 5-year NH&MRC funded collaborative research program between the University of Queensland (UQ) and the University of Melbourne (UoM). From 2007, the University of Melbourne team has relocated to Deakin University (DU). The project aims to assess the cost-effectiveness of 100 preventive interventions across the range of non-communicable diseases and associated risk factors to help determine potential “best buys”. In addition, the projects also includes analyses of another 50 interventions for the treatment of non-communicable disease and control of infectious disease. The results will be brought together to provide a ‘league table’ in which all of the interventions are ranked in order of their economic merit. How interventions combine to form a coherent strategy for prevention, as well as their merit as individual interventions, will also be considered. Another important aspect of the research is that it is evaluating interventions for application to both Indigenous and non-Indigenous populations.

This protocol details the economic evaluation methods that will be used in the ACE-Prevention study. It has two principal purposes: i) to provide transparency for stakeholders and potential users of the research results, allowing scrutiny of the scientific merit of all elements of the study; and ii) to provide guidance to the ACE-Prevention research team in the development and application of their individual studies to ensure comparability of results. In addition to addressing general methodological issues, therefore, the protocol will from time to time provide steps to guide researchers in the application of chosen methods. It is also important to note that this protocol is a ‘living document’ and will be revised and updated until the evaluation methods are fully developed, illustrated and implemented in a satisfactory way.

This ACE-Prevention study builds on earlier ACE studies that are part of a broader body of work on priority setting in the health sector. An overview of the “ACE approach” and of previous ACE studies in cancer, mental health, heart disease and obesity, is provided at Appendix A. While this protocol has many similarities with the protocols for previous ACE studies, it also has differences that reflect the specific context of this study – that is, the prevention of non-communicable disease in application to both non-Indigenous and Indigenous populations. We believe there is sufficient consistency in methods across all the ACE studies to allow comparison across the cost-effectiveness results.

The use of standardised evaluation methods described in this protocol is designed to address the reservations expressed by many economists about the simplistic use of league tables based on evidence of cost-effectiveness from studies lacking in comparability (due to differences in methods, context and setting). The following key features of this protocol will help to avoid methodological confounding and promote a balanced approach to priority setting:

- The economic evaluations are undertaken as an integral part of the priority setting task (rather than relying on pre-existing studies in the literature);
- The economic evaluations are based on modelling best available information on costs and benefits from a range of data sources for demography, health system costs and cost offsets, disease incidence/prevalence, risk factors and disease burden that best describe the context of Australian health services;
- A common setting (the Australian population) and decision making context (recommendations for change that can be applied on an Australia-wide basis) is applied across all interventions and options for change;
- The rationale for selecting interventions is clearly specified and consistently applied (a key factor in operationalising “opportunity cost”);

- The economic evaluation methods are standardised, documented and open to scrutiny, including:
  - the choice of ‘health sector’ as study perspective, with a particular focus on government as third-party payer;
  - the definition of ‘benefit’ as reflected in the cost-effectiveness ratios (CERs) and the qualification of these benefits by applying a number of ‘other’ policy relevant criteria (such as equity, feasibility and acceptability) in a ‘2nd stage filter analysis’;
  - the choice of ‘current practice’ in Australia as the base case comparator in the CERs when interventions are ranked in a cost-effectiveness league and the choice for a ‘do nothing’ comparator when evaluating the most efficient mix of preventive interventions to address a particular health problem in comparison to the current set of preventive interventions (while striving for consistency between the two approaches);
  - the identification, measurement and valuation of costs, cost offsets and consequences;
  - the choice of discount rate (3 per cent) and reference year (2003);
  - extensive uncertainty testing<sup>1</sup>, together with sensitivity analysis, and reporting of a range of results (not just point estimates);
- Information is assembled by a multidisciplinary research team; preparing briefing papers to a standardised format (refer Appendix F).
- The technical cost-effectiveness results (i.e. cost per health-adjusted life year gained—which we use synonymously as the cost per DALY averted) are presented within a broader decision-making framework provided by the PBMA<sup>2</sup> approach;
- Our concern for technical rigour and an evidence-based approach in the economic evaluations is balanced by our awareness of the need for due process in priority setting. This is achieved by:
  - involving stakeholders in the project steering committees and advisory panels;
  - taking into account broader considerations that impact on resource allocation decisions but are less amenable to quantification (through our 2nd stage filter approach); and
  - giving all stakeholders on the steering committees and advisory panels the opportunity to express their views and seeking consensus decisions after informed discussion.

In the ACE studies, emphasis is placed on utilising best available evidence. For each application of the ACE approach, therefore, careful consideration has been given to clearly defining the concept of ‘evidence’ being applied. This was an important issue in ACE–

---

<sup>1</sup> ‘Uncertainty testing’ is defined to cover variation in those technical parameters (usually economic and epidemiological inputs) that impact on disease incidence/prevalence, efficacy/effectiveness, attendance, compliance rates, complication rates, unit costs and so on. ‘Sensitivity testing’, on the other hand, is defined to cover variation in social value parameters and/or the scenario under evaluation. Variations in the scenario might include changes in the study perspective, in the choice of comparators or inclusion of contentious cost impacts (such as production losses). Social value parameters include issues such as the choice of discount rate (social rate of time preference), weighting the health gain for equity (who receives the health gain) or for those most in ‘need’ (having regard to those severely ill and their fate if left untreated). It is useful to separate the technical calculation of the anticipated health gain from the social valuation placed on the anticipated health gain.

<sup>2</sup> PBMA: Program Budgeting and Marginal Analysis (refer Appendix A for an explanation of the PBMA approach).

Obesity, for example, where the level and quality of evidence available was less than for some of the other ACE studies (such as in cancer and mental health). The approach to “evidence” is also an important issue for ACE–Prevention, particularly having regard to the paucity of high quality efficacy/effectiveness data for interventions addressing Indigenous health problems.

ACE–Prevention is overseen by a Project Steering Committee (PSC), with the Indigenous component overseen by an Indigenous Steering Committee (ISC). The PSC includes senior federal and state health policy representatives, as well as public health experts and key stakeholders of various government and non-government organisations. The ISC is also constituted to include important stakeholders and is being supported by the Cooperative Research Centre (CRC) for Aboriginal Health, which has endorsed ACE–Prevention as an “in-kind” project.

The key tasks of the two steering committees are:

- i) to provide agreement on a broad program of work (including topic areas and timing);
- ii) to advise on data availability
- iii) to advise on the interpretation of best available information
- iv) to advise on how best to communicate the findings of the project to all key stakeholders (including policy makers, service providers and health NGOs); and
- v) to specify and interpret the 2<sup>nd</sup> stage filters.

It is not expected, however, that the Steering Committees will provide detailed advice on each of the selected topic areas. It is for this reason that Technical Advisory Panels (TAPs) have been established. The role of the TAPs is to:

- Provide assistance with the choice and specification of interventions within the agreed topic areas;
- Advise on the research methods within their areas of expertise;
- Advise on and facilitate access to data sets that can contribute to more accurate cost-effectiveness estimates; and
- Report back to the PSC/ISC on technical issues related to each topic.

Appendix B provides the terms of reference and membership for the two steering committees, as well as the TAPs. The topic areas and assignment of researchers agreed to date is set out in Attachment C.

## 2 Overview of key issues in economic evaluation

Before undertaking an economic evaluation it is wise for the analyst to stop and confirm that the overall approach to the study is appropriate, in order to ensure all key aspects of the assessment task are aligned and fully integrated.

Gold et al (1996) describe two key tasks: i) deciding the “study frame”; and ii) deciding the “study design”. These two tasks involve a series of decisions that are briefly outlined below. The decisions taken collectively both define and describe the economic evaluation to be undertaken, including the approach to data collection and analysis.

For the ‘study frame’, the key decisions are:

1. Decide the study objectives, paying careful attention to the decision context and intended audience for the study;
2. Decide the perspective or perspectives from which the study will be undertaken;
3. Consider the available forms of economic evaluation and decide which type or combination best matches the research question;
4. Select and define the intervention and comparator in a clear and concrete way;
5. Define the target population;
6. Consider the boundaries to the study; and
7. Decide on the appropriate time horizon.

For the ‘study design’, the key elements involve designing the data collection and analytic plan. These involve three basic decisions:

1. Decide on the conceptual model that best describes the intervention and its effects on health outcomes (i.e. whether to structure the analysis around an event pathway; a decision tree; a Markov model; a simulation, etc);
2. Determine how the data will be collected on activities, costs and outcomes for the intervention and the comparator(s); and
3. Consider the most appropriate analytic methods to combine the information in an economic evaluation.

Framing and designing the economic evaluation are the crucial first steps in undertaking an economic analysis and is the job of this protocol. Framing involves a series of decisions, which lay out in broad outline the intended methodology of the study. Designing the study requires the analyst to fill in the study frame, making the practical decisions that will determine the structure of the analysis and the data to be used. It is important to note that these earlier ‘study frame’ decisions need to be consistently implemented as we move into the measurement of costs and benefits in the ‘study design’. To use an analogy, the measurement of costs and benefits is very much the engine room of economic evaluation, but you need to have a driver and a sense of direction to get the car to go in the direction you intend. The ‘study frame’ provides this.

A discussion on each of the above listed decisions about the ‘study frame’ and ‘study design’ follows.

### 2.1 The research question

#### Background

The NHMRC grant application describes the ACE–Prevention research question in the following general terms:

*“The primary aim of this research is to provide a comparative assessment of the cost-effectiveness of various preventive strategies for non-communicable diseases in improving the health of Australians, both for the population as a whole and for Indigenous Australians.”*

More specifically, the aim is to assess the comparative cost-effectiveness of 100 preventive interventions for non-communicable disease for both Indigenous and non-Indigenous Australians, benchmarked against 50 major curative/infectious disease interventions, both individually and as components of disease/risk factor strategies.

Clear definition of the research question is an important first step in any form of economic evaluation. More specifically, definition of the research question should address three key issues:

- the study purpose, setting and decision-making context to guide the selection of appropriate methods (section 2.2);
- the clear specification of study comparators (section 2.3); and
- the study viewpoint or perspective to guide the identification of relevant costs and outcomes for inclusion in the study (section 2.4).

## **2.2 Decision context:**

Setting and context are important because they inform judgements about the relevance of the study to particular users, together with critical appraisal of the appropriateness of the methods used. It is important to realise that ACE-Prevention as an NHMRC-funded project, clearly has not been initiated directly by government for specific policy purposes. However, the central knowledge transfer objective of ACE-Prevention is to inform government decision-making. This is reflected, for example, in the number of government stakeholders on the grant application, as well as their subsequent membership on the Steering Committees. Apart from government decision-making, the project also aims to inform other key players involved in shaping decisions around preventive health services, such as non-Governmental health organisations and health service providers.

The setting for the evaluations in ACE-Prevention is thus clearly intended as the possible adoption in Australia, preferably on a national basis, of the options for change for the target populations. More specifically, ACE-Prevention will identify whether there are options for change that could improve the effectiveness and efficiency of current Australian health services for the prevention of non-communicable disease by:

- Directing available resources towards “best practice” cost-effective services;
- Providing best practice cost-effective services that address “unmet needs” in the Australian community;
- Modifying cost-ineffective services to improve their cost-effectiveness;
- Discontinuing inefficient use of resources in prevention;
- Targeting services to those in need as opposed to people with low risk profiles who are unlikely to benefit in a cost-effective manner; and
- Informing policy makers in the area of prevention about the best bundle of interventions, given alternative levels of budget availability.

The Project and Indigenous Steering Committees have confirmed that the study should have an Australia-wide focus and employ national level data. The exception would be where an intervention would be applied differently between States or where a State government stakeholder wished to resource a particular policy question of interest (either financially or

through secondment of a researcher to the project). The 2nd stage filter criteria may play a role in distinguishing packages of interventions relevant at a nationwide level versus a State/Territory or regional/community level.

As a university-based project, the ACE–Prevention decision context is also one of providing rigorous advice based on best available evidence on efficacy/effectiveness and ‘value-for-money’, coupled with broader considerations addressed by our ACE 2<sup>nd</sup> stage analysis. Attention will, therefore, be devoted to assessing the degree of confidence that can be placed in the findings, by clearly documenting the strength of the available evidence, the data gaps and associated uncertainties.

Next, it is important to note that both Indigenous and non-Indigenous target populations are involved in the ACE–Prevention research. There are important historical and policy reasons for having dedicated health services for Indigenous Australians and for questioning the effectiveness of mainstream services in catering for the needs of Indigenous Australians. Putting these arguments to one side for the moment, there is also an important economic evaluation rationale for evaluating interventions provided to Indigenous Australians separately to interventions provided to non-Indigenous populations. The economic evaluation rationale reflects the likelihood that the target disease burden, the prevalence and distribution of harmful exposures, the design and effectiveness of intervention strategies and the cost of implementing effective interventions will all be substantially different for the Indigenous and non-Indigenous populations. It follows that the cost-effectiveness of interventions will also vary considerably for Indigenous and non-Indigenous Australians and that separate evaluations are therefore justified.

Undertaking this research for both Indigenous and non-Indigenous populations also raises special issues for consideration in the selection of appropriate methods. On the one hand, comparable objective information about the potential cost-effectiveness of intervention options is an important component of the ACE approach to avoid methodological confounding. This suggests that basic components of the technical analysis that generate the incremental cost-effectiveness ratios (ICERs) need to be applied **consistently** to both the Indigenous and non-Indigenous applications. On the other hand, we are conscious that standard concepts of benefit based on health gain experienced by individuals (such as the ‘QALY’ or ‘DALY’), may not give justice to how benefit may be perceived by Indigenous Australians, who may give particular importance to the health of the extended family and/or community and health services that foster cultural security. Supplementary analysis will be undertaken that defines and applies a different concept of ‘benefit’ that is acceptable to Indigenous Australians. However, the main aim of the Indigenous component of ACE–Prevention is to estimate the difference in cost-effectiveness between preventive interventions addressing Indigenous and non-Indigenous Australians and thus, advise on the appropriate distribution of resources to preventive services for Indigenous and non-Indigenous people. That can only be achieved if benefits are measured in the same way in the Indigenous and non-Indigenous analyses. Separately, using a different concept of benefit as defined by Indigenous people, cost-effectiveness ratios can be recalculated accordingly. This may help to answer the question: “what are the most cost-effective preventive interventions for non-communicable disease from an Indigenous perspective and how should resources be allocated within Indigenous health services?”

A similar issue arises in relation to the selection and specification of interventions for evaluation. In addition to the Indigenous/non-Indigenous component of this research, another important component was “...a focus on delivery of interventions through primary care”, on “...how and by whom” interventions will be delivered, and on how the delivery mechanisms interact with cost-effectiveness. As with the concept of benefit, this raises options from applying “standard” interventions suitable for non-Indigenous populations to Indigenous populations; to adapting the design of these interventions to meet guidelines on what constitutes a culturally sensitive service; through to fundamental re-design of the way in which we provide services to the Indigenous community. This suggests that careful



thought needs to be given to how interventions are defined as we move from the non-Indigenous population through to the Indigenous population. It is important to note that while key issues of intervention design will be addressed, the central focus of a priority setting project like ACE-Prevention is on allocative efficiency<sup>3</sup> ('what to do?'), rather than on technical/productive efficiency ('how to do it?'). This distinction will influence the choice and level at which components of the evaluation are undertaken (e.g. how fine-grained the cost analysis is).

### Summary Box 1. Decision Context

1. the knowledge transfer objective of the project is to assist policy-makers and health service managers in making practical decisions about what prevention services to provide;
2. the main aim of the Indigenous component of ACE-Prevention is to estimate the difference in cost-effectiveness between preventive interventions addressing Indigenous and non-Indigenous Australians and thus, advise on the appropriate distribution of resources; for this purpose, costs and benefits will need to be measured in a consistent manner in the Indigenous and non-Indigenous analyses; separately, the Indigenous cost-effectiveness ratios will be recalculated using a broader concept of benefit as developed in consultation with the Indigenous Steering Committee;
3. specific consideration needs to be given to the definition and use of evidence, particularly in the context of Indigenous health; and
4. there is a balance in the choice of methods between the level of detail required to pursue the technical efficiency objective (how best to provide chosen interventions) and the number of interventions that need to be evaluated to pursue the allocative efficiency objective (what to provide). The emphasis in ACE-Prevention is on allocative efficiency. Issues around technical efficiency will be explored for a subset of the interventions where the research team decides it will provide important additional information to policy makers.

---

<sup>3</sup> Economics generally distinguishes three concepts of efficiency. The first two address the supply side and are sometimes rolled into one in introductory textbooks. 'Technical efficiency' is achieved when production is organised so that maximum output is produced with the resource inputs available. 'Productive efficiency' (sometimes called 'cost-effectiveness efficiency') is achieved when production is organised to minimise the cost of producing a given output. It takes into account both the production function and prevailing factor input prices (salary and wages, rent, interest and normal profit). The third and most important concept of efficiency, particularly for strategic planning and priority setting, is 'allocative efficiency', which incorporates the demand side and is achieved when resources are allocated to produce the optimal level of each output in line with the value consumers/society place on them.

## 2.3 Choice of comparator(s)

The choice of the comparator(s) is important in economic evaluations because study results are determined by the *net additional cost* of the new intervention (that is, option for change) in relation to the *net benefits* achieved relative to the comparator.

In the methods adopted in traditional economic evaluations, the usual comparator to the interventions under study is the 'status quo' or 'current practice' –often called the 'base case'. This is because one of the fundamental questions for economic evaluation to answer is what difference the option for change makes to current policy. This is often referred to as 'incremental analysis', with the resulting cost-effectiveness ratios referred to as incremental cost-effectiveness ratios (or 'ICERs'). The research question being addressed is: "What is the cost-effectiveness of replacing existing practice for dealing with the health problem with the new intervention?" This relates most closely to short term policy decision making: "what difference will it make if I do or do not invest resources into this intervention at this moment and in this context?" Such results are best presented in a cost-effectiveness league table at the end of the ACE-Prevention project indicating to policy makers what the best opportunities for improving the efficiency of the preventive health services for non-communicable disease in Australia.

If the intervention of interest is currently not implemented, the comparison would be between a 'base case' of existing prevention and treatments for the affected health problems and that 'base case' with the addition of the new treatment **or** the new treatment replacing an intervention within the current mix of interventions. Example of the latter case would be the replacement of one type of drug for another (e.g. 'new' atypical anti-psychotic drugs for the older typical drugs) or the replacement of current QUIT-line to assist smokers to stop with an extended version of the same intervention in which clients are called back after making initial contact. In these two examples, the new intervention replaces the old one as they are mutually exclusive (in case of the usual and extended QUIT line service) or against best practice rules (in case of the two types of anti-psychotic drugs which should normally not be combined).

One argument for the use of "current practice" as the comparator is that the result could be misleading if a "do nothing" comparator was modelled, where in reality current practice was already moderating the health problem. A counter-argument is that sometimes the reverse may be true if current practice is very inefficient. This would make a 'new' intervention look unduly favourable and it may be more informative to model a "do nothing" comparator. A 'new' intervention may appear cost-effective in comparison to an inefficient current practice (due to low effectiveness and/or high cost) as the cost-effectiveness ratio is determined by the difference in costs and benefits between intervention and comparator. The same intervention analysed against a 'do nothing' scenario may have very different cost-effectiveness credentials. Analysts need to be on the watch for "straw men" comparators – that is, comparators that artificially inflate the cost-effectiveness of the intervention of interest, because they are so cost-ineffective.

The notion of a 'do nothing' comparator has been developed and applied in a comprehensive fashion in a new approach to economic evaluation developed by the WHO, called the 'Generalised Cost-Effectiveness Approach' (Tan-Torres Edejer et al. 2003). In this approach the fundamental question being answered is not what difference the option for change makes to current practice, but rather what is the optimal intervention (or mix of interventions) if we could start from scratch (called the 'null comparator'). The GCEA approach assumes that 'current practice' is not in place, and compares the costs and benefits stemming from the introduction of a new intervention with the costs and benefits in the absence of all interventions (the 'null set'). In other words, all costs and outcomes will

be attributed to the intervention (involving average cost-effectiveness ratios, not incremental cost-effectiveness ratios), as it assumes there is no existing program in place. Modelling of the null set is an important aspect of this approach and requires back-calculations to life without current policy interventions. While such calculations can be quite complex, the advantage of this method is that it allows for the identification of current inefficient allocations of resources, as well as opportunities presented by new interventions (Tan-Torres Edejer et al. 2003).

The research question being addressed by the WHO-CHOICE methods is: "What is the most cost-effective mix of interventions for dealing with the health problem?" This relates more to long-term policy decision making: "what is the most efficient approach to dealing with this health problem?"; "how far removed from this ideal is current practice?"; and "how can we strategically redirect resources and/or build in incentives and disincentives to move from current practice closer to ideal practice?". Such results are best presented in an intervention expansion pathway that shows the most cost-effective mix of interventions given a certain budget level and plots the position of current practice in terms of cost-effectiveness in relation to the ideal mix.

In reality, in large projects like WHO-CHOICE and ACE-Prevention when many interventions addressing a wide range of health problems are evaluated it is impossible to calculate a 'true' null option of no health service intervention. Instead the pragmatic approach is to define a 'partial null' that is the theoretical level of disease that would be present if none of the interventions under scrutiny that address a particular health problem were in place.

The approach in ACE-Prevention will be to bring the two methods together in a consistent manner. The back-calculation from current burden of disease to the 'partial null' is done using the same assumptions on effective coverage, effectiveness and costs of current practice as would be used to calculate the incremental cost-effectiveness of changing current practice by adding or replacing one or more interventions. In other words, the modelling from current practice back to the 'partial null' should mirror the reverse of modelling costs and benefits from the partial null to current practice.

Given such issues surrounding the specification of the comparator, the choice of comparator will always be clearly specified in the ACE-Prevention study. Where appropriate, both a 'current practice' and a 'do nothing' comparator will be specified. The intervention and comparator(s) will be clearly specified and illustrated using a decision tree approach and/or a series of boxes that describe the major components of the intervention, together with the activities within each component. These intervention/comparator diagrams may well be accompanied by target population flowcharts as required (further detail provided later). The detailed specification of the activities for the average client ("who does what, to whom, when, where, and how often?"), provides the foundation to identify data needs and to decide how data will be organised and collected. This process is often called 'pathway analysis' in economic evaluation, but we won't be using this term as it may cause some confusion with another term ('expansion path analysis') that will be used in considering the optimum mix of interventions.

## Summary Box 2. Choice of comparator(s)

**Thus in ACE: Prevention, the approach to the selection of comparators is as follows:**

1. for all interventions we specify current practice and a 'do nothing'; as the comparator;
2. we define the 'do nothing' comparator as the 'partial null' back-calculated from current burden of disease to a hypothetical situation where all relevant interventions under scrutiny in the analyses for the same health problem are absent;
3. the back-calculation to the 'partial null' should describe current practice in terms of expenditure and health outcomes when the same intervention parameters of costs, effective coverage and effectiveness are applied;
4. results of the incremental analyses against current practice will be used in the league table of cost-effectiveness ratios for all interventions analysed within ACE-Prevention; and
5. the results of average cost-effectiveness ratios will be used in defining the optimum mix of preventive interventions in an intervention expansion pathway in contrast to current practice.

## 2.4 Choice of study perspective

Clear specification of the study perspective provides the foundation for identifying and measuring the costs and consequences that come together in the cost-effectiveness<sup>4</sup> indices (the cost per DALY results). This is considered further under the cost measurement and benefit measurement sections of the protocol.

A health sector perspective has been adopted for the economic studies undertaken in ACE-Prevention, with a major focus on impacts for government as third-party funder. A full health sector perspective includes:

- the government as health service funder (Commonwealth, States and Territories); together with
- impacts of the interventions on clients and their families (including out-of-pocket costs; travel costs; time costs involved in travel, waiting, treatment and recuperation; and carer costs).

An important alternative is to adopt a narrower perspective, which considers only the impacts on 'government as third-party funder'. This government perspective is adopted in

---

<sup>4</sup> Some economists make a distinction between cost-effectiveness analysis (CEA) and cost-utility analysis (CUA). The difference is that in CEA benefits are measured in a single physical dimension (for example, deaths saved, life years saved or cases of disease averted) while in CUA the measure of health benefit is a utility or preference-based measure (such as the QALY or DALY) combining a mortality and a morbidity component. Throughout this document and the ACE-Prevention study, the nomenclature of Gold et al. (1996) will be followed where CEA will be assumed to be a generic term that includes CUA and CEA is distinguished from cost-benefit analysis (CBA) and financial appraisal where outcomes are measured in dollar terms. At times, however, the more specific CEA and CUA terms will be used, particularly where specific cost outcome ratios are being referenced.

many economic studies and reflects issues of data availability, time available to researchers to complete their appraisals, and decision context. While it has pragmatic appeal, this perspective omits important issues for prevention policy in Australia. It does not, for example, include the cost impacts on the private sector, including clients and their families/carers and non-government organisations, which together account for approximately 33% of current health expenditure. Cost impacts on clients may also have major impacts on participation rates and are sometimes important aspects of intervention design.

It is acknowledged that taking a full 'health sector' perspective usually involves far more work than the narrower 'government as third party funder' perspective. Nonetheless, it is proposed to take a full 'health sector' perspective in ACE-Prevention, subject to the availability of data and resources. This means that all health sector impacts will at least be identified and then either included or excluded in the measurement stage in accordance with specified criteria. Layering of the different perspectives would also enable separate reporting of the costs in accordance with who bears them.

An alternative would be to broaden the perspective and adopt a 'societal perspective', which includes all costs and impacts across all sectors of the economy, irrespective of who is affected. A societal perspective would thus include impacts on non-health sectors (such as education, food supply, police/courts, production gains/losses in the broader economy, etc), in addition to the health sector impacts specified above. This may be important for some diseases/risk factors that involve interventions outside the health sector (e.g. illicit drugs, alcohol or obesity prevention). When non-health sector impacts are important to an intervention (either on the cost or outcome side), it will be important to: i) flag the issue; and ii) undertake sensitivity analysis to assess the significance of adopting the broader perspective. The alternative of adopting a societal perspective for the whole ACE-Prevention study is simply not feasible, given the size of the evaluation task and the limited time and resources available.

### **Summary Box 3. Choice of Study Perspective**

**Thus in ACE: Prevention, the approach to study perspective is as follows:**

1. all analyses will be carried out from the health sector perspective;
2. costs to government and the 'private sector' will be kept separate in the analyses to allow commenting on the contribution of each to overall costs; and
3. for relevant interventions with large non-health benefits or costs, a broader societal perspective will be adopted, separately.

## 2.5 Study design

A number of economic evaluation methods are available. These include cost-minimisation analysis (CMA); cost-effectiveness analysis (CEA); cost-utility analysis (CUA); and cost-benefit analysis (CBA). As a general guide the complexity of the economic method should match the complexity of the research question, and usually this is heavily influenced by how "benefit" is to be defined.

The focus of ACE-Prevention is on evaluating multiple interventions (i.e. allocative efficiency) and therefore requires a form of analysis which provides a way of measuring outcomes in a consistent way across multiple risk factors and diseases, with varying impacts on morbidity and/or the prevention of premature mortality. To achieve this, the primary analytic method chosen is cost-utility analysis (CUA), with outcomes expressed in terms of 'cost per health-adjusted life year saved'. As the health adjustment relies on the same disability weights used in constructing disability-adjusted life years (DALYs) in burden of disease analyses, we equate the cost per health-adjusted life year saved to a cost per DALY averted. It should be noted that there is a fundamental differences in the construction of a DALY used in burden of disease and the concept of health-adjusted life years saved. In burden of disease studies loss of health due to mortality is valued as Years of Life Lost against an ideal that everyone in a population lives into old age (and 'free of disease'). To reflect that ideal each death is valued as the remaining life expectancy for that age from a standard life table. In economic evaluations the difference in years lived between an intervention and comparator scenario is estimated based on actual mortality rates (and hence life expectancy). Despite these differences, it has become the custom in the literature to use the term DALY both in burden of disease studies and economic evaluations interchangeably and that is what we will do in ACE-Prevention as well.

CUA measures the additional costs imposed by an intervention over the comparator, expressed as a ratio in relation to the additional health benefits achieved, with the focus on the length and quality of life outcome produced or averted by the intervention. The DALY is chosen as the quality of life measure because DALY information is available using consistent methods across a wide range of diseases and risk factors, for both non-Indigenous and Indigenous Australians.

There is the potential to also report secondary outcomes from cost-effectiveness analysis (CEA), where the results are expressed as a '\$ cost per physical unit of effect'. In this study, '\$ cost per life year' is a secondary outcome which could be presented, together with outcomes specific to the intervention/risk factor or intervention/disease context (such as \$ per small cancer detected; cost per BMI change; cost per adverse event avoided; etc).

The Australian Government has also expressed an interest in using cost-benefit analysis (CBA) in previous ACE studies, which values the consequences of an intervention in monetary units to determine whether an intervention's benefits exceeds its costs and gives rise to a net social benefit. This requires a \$ value to be placed on life, which normally involves approaches such as 'willingness-to-pay', 'revealed preference' or 'human capital'. Some of these techniques require empirical data collection which is beyond the capacity of ACE-Prevention to undertake and there is no similar pre-existing data base analogous to the DALY information. A crude form of CBA may be undertaken in ACE-Prevention if time permits, but would utilise literature-based estimates, rather than involving any empirical data collections. The simplest approach would be to place a \$ value on the DALY, such as the \$30,000 per DALY used in *Best investments* (Swinburn & Gill 2002) or the \$50,000 per DALY used as the decision threshold in previous ACE studies or the 3 x GDP threshold used by WHO-CHOICE and the Commission on Macroeconomics and Health (<http://www.who.int/macrohealth>).

The selected study design (CUA/CEA) will be employed within a Program Budgeting and Marginal Analysis (PBMA) framework. Placing conventional economic evaluation within the PBMA framework offers a number of advantages, including:

- a broader concept of 'benefit' that can incorporate policy objectives important to decision-makers;
- the involvement of stakeholders via the Working Group approach;
- a systematic approach to the selection of options for change; and
- greater potential for ownership of results by stakeholders.

The role of PBMA within the ACE approach is further explained in Appendix A.

In addition to placing the economic analysis within a PBMA framework, a two-stage approach to CEA may be taken for some interventions with detailed information on a single trial rather than relying on meta-analysis of many trials, very akin to the process adopted by the Pharmaceutical Benefits Advisory Committee (PBAC). The first part in this approach is to model *cost-efficacy* for the Australian population based on the same assumptions on costs and efficacy as in the trial data. In other words, it will reflect the trial conditions as closely as possible. The second part is a *cost-effectiveness* evaluation with assumptions modified to model the introduction of the interventions under real life conditions. This involves a stepwise introduction of changes to the cost-efficiency model including:

- application of the intervention to the population of interest (and not just the trial population);
- extrapolation of evidence beyond the time horizon of the trial;
- an adjustment of the expected impact and coverage in the context of health services in Australia; and
- modification of resource use to reflect Australian routine health service conditions

An advantage of this two-stage approach is that the source and uncertainty of data assumptions are made explicit. To see worked examples of this two-stage approach, ACE-Prevention researchers should refer to distributed briefing papers from the ACE-Obesity project.

#### Summary Box 4. Study Design

**So in summary, the study design for ACE-Prevention will involve:**

1. Cost-utility analysis (\$ per health-adjusted life year saved or DALY averted) is the primary form of analysis, sometimes supplemented by cost-effectiveness analysis (\$ per life year; \$ per unit of effect);
2. A cost-benefit analysis may complement the cost-utility analysis of interventions for selected topic areas with a large expected non-health cost and/or benefit component; and
3. The CUA/CEA analysis will be placed within an overarching PBMA framework provided by the Steering Committees and 2<sup>nd</sup>-stage filter analysis.

## 2.6 Selection & description of competing alternatives

### Criteria for selection of the interventions

A fundamental question in critically assessing economic evaluations is whether any important alternatives have been omitted. The weight given to the selection of comparators is a direct reflection of the importance placed on the concept of 'opportunity cost' in economics (that is, the notion that the real cost of an intervention is the benefits forgone in alternate use of the resources involved). The issue relates both to the correct specification of the options for change, as well as to the specification of the 'base case' ('current practice' and/or 'do nothing' comparators). In the context of an economic evaluation addressing a single topic or problem, there is usually a reasonably limited set of possibilities.

In the context of a study addressing priority setting across an area as broad as the prevention of non-communicable disease, there is a wide range of possibilities. This makes the process by which the options for change are selected both an important theoretical issue and an important policy issue. On the other hand, the intention of ACE-Prevention is to provide policy makers with a comprehensive overview of the cost-effectiveness of prevention for non-communicable disease by analysing 100 interventions. That number of 100 interventions was chosen to approximate the intended comprehensive overview. Of course, one can debate what constitutes an separate intervention or would rather be considered a 'subset' of an intervention. For instance, dietary counselling to address blood cholesterol can be provided by a GP or a nutritionist. This could be construed as a single intervention with two types of provider or as two separate interventions. If the former rule is taken it could well be that 100 interventions do not cover the full gamut of prevention for non-communicable disease. That would also mean that there is more leeway for the Project Steering Committee (PSC) to influence the choice of interventions.

With that in mind, after discussion, the PSC agreed at its first meeting in August 2005, to the following criteria to guide the selection of interventions:

1. Size of the problem addressed;
2. Importance in terms of current investment;
3. Relevance to current policy decision-making;
4. Availability of evidence of efficacy/effectiveness to support the analyses;
5. Indications that additional investment for an intervention would lead to significant health gain or, conversely, that decreased investment would lead to little or no reduction in health outcomes;
6. The ability to specify the intervention in clear concrete terms to facilitate meaningful evaluation;
7. The inclusion of a mix of interventions from across the prevention pathway and from a range of settings;
8. Specific relevance to the health of Indigenous Australians (for the Indigenous component of the analysis); and
9. Considerations of program logic (for example, the inclusion of an option that is ineffective as a stand-alone intervention or for which there is poor effectiveness evidence, but which is integral to the success of a package of interventions).

It was further agreed that the application of these criteria would take place as a two-step process: i) the selection of broad topic areas by the PSC (added to as appropriate by the



Indigenous Steering Committee (ISC); and ii) the selection of specific interventions within the topic areas by the Technical Advisory Panels (TAPs) and by the ISC<sup>5</sup>.

## **Topic areas selected by the Project Steering Committee (PSC)**

At its initial meeting in August 2005, the PSC selected the following topic areas as priority areas to start work on:

1. lifestyle risk factors: physical activity, nutrition, cholesterol, blood pressure, obesity, illicit drugs, alcohol, tobacco, UV exposure, osteoporosis and falls prevention;
2. cancer screening;
3. mental health and suicide prevention;
4. musculoskeletal disorders;
5. cardiovascular disease;
6. diabetes;
7. oral health (caries and periodontitis); and
8. SIDS (for Indigenous population and as part of smoking).

The following areas were agreed as having lower priority:

1. air pollution;
2. vision and hearing; and
3. congenital.

The following areas were tentatively agreed for inclusion in the list of benchmark interventions (although this list was flagged as requiring further discussion):

1. hepatitis B vaccination;
2. knee and hip replacements;
3. epilepsy treatment;
4. asthma;
5. childhood vaccinations and
6. at least one curative intervention from each non-communicable disease topic area (for comparison).

In making these initial selections, the PSC and research team were conscious that the selections reflected the membership of the PSC at its initial meeting (some members were absent) and may need to be re-visited as the project proceeded.

---

<sup>5</sup> Note that the ISC functions at both the Steering Committee level and at the TAP level for the Indigenous component of the analysis. While ISC members will sit on the various TAPs and offer their guidance on the Indigenous component of ACE: Prevention, the ISC will also need to consider interventions that originate as Indigenous-specific interventions. So the ISC will consider interventions that originate from the TAPS selected initially for reasons associated with application to non-Indigenous Australians, as well as interventions it may select itself.

## Technical Advisory Panels selected by the Project Steering Committee

The Steering Committee also agreed at its initial meeting to the creation of the following Technical Advisory Panels (TAPs) to select specific interventions within the broad topic areas using the agreed selection criteria. The agreed TAPs were:

1. obesity, physical activity, and nutrition.
2. alcohol, illicit drugs, and tobacco.
3. cancer screening.
4. mental health.
5. cardiovascular disease, blood pressure, and cholesterol.
6. musculoskeletal disorders and osteoporosis.

Some of the TAPs cover multiple areas. Initially, it is probably a good idea to hold separate meetings for specific topics (e.g. alcohol, and tobacco) and towards the stage where multiple interventions have been analysed a broader TAP to finalise second filter criteria and recommendations would have value.

Going on the experience of previous ACE studies, the TAPs may find it challenging to apply the selection criteria. It is anticipated that this might be due to: i) a lack of sufficient information on the criteria in relation to what the various possible interventions might involve; ii) the number of possible interventions to choose from; and/or iii) different perspectives among TAP members about the relative importance of the different criteria.

It is recognised that selection of interventions inevitably involves a subjective judgement, but we can try to take it in a 'scientific' way. Under the ACE approach, this issue has been addressed previously in two ways: i) classify possible interventions into groups reflecting their evidence base; and ii) develop 'scoping papers' on interventions that have potential, in order to clearly specify the intervention and to summarise available information on their efficacy/effectiveness and likely cost.

In relation to classifying potential interventions by their evidence base, Table 3.1 provides a useful template, based on previous ACE studies. Once put into the appropriate groups, the criteria can then be used to govern the selection of the interventions from group 1 or 2. Most interventions would come from group 1, but there may be sound reasons for choosing the odd one from group 2 (e.g. policy relevance; size of the potential impact, etc).

**Table 2.1: Categorisation of interventions to assist selection process by TAPs & researchers**

Rank order	Group	Definition
1	Likely options for change	<ul style="list-style-type: none"> <li>Interventions where sufficient 'evidence'<sup>6</sup> exists to indicate that strategies involving additional investment would be associated with significant reductions in the prevalence of the risk factor/disease and little or no likelihood of causing harm.</li> <li>Interventions where sufficient 'evidence' exists to indicate that strategies involving reduced investment would be associated with no increases or insignificant increases in the prevalence of the risk factor/disease and little or no likelihood of causing harm.</li> </ul>
2	Possible options for change	<ul style="list-style-type: none"> <li>This group includes: <ul style="list-style-type: none"> <li>(i) interventions where some evidence exists to indicate that strategies involving additional (less) investment would be associated with significant reductions (insignificant change) in the prevalence of the risk factor/disease and little or no likelihood of causing harm.</li> <li>(ii) interventions for which it is difficult to conduct rigorous trials, but program logic strongly suggests their likely effectiveness and/or their place within a coherent package of interventions.</li> </ul> </li> </ul>
3	Other options for change which require monitoring	<ul style="list-style-type: none"> <li>This group includes: <ul style="list-style-type: none"> <li>(i) ideas for action that are considered to have merit but are too broad and abstract to evaluate (and for which specific research work has not been developed), or are politically sensitive.</li> <li>(ii) interventions that are currently being worked on and/or implemented in another context, or which require more research before they can be evaluated, that is, evidence does not exist to sustain their efficacy/effectiveness credentials and a clear intervention cannot be specified.</li> </ul> </li> </ul>

## Interventions selected for analysis

Appendix C shows a detailed list of interventions that are currently being analysed or that are being considered for analysis. These interventions represent a wide range of non-communicable disease areas or risk factors for non-communicable disease. Benchmark

<sup>6</sup> For a definition of 'sufficient evidence', see Table 7.1

interventions are chosen (a) to have at least one curative intervention in each major disease area; (b) to include prominent curative interventions for non-communicable and communicable diseases; and (c) to cover major infectious disease control interventions. The list currently contains 144 preventive interventions for non-communicable disease and 45 'benchmark' interventions. This is well above the total of 150 interventions. Some of the topic areas have not yet been fully explored and this may influence the details in the list. To some extent there is also a requirement to group a number of interventions to provide a reasonably homogeneous set of interventions for a particular researcher, particularly for those who are doing the analyses as part of their PhD studies. This means that there will remain uncertainty about the exact number and type of interventions that are part of ACE-Prevention's work plan until close to the end of the project.

## 2.7 Target population

The target population is the population for whom the program impact is modelled. There is a range of approaches adopted in economic evaluation for specifying the population and target groups for which the interventions are intended. These include the use of a notional population of 100,000 people; a standardised population; an actual national or state population; a steady-state population or one that incorporates anticipated demographic changes through time; a simulated population; or combinations thereof. The choice of target population is often integrated, therefore, with the 'study design' decisions addressed later in the protocol.

In this study, all models will take the actual Australian population in the year 2003 as a starting point. The preferred approach is to model those in the 2003 population who are eligible for the intervention over their remaining lifespan, rather than to add new cases year by year to the 'eligible population' to mimic changes over time in the Australian population.

The target groups within the Australian population for which the interventions are intended need to be clearly specified. The relevant target group will vary depending on the specific intervention; it may target all persons in a particular age group, or a specific group of people. Depending on the intervention this may be people of a given age or sex, people with a particular disease or risk factor, or people defined by a cluster of such characteristics. This requires clear illness descriptors (using ICD-10 criteria) for interventions directed at diseased populations or a precise description of the 'at risk' target population for preventive interventions.

### Summary Box 5. Target population

**So in summary, the target population for ACE-Prevention will involve:**

1. Identification of the target group in the Australian population in the year 2003 as the potential recipients of each intervention to be modelled; this can be: a) the whole population such as for population-wide health promotion campaigns; or b) a sub-population based on characteristics such as age, sex, risk factor profile or disease
2. The target groups within the Australian population for which the interventions are intended need to be clearly specified and justified; this should flow from the clear description of the interventions (see section 3.1).

## 2.8 Study boundaries

In framing the economic evaluation, the analyst must also consider the boundaries, or scope, of the study. Defining the boundaries of a study can be thought of as drawing a circle around the study to contain it. As Gold et al (1996) argue:

*"In circumscribing the study, the analyst must attempt to balance the need to capture all significant effects of the intervention that will be relevant to the decision-maker with the need to contain the study to the form of a manageable and feasible study." (p.68).*

Spill-over effects ripple out from every intervention and the question is how far to follow them. These may include: production gains/losses in the general economy; side effects related to the intervention; and all-of-life effects (i.e. whether to count the unrelated ongoing health care costs of people who are alive because of the intervention). Part of the framing task is to draw practical limits around the analysis. Two generic aspects of scope can be differentiated. The *first* concerns the groups of people to be considered in the analysis. There is clearly some overlap on this issue with section 2.7 above. One way to deal with this is to think of section 2.7 as specifying the target population in terms of the intervention, while this section adds protocol-driven limitations, mostly from the study perspective (e.g. inclusion or exclusion of cost impact on patients and their family). Whatever the judgement, the analyst should clearly delineate the groups of people who are included and explain the exclusion of other affected groups. Although economic theory prescribes the inclusion of consequences for all (most) affected parties, it has not in fact been common practice to include health-related quality of life effects on persons other than the patients or those affected directly by the intervention. Gold et al (1996) note that little research has been done on health-related quality of life impacts for family members and/or carers, for example, and that little precedent exists for including these effects in economic analysis. The approach in ACE-Prevention is to exclude the impact on (health-related) quality of life of carers and family members but to include carer and family time in our costing. With the adoption of a health sector perspective for the main ACE-Prevention analyses, it follows that we will not include production losses and other non-health sector outcomes and costs. However, for selected interventions in separate analyses we will explore a broader, societal perspective. This will be particularly relevant for topic areas such as alcohol, illicit drugs and conduct disorder.

A *second* aspect of scope involves the type of health outcomes to be counted. In framing the study, the analyst must consider which types of outcomes are important and which type of economic evaluation is appropriate. In ACE-Prevention, we use cost-utility analysis (section 2.5) measuring health outcomes (including the benefits and harmful side-effects) as a difference in health-adjusted life years lived between the intervention and comparator scenario as described in section 2.5. This includes an adjustment for expected levels of disability by age and sex also for conditions not immediately affected by the intervention(s) of interest. While adjusting the denominator for future health loss unrelated to the intervention(s) of interest we decide not to add the costs of treating unrelated future health loss in our numerator of the cost-effectiveness ratio.

Finally, the analyst also needs to be aware that the decision-maker often has to consider a broader range of factors than just economic efficiency (defined as cost per health gain). These may include, for example, the equity implications of the options for change; the size, significance and severity of the health problems addressed; the affordability and feasibility of the intervention on a national scale; and their acceptability to a range of stakeholders. The relevance of these broader considerations has led to the development of different types of economic evaluation, which have the potential to take them into account. Some of these approaches adopt a balance sheet approach to complement the arithmetic (such as the cost-consequences approach); while others try to broaden the arithmetic to include them (such as the cost-value approach (CVA) of Eric Nord or the Options Appraisal of Gavin Mooney). The basic approach used in ACE-Prevention will be to complement the arithmetic by using our 2<sup>nd</sup>-stage filter approach. We will give consideration in the Indigenous analysis

to modifying the '\$ per DALY' ICERs by incorporating additional considerations of health gain such as 'community health gain', cultural security and equity.

### Summary Box 6. Study Boundaries

**So in summary, the study boundaries for ACE-Prevention will involve:**

1. inclusion of health benefits for participants only (i.e. not for carers or family members – but subject to review and inclusion via sensitivity analysis [note where relevant the impact on carers/families will be included as time costs]);
2. inclusion of positive and harmful effects directly related to the intervention(s) of interest;
3. adjustment for non-fatal health loss unrelated to diseases affected by the intervention of interest but exclusion of the cost of treating future 'unrelated' disease;
4. exclusion of production gains and losses and other non-health sector impacts (subject to review and inclusion in the sensitivity analysis);
5. consideration of second filter criteria to complement the cost-effectiveness results; and
6. development of an Indigenous concept of benefit to be used as an alternative to the results presented as a cost per individual health gain captured in the DALY

## 2.9 Time horizon

The time horizon for the provision of each intervention will vary according to the nature of the intervention. For instance, the duration of nicotine-replacement for smoking cessation (recommended for up to three months of use) is very different from that of anti-retroviral drugs in the treatment of HIV/AIDS or insulin in type 1 diabetes which need to be taken over a lifetime in order to continue to reap benefits. The time horizon for tracking the associated costs/cost offsets should reflect a period as long as costs and benefits continue to accrue. The consequences of preventive intervention will often extend over a lifetime of the target population.

Our choice of time horizon is guided by the following principles:

- the intervention time horizon realistically reflects how the intervention would be applied in real life;
- follow-up time allows all relevant costs and benefits attributable to the intervention runtime to be measured;
- all relevant time lags (i.e. between episodes of illness/care or before health gains are experienced) are incorporated and duly referenced;
- avoiding bias is more important than standardisation for its own sake.

The analyst should document the intervention time horizon chosen, together with the underlying rationale. In other economic studies choices have ranged from assuming the intervention to be in 'steady-state'<sup>7</sup> and modelling over a representative period (possibly

<sup>7</sup> We use the term 'steady-state' to characterise an intervention that is fully implemented and operating in accordance with its effectiveness potential. The question we would be answering is: "What is the cost-effectiveness of this intervention if it achieves its full potential as per the evidence on effectiveness".

one-year), to modelling a full national rollout with an associated 'learning curve' and capacity utilisation assumptions, over say a five year period. In ACE-Prevention it will be assumed that all interventions are in 'steady state' operation. Thus, the CEA analyses will not include a roll-out phase and will not address implementation and learning curve issues. However, issues related to feasibility of implementation will be considered in the 2<sup>nd</sup> stage filter analysis.

It is important to note that the 'learning curve' issue can be quite significant for some interventions, such as screening, where high recall rates can be expected until screen readers gain confidence in calling 'positive' screens. But, while useful from a practical policy and clinical perspective, economic evaluation of a full roll-out period can take considerable time and resources to complete and may not be possible for a project like ACE-Prevention. On the other hand, we have to be conscious of trial data that are based on the performance of interventions before they reach this 'steady-state'. This occurred, for example, with the walking school bus intervention in the ACE-Obesity study, where part of its poor C/E performance was postulated to be due to very poor capacity utilisation documented at the early stage of its implementation. When we make the 'steady-state' assumption, we need to make sure we are being fair to the intervention in terms of the cost/throughput/outcome assumptions we make and the corresponding data we use – neither unduly over-optimistic nor unduly pessimistic. This can be a judgement call, particularly when the evidence base is weak.

#### **Summary Box 7. Time Horizon**

**So in summary, our choice of time horizon should be guided by the following principles:**

1. The intervention time horizon should realistically reflect how the intervention would be applied in real life; and the follow-up time should allow all relevant costs and benefits attributable to the intervention to be measured.
2. All relevant time lags should be incorporated and duly referenced.
3. Avoiding bias and including relevant effects is more important than standardisation for its own sake.

**The choice of time horizon in ACE-Prevention will involve:**

1. Modelling all interventions in 'steady state' operation.
2. Tracking the costs and outcomes for as long as they accrue through time in the 'eligible' population in 2003; often this means modelling the target population through to death or 100 years of age; and
3. Issues related to feasibility of implementation will be considered in the 2<sup>nd</sup> stage filter analysis.

## **2.10 From conceptualising to undertaking the CEA**

The early conceptualisation and planning steps covered so far are essential for focusing the study on relevant research questions, maintaining the focus as the study progresses, and avoiding analytical pitfalls midway through the analysis. Now

we need to sort out the 'study design' questions and appreciate the practical steps in conducting a CEA.

For the 'study design', the key elements involve designing the data collection and analytic plan. These involve three basic decisions:

1. Decide on the conceptual model that best describes the intervention and its effects on health outcomes (i.e. whether to structure the analysis around an event pathway; a decision tree; a Markov model; a simulation, etc);
2. Determine how the data will be collected on activities, costs and outcomes for the intervention and the comparator(s); and
3. Consider the most appropriate analytic methods to combine the information in an economic evaluation.

These questions will be addressed as we move into the task of assessing costs (Section 3) and outcomes (Section 4) for our chosen interventions.

The measurement of costs and benefits in economic evaluation is usually specified (Gold, Siegel et al. 1996; Drummond, O'Brien et al. 1997) in terms of the following three steps:

- **Identification** of the appropriate costs and benefits to include in the appraisal (i.e. the inclusion/exclusion criteria);
- **Measurement** of the resources used and saved by the program alternatives and the outcomes produced by each; and
- **Valuation** of the resources used (and saved) and outcomes produced in appropriate units for comparison

## The Identification Step

Drummond and colleagues (1997) provide useful guidance on principles/criteria to guide the selection of which costs/consequences to include in the study:

- inclusion/exclusion should mirror closely the study perspective(s) (section 2.4), a precise definition of the intervention(s) and reflect all important consequences for costs and health outcomes including positive impact and adverse effects;
- it is not worth investing a great deal of time and effort considering costs/consequences that, because they are small, are unlikely to make any difference to the study (such costs/consequences may be approximated or identified as relevant but not measured). This criterion can be broadened to include the notion of attribution (i.e. the size of the cost/consequence change that can be attributed to the intervention);
- it is not worth including costs/consequences that are merely likely to confirm a result that would be obtained by a consideration of a narrower range of key costs/outcomes (this may apply where costs/consequences are linearly related e.g. costs to service providers may be closely related to costs to patients);
- costs/outcomes common to both the comparator and intervention may be omitted; if they are the same with and without intervention they do not occur as a consequence of the intervention but rather the disease process and hence can safely be omitted;
- the overarching importance of opportunity cost, i.e. if there is no opportunity cost involved then there is no economic cost; similarly if there is an opportunity cost involved, then the cost should be included irrespective of whether market prices are involved (e.g. donated space, volunteer time, patient's time).

Each briefing paper should specify the inclusion/exclusion criteria being employed, i.e. provide for all relevant costs and consequences a clear statement as to which ones are



included/excluded and on what basis. This helps to achieve consistency across the evaluations.

## **The Measurement Step**

Key issues for the measurement of resource usage/outcomes produced are to ensure that:

- the target populations are clearly specified;
- costs and outcomes are measured accurately in appropriate physical units;
- the time horizon for the intervention itself and for the tracking of the cost and outcome profiles are clearly specified and justified;
- special circumstances (e.g. joint use of resources) are mentioned and the measurement approach (e.g. attribution rule) is clearly specified;
- rigour is applied to the cost data collections in an analogous manner to outcomes. Issues that arise include whether top-down or bottom-up costing has been followed; whether point estimates or stochastic data are available; whether real or modelled costs are being used; whether administrative collections, price schedules or real data are being used.
- appropriate distributions are defined around critical input parameters for the uncertainty modelling.

## **The Valuation Step:**

Key issues in the valuation step on the benefit side are:

- The choice of utility instrument to value impacts of the intervention on length and quality of life;
- Discounting;
- The application of community "value" weights for age, equity, severity or other aspects of social justice.

Key issues in the valuation step on the cost side are:

- The choice of unit prices to assign to the activities (or expenditure categories where more detailed costing is attempted);
- The decision as to whether available price are adequate or whether "shadow pricing" is necessary (Drummond, O'Brien et al. 1997, pp 55-56);
- Discounting and annuitisation of capital assets;
- Distinction between real and current prices. The analysis should be in real prices for the reference year 2003.
- Where overseas comparisons are made, the purchasing power parity method, not exchange rate values should be used.

The following two sections discuss these three steps of evaluation in greater detail for cost (section 3) and benefits (section 4).

### 3 Measurement of costs

The ACE–Prevention project adopts an economic approach to costing, which while similar to financial costing in many respects, also involves some important differences.<sup>8</sup>

As we introduced in section 2.10, there are three basic steps involved in costing from an economic perspective. These are:

- i) Identification of the appropriate costs to include in the evaluation;
- ii) Measurement of resources used and saved by the program alternatives; and
- iii) Valuing the resources used and saved by the program alternatives.

Fundamental to the identification and measurement of costs is a clear thorough understanding of exactly what activities the intervention and its comparator(s) involve. This is often characterised in economic texts with a phrase along the lines of: “who, does what, to whom, when, where and how often”. Because of the centrality of understanding this step in economic appraisal, section 3.1 (‘Describing the intervention and comparator’) is devoted to this issue. The sections 3.2 to 3.4 then explain how each of the three steps will be applied in ACE–Prevention followed by advice on how the cost results should be reported in the briefing papers (section 3.5).

#### 3.1 Describing the intervention and comparator

Fundamental to costing (and modelling of benefits) in an economic evaluation is a clear description of the course of events with the intervention compared to that without the intervention (sometimes called the ‘conceptual model’). In concrete and well-defined steps, this description outlines an ‘event pathway’ that generally includes the following elements:

1. Recruitment to the intervention (+/- training of providers);
2. Provision of the key elements of the intervention (e.g. advice, consultations, care, immunisation, etc);
3. A routine level of monitoring, evaluation & support; and
4. Downstream effects.

Depending on the particular study, there may also be a ‘patient flowchart’ that describes how we get from the target population to those who actually participate in the activities of the intervention (see Appendix D for a worked example), or even a ‘practitioner recruitment flowchart’ that describes how we recruit and train the practitioners who will provide the intervention (e.g. recruitment of GPs through Divisions of General Practice). Spreadsheet analysis, a decision tree, or more complex models may complement these flowcharts, as the analyst develops the ‘study design’.<sup>9</sup>

The intervention to be analysed should be clearly specified as early as possible.

---

<sup>8</sup> Those unfamiliar with economic costing are referred to basic economic evaluation texts particularly by Drummond et al (2005) ‘Methods for the economic evaluation of health Care Programs 3<sup>rd</sup> ed’ and the text by Gold et al (1996) ‘Cost-effectiveness in health care and medicine’.

<sup>9</sup> The analyst can then decide how to collect data for the activities described in the conceptual model. The tasks required here vary greatly depending on whether, and to what extent, the analysis will collect primary data; use existing data (e.g. perform secondary data analysis from administrative data bases or published reports); or estimate parameters using mathematical models. In ACE–Prevention our economic analyses will be based on analysis of existing data, because the size of the task generally excludes primary data collection (with the possible exception of some of our PhD students).

The key steps for the ACE Researchers in specifying the intervention are:

1. Collect all the efficacy/effectiveness literature on your intervention (not just the review articles and not just the cost-effectiveness literature);
2. Summarise the basic activities for the intervention and key design issues in each of the key papers (in review articles assure yourselves that the effect sizes relate to studies where the interventions are sufficiently homogeneous in terms of the resource utilisation);
3. Specify the version of the intervention where 'best evidence' exists and where national implementation has a chance of being acceptable.

Literature on the intervention may include a number of possible variations; for example, variations in the frequency of receiving the intervention, in the ages and types of patients involved or in the presence of co-morbidities or risk factors. Aspects of program characteristics that will be important are: the 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. The target population is also a critical aspect of the program definition.

Thus, there may be many versions (or service intensities) of the intervention that are evaluated in the literature. This is entirely appropriate, as one of the strengths of economic appraisal is its ability to demonstrate the relative cost-effectiveness of interventions given a range and variety of design options. However, it is clearly essential in framing the study for evaluation in ACE-Prevention, to define precisely what interventions and intervention variations are to be included. In an allocative-efficiency based evaluation like ACE-Prevention, there is less room (i.e. less time/resources) for extended evaluation on any one intervention and so the focus is more on selecting the 'best' version of the intervention to model from the outset, rather than building it up through marginal analysis. Now the problem here is that 'best' can be specified in different ways. In previous ACE studies we have interpreted 'best' to mean the intervention for which there is 'best evidence' and this was judged from the literature (with careful attention given to our definition of 'evidence'). In addition, we have tried to be alert to the possibility of using the ACE studies to design interventions that will be good policy choices.

The last point raises an important judgement issue for researchers, viz: 'feasibility' should not pre-empt selecting the intervention design with 'best' evidence (as the 2<sup>nd</sup> stage filters are designed to be applied after the technical analysis of the 'best' intervention design), but all else being equal, feasibility may be a factor that helps researchers make decisions about their intervention design. An example might be where the intervention in question uses nurse practitioners (UK) or specialists (USA) to provide a service, but where we know in Australia only GPs would be acceptable. In this instance, we might model our intervention with GPs, but still use the efficacy data from the trial (i.e. we make the assumption [and document it] that GPs are as efficacious as nurse practitioners or specialists for the task in question).

Related to this issue is the requirement that the intervention is modelled as part of **current** Australian health services. Even though the project reference year is 2003, this simply means that all prices are expressed in 2003 dollars. The organisation of health services however should reflect current practice (e.g. if a drug is recently listed on the PBS but was not listed in 2003, then this drug should be priced as listed on the PBS with prices deflated back to 2003).

In general, the definition and specification of the intervention should make clear to potential readers whether or not the cost-effectiveness results will apply to specific real life settings (this is an important aspect of external validity).

## Summary Box 8. Describing intervention and comparator for costing purposes

**Interventions should be described in concrete and well-defined steps that generally include:**

1. A description of the 'event pathway';
2. A 'patient flowchart' that describes how we get from the target population to those who actually participate in the activities; and possibly
3. A 'practitioner recruitment flowchart' that describes how we recruit and train the practitioners who will provide the intervention.

**The 'event pathway' will generally have the following elements:**

1. Recruitment element (+/- training of providers);
2. Intervention elements (e.g. advice, consultations, care, immunisation, etc);
3. Monitoring, evaluation and support elements; and
4. Downstream effects.

**Key steps for ACE: Researchers in specifying the intervention are:**

1. Collect all the efficacy/effectiveness literature on your intervention (not just review articles and not just the cost-effectiveness literature);
2. Summarise the basic activities for the intervention and key design issues in each of the key papers (consider link between resource utilisation & outcomes); and
3. Specify that version of the intervention where 'best evidence' exists or those which are most appropriate in the Australian health service context.

**Aspects of program design characteristics that will be important are:**

1. the target population;
2. specific technologies used;
3. the type of personnel delivering the service or treatment;
4. the site of delivery;
5. whether the service is bundled or piggy-backed with other services; and
6. the timing of the intervention.

## 3.2 Identification of relevant costs for inclusion

In the 'identification' phase of cost analysis, all the important costs and cost offsets relevant to the study perspective are identified and included in the study (stemming from the preceding section). Equally important, any costs excluded should be clearly specified and justified. Costs can be excluded on 'protocol grounds' (i.e. not relevant to the study perspective; not part of the specified 'event pathway') or 'practicality grounds' (i.e. there is no data available to estimate the cost; its likely importance does not warrant the effort required to model it; and/or a judgement that its omission will not bias results). It is important to note that:

- The identification phase is done for both the intervention and its comparator(s) and that consistency in approach is essential to avoid bias;
- The clear specification of the intervention and comparator 'event pathway' provides useful input in clarifying which costs are on the agenda for consideration;
- A similar identification process is taken with the measurement of benefits and symmetry in approach between cost and outcome assessment is desirable, but not essential (although any departures from symmetry should certainly be noted and justified).

In ACE-prevention the '**health sector**' perspective is mostly adopted. This means that costs and cost offsets that have an impact on both public providers (Commonwealth Government, State and Territory governments) and the private sector (clients/participants, their family/carers, non-government bodies such as health insurance funds or disease advocacy/patient support groups) are included; but not costs to sectors other than health (for example, education and housing). However some interventions will necessitate a broader societal perspective (e.g. interventions for conduct disorder or illicit drug use) whereby significant costs accrue to other sectors such as the criminal justice or the educational sector and to exclude such costs would severely bias results. Even if that is the case, the main results for interventions addressing these health problems will still be presented from a health sector perspective to make valid comparisons with other interventions in ACE-Prevention. The results from a societal perspective will be contrasted with the results from a health sector perspective as additional information to policy makers.

In addition, some decision makers, particularly the Commonwealth Government, are interested in the potential impact of interventions on production in the general economy due to early return to work or reduced disease incidence. This impact is referred to as "production losses/gains". The intention is to flag these impacts where they may be significant and to include production gains/losses in the sensitivity analysis if time permits. The possible inclusion of unrelated health care costs in additional years of life conferred by an intervention is a contentious issue amongst economists (Gold et al. 1996, Drummond et al. 1997) and such costs will not be included in the ACE-Prevention study.

With respect to the worked example in Appendix D the included and excluded costs are:

### **Costs included:**

Costs to the health sector, patients and families involved in the delivery of the intervention are included in the modelled cost-effectiveness evaluation. This includes the costs associated with:

- Central coordination of the intervention program by the State offices of Divisions of General Practice;

- Organisation and coordination of the GP recruitment and training by local Divisions of General Practice;
- Recruitment of GPs to the program;
- Training of participating GPs with respect to the intervention and its delivery, and associated time and travel costs;
- Conduct of training by a psychiatrist qualified in family therapy;
- Printing of training materials for GPs;
- Equipment required by each GP for the measurement of BMI;
- Costs of GP consultations associated with the delivery of the intervention including patient out-of-pocket gap payments;
- Printing of the family materials given to participants and their parents by the GPs;
- Travel costs of families in attending the GP consultations;
- Time costs of families in attending the GP consultations (including time spent in travel, waiting and the consultation). Results are reported both with and without these costs.
- Routine monitoring and evaluation by Divisions of General Practice.

**Costs excluded:**

The intervention is assumed to be operating in 'steady state'. It is assumed to be working in accordance with its efficacy potential as established by the trial, and that trained personnel are available to deliver the intervention and that the infrastructure is available. Given this, the following costs are excluded:

- Costs associated with the research, development and maintenance of the materials to be used in the intervention. This includes the materials to be used in the training of participating GPs as well as the purpose-designed materials to be employed by the GPs in delivering the intervention to families. Both of these were developed and piloted by the LEAP study team and it is assumed that they would be available for use if the intervention were implemented out on a larger scale. Our research question is the C/E of an established intervention with existing resource material.
- Costs at the Commonwealth government level. It is assumed that any activity is handled within existing capacity.
- Costs associated with training the trainer i.e. the psychiatrist who will deliver the training to participating GPs.
- Costs associated with the development and education of an adequate GP workforce. The intervention assumes that there is an adequate supply of qualified GPs who are capable of delivering the intervention, following training specifically related to the intervention.
- Costs associated with changes in physical activity or eating patterns of participating families as a result of the intervention (e.g. food costs, sports equipment, time costs). The evidence from the LEAP Study based on information from parents suggests no significant difference in shopping/food preparation costs (in expenditure and time) or in physical activity costs at the longer-term follow-up.

- Production gains and losses other than time cost of participation of parents.
- Time costs of children.
- Monitoring and evaluation above more than a routine level.

### Summary Box 9 Identification of Costs (Step One in Cost Analysis)

**All costs relevant to the intervention and its comparator should be clearly identified and inclusion/exclusion justified using the following guidelines:**

1. A consistent approach is taken for both the intervention and the comparator(s) to avoid bias;
2. 'C<sup>1</sup>' and 'C<sup>3</sup>' costs are included ;
3. 'C<sup>2</sup>' and 'C<sup>4</sup>' costs are excluded, but important impacts flagged and included in the sensitivity analysis;
4. 'C<sup>1</sup>' and 'C<sup>3</sup>' costs can be excluded on 'protocol grounds' (i.e. not relevant to the study perspective; not part of specified event pathway) or 'practicality grounds' (i.e. insufficient data; omission will not bias results);
5. All-of-life costs are excluded.

C<sup>1</sup> costs refer to government health sector such as medical, pharmaceutical, hospitalisation etc. costs.

C<sup>2</sup> costs refer to costs in other sectors such as welfare organisations, forensic services, educational services etc.

C<sup>3</sup> costs refer to any out of pocket expenses incurred by patients and their families such as travel, co-payments (e.g. for medical services or drugs) expenditure in the home and time.

C<sup>4</sup> costs refer to productivity costs. Source: Drummond et al 2005

### 3.3 Measurement of resource usage

Having determined which costs will be included or excluded, the next phase in cost assessment is to measure the extent of resource usage. Simply put, this means that the frequency of use of each cost component needs to be determined.

The "cost" of each factor of production (or service) is measured by multiplying the quantity of the factor consumed ("q") by its relevant price ("p"). In the measurement phase we assess the 'q', while in the valuation phase (section 3.4) we assess the 'p'. Sometimes intermediate products (such as radiology, pathology or pharmaceuticals) are consumed in the production of a more complex product (such as an episode of care in hospital). Sometimes consumers contribute to the production process with their own time or resources.

Consider, for example, when a patient with a high temperature visits a physician. Various activities may be undertaken to define the patient's problem and render treatment. During this process (i.e. event pathway') various resources are expended and might include:

- The physician's time;
- Activities of the physician's ancillary staff;
- Use of medical office space;
- Laboratory and/or radiology facilities, staff and supplies;
- Medication;

- Pharmacist's time;
- Patient's time and transport.

You can see from this simple example, that again the 'event pathway' plays a pivotal role in assessing the activities, and from there, we can estimate the resource utilisation by working out:

- What happens (the activities; who provides them; where);
- The probability of the various events happening (we use probabilities to work out what happens to the 'average' participant or the 'average' patient); and
- How many times does each activity happen?

There are various approaches to assessing resource usage that range from 'micro costing' (where individual patients/participants are followed through time, as in an RCT, time & motion study or data collection based on patient diaries); through 'activity costing' based on assigning unit costs to events in the intervention 'event pathway'; to 'macro costing' where costs are assigned to the whole intervention or major components thereof (e.g. hospital episode). One of the key tasks for researchers will be to work out which approach (or combination of approaches) is appropriate for their intervention and its comparator(s). It is expected that the core approach will be the 'activity costing' approach, working off the clear specification of the intervention/comparator 'event pathway'.

Under this approach the activities and associated providers for each segment of the intervention (i.e. recruitment element; prevention/care element; etc) are specified and measured as illustrated in the worked example discussed earlier (refer Appendix D for further details). It is anticipated that recourse will be made to 'macro costing' for some elements of the intervention (e.g. hospital episodes, a doctor's visit and cost offsets) and also sometimes to 'micro costing' where an RCT provides such detail. If unit record data across a larger number of intervention participants is available, mean/median cost estimates plus the associated confidence intervals can be calculated.

Linked to the 'micro costing' approach, is a further classification of costs by expenditure category. The expenditure categories can vary a little between studies, but usually cover five basic categories:

- Salary and wages (i.e. labour);
- Capital (i.e. land, buildings & equipment);
- Consumables;
- Overheads; and
- Other.

Notice that these expenditure categories reflect the factors of production, resource inputs and intermediate products used in the production process. This typology provides a bridge to the line codes and associated cost centres in the financial accounts. The expenditure categories help identify the cost-drivers and provide an important format for data collection. While researchers will use elements of this typology (particularly the staff and capital items), it is not anticipated that this level of fine-grain costing will generally be required. For example, if we consider the example above of a patient with a high temperature, the resource items of physician time, physician ancillary staff, medical office space etc are all encompassed by the MBS item category (e.g. item 23 for a standard GP visit), therefore detailed micro-costing is not required. The measurement of this activity is likely to be 1 or 2 item 23 GP visits. If however the person also requires subsequent blood tests or X-rays, these need to be identified and measured separately.

Micro-costing at the factors of production level will only be required in rare instances where activities are not already defined (such as the development of a new service from scratch which is not funded by the MBS). In such instances all researchers need to seek advice from



the three senior health economists on ACE-Prevention (Rob Carter, Chris Doran or Cathy Mihalopoulos)

The following table provides a general guideline on how common health activities should be measured in ACE-Prevention

**Table 1 Measurement of resource use**

<b>Activity</b>	<b>Unit of Measurement</b>
Community* Medical Consultations	Frequency of appropriate MBS item (e.g. item 23 for a surgery based GP consult lasting less than 20 minutes)
Community* Pathology Services	Frequency of appropriate MBS item (e.g. item 66542 for a oral glucose tolerance test for diabetes)
Community* Diagnostic Services	Frequency of appropriate MBS item (e.g. item 63000 for a MRI of the head)
Hospital Admissions	Episode of care using AR-DRG (NHCDC, R8) categories (e.g. AO57 is an admission for a heart transplant)
Community Allied Health – including psychology, counselling, dieticians, occupational therapy, physiotherapy, podiatry, Social Work, speech pathology	Frequency of Consultations (e.g. Cognitive behavioural therapy requires 12 consultations with a clinical psychologist)
Pharmaceutical	Average dose per day
Ambulance	Numbers of trips (differentiated by urban or rural if possible)
Nursing Home	Length of stay (in terms of days) differentiated by the Resident Classification scale – further details available from CDHA)
Home nursing (e.g. services provided under the Home and Community care program)	Time requirement of services
Time	Hours (or part thereof) involved
Travel	Numbers of kms per trip as well as numbers of trips
Other professional services	Hours of professional time involved (e.g. teacher time required for an educational intervention)

\* The term community here refers to 'out-of-hospital' services

## Summary Box 10, Measurement of Costs (Step Two in Cost Analysis)

**Resource utilisation associated with the intervention and its comparator are measured using the following guidelines:**

1. 'Activity costing' based on the 'event pathway' is the primary method, complemented by occasional recourse to 'macro costing' and 'micro costing';
2. Resource utilisation will be measured from the 'event pathway' by working out:
  - what happens (the activities; who provides them; where);
  - the probability of the various events happening (we use probabilities to work out what happens to the 'average' participant or the 'average' patient); and
  - how many times does each activity happen?

### 3.4.1 Valuation of costs

Individual unit cost data for all resources associated with an intervention should be obtained from the "ACE-Prevention Table of Unit Costs<sup>10</sup>", which will be based on the most up-to-date and accurate sources. The unit cost table is largely based on the Manual of Resource Items and Their Associated Costs (CDHA, 2002). Researchers should also refer to table 3.3 (at the end of this section) for guidance on where to find unit costs which are not contained in the unit cost table. Note that there may be some resources which are not specified in the manual or table and in this instance researchers must consult the senior health economists on the project for guidance on the appropriate valuation.

Costs and cost offsets for the health sector are measured in real prices for the reference year (2003). The AIHW health sector deflators are used to adjust prices to the reference year (available in the table). Where interventions fall totally outside of the health sector, adjustment are made using the relevant Consumer Price Index (Australian Bureau of Statistics, Consumer Price Index, Australia, Cat. No. 6401.0) (available in the table). Given that the task is to as accurately as possible reflect costs in the reference year, where unit cost data are available for 2003, they should be used rather than deflating more recent prices back to that year. For example, any interventions entailing medical services or pharmaceuticals, unit cost data should be accessed directly from the 2003 versions of the Medicare Benefits Schedule (MBS) and Pharmaceutical Benefits Schedule (PBS) which are both available on the FTP website.

### 3.4.2 Attribution of joint or common costs

One issue of concern to the accurate measurement of costs is the attribution of costs to an intervention in situations where resources are jointly used by one or more programs. For example, in interventions in the school setting in ACE-Obesity, the costs of teachers, materials and equipment could be shared by several programs. Any criteria used to distribute the 'common costs' need to be clearly tabulated so that users of the study results can satisfy themselves that they are reasonable. There are a variety of approaches here, that rest on objectives of either: i) cost recovery; ii) equity or fairness; and iii) efficiency. Often there are commonsense 'rules of thumb', such as 'allocating according to floor space' or allocating 'according to patient usage' that are routinely employed in economic evaluations.

---

<sup>10</sup> This is an EXCEL based work-book which is maintained by the senior health economists at Deakin University and can be accessed on the ACE-Prevention ftp site in the folder called unit costs. This folder also contains some of the government documents with costing information on MBS, PBS and DRG hospital costs.

### **3.4.3 Staff costs and associated expenses**

In costing staff employed by an intervention, a factor of 1.6 will be applied to the base salary to cover: salary on-costs (Workcover, annual leave, superannuation, long service leave etc.); a loading to cover administrative assistance, office space and utility services (e.g. heating, electricity), workstation and other equipment, consumables (such as stationery, travel); and routine monitoring, support and evaluation. This loading will only be applied to new staff employed directly as a consequence of an intervention (such as staff employed at the state or national coordination level in the coordination). It will not be applied to staff who are already employed by an organisation and have access to office space, equipment etc as part of that position. In this instance, the loading should be reduced to 30% to cover the salary oncost component only. A range of uncertainty around both of these loadings will be included as part of the uncertainty testing.

The fundamental principle which should be employed in the pricing of staff time allocation to an intervention is that of 'opportunity cost'. Two questions need to be answered. Firstly, for the specific activity to take place, what additional staffing resources are involved, and secondly, what dollar value do you place on their time. This issue arose, for example, in ACE-Obesity with respect to teachers in the TravelSMART intervention. When teachers attend training, their positions are required to be back-filled by casual staff, which entails a cost. However, on the other hand, when they conduct the TravelSMART lessons, they would have been teaching the particular class anyway, so no additional cost was incurred.

Where a particular item which normally would be accounted for in the staff salary loadings discussed is fundamental to an intervention and constitutes a significant expense, the item should be separately costed. For example, in ACE-Obesity, the school-based focused nutrition intervention, the field presenters were constantly 'on the road' meaning that their travel costs were a significant component of total costs, which warranted their separate itemisation. Likewise, any items that are unique to a particular intervention, such as behaviour change manuals, should be costed as separate items.

### **3.4.4 Time costs for carers/family**

The term 'cost' in economic analysis is broader than just the financial cost of each intervention. Economic analysis focuses on the real resources used in the provision and consumption of a health service. Resources include not only the basic factors of production for health providers (staff, capital, land, enterprise and intermediate products), but also the resources provided by consumers. In this context, the time of parents/carers is an important resource and has a clear opportunity cost (that is, time can be used for alternative purposes). While there is clear guidance (Gold et al. 1996, Drummond et al. 1997) that time costs associated with providing an intervention should be included in economic appraisals, there are few studies that do so in practice.

Time costs can be divided into two broad categories:

- time costs that are an integral part of providing the health service itself (such as travelling time, waiting time, treatment time); and
- time costs that are a consequence of providing the intervention (such as time in taking children to physical activities, time in preparing healthier meals etc).

Depending on the purpose and perspective of the evaluation, quite different valuations of time will be used. Production losses or gains (usually defined as time away from paid work or alternatively early return to work) in terms of losses and gains to economic production indicators such as Gross Domestic Product, can only be justified if there is good evidence that the disease and its treatment lead to real production losses or gains.

It is assumed that it is the valuation of time per se which is important in the current study. Therefore both time costs that are integral to the intervention shall be counted as well as

time costs that are a consequence of the intervention. However production effects will not be included in ACE-Prevention. There is no set method of valuing time (Jacobs and Fassbender 1998). The simplest method uses the hourly wage rate as a proxy for the value of time. A recent review of indirect costs states that a common convention of valuing leisure time is 25% of the wage rate (Jacobs and Fassbender 1998) and that the value of working time should be age/sex adjusted.

The method used for valuing time costs in the ACE-Prevention study is based on the above principles. National information on wage rates and employment rates are available from the Australian Bureau of Statistics in the following publications:

- Average Weekly Earnings (Report Nos: 6302.0, published Aug 2003)
- Persons not in Labour Force (Report Nos: 6220.0, published 2005)

Table 1 contains the parameters used to determine the average hourly wage rate for males and females

**Table 3.2 Hourly wage rates and leisure time rates used in the determination of time costs.**

	Weekly ordinary time earnings	Per hour <sup>1</sup>	Leisure time rate <sup>2</sup>
Males	\$ 961.30	\$ 25.03	\$ 6.26
Females	\$ 786.00	\$ 20.47	\$ 5.12
Persons	\$ 904.30	\$ 23.55	\$ 5.89

Sources: (ABS, 2003;)

Notes:

<sup>1</sup> Calculated by dividing the weekly ordinary time earnings by ordinary time (38.4 hours)

<sup>2</sup> Calculated by multiplying the hourly rate by 25%

A weighted average hourly rate is then determined by considering the proportions of the populations who are employed, unemployed or 'not in the work-force'. For the general population these rates come from the ABS publication "Persons not in the Labour Force" (ABS 2005). The ACE-Prevention common costing excel file presents these calculations and derives an average hourly time cost of \$17.44 which we then use for all time costs in adults regardless of age and sex. Note that we have already decided not to calculate time costs for children (section 3.2).

Once the value of hourly time rates are derived, the calculation of the time costs of each intervention can then be determined. The assumptions required to perform these calculations for each intervention may differ. These will ultimately depend on lengths of contacts (consultations), travel time and waiting time. For example contacts with GPs costed at item 23 are quite short but may involve significant waiting times.

## Summary Box 11. Time costs

**In summary, the following steps are undertaken with regards to time costs:**

1. Determine a weighted average hourly cost of time utilising the base costs in table 1 adjusted for the proportion of the population of interest which are employed/unemployed.
2. Determine the number of health service contacts required for each intervention (as per the numbers used in the estimation of the costs of each intervention), the length of time of each contact, any travel and waiting time associated with each contact. Data sources for this information include:
  - Any published empirical information
  - In the absence of any published information, consultations with key experts regarding realistic estimates.
3. Calculate cost-effectiveness ratios with and without time costs to determine the effect of time cost inclusion.

### 3.4.5 Cost of non-adherence

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 seeking behaviour and associated costs of non-adherence. 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 seeking behaviour (and associated costs) as those currently not receiving the evidence-based intervention.

### 3.4.6 Cost offsets

If an intervention prevents future disease or treats current disease so that future complications are avoided, the projected health care costs are estimated in the intervention and comparator scenarios. The difference in cost offsets between the intervention and comparator may arise from a reduction in incidence, duration and/or severity of disease or in some cases an improved remission (or cure rate).

The rule of thumb in ACE-Prevention is to:

- calculate the disease cost per prevalent case in 2001 from the 2001 DCIS data in the numerator (ignoring the small component for research) and 2001 prevalent cases from the AusBoD results spreadsheet in the denominator; *for each year* lived by any individual with the disease of interest in our models, the cost per prevalent case is awarded; note that for some diseases you will need to do something different; i.e. the prevalence figures in AusBoD for IHD are those with symptomatic disease of any of the sequelae (heart failure, angina or myocardial infarction); in ACE-Prevention we will usually define prevalence of IHD as anyone who has had a first event regardless of whether symptomatic or not; therefore for IHD we make use of a separate DisMod model with prevalence as an output defined as for our purposes (and this file is available on the common ACE-Prevention drives and ftp site);
- when AIHW expenditure estimates for the year 2003 becomes available we will update the information in our models;
- we will incorporate information on trends in expenditure by disease with help from AIHW;

- in some instances where prevalence is hard to define (e.g. the prevalence of a wrist fracture) and costs are (almost) exclusively clustered around an incident event we will apply a cost per incident case and apply this as a *one-off* cost in our models.

### 3.4.7 Government versus patient costs

All costs need to be divided into costs accruing to the government and costs accruing to patients. Generally the following principles should be used as guidance for attribution of costs to government and patients

- For medical, pathology and diagnostic services the cost to government is 85% of the scheduled fee. The simplest way to estimate the cost to patients is to assume the remaining 15% of the scheduled fee is charged to patients. This may be an overestimate for some consultations and procedures due to bulk-billing. Conversely, it may also be an under-estimate for other consultations and procedures due to above-fee scheduling (particularly for specialist medical services and some imaging procedures). While detailed data on patient contributions is available from the HIC at an individual item level there is an associated cost to processing such information (both financially as well as time costs!!). It is therefore recommended that for ACE-prevention the cost to patients is calculated as 15% of the scheduled fee with this proportion varied in the uncertainty analysis (as a triangular distribution of 10%, 15%, 20%).
- For allied health services (except for some psychology services and more recently exercise physiologists) assume all costs are borne by patients; there are limited psychology services funded by the MBS particularly for the treatment of depression and anxiety using cognitive behavioural techniques. In this instance some psychology consults will accrue to the government (as specified in the MBS) and the rest will be patient costs (see Cathy Mihalopoulos for further details)
- Hospital costs will be modelled as public admissions meaning that the cost accrues to the government (since there is usually insufficient data to divide disease level hospitalisation costs into public and private hospitalisations). However if a researcher has sufficient data to divide hospital costs into public and private admissions this may be desirable to do so however needs to be assessed on a case by case basis with sensitivity testing around the division.
- Travel costs may be patient or government costs. This needs to be decided case by case (e.g. some government-funded interventions may fund travel of health professionals out to remote areas whereas travel costs of patients are financed by the patient)
- Time costs should be reported separately to other government and patient costs as these tend to reflect the opportunity cost of time rather than a monetary cost.

### 3.4.8 Discounting

Discounting is applied to both costs and benefits. This reflects the fact that, individually and as a society, we prefer to have dollars or resources now as opposed to later, because we can benefit from them in the interim<sup>11</sup>. Similarly, we prefer to have benefits now rather than later. A 3 per cent discount rate is applied to match the rate chosen in the Australian burden of disease studies. It is also the rate of discounting recommended by a consensus panel of health economists in the US (Gold et al. 1996). This rate also approximates the long term bond rate, the rule of thumb often used in selecting the appropriate social discount rate.

Sensitivity analysis may be used to test discounting scenarios other than the 3 per cent rate. The interventions for example can be run with 0, 5 and 7 per cent discount rates in place, to ascertain the impact on the results of changes in the discount rate.

---

<sup>11</sup> For further discussion of discounting, refer to Drummond et al. (1997) and Gold et al. (1996).

**Table 3.3 General Sources of Valuation for Resources**

<b>Activity</b>	<b>Source of Valuation</b>
Community Medical Consultations	MBS (on FTP site)
Community Pathology Services	MBS (on FTP site)
Community Diagnostic Services	MBS (on FTP site)
Hospital Admissions	AR-DRG public hospital cost weight from NHCDC (on FTP) for a total episode of care, (only variation is if a intervention reduces length of stay where a bed day cost should be used – refer to a senior health economist)
Community Allied Health – including psychology, counselling, dieticians, occupational therapy, physiotherapy, podiatry, Social Work, speech pathology	Manual of resource items and their associated costs (on FTP site)
Pharmaceutical	PBS, 2003 (on FTP site)
Ambulance	Manual of resource items and their associated costs (on FTP site called manual.pdf)
Nursing Home	Manual of resource items and their associated costs (on FTP site)
Home nursing (e.g. services provided under the Home and Community care program)	Manual of resource items and their associated costs (on FTP site)
Time	Wage rates as specified above
Travel	To be finalised
Other professional services	Varied – need to consult a senior economist

### 3.5 Reporting costs

Intervention costs will be reported in gross and net form, that is, with and without the estimated cost offsets. A unit price for each of the activities, together with the data source, will be specified in briefing papers for each intervention. Detailed information on the composition of costs by expenditure type (such as capital, staff, consumables and overheads) will generally not be provided. This reflects the focus of the study on 'allocative efficiency' rather than 'technical efficiency'.

As noted costs to government and patients should be prepared separately as well as in aggregate. Intervention results should be reported both with and without the time costs.

There is a practical financial issue for governments that warrants separate reporting of the potential cost offsets. The potential cost offsets are opportunity cost estimates - that is, they are estimates of resources devoted to the treatment of preventable diseases (or early intervention programs) that could be available for other purposes. Conversion of opportunity cost savings into actual financial savings involves a number of practical and theoretical considerations and cannot be taken for granted. For instance, avoiding a heart attack will avoid the costs associated with hospital treatment of that condition in that

individual. However, the hospital is likely to have incentives to fill that intensive care bed with another patient and hence there may not be any financial savings.



## 4 Measurement of health benefits

Set out below are more detailed notes that specify how the three steps of identification, measurement and valuation are applied to the estimation of health benefits in this project.

### 4.1 Identification of Benefits

A guiding rule in identifying dimensions of benefit is that they are mutually exclusive so that the elements do not overlap and are not counted more than once. The principle dimension of health benefit for this study is the estimated "size of the health gain" associated with each option. The size of health gain is a quantitative measure, the calculation of which is evidence-based using a combination of the scientific literature, analysis of available routine and survey databases and expert opinion.

We choose to measure health gain chosen in 'health-adjusted life years' where the loss of health due to non-fatal health states is valued 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, it is important to realise that there are philosophical differences between the two. First, in a burden of disease study we estimate the health status of a population in a particular year. It is therefore, a cross-sectional measure even if the non-fatal component is measured as the loss of health estimated to arise from incident events. Economic evaluation methods always have a time dimension: 'what happens over time if a target population is exposed to an intervention of interest or a comparator?' 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 'horizon').

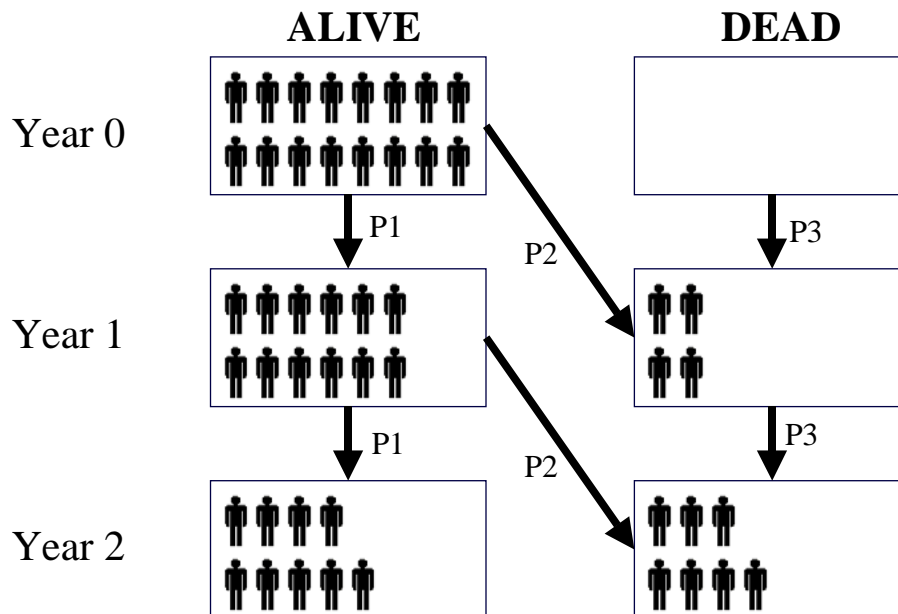
Second, in burden of disease the DALY is constructed as a 'health gap' measure, i.e. we set an ideal ("everyone ought to live into old age free of disease") and contrast the current health status of a population with that ideal. Thus, Years of Life Lost (YLL), the mortality component of DALYs, are calculated as the difference between age at death and a 'standard' life expectancy at that age for each death. It is best to view these conversions of counts of deaths into YLL as a weighting of deaths by age. Young deaths accrue more YLL than old deaths. In economic analyses, we do not use the standard life table to give a value to loss of life. Instead, we keep track of a target population over time and count the years of life lived in intervention and comparator scenarios assuming 'realistic' mortality risks as people age. If we assume no trends in mortality this would equate to giving a death the value of the equivalent life expectancy for the age at death from the population's 'period life table'. If we apply mortality trends in our models, it is equivalent to awarding remaining life expectancy from a 'cohort life table' to each death.

### 4.2 Measurement of Benefits

#### 4.2.1 Modelling benefits

In ACE-Prevention we use mathematical models to predict the costs and benefits that are relevant to an intervention by combining information, often from disparate sources, on disease epidemiology, effectiveness and costs. In predicting population-level costs and consequences of health interventions, there are a variety of modelling techniques available. Most analyses in this project rely on the principles of Markov models (Sonnenberg and Beck 1993). Markov models are distinguished by three concepts. Firstly, conditions in a model are comprised of a number of mutually exclusive health states. For example, in Figure 4.1

we have a Markov model with 2 health states: Alive and Dead. Models may have any number of health states, e.g. Alive, Diseased, Dead.



**Figure 4.1 Graphical description of a Markov process.**

Secondly, Markov models assume that time is broken into discrete intervals or cycles. In this example, we have a model with a cycle length of 1 year.

Finally, events are modelled as transitions between states, with these transitions occurring only once at the end of each cycle. These transitions are governed by a set of transition probabilities. In other words, transition probabilities denote the probability of being in a given state in the next cycle conditional on membership in a particular state in the current cycle. In the current example, we have three transitions: Stay Alive, Die, and Remain Dead with P1, P2, and P3 denoting the transition probability of each of these three states. As the cycle length is 1 year, our transition probabilities denote an annual probability.

Note that the sum of the transition probabilities for a given state must equal 1. For example, the sum of the two transition probabilities - P1 (Probability of Staying Alive) and P2 (Probability of Dying) for the Alive State - must be equivalent to 1. P3 is necessarily equal to 1 as once dead you will continue to remain so.

Finally, in Markov models rewards or utilities are assigned according to the time spent in each health state. The value of each incremental reward reflects the cycle length of the model. For example, as the current example is a 1-year cycle length model, a yearly reward of 1 year of life in full health would be assigned to each individual in the Alive state with a utility of 0 for each individual in the Dead state.

### **Deterministic Markov models**

The health experience (e.g. life years lived) of a population in a Markov model is usually calculated using cohort simulation. This type of Markov model is sometimes called a Deterministic Markov model. In cohort simulation, a cohort is assumed to begin in a particular state or states at time 0, and portions of the cohort will make transitions to other states in the process for the next cycle according to the transition probabilities between the states. This will determine, for each time period or cycle, the portion of the cohort that inhabits each state. The average life years lived or remaining life expectancy can then be calculated as the sum over all time of the portion of the cohort in each state multiplied by the 'reward' (or 'utility') associated with being in each state. As Markov models are

iterative, a termination rule defines when the analysis stops, e.g. when a certain number of years or cycles have elapsed.

## **Stochastic Markov models**

The limitation of deterministic Markov models is that transition probabilities depend only on the current state of membership. The path by which portions of the cohort arrive in different states is not available for calculating transition probabilities. This is termed the “no-memory” property of Markov models. A solution to this is to use Stochastic Markov models that track individuals of the cohort rather than as a whole group. By allowing one member of the cohort into the model at a time, one can keep track of the complete history of state membership as the individuals move through the model. Using Monte Carlo simulation<sup>12</sup> this is repeated many times.

The advantage of this is that counts of specific events or time (e.g. number of times hospitalised, length of time in a particular state) are available and subsequent transitions may depend on these values. This is particularly useful for the evaluation of smoking cessation interventions where the reduction in risk following cessation of smoking is dependent on the time since quitting. This approach also allows the generation of a probability distribution that represents between-patient variance. The major disadvantage of stochastic models is they are computational intensive and often require long evaluation times. When the no-memory property of deterministic Markov models is not an issue, stochastic Markov models will approximate the same answer. Therefore, deterministic Markov models should be used when the no-memory property is not an issue.

## **Systematic review of natural history of disease and risk factor mapping**

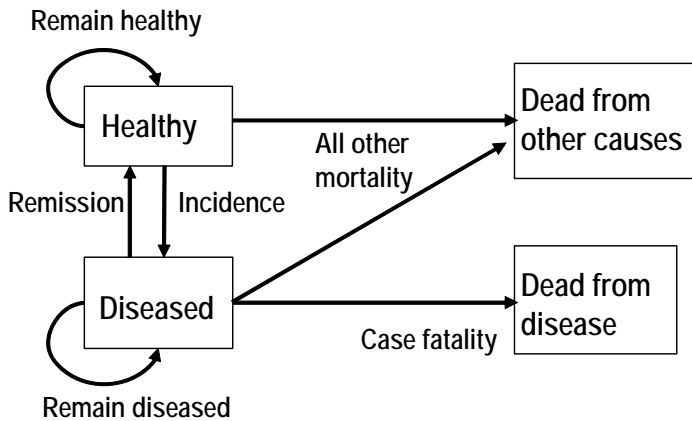
In order to define a model structure and determine transition probabilities that appropriately model the effectiveness of interventions, an important first step is to systematically review the available epidemiological data, and to describe the natural history of the disease(s) and/or risk factors under study.

It is helpful to start with a graphic representation of diseases and risk factors that are modified by an intervention with boxes for each health state and arrows for transition probabilities (Figure 4.2). Typically, for each health state there are one or more of the following transition probabilities:

1. move to another health state, e.g. from healthy to diseased (‘incidence’);
2. remain in health state;
3. get better or remit to a non-diseased health state (‘remission’);
4. die from disease (‘case-fatality’ which we will define as the risk of death due to being in a particular health state in excess of the average risk of dying in the population);  
and
5. die from all other causes (‘background mortality rate’).

---

<sup>12</sup> In order to distinguish this type of Monte-Carlo simulation from that used to evaluate uncertainty, Monte Carlo simulation used in stochastic models is sometimes termed 1<sup>st</sup> order Monte Carlo simulation while that used in evaluating uncertainty is termed 2<sup>nd</sup> order Monte Carlo simulation.



**Figure 4.2 Disease diagram with health states and transition probabilities**

## 4.2.2 Effectiveness and Safety of Interventions

### Efficacy versus effectiveness

There are two ways of measuring the impact of interventions. For most interventions the evidence of the health impact has been examined under carefully set up trial conditions. This is called efficacy and determines the validity of an intervention, i.e. to establish causality and confidence that the interventions can achieve health gain when implemented under controlled conditions. As we try to inform policy making about health interventions being implemented under routine health service conditions, in ACE-Prevention we are more interested in the effectiveness credentials (i.e. impact in real life application) of the options for change. Effectiveness reflects the generalisability or the potential health gain in practical application.

Impact is most commonly measured as efficacy. In ACE-Prevention we will give more more emphasis to evidence on effectiveness, i.e. measures of impact that have been realised under routine health service conditions. If only efficacy data are available, judgment is made on the applicability of such measures under routine health service conditions within the context of Australian health services. For instance, adherence to a treatment may be substantially lower under usual health service conditions compared to adherence under trial conditions. A lesser impact may also be expected because of a lesser quality of implementation of the intervention under routine service conditions. Where possible, such judgments are based on evidence (for instance comparing pooled data on adherence from trials with adherence rates from observational studies of Australian health services), but not infrequently it is necessary to rely on expert opinion in the absence of evidence. The members of the Technical Advisory Groups and in some cases of the Project Steering Committee have an important role in advising the researchers on these issues and in suggesting additional expertise on which to draw.

### Safety

When determining the beneficial effects of interventions it is equally important to determine whether there are possible adverse effects, as these may outweigh the positive benefits of an intervention. Ideally, all possible positive and negative effects should be tabulated a priori, regardless of data availability (Glasziou and Sanders 2002).

Clinical trials are not the ideal methodology for measuring adverse effects. Trials are statistically powered to determine beneficial effects on an outcome of interest, and may not show increased risk of less common adverse events. The limited follow-up period of trials also prevents them from properly assessing adverse effects that have long lag periods between exposure and outcome, e.g. carcinogenic effects. Given these limitations it may be more appropriate to derive estimates of adverse events from observational studies such as case-control studies, post-trial marketing surveillance data or analyses on linked routine databases (that are common in Scandinavian countries; for instance the risk of suicide while taking anti-depressants has been evaluated in Sweden linking prescription data to death data base).

## **Systematic review and meta-analysis**

Given the critical nature in cost-effectiveness analyses of the size of both positive and negative impacts of an intervention, considerable effort should go towards identifying the best available evidence on efficacy/effectiveness and safety. The identification of information to estimate the efficacy/effectiveness and safety of interventions is preferably done by systematic review. Systematic review reduces bias and provides a standardised way of collecting information on the impact of different interventions (Egger, Davey Smith et al. 2003). This enhances the comparability of different economic evaluations.

Estimates of the impact of an intervention are often spread across a number of trials or studies. Meta-analysis is the preferred method of combining information from multiple trials to derive an aggregate measure. This increases the precision of the estimate, thereby reducing the uncertainty associated with the impact of the intervention (Egger, Davey Smith et al. 2003). Empirical research has shown that meta-analyses are comparable to or better than small trials at predicting the results of subsequent large trials (Saint, Veenstra et al. 1999).

Statistical techniques related to meta-analysis (e.g. meta-regression, trial stratification, sub-group analysis) allow sources of heterogeneity (arising from differences between trials, e.g. in follow-up periods or intensity of the intervention) to be formally examined (Glasziou and Sanders 2002). Pooled estimates from meta-analysis estimates may not be appropriate if significant heterogeneity in the intervention effect remains even after trial stratification or subgroup analyses. In these situations, the final choice of the source on efficacy/effectiveness or safety may be guided by the following considerations:

- One or more sources are of the greatest size or are of superior quality.
- One or more sources best represents the population to which the results will be generalised. When the results of a cost-effectiveness analysis are generalised to a particular population, it is relevant to use information that best represents the population in the analysis.

## **Using published systematic reviews**

As comprehensive systematic reviews require considerable resources to undertake, the first strategy is to look for published systematic reviews on intervention impacts. However, it is often necessary to update search strategies to cover recent studies not included in the published review, as it is possible for the results of even a single recent study to significantly affect the results of a meta-analysis. Published reviews may also not cover all the outcome measures necessary for a cost-effectiveness analysis.

## **Choice of outcome measures**

While Gold and others (1996) recommend the use of reductions in all cause mortality as estimates of effectiveness, risk reductions in all-cause mortality in one population cannot be extrapolated directly to another population if the relative contribution of cause-specific mortality to total mortality is different between these populations. For example, in a population with a high cardiovascular or HIV/AIDS mortality rate dominating all-cause

mortality, estimates of reductions in all-cause mortality in this population cannot be extrapolated to another population with a low cardiovascular or HIV/AIDS mortality rate. This is particularly important in this study as the bulk of the evidence for interventions come from populations outside of Australia.

To illustrate the shortcomings of applying RRs on total mortality to a different context consider a population A with total mortality rate of 20 per thousand, with 10 per thousand for HIV/AIDS mortality and 1 per thousand for diabetes mortality. A RR reduction in total mortality of 0.95 for a diabetes intervention would lead to a mortality reduction of 1 per thousand. In population B with same level of diabetes mortality, no HIV/AIDS mortality and a similar level of mortality from all other causes (i.e. total mortality is 10 per thousand), the same RR of 0.95 would lead to a 0.5 per thousand reduction in mortality.

For this reason in ACE-Prevention the preferred outcome measure for the effectiveness of an intervention is a relative risk reduction in cause-specific mortality and/or event rate. For many interventions that impact on more than one disease outcome, a range of outcome measures may need to be identified. For example, aspirin has beneficial effects on CHD and ischaemic stroke but negative impacts on hemorrhagic stroke and gastrointestinal bleeding.

### **Calculating pooled estimates**

Pooled estimates using meta-analysis are always presented using random effects models, as this method includes both "between trial" error and "within trial" error. Fixed effects models limit themselves to an analysis of "within trial" error and may underestimate overall uncertainty associated with the treatment effects (Deeks, Altman et al. 2003). If all the pooled trials have similar results random and fixed effects models give the same outcome. In other cases, a random effects model will give a wider confidence interval around the mean estimate. Note that the use of a random effects model does not solve the problem of significant heterogeneity between trials!

The choice of summary statistic for meta-analysis of binary outcome data is also important, as the statistic used should be the one that gives the most consistent results across trials. Empirical research on a range of meta-analyses demonstrates that relative risk measures such as the Risk Ratio (RR) or Odds Ratio (OR) are more stable than the Risk Difference (RD) (Deeks 2002). Care must be taken when using the OR to estimate relative risk of an intervention. If the outcome of interest is a rare event the OR and RR will be approximately similar and can be used interchangeably. If the outcome of interest is more common, using the OR as a RR will overestimate the intervention impact. The example below illustrates this:

	Common disease outcome			Rare disease outcome			
	+	-		+	-		
Intervention A	10	90	100	Intervention A	1	99	100
Intervention B	20	80	100	Intervention B	2	98	100
Risk outcome intervention A = 0.1			Risk outcome intervention A = 0.01				
Risk outcome intervention B = 0.2			Risk outcome intervention B = 0.02				
RR (A vs B) = 0.5			RR (A vs B) = 0.5				
Odds of outcome with A = 1/99			Odds of outcome with A = 1/99				
Odds of outcome with B = 2/98			Odds of outcome with B = 2/98				
OR = $(80 * 10) / (90 * 22) = 0.444$			OR = $(98 * 1) / (99 * 2) = 0.495$				

Jan Barendregt has developed a small add-in program to Excel that allows you to transform an OR into a RR provided you have information on the incidence of the health problem in the population.

### Exploring heterogeneity in the treatment effect

A Galbraith plot or the chi-square statistic can give a preliminary indication of between-study heterogeneity. Outliers on a Galbraith plot are the studies contributing most to the heterogeneity and can point in the direction of study characteristics contributing to heterogeneity that are worth exploring (Thompson 2003). The absence of heterogeneity as indicated by the Q statistic or a Galbraith plot, however, does not necessarily preclude further investigation by techniques such as meta-regression or sub-group analysis (Thompson and Higgins 2002). Meta-regression is a method where characteristics of individual trials the results of which are being pooled in a meta-analysis are correlated with the study outcome. For instance, a meta-regression of Cognitive Behavioural Therapy interventions for anxiety and depression identified, among others, the type of therapist, the language used and the number of therapy sessions as potential explanations for differences in outcomes. These factors were identified a priori and then regressed against the intervention outcomes (Haby 2006). Meta-regression covariates, and subgroup analyses should be stated a priori and should not be overdone (Davey Smith and Egger 2003). The chances of finding a significant result increases with the number of covariates or subgroup analyses performed. These techniques should be used and interpreted with care as the comparisons that are made are not randomized and are prone to confounding.

Glasziou and Sanders (2002) distinguish between artefactual causes and real causes of heterogeneity in the treatment effect (Table 4.1). Artefactual causes are related to the design and conduct of trials. For example, trial stratification by length of follow-up (an artefactual cause) in a meta-analysis of interventions to reduce dietary fat intake demonstrates significant differences only when analysis is restricted to long-term trials (Hooper, Summerbell et al. 2001).

Real causes of variation in the treatment effect, particularly at the patient-level are important to investigate, as the treatment effect may differ in particular sub-groups of the population. Subgroup analyses should be conducted according to, for example, age, sex, presence of previous disease or disease severity, or concomitant medication. If the subgroup analyses indicate similar results there is a strong case to use a single impact measure, as this will increase the precision of the estimate, i.e. reduce uncertainty. For example, the relative reduction in coronary heart disease events and mortality due to

statins shows very little differentiation by age, sex or presence of previous disease (LaRosa, He et al. 1999).

**Table 4.1 Real and artefactual causes of variation in treatment effect**

	Real	Artefactual
Patient	Disease severity	Improper randomization
	Age	Differential follow-up
	Co-morbidity	
Intervention	Time	Non-compliance
	Duration	Cross-over
	Dose	
Co-intervention	Drugs	Undetected co-interventions
	Therapy	
Outcome	Timing of outcome	Differential and non-differential measurement error, e.g. lack of blinded outcome assessment or event verification
	Event type	

Source: (Glasziou and Sanders 2002)

### Extrapolating surrogate outcomes

Not all studies measure changes in event or mortality rates, and those that have, may not have an adequate follow-up period to ascertain significant differences. In cases such as this, treatment effects may need to be measured as a change in surrogate endpoints, e.g. cholesterol level or energy expenditure, and surrogate measures can then be extrapolated using information from other studies that determine the proportional changes in mortality or event rates by units of the surrogate measure.

In ACE-Prevention with a large focus on preventive interventions that primarily address risk factors rather than disease outcomes, we commonly model outcomes via a surrogate measure. We then make use of meta-analysis data on the relationship between the surrogate measure, exposure to a risk factor, and disease outcomes. Much of this information comes from work done for the Comparative Risk Assessment component of the Global Burden of Disease study (ref to CRA). For continuous risk factors we apply a RR of disease occurrence or mortality to a change in the mean level of exposure to estimate a proportional reduction in disease. It is more complicated to apply intervention effects expressed as a change in a continuous measure if the data on the relationship between risk factor and disease is expressed as a RR for discrete exposure categories. For instance, this is the case for physical activity (evidence for change in energy expenditure as a continuous variable while prevalence data and RRs refer to discrete exposure categories of sedentary, low and adequate physical activity) and alcohol (evidence for change in mean grams of alcohol per week, but prevalence and RRs by no, low, hazardous and harmful drinking levels). Simply applying the mean change in energy expenditure and grams of alcohol consumed to individual record survey data and recalculating the prevalence figures by



categories of exposure makes the overall impact of the intervention rely on a small number of survey respondents who will 'tip over' a threshold into a lower exposure category. This is likely to lead to inaccurate estimation of the true effect. Instead, in ACE-Prevention we use an alternative method where the impact on a continuous measure is translated into a change in the RR for an exposure category. To illustrate this method, the alcohol model uses the following steps for each age group and separately for males and females:

1. in National Health Survey data determine the mean grams of alcohol consumed per week in the four exposure categories (abstinent, low, hazardous and harmful);
2. plot the RRs of disease outcomes for the four exposure categories against the mean grams of alcohol/week and linearly interpolate a change in RR per grams of alcohol/week;
3. reduce the RR in each exposure category by the product of the mean change in grams of alcohol/week from trial data and the per unit change in RR calculated in point 2; and, lastly,
4. in the intervention scenario apply the lower RR and in the comparator scenario the original RR in each exposure category.

### **Extrapolating treatment effects over time**

In modelling health outcomes the researcher has to confront the issue that trials measure outcomes over a limited time period while the interest is in the true impact on disease outcomes and costs that would arise under routine intervention implementation circumstances. One option is to limit the modelling to the duration of the trial but this may not adequately reflect reality. The alternative is to make assumptions about the impact beyond the duration of the available trials: a continued impact over time, a lessening of the impact over a period beyond the known impact time from trials or the abrupt disappearance of the impact. It depends on the intervention in question what the most plausible way of modelling is and it is typically something to discuss with your technical experts. Often, however, there is no clear choice and the solution may be to present results as discrete scenarios using different choices as a sensitivity analysis. For instance, we assume an annual decay of the impact of GP mediated physical activity interventions of 50% and vary this between 25% and 75% in a sensitivity analysis as three distinct scenarios.

### **Summary Box 13 Estimating effectiveness and safety**

#### **Summary of key steps for estimating effectiveness and safety of interventions:**

1. Tabulate all potential positive and negative effects of the intervention
2. Identify published systematic reviews of the intervention.
3. If an appropriate published review is identified, update the review and/or incorporate additional outcomes where necessary. If no appropriate review is identified conduct a full systematic review.
4. Use meta-analysis to determine a summary effect of the intervention if appropriate.
5. Identify potential sources (real and artefactual) of heterogeneity in the intervention effect.
6. If the effects of the intervention are measured according to a surrogate outcome, extrapolate surrogate outcomes using available data (e.g. using RRs of relationship between risk factor and disease outcomes from observational studies).
7. If appropriate, determine adverse effects of an intervention from observational data.
8. Identify potential effect modification of the intervention over time.
9. Identify potential effect modification of the intervention under routine health service conditions in Australia.

### 4.2.3 Current coverage of interventions

In order to determine the relative (in)efficiency of currently implemented interventions, it is necessary to determine the current levels of implementation of the chosen interventions. As the absolute effect of many interventions is dependent on baseline risk of disease (e.g. risk factor levels), estimates of current coverage should be as specific as possible. Coverage estimates should therefore ideally come from unit record data of representative surveys that have information on intervention use by level of risk factor exposure and/or prevalence/incidence of disease. If that information is not available, look for routine health service data collection systems (e.g. PBS data, hospital data or the GP collection data from BEACH) that would give you an indication of the numbers of consultations with a particular intervention mentioned (e.g. nutritional counselling by GP for people at high risk of CVD), procedures or drugs prescribed which you can then divide by the relevant 'target population' to estimate coverage. This may not always be so easy as the data is often not presented in the right way and a judgment call may be necessary by a Technical Advisory Panel.

### 4.2.4 Trends in disease incidence and case fatality

#### Methods

The Australian Burden of Disease study estimated trends in disease incidence and case fatality from time series of cause of death data. This is because mortality is the most reliable time series data to analyse trends over time and from mortality trends extrapolate to changes in incidence or case fatality. The drawback of this method is that it only applies to diseases that carry a significant risk of mortality. However, the problem is that for most diseases there is very little comparable information over a long enough time period to estimate trends in incidence. Thus, major causes of non-fatal health such as most mental, sense organ and musculoskeletal disorders are excluded from the trend analyses and, by default, incidence of these conditions is assumed to remain constant in the projections of disease burden.

The projection of diabetes is the exception. Mortality trends for diabetes are uninformative as a lot of diabetes-related mortality is coded under other causes and because there have been changes over time in the attribution of underlying cause to diabetes. However, we know there are significant trends in disease occurrence linked to the steady increase in body mass over the last 3 decades. Body mass index is the main risk factor for diabetes explaining over half of all disease incidence. However, the evidence base for the relationship between BMI and diabetes incidence is weak. The WHO CRA project used RRs from the Nurses' Health Study in the US. More recently, the Asia-Pacific Study Collaboration (2006) published a meta-analysis of 23 studies with measured BMI and diabetes mortality as an outcome. Only three studies had data on diabetes incidence as well and the Busselton study was by far the greatest contributor of cases. The RRs for mortality as an outcome and those for incidence only were very similar. They were quite a bit lower than those of the Nurses' Health Study and the authors argue that this is because that study had to rely on self-reported height and weight. As people with high BMI tend to overestimate height and underestimate weight more than others, their distribution of BMI would be biased towards a narrower range. If outcomes are regressed against a too narrow range of BMI values, the resulting RR estimate per unit BMI change would be overestimated.

#### Suggested application of trends in ACE-Prevention

Diseases with significant mortality trends include many of the cancers, COPD, diabetes, cardiovascular diseases, alcohol (including related cirrhosis) and most injury categories. We assume for most conditions that the trends apply to incidence. For cardiovascular disease we assume that 58% of mortality trends apply to incidence and 42% to case fatality as was found in a historical analysis of cardiovascular disease trends in England and Wales (Unal, 2004). For diabetes, we assume that half of the trend in case fatality from ischaemic heart

disease applies as a substantial number of excess deaths in diabetes are of cardiovascular origin and hence diabetics benefit from the favourable downward trends in cardiovascular case fatality.

The file 'full results Jan 2007.xls' on the ftp site in subfolder 'Aus BoD 2003' provides all trend information in the worksheet called 'Inc\_rankings\_males' and 'Inc\_rankings\_females'.

In ACE-Prevention we apply these trends for 20 years from 2003 to 2022 and keep the incidence and case fatality estimates constant at 2022 levels thereafter. There may be instances where the intervention(s) of interest are deemed to drive a most of the trends. In that case, the advice is to not apply the trends from the burden of disease study but to let the model drive those trends explicitly by the changing coverage, intensity or mix of interventions studied. Researchers are advised to consult Jan or Theo if this applies to their analyses.

## **4.3 Valuation of Benefits**

### **4.3.1 Utility measure**

As we move people through our models in ACE-Prevention we give a 'reward' or utility for time spent in each health state. We use disability weights from the Australian Burden of Disease Study in two ways:

1. to adjust time lived for the loss of health due to disability from the disease of specific interest in the model; and
2. to value the loss of health due to disability from all other causes .

The latter is estimated as the total of Prevalent Years Lived with Disability (PYLD) from the burden of disease study for all diseases not explicitly covered in the model divided by the population in each 5-year age and sex category. This information can be retrieved as an excel file (Full results Jan 2007.xls) from the project ftp site in the folder BoD Data. For diseases explicitly covered in a model the appropriate disability weight for that disease or associated health state applies. As the disability weights in the Australian Burden of Disease study are corrected for co-morbidity, we calculate the age and sex specific weights from the PYLD for that disease divided by the prevalence. For diseases with multiple disabling sequelae which we do not want to model explicitly, the sum of PYLD across all disabling sequelae of the disease divided by the prevalence gives the average probability of disability for each case of disease by age and sex.

We use a multiplicative model to 'combine' weights. This is done to avoid having combined disability weights greater than 1 if multiple severe disabilities are present in one person. As an example how to calculate combined weights, if the PYLD rate is 0.10 and the disability weight for a disease is 0.4, the combined weight is  $1-(1-0.1)*(1-0.4)=0.46$ . This weight reflects the loss of health due to co-morbidity between the disease of interest and the average of all other causes of disability (for a person of that age and sex). The utility or reward for a year lived in this example is  $1-0.46=0.54$ . In other words, a year lived with those two diseases in the model is valued at 0.54 years.

### **4.3.2 Discounting**

Similar to costs, a 3% discount rate is applied to health benefits.

### **4.3.3 Current practice and the null scenario as comparators**

In ACE-Prevention, we will evaluate all interventions against two comparators: 'current practice' and a 'hypothetical null option'. These concepts have been introduced in section

2.3 and are further elaborated upon here. Comparison with current practice gives most appropriate information to policy makers on the shift in costs and benefits that can be expected from implementing an option for change. This is the traditional approach in economic evaluation and produces Incremental Cost-Effectiveness Ratios (ICERs). We will use these results in our cost-effectiveness league tables. There is a potential problem with this approach. If current practice is inefficient, a new intervention will look more favourable in comparison. The WHO-CHOICE project introduced an alternative method of comparing all interventions against a 'theoretical null option' in which no interventions are present. In practice, WHO-CHOICE researchers redefine this as a partial null, i.e. a situation in which there are no interventions in place for the particular health problem under evaluation. This approach is most useful when optimal packages of interventions for one or a set of health problems is the desired outcome. The partial null is then back-calculated from current disease parameters by applying estimates of effective coverage and effectiveness for all relevant interventions.

For example, for a health problem X there are five interventions that have an impact and effective coverage as listed in Table 6.2. In the example we assume that current incidence of X is 10 per thousand, the case fatality is 0.05 and the average disability weight for X is 0.4. Intervention 'a' effectively reduces incidence by  $20\% \times 40\% = 8\%$ . Likewise, intervention b reduces incidence by  $50\% \times 40\% = 20\%$ . We assume a multiplicative impact of interventions a and b. That means the two interventions combined reduce incidence by  $1 - (1 - 8\%) \times (1 - 20\%) = 26.4\%$ . Therefore the partial null incidence equals 10 per thousand divided by  $(1 - 26.4\%)$  or 13.6 per thousand.

Table 6.2 Hypothetical example of impact and effective coverage for 5 interventions addressing health problem X

Intervention	Impact	Effective coverage
a	reduces incidence by 20%	40%
b	reduces incidence by 40%	50%
c	improves survival in cases by 10%	60%
d	improves survival in cases by 20%	50%
e	reduces severity by 50%	60%

In a similar vein the case fatality rate in the absence of interventions c and d would have been  $0.05 / (1 - (1 - (1 - 0.1 \times 0.6) \times (1 - 0.2 \times 0.5))) = 0.059$ . Also, if nobody with X would receive intervention e the average disability weight would be current average DW /  $(1 - cov + cov \times red)$  or  $0.4 / (1 - 0.6 + 0.6 \times 0.5) = 0.57$  where 'cov' is the effective coverage and 'red' is the reduction in severity.

## Summary Box 14 Determining health benefits

### Summary of key steps for determining the health benefits of interventions:

1. Systematically review natural history of the diseases and risk factors that are influenced by the intervention (often this information is available from the burden of disease study).
2. Identify and obtain Australia-specific data sets on population-level disease parameters: incidence, prevalence, case-fatality (again in most cases you will get this from the Australian Burden of Disease study)
3. Trends in incidence and case fatality as estimated in burden of disease study
4. If the intervention primarily addresses a risk factor, identify and obtain Australia-specific data on disease-risk factor associations. Select best estimates and check for consistency of estimates.
5. apply the measures of impact to the relevant transition probabilities (preventive intervention impact on incidence, curative intervention impact on remission, life saving interventions impact on case fatality and rehabilitative interventions impact on the disability weights);
6. To calculate the partial null, back calculate hazards by applying measures of impact and current exposure levels to interventions for disease/risk factor.
7. Construct mathematical models for projecting health benefits of interventions and assess intervention benefits relative to the null scenario.

## 5 Cost-effectiveness ratios

### 5.1 Incremental and marginal cost-effectiveness

Initially all analyses are undertaken on an incremental basis, where the incremental change in costs of the intervention compared to the 'current practice' base case are compared to the incremental change in benefits. The incremental cost-effectiveness ratio (ICER) uses the formula:

$$\Delta C/\Delta E$$

$\Delta C$  refers to the incremental net cost in Australian dollars of delivering the intervention over the comparator/base case.

$\Delta E$  refers to the incremental net health benefit in health adjusted life years saved.

Where appropriate, interventions will also be assessed using marginal analysis. This enables increasing amounts of investment in the chosen intervention to be compared with the additional benefits conferred. For instance, this may concern the addition of additional age groups or different 'at risk' groups to the target population for intervention or a stepping up of the intensity of the intervention effort. In a marginal analysis the difference in costs and outcomes is then calculated between the scenario with the initial target population/intensity and that of the expanded target population/intervention effort. This answers the research question: "is expansion of the target population cost-effective?"

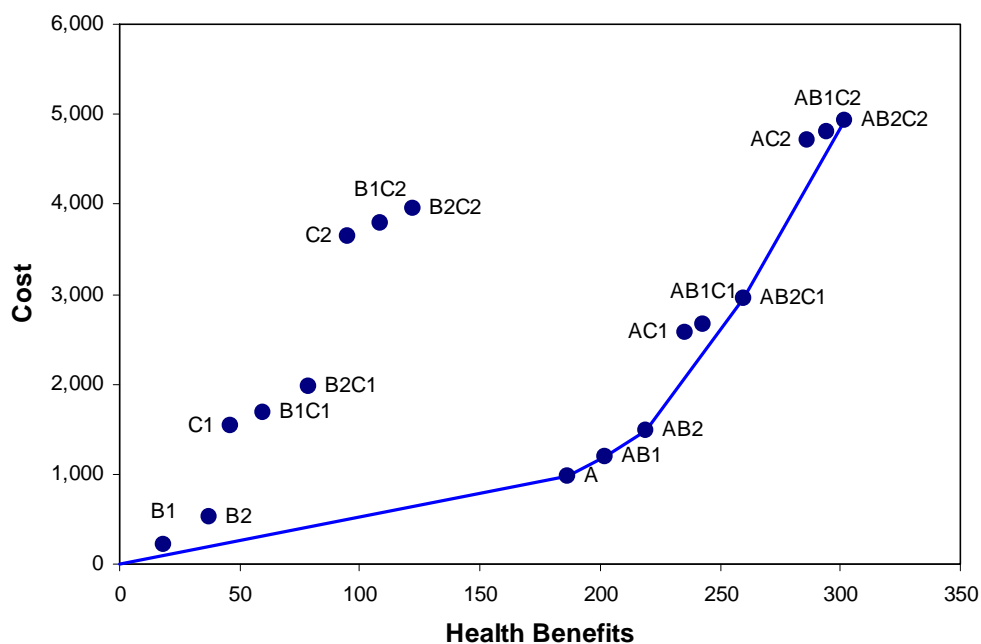
### 5.2 Average cost-effectiveness and the optimal expansion path

The WHO-CHOICE project uses the Generalised Cost-Effectiveness Analysis framework of assessing interventions comparing all interventions against the 'null scenario' to identify the optimal expansion path, in terms of costs and effects for a particular set of interventions (Murray, Evans et al. 2000). Table 5.1 illustrates the hypothetical results of a cost-effectiveness analysis of 3 separate interventions (denoted by letters A through to C). A is a population-wide intervention, while B and C are individually directed intervention that are implemented at 2 different coverage levels (denoted by numbers 1 and 2). For this set of interventions the costs and benefits of each individual intervention and each possible combination of interventions is assessed (total of 17 interventions or options for change). Here, the term average cost-effectiveness ratio is used to denote the cost-effectiveness ratio of an intervention in relation to the null scenario.

The intervention with the lowest cost-effectiveness ratio (A) is the most efficient and should be chosen first if resources are available. From this point, the incremental costs and effects of the remaining interventions are assessed, and the most cost-effective option based on the incremental cost-effectiveness ratio is then chosen. Note, that the incremental cost and impact of implementing the second intervention in the expansion pathway is calculated as the difference in the combined costs and outcomes of the first two interventions and the costs and outcomes of the first intervention only. Repeating this process until the set of options for change is exhausted determines the expansion path of the most cost-effective options. This process results for the example above in the following expansion path (Figure 5.1): Intervention A, combine A with B at coverage level 1 (AB1), extend B to coverage level 2 (AB2), combine AB2 with C at coverage level 1 (AB2C1), extend C to coverage level 2.

**Table 5.1 Hypothetical costs and health benefits of 17 interventions.**

Intervention	Costs	Effects	Average cost-effectiveness ratio
A	985	187	5,275
B1	219	18	12,173
B2	518	37	13,815
C1	1,549	46	33,611
C2	3,646	95	38,371
AB1	1,199	202	5,925
AB2	1,492	219	6,817
AC1	2,568	235	10,910
AC2	4,705	287	16,417
B1C1	1,684	59	28,302
B2C1	1,982	79	25,104
B1C2	3,781	108	34,869
B2C2	3,948	122	32,261
AB1C1	2,670	243	10,986
AB1C2	4,806	294	16,338
AB2C1	2,963	260	11,416
AB2C2	4,934	302	16,337



**Figure 5.1 Expansion path of the optimal set of interventions according to costs and effects.**

Furthermore, by determining the costs and benefits of current practice, the relative efficiency of the set of currently implemented interventions can be compared to the optimal set of interventions. For example, if current practice was at point B2C1, the various health gains and/or resource savings that could be made by moving to one of the optimal set of interventions (Table 5.2).

**Table 5.2. Incremental costs and benefits of moving from current practice (B2C1) to one of the optimal set of interventions.**

<b>Intervention</b>	<b>Incremental costs</b>	<b>Incremental benefits</b>
B2C1 → A	-997	108
B2C1 → AB1	-783	123
B2C1 → AB2	-490	140
B2C1 → AB2C1	981	181
B2C1 → AB2C2	2,952	223

In other words, a decision maker may choose to reallocate resources from B2C1 to A, with cost savings of 997 and health gains of 108. Moving from current practice B2C1 to the optimal mix AB2C2 would cost an additional 2,952 for 223 health gain.

In the WHO-CHOICE project all possible combinations of interventions are analysed to derive the most cost-effective mix of interventions. In ACE-Prevention we will adopt a more targeted approach to analysing combinations of interventions by always taking the most cost-effective intervention as a starting point and expanding the 'ideal' package with each successive next cost-effective intervention. If the next intervention to be included is mutually exclusive with a previous intervention (but provides greater benefit at greater cost) it will replace the other intervention in the package.

#### **Summary of key steps for determining the cost-effectiveness of interventions**

1. For the main results to be tabulated in a league table, calculate the incremental cost-effectiveness ratio by dividing the difference in cost between intervention and the 'current practice' comparator by the difference in health outcomes between intervention and comparator;
2. Calculate marginal cost-effectiveness ratios for 'expanded' coverage or target populations for the same intervention if applicable;
3. Determine the expansion path by estimating costs and effects of all interventions for one or more health problems (including current practice) against the 'partial null' (i.e. no interventions in place that address these health problems) and combining interventions in the order of their average cost-effectiveness ratios
4. Determine incremental costs and effects from current practice.



## 6 Sensitivity and uncertainty analyses

It is useful to make a distinction between sensitivity and uncertainty analyses. We define 'uncertainty testing' to cover variation in those technical parameters (usually economic and epidemiological inputs) that impact on disease incidence/prevalence, efficacy/effectiveness, attendance, compliance rates, complication rates, unit costs and so on. We define 'sensitivity testing', on the other hand, to cover variation in social value parameters and/or the scenario under evaluation. Variations in the scenario might include changes in the study perspective, in the choice of comparators or inclusion of contentious cost impacts (such as production losses). Social value parameters include issues such as the choice of discount rate (social rate of time preference), weighting the health gain for equity (who receives the health gain) or for those most in 'need' (having regard to those severely ill and their fate if left untreated). It is often useful in economic evaluation to separate the technical calculation of the anticipated health gain from the social valuation or decision context placed on the anticipated health gain. Both our approach to sensitivity analysis and uncertainty testing is set out below.

### 6.1 Sensitivity Analysis

Largely under the banner of the Capacity Building Grant we will examine the impact of variations in social value parameters and the inclusion or not of certain cost elements for selected analyses. Potential topics for such sensitivity analysis include:

- Variations in the discount rate (0%, 5% and 7%);
- Inclusion of volunteer time costs;
- Cost offsets for diseases not explicitly modelled as being influenced by the intervention;
- Inclusion of production gains and losses; and
- Inclusion of placebo effects.

### 6.2 Uncertainty analysis

In the ACE-Prevention study, greater emphasis is placed on uncertainty analyses to evaluate the impact of uncertainty around the epidemiological and costing estimates on the final results. In the primary analyses, point estimates are calculated to measure benefits, costs and, ultimately, the incremental cost-effectiveness ratios for the various interventions. While the best evidence available is used, there is always a level of uncertainty associated with cost and outcome estimates. Even data from randomised controlled trials (RCTs), may not be easily transferable to the Australian setting or to the proposed intervention. To examine the impact of uncertainty on the study results, simulation modelling techniques are used to present uncertainty around each incremental cost-effectiveness ratio that reflects all the main sources of uncertainty in the calculations. This uncertainty can be presented numerically as a range of values around the point estimate or graphically in a cost-effectiveness plane or acceptability curve.

The research team uses the commercial package @RISK or Ersatz created by Jan Barendregt, both of which are add-in software programs to commercial spreadsheet packages, to conduct Monte Carlo simulation of uncertainty. It allows estimates and assumptions to be entered as probability distributions in a spreadsheet. The program then recalculates the spreadsheet many times over—each time picking a value out of all defined probability distributions—and provides summary statistics across all iterations (usually 2000 or more) for selected outcome variables. From the values generated by the iterations of the simulation, a 95% uncertainty interval can be calculated by taking the 2.5 and 97.5 percentiles to mark the lower and upper bounds. This uncertainty interval can be interpreted as the range within which the true result lies with 95% certainty. An uncertainty

interval differs from a confidence interval in that it includes both type I and type II errors. The probability distributions around the input variables are based on standard errors quoted in, or calculated from, the literature; the range of parameter values quoted in, or calculated from, the literature; and from expert advice on the likely scenarios under Australian conditions.

Table 6.1 shows the recommended uncertainty distributions to be used around common model input parameters.

**Table 6.1 Parameters and distributions included in uncertainty analyses in the ACE-Obesity study**

Parameter	Uncertainty distribution	Comments
<b>Effectiveness</b>		
RR or OR as effect size – calculated from RCTs	the exponential of the normal distribution defined by the natural log of the RR and the standard error of the logged RR	if 95% CIs are presented in a paper the SE is calculated as the average of $\ln RR - \ln(LL)$ and $\ln(UL) - \ln RR$ where UL and LL are upper and lower limit of confidence interval
Effect size measure as difference in mean (e.g. difference in mean number of daily drinks of alcohol per day)	a normal distribution defined by the mean difference and the standard error thereof	in some cases you may want to truncate the distribution at 0 to avoid a negative impact if that is considered implausible
<b>Adherence</b>		
Adherence with the intervention – if RCT available	triangular distribution	Max is mean from the RCT, min is 50% (or other value determined in consultation with expert panel) and top of triangle in middle
<b>Costs</b>		
PBS and MBS listed drug costs and medical visits	none	
uncertainty around patient contribution to drugs and medical visits	triangular distribution	(10%, 15%, 20%)

It is important to avoid entering the same distribution key input variables twice in the same model. For instance, when the same measure of impact is used across all age groups, for each age group it must refer to the same uncertainty distribution, i.e. for each iteration of the simulation the same value must be chosen for each of the age groups. A simple cell reference to the single spot in an excel model can be used to avoid such duplication. It is good practice in excel based models to cluster all uncertainty distributions in one spot. The effect of duplication of uncertainty distributions is that the uncertainty around the final results of cost-effectiveness will be made wider than necessary.

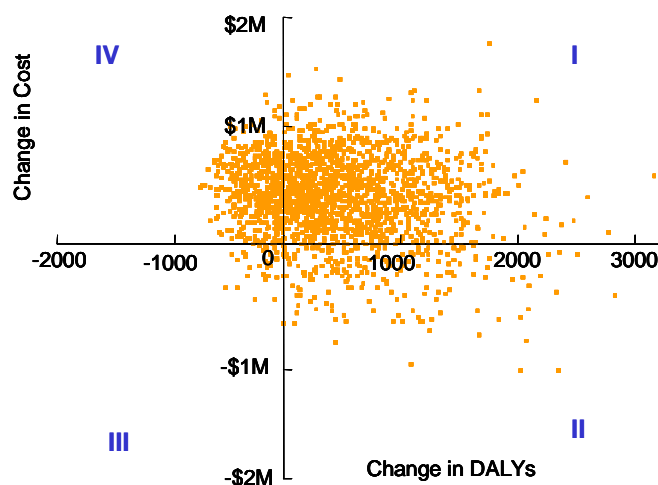
The uncertainty around different input variables may be correlated. For instance, severity of disease and cost of treating the disease are likely to be correlated. If a model has a distribution around the range of severity of disease and treatment costs, you would like to see a greater probability of a high cost estimate being chosen for a more severe presentation of the disease. In such cases, @RISK or Ersatz allow you to correlate the two uncertainty distributions. A correlation of 1 means total correlation: each time it will choose a value at the same percentile of the distributions of both input variables. Correlation values between 0 and 1 will mean a greater or lesser likelihood of similarly high or low values being chosen between the two distributions.

So far we have dealt with parameter uncertainty in this section. Another source of uncertainty is 'structural uncertainty'. That is uncertainty introduced by the structure of the model. It generally is not so easy to quantify this type of uncertainty but there are examples that this may be considerable. For instance, in an earlier ACE study we evaluated a change of the interval for cervical cancer screening from 2 to 3 years and concluded that there would be minimal health loss and considerable cost savings. However, when in this project we examined the same research question in a micro-simulation model which allowed retaining information on the distribution of the speed at which different cancers grow (rather than assuming average growth for all cancers) we came up with a very different answer. As more aggressive growing cancers would be missed with a longer interval, the health loss associated with a 3 year interval between PAP smears was much more than trivial and altered our conclusion recommending the longer interval. This topic has not yet been studied a lot in economic evaluation, yet is critical. We plan to do more work in this area under the Capacity Building Grant.

### 6.2.1 Presentation of uncertainty

There has been considerable debate in the literature over the presentation of uncertainty around the incremental cost-effectiveness ratio (ICER) (Hunink, Bult et al. 1998; Briggs and Gray 1999; Craig, Black et al. 2000; Halpern, Weinstein et al. 2000; Fenwick, Claxton et al. 2001; Hutubessy, Baltussen et al. 2001). To illustrate some of the problems of interpreting uncertainty around the ICER, Figure 6.2 shows a scatter plot of the incremental cost (IC) against the incremental effectiveness (IE) for a hypothetical intervention, with each point representing an iteration of a simulation (i.e. one recalculation of the model).

Typically, new interventions are more effective and more costly (IC>0, IE>0) than current practice, and therefore results in the right upper quadrant (Quadrant I) are the most common. Results in the right lower quadrant (Quadrant II) are a "win-win" situation with health benefits at a net cost saving (IC<0, IE>0). The left lower quadrant (Quadrant III) reflects health loss as well as cost savings (IC<0, IE<0). The left upper quadrant (Quadrant IV) represents adverse outcomes at a cost (IC>0, IE<0), the most unfavourable outcome.



**Figure 6.2 Example of an incremental cost-effectiveness scatter plot.**

Due to the mathematical properties of the ICER, there are several problems with its interpretation. As the ICER is a ratio of the incremental cost (IC) over the incremental effectiveness (IE), ICERs in Quadrant II (where the incremental cost is negative) are indistinguishable from ICERs in Quadrant IV (where the incremental effectiveness is

negative). For example, an intervention that saves \$50 (IC = - 50) with a health gain of 5 DALYs (IE = 5) has the same ICER (ICER = -10) as an intervention that costs \$50 (IC = 50) with 5 additional DALYs incurred (IE = -5, ICER = -10). This is problematic because these two situations are opposing extremes – ICERs in Quadrant II are highly favourable, whereas ICERs in Quadrant IV are highly unfavourable. ICERs in Quadrant III are positive but present another problem when compared with the positive ICERs in Quadrant I. Whereas, a small ICER in quadrant I indicates a favourable result (lots of health benefits for low cost), the direction of ICERs in quadrant III goes in the opposite direction: a very large ICER is favourable as it mean a lot of cost saving for only a small amount of health loss.

Additional problems arise when results that are 'dominant' (i.e. have a negative IC but positive IE) are examined more closely. For example, consider the three outcomes outlined in Table 6.2.

**Table 6.2 Interpreting the magnitude of negative cost-effectiveness ratios (ICERs)**

	IC	IE	ICER
<b>Iteration 1</b>	-10	10	-1
<b>Iteration 2</b>	-10	5	-2
<b>Iteration 3</b>	-5	5	-1

In iteration 1, the ICER = -1, while in simulation 2, the ICER = -2. In cost-effectiveness terminology, low ICERs are considered more favourable. Therefore, iteration 2 with an ICER of -2 should be better than iteration 1 with an ICER of -1. However, it is clear that iteration 1 is the more favourable outcome, as it saves more years of life (10 vs 5) at the same cost saving (-10). In this situation, the outcome with the higher ICER is more favourable than the outcome with the lower ICER. This result might suggest that negative ICERs are ranked inversely to positive ICERs, however, it is not as simple as this. iteration 3 has a higher ICER (-1) than iteration 2 (-2), but it is clear that iteration 2 is more favourable than iteration 3 (it saves more money with the same benefit). Similarly, even though the ICERs of iterations 1 and 3 are equal, iteration 1 is the more favourable outcome (saving money with additional benefit). The magnitude of a negative ICER, therefore, does not convey any useful information and including it in the calculation of an uncertainty interval leads to misleading results. The underlying problem is that "for a dominant intervention, a large magnitude is desirable in both the numerator and the denominator, yet these 2 desirable features drive the magnitude of the ratio in opposite directions" (Stinnett and Mullahy 1997). Problems with negative ICERs are further highlighted by the complex distribution of ICER, which is apparent when denominator values close to zero are examined (small negative IEs lead to highly negative ICERs). If all points lie in Quadrant II, a correct interpretation of the relative value of each point against others can be made using the product of IE and IC rather than the ratio. However, there is no intuitive way to interpret such a cost-effectiveness product.

The solution is to avoid presenting numerical values in uncertainty ranges when there are ICER values in Quadrants II, III or IV. Instead, all values in Quadrant II should be called 'dominant' and all values in Quadrant IV 'dominated'. If any ICER value is outside the usual quadrant I you should not rely on the simulation package (e.g. @RISK) to calculate a median or mean value for you because as we saw above the numerical values of ICERs in other quadrants cannot be interpreted. The cost-effectiveness plane is an alternative way of presenting uncertainty particularly when you have values in more than one quadrant. It helps in such a case if you add the proportion of iterations that produce a value in each quadrant and for the values in quadrant I what proportion fall below and above the 'accepted' willingness to pay for a DALY threshold. By doing that you can tell policymakers that there is a probability of x that your ICER is favourable and a probability y that the intervention may do more harm than good. Aspirin in the prevention of cardiovascular disease is a good example where you may find ICER results spanning Quadrants I, II and

IV. If you find, for example, that 20% of values are in quadrant II, 60% of values are in Quadrant I and fall below your threshold, 15% of values are in Quadrant I but are higher than the threshold; and 5% of values fall in quadrant IV your results could be formulated as follows: “there is an 80% probability of a favourable ICER; the remaining 20% probability that the ICER is not favourable includes a 5% chance that the intervention actually could do more harm than good”.

Another graphical way of presenting your uncertainty is in an acceptability curve. ON the x-axis it shows a range of ‘willingness to pay threshold values’ for the cost per DALY (or life year save) and plots on the y-axis the probability that the ICER value lies below each threshold value. In the example of figure 6.3 the simulation results indicate that there is a slightly more than 90% probability of a cost-effective result below the threshold of \$50,000 per DALY (of for that matter a probability slightly below 90% that the ICER is less than \$20,000 per DALY. However, the acceptability curve asymptotes as 91% indicating that for this intervention 9% of the iterations of the uncertainty simulation gave values in Quadrant IV (i.e. health loss at a cost). If there would have been values in Quadrant II, the curve would not have started at probability 0 for a threshold value of 0 but at a level on the y-axis corresponding with the proportion of ICER values in Quadrant II.

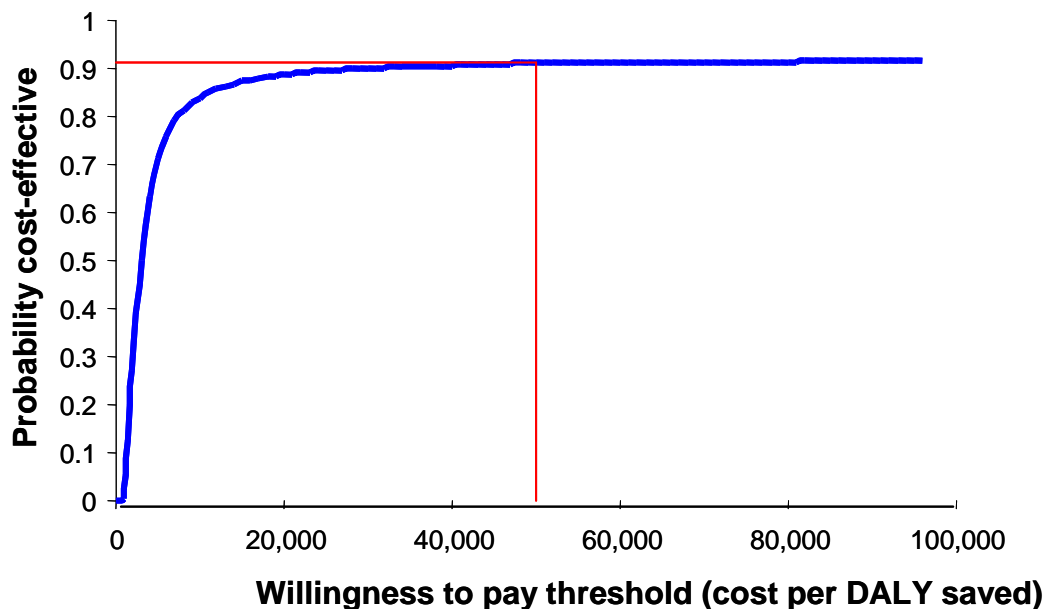


Figure 6.3 Example of an acceptability curve

### 6.2.2 Expansion path uncertainty

As GCEA evaluates interventions in respect to the null scenario, it has been argued that GCEA removes the problem of negative cost-effectiveness ratios (Hutubessy, Baltussen et al. 2001). In other words, as cost-offsets are not considered costs will not be negative and as interventions are on the whole beneficial there will not be negative effects relative to the null scenario. While this is largely true, certain interventions may in the presence of parameter uncertainty result in net negative benefits or harm. For example, aspirin reduces the risk of CHD and ischemic stroke events, but increases the risk of gastrointestinal bleeding and hemorrhagic stroke. Under uncertainty analysis, aspirin may result in net harm rather than net benefit particularly in those with low absolute risk of a CVD event. Furthermore, if cost offsets are included, as they are in this project, negative costs are still

possible. This problem is easily solved by incorporating decision rules into the determination of the optimal expansion that state that if an intervention has negative costs (i.e. resource savings) and positive benefits, it is included in the optimal package. On the other hand, if an intervention has negative benefits (i.e. net harm) and positive costs it is excluded from the optimal package. Given that interventions are assessed against the null it is unlikely that interventions will have both negative effects and negative benefits.

Uncertainty in the optimal expansion path for GCEA analyses have been presented as “cloud graphs” (Murray, Lauer et al. 2003). Figure 6.4 demonstrates a cloud graph of the effect of parameter uncertainty on the costs and effects of the set of interventions outlined in Table 5.1.

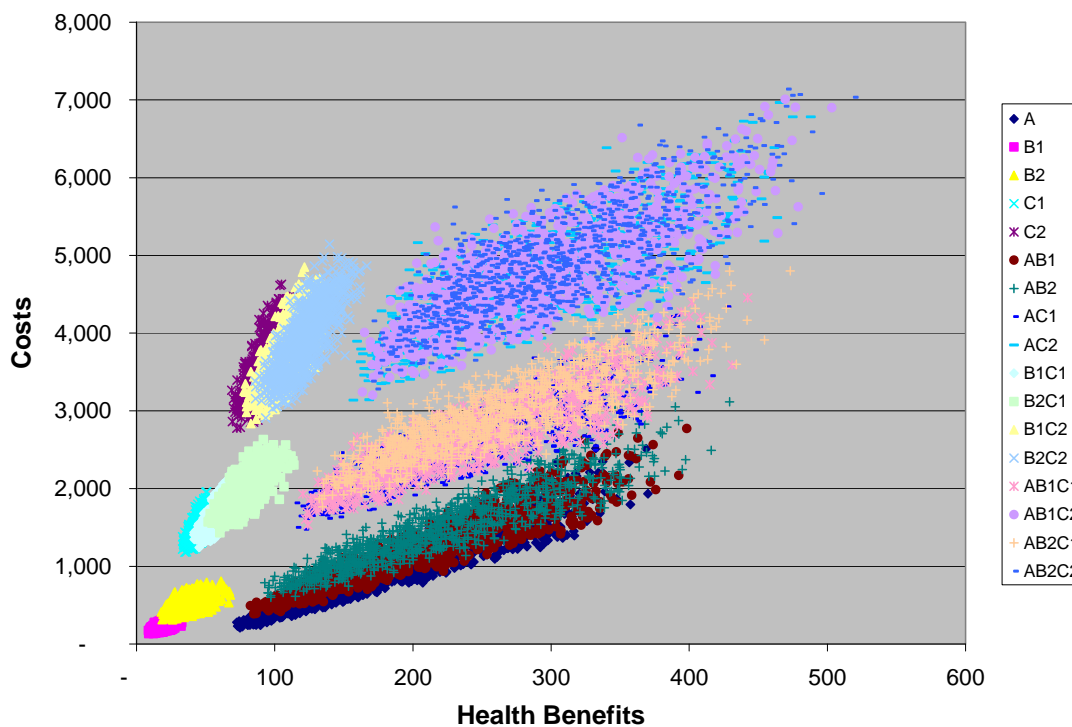


Figure 6.4 Cloud graph based on example from Table 5.1

As can be seen from the graph there is substantial overlap in the distributions or “clouds” and this suggests that there is uncertainty as to the choice of interventions. As is acknowledged by the GCEA guidelines (Tan-Torres Edejer, Baltussen et al. 2003), this interpretation is problematic as the costs and effect of interventions are often highly correlated. For example, the uncertainty in the cost of health centre visits will be common to all primary care based interventions<sup>13</sup>.

The solution proposed by the GCEA guidelines is to use a stochastic league table (discussed below). Stochastic league tables, however, incorporate a health maximising condition at different budget levels that may obscure the choice of expansion path and make

<sup>13</sup> As a result of these common uncertainty inputs, it is necessary to simulate uncertainty in costs and effects using the same value of the distribution for each intervention that uses the input. Available software for constructing stochastic league tables (MCLeague) does not contain this feature. The example provided in this document was simulated using the same value of the distribution for each intervention, by assessing uncertainty using @RISK and excel, with each spreadsheet linked to a common input parameter.

interpretation difficult. A useful intermediate step is to determine the probability that a particular expansion path is chosen (Table 6.3).

**Table 6.3 Probability of particular expansion paths being chosen based on the uncertainty in costs and benefits.**

Expansion path	Probability
A→AB1→AB2→AB2C1→AB2C2	96%
A→AB1→AB1C1→AB2C1□AB2C2	4%

Presentation of uncertainty results in this manner demonstrates that the choice of expansion path is clear. 96% of the time the expansion path of Intervention A, combine A with B at coverage level 1 (AB1), extend B to coverage level 2 (AB2), combine AB2 with C at coverage level 1 (AB2C1), extend C to coverage level 2 is chosen. The other 4% of the time, AB1 is combined with C1 first before extending coverage of B to coverage level 2.

### 6.2.3 Stochastic League Tables

Stochastic league tables (Hutubessy, Baltussen et al. 2001; Tan-Torres Edejer, Baltussen et al. 2003) assess the impact of parameter uncertainty on the choice of the optimal package of interventions according to cost-effectiveness but with two additional conditions. Firstly, the choice of intervention packages is constrained by a specified budget level. Secondly, for the given budget level the most cost-effective option is not necessarily chosen, but the intervention that maximises health benefits. Table 6.4 demonstrates a stochastic league table for the set of interventions from Table 5.1.

This table demonstrates that at low budget levels (500 and 750) the decision maker may not be able to purchase the first option (A) determined from the expansion path, therefore the more affordable but less cost-effective interventions (B1 and B2) are chosen the majority of the time. This is important information that is otherwise not conveyed by the usual expansion path.

Interpretation of results for this stochastic league table can be confusing, and aspects of this have been covered in the GCEA guidelines (Hutubessy, Baltussen et al. 2001; Tan-Torres Edejer, Baltussen et al. 2003). Another aspect of stochastic league tables that makes interpretation difficult, is that in order to determine whether a specific intervention (A, B, or C) is included at a given budget, it is necessary to sum the probabilities for the packages that include that intervention. For example, at budget level 1000, to determine that intervention A is part of the package, it is necessary to add the probabilities for A at 16%, AB1 at 23% and AB2 at 16%, for a total probability of 53%. This can be facilitated by presenting the probabilities according to the specific interventions (Table 6.5) but information is then lost as to whether each intervention is combined with others, or implemented alone.

**Table 6.4. Stochastic League Table by Intervention Packages.**

	Budget Level									
	500	750	1,000	1,250	1,500	2,000	2,500	3,000	4,000	5,000
<b>A</b>	13	18	<b>16</b>	14	11	4	-	-	-	-
<b>B1</b>	<b>48</b>	-	-	-	-	-	-	-	-	-
<b>B2</b>	36	<b>63</b>	45	28	11	-	-	-	-	-
<b>C1</b>	-	-	-	-	2	-	-	-	-	-
<b>C2</b>	-	-	-	-	-	-	-	-	-	-
<b>AB1</b>	3	17	23	<b>22</b>	20	9	2	-	-	-
<b>AB2</b>	-	2	16	36	<b>54</b>	<b>72</b>	<b>47</b>	20	1	-
<b>AC1</b>	-	-	-	-	-	4	8	5	-	-
<b>AC2</b>	-	-	-	-	-	-	-	-	4	6
<b>B1C1</b>	-	-	-	-	2	1	-	-	-	-
<b>B2C1</b>	-	-	-	-	-	1	-	-	-	-
<b>B1C2</b>	-	-	-	-	-	-	-	-	-	-
<b>B2C2</b>	-	-	-	-	-	-	-	-	-	-
<b>AB1C1</b>	-	-	-	-	-	7	22	17	3	-
<b>AB1C2</b>	-	-	-	-	-	-	-	-	4	8
<b>AB2C1</b>	-	-	-	-	-	1	21	<b>59</b>	<b>81</b>	29
<b>AB2C2</b>	-	-	-	-	-	-	-	-	6	<b>58</b>

\*Numbers in bold denote intervention options chosen without uncertainty.

Another aspect that obscures the cost-effectiveness results is the health maximising condition. At budget levels of 2000, 2,500 and 3,000 the intervention package there is a small probability of choosing AC1. As there are two factors (cost-effectiveness and health maximisation) that determine choice of an intervention at a given budget level, it is uncertain whether this package is chosen because AC1 is more cost-effective than AB2, or whether AC1 is health maximising (but less cost-effective than AB2). By looking at the uncertainty results presented according to the probability of a particular expansion path being chosen (Table 6.3), it is clear that it is because AC1 is health maximising, but less cost-effective than AB2 at those budget levels.

**Table 6.5. Stochastic League Table by interventions.**

	Budget Level									
	500	750	1,000	1,250	1,500	2,000	2,500	3,000	4,000	5,000
<b>A</b>	16	37	55	72	85	98	100	100	100	100
<b>B1</b>	51	17	23	22	22	17	23	17	7	8
<b>B2</b>	36	66	61	64	65	74	68	78	89	87
<b>C1</b>	-	-	-	-	5	14	51	80	85	29
<b>C2</b>	-	-	-	-	-	-	-	-	15	71

With these interpretation issues, this project advocates the presentation of expansion path uncertainty using the method described in Table 6.4 as the preferred method of communicating uncertainty results. Stochastic league tables can be presented as information not contained in the usual expansion path may be gathered on intervention options under certain budget constraints (e.g. the use of B1 and B2 in low budget settings).



Care must be taken, however, in the interpretation of stochastic league tables. Researchers under the Capacity Building Grant may best explore this

#### **6.2.4 Determining the most critical input parameters**

Multivariate stepwise regression of the 2000 or so iterations of cost-effectiveness ratio using the sampled values of the model input parameters as covariates allows the assessment of those parameters which impact most on the cost-effectiveness of an intervention. This can indicate future research priorities if greater accuracy of results is desired. Parameters with regression coefficients greater than 0.25 (or less than -0.25) are reported.

##### **Summary of key steps for uncertainty and sensitivity analysis:**

1. Sensitivity analysis to determine the effect of different social value parameters and inclusion criteria for costs, e.g. discount rate, age-weighting, inclusion of production losses will be largely the domain of researchers under the Capacity Building Grant.
2. Use multi-way uncertainty analysis (Monte-Carlo simulation) for determining the effect of input parameter uncertainty.
3. Determine the probability that a particular expansion path is chosen.
4. Use of a stochastic league table to determine the effect of budget constraints will largely be a topic to be explored by researchers under the Capacity Building Grant.
5. Determine critical input parameters using multi-variate stepwise regression.

## 7 Second stage filter criteria

Important dimensions of the PBMA approach to be adopted in the ACE–Prevention study include its potential to engage stakeholders and to broaden the concept of benefit to reflect the underlying goals, objectives and principles of the organisation or program wishing to employ the technique. There is an increasing awareness in the literature on priority setting of the need to combine technical approaches—such as economic evaluation—with approaches that facilitate due process (Carter et al. 2000, Carter 2001). While evidence on cost-effectiveness is the main focus of activity in ACE–Prevention, other criteria of benefit that can influence the priority ranking of the selected interventions will be recognised. These additional criteria function as a second filter by which each of the interventions are judged before recommending allocation of more or less resources. The ACE studies consciously adopt an explicit approach to priority setting where visibility of the cost-effectiveness estimates, of judgements about the second stage filters, and about the processes employed, are all emphasised.

The selected filters should be in alignment with government objectives and principles for obesity prevention and management. In the ACE-Obesity study for example, *Healthy weight 2008* (National Obesity Taskforce 2003) was found to embody the principles of effectiveness, efficacy, efficiency, equity, sustainability, empowerment, affordability, ethics and feasibility. The selected filters should also reflect previous ACE studies.

The Project Steering committee will be tasked to agree on the selection and definition of the second stage filter criteria to be employed in the ACE–Prevention study. These are likely to mirror those used in the ACE-Obesity study as defined below with some modifications to reflect the current study:

### 7.1 Capacity of the intervention to reduce ‘inequity’

This criterion was defined as ‘the impact of the intervention on inequity in the distribution of disease and health status and access to, or utilisation of, specific intervention(s)’. This equity criterion may include an analytical component. Presented with numerical evidence of inequalities in current health status and/or access/utilisation of services, the Project Steering Committee is likely to need to use a qualitative judgement to apply the equity criterion. There are various possible methods to combine cost-effectiveness ratios and equity considerations in a numerical way. These include weighting the DALY measure; using decision theory to create a new index score; hand shuffling the ranked interventions to reflect the equity results; cost value analysis developed by Nord (1999); or cost consequences analysis (Drummond et al. 1997). Researchers under the Capacity Building Grant will explore this further.

### 7.2 Strength of evidence

Evidence impacts on the ACE–Prevention study in a number of ways: i) through the selection of the options for change and the availability of evidence to underpin them; ii) through the sensitivity and uncertainty analysis, and iii) through the confidence that can be placed in the ICERs.

There is an emerging view that there is a single framework within which evidence on clinical, public health and behavioural interventions can be assessed. While the nature of the evidence for different kinds of health interventions inevitably varies, and the evidence for public health and social science interventions often is weaker than that for clinical interventions, the logic used to assess the evidence is the same for all of them. In the first instance during the ACE-Obesity study, it was proposed to use the existing ACE approach

(Carter et al. 2000) based on the grading of the level of evidence into three categories: 'sufficient evidence'; 'limited evidence' and 'inconclusive evidence' as set out in Table 7.1.

**Table 7.1 Classifying the strength of the evidence ~ initial approach**

Strength category	Strength of the evidence
<p><b>“Sufficient evidence of effectiveness”:</b> Effectiveness is demonstrated by sufficient evidence from well-designed research.</p>	<ul style="list-style-type: none"> <li>• The effect is unlikely to be due to chance (e.g. P is &lt; 0.05) and</li> <li>• The effect is unlikely to be due to bias (e.g. evidence from:               <ul style="list-style-type: none"> <li>▪ a level I<sup>14</sup> study design;</li> <li>▪ several good quality level II studies; or</li> <li>▪ 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)</li> </ul> </li> </ul>
<p><b>“Limited evidence of effectiveness”:</b> Effectiveness is demonstrated by limited evidence from studies of varying quality</p>	<ul style="list-style-type: none"> <li>• The effect is probably not due to chance (e.g. P is &lt; 0.05) but</li> <li>• Bias, while not certainly an explanation for the effect, cannot be excluded as a possible explanation (e.g., evidence from:               <ul style="list-style-type: none"> <li>▪ one level II study of uncertain or indifferent quality;</li> <li>▪ evidence from one level III-1 or III-2 study of high quality;</li> <li>▪ evidence from several level III-1 or III-2 studies of insufficiently high quality to rule out bias as a possible explanation; or</li> <li>▪ evidence from a sizeable number of level III-3 studies which are of good quality and consistent in suggesting an effect).</li> </ul> </li> </ul>
<p><b>“Inconclusive evidence of effectiveness”:</b> Inadequate evidence due to insufficient or inadequate quality research.</p>	<ul style="list-style-type: none"> <li>• No position could be reached on the presence or absence of an effect of the intervention (eg. no evidence from level I or level II studies and level III studies are available, but they are few and of poor quality, or only level IV studies are available.)</li> </ul>

Source: (Carter et al. 2000)

At a later stage, however, it was decided that further work was necessary on this filter because it was inappropriate for obesity interventions as it would exclude most evidence as being limited or inconclusive. The working group endorsed the need for a category between 'limited' and 'inconclusive'. Following work on alternative classifications, the working group resolved at its March 2005 meeting to accept a new classification system which sought to combine the traditional classification of evidence based on epidemiologic study design with indirect and parallel forms of evidence that would not ordinarily be captured. This classification draws on the work of Hawe and Shiell 1995, Swinburn et al. 2005 and also reflects aspects of other evidence frameworks (NHMRC National Breast Cancer centre Psychosocial Working group 2000; Loxley et al. 2004; Rychetnik and Frommer 2002). This proposed approach is set out in Table 7.2

- 
- <sup>14</sup> 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 with historical control, two or more single-arm studies, or interrupted time series without a parallel control group.
  - IV Evidence obtained from case series, either pre-test and post-test.

Source: National Health and Medical Research Council 1999

**Table 7.2 Classifying the strength of the evidence ~ proposed approach**

Evidence from Level I-III study designs	Evidence from Level IV studies, indirect or parallel evidence and/or from epidemiological modelling using a mixture of study designs
<p><b>“Sufficient evidence of effectiveness”</b></p> <p><i>Effectiveness is demonstrated by sufficient evidence from well-designed research, viz:</i></p> <ul style="list-style-type: none"> <li>• The effect is unlikely to be due to chance (e.g. P is &lt; 0.05) and</li> <li>• The effect is unlikely to be due to bias (e.g. evidence from<sup>3</sup>:               <ul style="list-style-type: none"> <li>• a level I study design;</li> <li>• several good quality level II studies; or</li> <li>• 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)</li> </ul> </li> </ul>	<p><b>“Likely to be effective”</b></p> <p><i>Effectiveness results are based on:</i></p> <ul style="list-style-type: none"> <li>• Sound theoretical rationale and program logic; and</li> <li>• Level IV studies, indirect<sup>1</sup> or parallel evidence<sup>2</sup> for outcomes; or</li> <li>• Epidemiological modelling to the desired outcome (BMI) using a mix of evidence types or levels.</li> </ul> <p>The effect is <u>unlikely</u> 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).</p> <p>Implementation of this intervention should be accompanied by an appropriate evaluation budget.</p>
<p><b>“Limited evidence of effectiveness”</b></p> <p><i>Effectiveness is demonstrated by limited evidence from studies of varying quality, viz:</i></p> <ul style="list-style-type: none"> <li>• The effect is probably not due to chance (e.g. P is &lt; 0.10) but</li> <li>• Bias, while not certainly an explanation for the effect, cannot be excluded as a possible explanation (e.g., evidence from<sup>3</sup>:               <ul style="list-style-type: none"> <li>• one level II study of uncertain or indifferent quality;</li> <li>• evidence from one level III-1 or III-2 study of high quality;</li> <li>• evidence from several level III-1 or III-2 studies of insufficiently high quality to rule out bias as a possible explanation; or</li> <li>• evidence from a sizeable number of level III-3 studies which are of good quality and consistent in suggesting an effect).</li> </ul> </li> </ul>	<p><b>“May be effective”</b></p> <p><i>Effectiveness results are based on:</i></p> <ul style="list-style-type: none"> <li>• Sound theoretical rationale and program logic; or</li> <li>• Level IV studies, indirect<sup>1</sup> or parallel evidence<sup>2</sup> for outcomes; or</li> <li>• Epidemiological modelling to the desired outcome (BMI) using a mix of evidence types or levels.</li> </ul> <p>The effect is <u>probably not</u> due to chance but bias, while not certainly an explanation for the effect, cannot be excluded as a possible explanation.</p> <p>Would benefit from further research and/or pilot studies before implementation.</p>
<p><b>“Inconclusive evidence of effectiveness”</b></p> <p><i>Inadequate evidence due to insufficient or inadequate quality research.</i></p> <p>No position could be reached on the presence or absence of an effect of the intervention (eg. no evidence from level I or level II studies and level III studies are available, but they are few and of poor quality.)</p>	<p><b>“No evidence of effectiveness”</b></p> <p>No position could be reached on the likely credentials of this intervention. Further research may be warranted.</p>

**Notes**

1)“Indirect evidence”: Information that strongly suggests that the evidence exists (eg a high and continued investment in food advertising is indirect evidence that there is positive (but propriety) evidence that food advertisement increases sales of those products (Swinburn et al. 2005).

2) “Parallel evidence”: Evidence of intervention effectiveness for another public health issue using similar strategies (eg the role of social marketing, regulation or behavioural change initiatives in tobacco control, sun exposure, speeding, etc) (Swinburn et al. 2005)

3) These evidence classifications are based on the NHMRC

- 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 with historical control, two or more single-arm studies, or interrupted time series without a parallel control group.
- IV Evidence obtained from case series, either pre-test and post-test.

Source: National Health and Medical Research Council 1999

The categories in Tables 7.1 and 7.2 largely grade the potential for bias by the type of study design. In addition, there are other important biases to consider, including the selection of respondents in trials (for example, biased towards the young, males or less severe cases); the professional and financial allegiances of researchers; and publication bias (towards positive findings). A good systematic review of the literature should examine the potential for such biases. Researchers will present the Project Steering Committee with a summary statement on the level of evidence used for each analysis based on the categories in Table 7.2, but with a qualification noting additional sources of potential bias.

To a large extent, this strength of evidence criterion will be presented quantitatively in the uncertainty analyses. In other words, if the evidence on effectiveness is weak, large uncertainty around the size of the impact measure will be entered into the simulation modelling of uncertainty. However, there may be interventions for which evidence is so limited that the Project Steering Committee would not want to make a firm recommendation to increase funding for the intervention even if the uncertainty analysis shows a favourable cost-effectiveness ratio. In such cases, a recommendation could be made to implement the intervention as a pilot project and to monitor the impact before recommending wider implementation.

### **7.3 Acceptability to stakeholders**

This criterion refers to the anticipated acceptability of proposed interventions to the various stakeholders affected by the intervention (patients, parents and carers, the general community, third-party funders, health service providers, government and the private sector). By its very nature, acceptability is a difficult criterion on which to find empirical data. It necessitates judgements being made by the Project Steering Committee and raises the issue of ensuring adequate stakeholder representation on the Project Steering Committee. The ACE-Obesity working group agreed that this criterion should be expanded to include non-government organisations and that the costs to government should be further broken down to distinguish between local, state and Commonwealth impacts. We will suggest the same to the ACE-Prevention Project Steering Committee.

### **7.4 Feasibility**

This criterion is concerned with the ease of implementing the intervention, considering factors such as the availability of appropriate expertise to implement the intervention on a national scale, the potential size of the financial commitment, and the time scale for implementation.

As with acceptability, feasibility is a criterion that involves judgement rather than technical decision rules. For this criterion, the Project Steering Committee may elect to restrict itself

to flagging issues that require attention and to presenting descriptive information to assist policy makers. Care needs to be taken to minimise the potential for overlap between the acceptability and feasibility filters.

## **7.5 Sustainability**

This criterion refers to the durability of the intervention considering such factors as the level of ongoing funding support required; the community empowerment and capacity building and level of policy support likely to be achieved; and the likelihood of required changes in behaviours, practices and attitudes being achieved on an ongoing basis.

## **7.6 Potential for other consequences**

This criterion refers to both positive and negative side effects arising from an intervention. These might include impacts such as other health consequences (for example, anxiety/depression stemming from stigmatisation); environmental consequences; social capital (for example, from empowered communities or improved social networks); increased household costs; or other economic consequences (for example, impact on industry). Care would need to be taken to ensure that any consequences noted under this filter were not already captured in the cost-effectiveness ratio, either on the cost side (cost impacts on families) or on the outcome side (in the DALY measure).

## **7.7 Other filters**

Second filter criteria are yet to be discussed with the Indigenous Steering Committee. It links in with the discussions on how best to define an Indigenous concept of health benefit. If the proposed elements of cultural security, community health gain and equity are incorporated in an alternative metric of health benefit, there is no need to discuss these again in the context of second filter criteria.

## **7.8 Application of the second stage filters**

The selected criteria will serve as the judgement filter in a two-stage approach to ranking the options. In the first stage, options will be ranked by those criteria directly related to determining the resources consumed or released by the option, together with the size and distribution of the anticipated health gain (based on the ICERs). The first stage will be characterised by aspects that lend themselves to logical decision rules, drawn essentially from the health economics discipline. The second stage incorporates aspects where it is very difficult to develop decision rules and decisions will rest heavily on judgement and due process.

An alternative to this two stage approach is to weight the ICER results and the second stage filters and then combine them into a single index score. However, steering committees in previous studies did not endorse a single index score approach, but favoured discussion of the second filters in a more qualitative manner. The method of reporting the second stage filters will be incorporated in a subsequent update of the protocol, but is likely to be similar to previous ACE studies.

## 8 Discussion and presentation of study results

A basic theme is whether cost-effectiveness ratios are interpreted in a mechanistic or intelligent fashion. Drummond et al. (1997) includes under this notion a concern for the equitable distribution of costs and consequences and issues relating to implementation. In this regard, the whole purpose of placing the cost-effectiveness ratios within the PBMA framework is to encourage such 'intelligent' interpretation. This is given explicit expression through the two-stage approach to ranking the interventions.

Drummond et al. (1997) also raises the issue of whether the results are compared with those of others who have investigated the same question (and whether appropriate allowances are made for differences in methods). Where conventional microeconomic evaluation results are available for the options assessed in ACE-Prevention, these will be reported and compared with our results.

In terms of presenting study results, the following reporting template (Table 8.1) is being used for each of the interventions analysed in the ACE-Prevention project, although it is possible that the template will be revised at a later date. As part of the standardised evaluation methodology, it is important that the results of interventions are reported in a consistent manner. Briefing reports using this template are prepared on each of the selected interventions for presentation to the working group.

Results from ACE-Prevention will be submitted as a series of peer-reviewed articles to appropriate academic journals. Normal authorship criteria will be applied to journal articles with acknowledgement of the working group members.

**Table 8.1 Reporting template for interventions analysed in ACE–Prevention**

**1. The intervention**

**Background** - Short justification for the inclusion of this intervention for analysis.

**Description** - A description of the intervention (and, if necessary, different components) and the comparator (current practice or a null option).

**2. The health benefit**

Evidence of impact for the intervention from the literature

Likely effectiveness of the intervention under routine health service conditions in Australia

Description of current practice

Description how benefits are modelled as the difference in outcomes (in DALYs) in the target population given the intervention of interest or the comparator intervention ('current practice' or the 'partial null' when analysing the optimal mix of interventions)

Table of sources of information and assumptions

**3 The health service costs**

Inclusion/exclusion criteria for costing

Sources of cost data

Cost of current practice

Cost offsets

Table of interventions costs, data sources and assumptions

**4 Uncertainty analyses**

present a table of the assumptions fed into the uncertainty simulation modelling; major assumptions driving uncertainty

**5 The incremental cost-effectiveness results**

Summarise benefits, costs and cost-effectiveness ratios in a table

If appropriate, present marginal cost-effectiveness ratios (comparing different intensities of the same intervention)

Present average cost-effectiveness ratios for individual interventions and incremental cost-effectiveness ratios for the most cost-effective mix of intervention addressing the same health problem

**6 The second stage filter analysis**

Describe how each of the second filter criteria apply to the intervention (this will largely depend on judgement by the Project Steering Committee and Technical Advisory Panels) but, where applicable, include evidence to support these judgements (for example, describe health inequalities and inequalities in service utilisation under the criterion of equity). The second-stage filter analysis is likely to be summarised in tabular form

**7 Discussion of results**

Compare results with other economic studies

Final conclusions and recommendations



## 9 References

(not all text references are listed yet)

- Campbell, K, Waters, E, O'Meara, S, Kelly, S & Summerbell, C 2002, 'Interventions for preventing obesity in children', *Cochrane Database Syst Rev* 2002(2):CD001871.
- Carter, R 2001, 'The Macro Economic Evaluation Model (MEEM): An approach to priority setting in the health sector' [PhD thesis]. In Faculty of Business and Economics. Monash University: Melbourne.
- Carter, R, Stone, C, Vos, T, Hocking, J, Mihalopoulos, C, Peacock, S & Crowley, S 2000, *Trial of Program Budgeting and Marginal Analysis (PBMA) to assist cancer control planning in Australia*. Commonwealth Department of Health and Aged Care: Canberra.
- Cole, TJ, Bellizzi, MC, Flegal, KM, Dietz, WH 2000, 'Establishing a standard definition for child overweight and obesity worldwide: international survey', *BMJ* 320 (7244): 1240–3.
- Deeks, JJ, Altman, DG & Bradburn, MJ 2001, 'Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis' in *Systematic reviews in health care: meta-analysis in context* (eds. Egger, M., Smith, G. D. and Altman, D. G.), pp. 285–312. BMJ Publishing Group: London.
- Drummond, MF, O'Brien, B, Stoddart, GL & Torrance, GW 1997, *Methods for the economic evaluation of health care programmes*. Oxford University Press: Oxford.
- Gold, MR, Siegel, JE, Russell, LB & Weinstein, MC (eds.) 1996, *Cost-effectiveness in health and medicine*. Oxford University Press: New York.
- Ham, C 1998, *Setting priorities for health care: Why government should take the lead*. Belfast, Northern Ireland Economic Development Office.
- Hunter, D 1996, 'Rationing and evidence-based medicine', *Journal of Evaluation in Clinical Practice* 1(4):134–136.
- Jacobs P & Fassbender, K 1998, 'The measurement of indirect costs in the health economics evaluation literature. A review', *International Journal of Technology Assessment in Health Care* 14(4): 799–808.
- Klein, R 1993, 'Dimensions of rationing: who should do what?', *British Medical Journal* 307:309–311.
- Klein, R 1995, 'Priorities and rationing: pragmatism or principles?', *British Medical Journal* 311:761–762.
- Mechanic, D 1995, 'Dilemmas in rationing in health care services: the case for implicit rationing', *British Medical Journal* 310:1655–1659.
- Muir Gray, JA, Hayner, RD, Sackett, DL, Cool, DJ & Guyat, GH 1997, 'Transferring evidence from research into practice:3', *Developing evidence-based clinical policy*. ACPJ Club 126(2):A14–16.
- Nord, E 1999, *Cost-value analysis in healthcare: making sense out of QALYs*, Cambridge University Press, Cambridge.
- National Obesity Taskforce 2003, *Healthy weight 2008 – Australia's Future – The National Action Agenda for Children and young people and their families*. Canberra: Department of Health and Ageing.
- Swinburn, B & Gill, T 2002. 'Best investments' to address childhood obesity. A scoping exercise. Final report. Deakin University and NSW Centre for Public Health Nutrition.
- Tan-Torres Edejer, R, Baltussen, R, Adam, T, Hutubessy, R, Acharya, A, Evans, DB & Murray, CJL (eds.) 2003, *Making choices in health: WHO guide to cost-effectiveness analysis*. World Health Organization, Geneva.

van Hout, BA, Al, MJ, Gordon, GS & Rutten, FF 1994, 'Costs, effects and C/E-ratios alongside a clinical trial', *Health Economics* 3, 309–319.

Vos, T & Begg, S 1999, *The Victorian Burden of Disease Study: morbidity*, Public Health and Development Division, Department of Human Services: Melbourne.

World Health Organisation 1995, *Physical status: the use and interpretation of anthropometry*, Geneva, WHO.