



Australian Government

**Australian Institute of
Health and Welfare**

Assessment of Global Burden of Disease 2010 methods for the Australian context

Australian Burden of Disease Study Working paper No. 1



Australian Government

**Australian Institute of
Health and Welfare**

*Authoritative information and statistics
to promote better health and wellbeing*

Assessment of Global Burden of Disease 2010 methods for the Australian context

**Australian Burden of Disease Study
Working paper no. 1**

February 2014

Australian Institute of Health and Welfare
Canberra

Contents

Acknowledgments.....	v
Abbreviations.....	vi
Summary	vii
1 Introduction.....	1
Background	1
Key considerations	6
Appendix 1A.....	11
Appendix 1B	12
Appendix 1C.....	13
2 Overview of methods	14
Outline of inputs and key decisions	15
Key issues in the Indigenous context	21
3 Cause and risk factor lists.....	22
Disease and injury causes	23
Risk factors	30
Intended approach.....	35
Appendix 3A.....	37
4 Mortality – Years of life lost.....	38
What does it measure and why is it important?	38
Overview of GBD 2010 data sources and methods	38
Methodological differences from the 2003 Australian and Indigenous studies.....	42
Applicability of GBD 2010 methods to the Australian and Indigenous contexts	44
Intended approach.....	52
Appendix 4A.....	53
5 Morbidity – Years lived with disability.....	55
What does it measure and why is it important?	55
Overview of GBD 2010 methods.....	56
Methodological differences from the 2003 Australian and Indigenous studies.....	63
Applicability of GBD 2010 methods to the Australian and Indigenous contexts	65
Intended approach.....	67
6 Burden of risk factors	68
Overview of GBD 2010 methods.....	69

Methodological differences from the 2003 Australian and Indigenous studies.....	72
Applicability of GBD 2010 methods to the Australian context.....	75
Intended approach.....	77
Appendix 6A.....	79
7 Data sources.....	84
Why are data sources important to BoD estimates?.....	84
GBD 2010 data sources	84
Methodological differences from the 2003 Australian and Indigenous studies.....	88
Applicability of the GBD 2010 data sources to the Australian context.....	89
Likely data improvements to Australian estimates	91
Intended approach.....	101
Appendix 7A: Data sources used in Australian 2003 BoD study.....	103
Appendix 7B: Data sources used in 2003 Indigenous BoD study	105
Appendix 7C: Population estimates	106
8 Logistics and implementation	107
Overview of GBD 2010 logistics.....	107
Differences from the 2003 Australian and Indigenous studies.....	109
Applicability of GBD methods to the Australian context.....	110
Design of the new Australian study	111
9 Intended approach	114
Methodological approach	114
Operational considerations.....	117
Appendix A New Zealand Burden of Disease study.....	118
Methodological overview	118
Comparing NZBD and GBD 2010.....	122
Strengths and weaknesses of the New Zealand approach.....	124
Applicability of New Zealand methods to the Australian context	124
Glossary.....	125
References	127
List of tables	129
List of figures	130

Acknowledgments

This working paper was prepared by the staff in the Health Group and the Indigenous and Children's Group of the AIHW, led by Dr Lynelle Moon and Ms Michelle Gourley. Staff members contributing to the report were Karen Bishop, Ilona Brockway, Tracy Dixon, Melissa Goodwin, Nick Mann, Michael McGrath and Jessica Zhang. Technical advice was provided by Dr Ken Tallis, engaged as a consultant to the Burden of Disease project.

The project was carried out under the guidance of Ms Lisa McGlynn (Head, Health Group), Dr Fadwa Al-Yaman (Head, Indigenous and Children's Group) and Ms Teresa Dickinson (then Head, Statistics and Communication Group). Funding for the project was provided by the Department of Health and the Australian National Preventive Health Agency.

Abbreviations

ABS	Australian Bureau of Statistics
AHS	Australian Health Survey
AIHW	Australian Institute of Health and Welfare
ANPHA	Australian National Preventive Health Agency
AATSIHS	Australian Aboriginal and Torres Strait Islander Health Survey
BoD	Burden of disease studies
COAG	Council of Australian Governments
CoD	Cause of Death
DALY	Disability-adjusted life year
DW	Disability weight
EAG	Expert Advisory Group
ERP	Estimated Resident Population
GBD	Global Burden of Disease
GHDx	Global Health Data Exchange
HALE	Health-adjusted life expectancy
ICD	International Classification of Diseases
IHME	Institute for Health Metrics and Evaluation
NZBD	New Zealand Burden of Diseases, Injuries and Risk Factors Study, 2006–2016
PAF	Population attributable fraction
RR	Relative risk
TMRED	Theoretical minimum risk exposure distribution
YLD	Years lived with disability
YLL	Years of life lost
WHO	World Health Organization

Summary

The Australian Institute of Health and Welfare (AIHW) has been funded by the Department of Health and the Australian National Preventive Health Agency to, first, review the Global Burden of Disease (GBD) methodology, and determine the best methodology for producing Australian Burden of Disease (BoD) estimates and estimates relating to Aboriginal and Torres Strait Islander people. To provide transparency of methods and decision-making processes, the AIHW intends to release a series of working papers over the course of the project. This first working paper documents various aspects of the latest Global Burden of Disease (GBD 2010) methods and an assessment of the applicability of these methods for estimating the burden of disease in Australia. As a working paper, the approach outlined in this document is preliminary and reflects the project status as at February 2014. As such, it may change as the project progresses. Any changes will be reflected in future publications in the working paper series.

This working paper describes the methods used in GBD 2010, and compares these to methods used in previous Australian burden of disease studies. It then assesses the applicability of the GBD 2010 methods to the current Australian and Indigenous contexts, and identifies the key decisions that need to be made for each step in the burden of disease estimation process. It then includes a methodological approach for the national and Indigenous components of the Australian study which has been discussed and refined with the project's Expert Advisory Group and the Indigenous Reference Group.

Overarching considerations for the Australian Burden of Disease Study that are referred to throughout this report include requirements for country-level analysis compared to global analysis; level of complexity and need for transparency; the need to produce Indigenous estimates; the need to produce sub-national estimates where valid; and the ability to update the work on a regular and ongoing basis.

Some of the more specific issues considered in this report include whether:

- to adopt the GBD 2010 cause and risk factor lists as is, or make minor adjustments to better suit the Australian context
- to adopt GBD 2010's approach of no discounting for time, and no age-weighting, or whether there is a need or capacity to build in different options for these in the analysis
- to produce uncertainty intervals for national and Indigenous Burden of disease estimates
- to use a modelling approach in estimating years of life lost for the Indigenous and sub-national populations with incomplete or small numbers of death registrations
- GBD 2010's redistribution of 'garbage codes' for mortality is appropriate for Australian mortality data
- to adopt GBD 2010's disability weights and method for co-morbidity adjustments
- to use the approach taken in GBD 2010 to use all available data sources, or to use the approach taken in previous Australian BoD studies using only the single best quality data source for each disease/risk factor. A related issue is how to take account of varying data quality in the data sources used.
- to use DisMod-MR as the software tool for estimating YLDs.

The national and Indigenous 'decision trees' summarise the key steps in the burden of disease estimation process, the main decisions that needed to be made regarding the

methodology to be used in the Australian study, and the key issues affecting these decisions for those areas of the methodology. These decision trees were developed as part of the assessment of GBD methods for the Australian study. The detail behind each part of the diagram is contained in the relevant chapter of this report, and the decision tree needs to be read in conjunction with that detail.

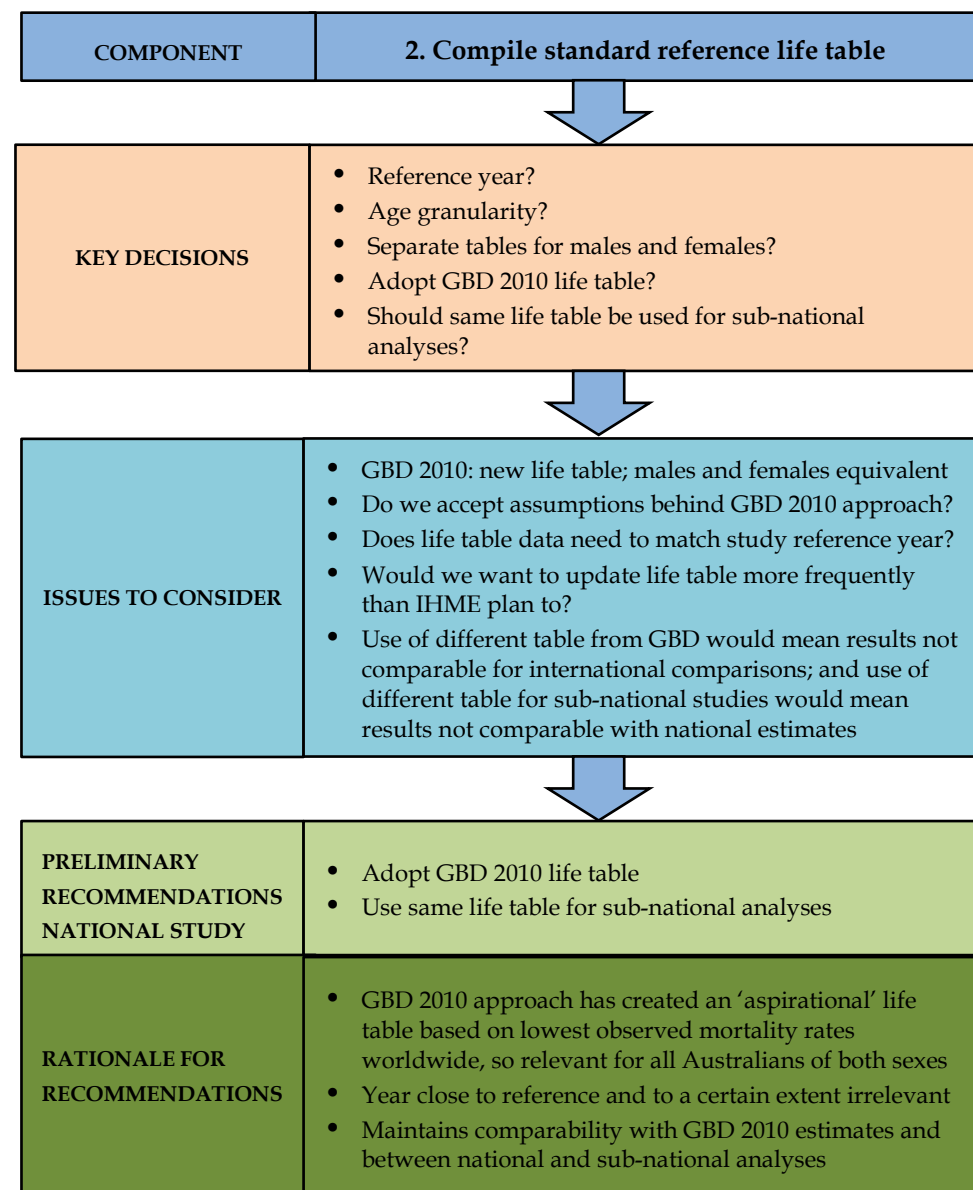
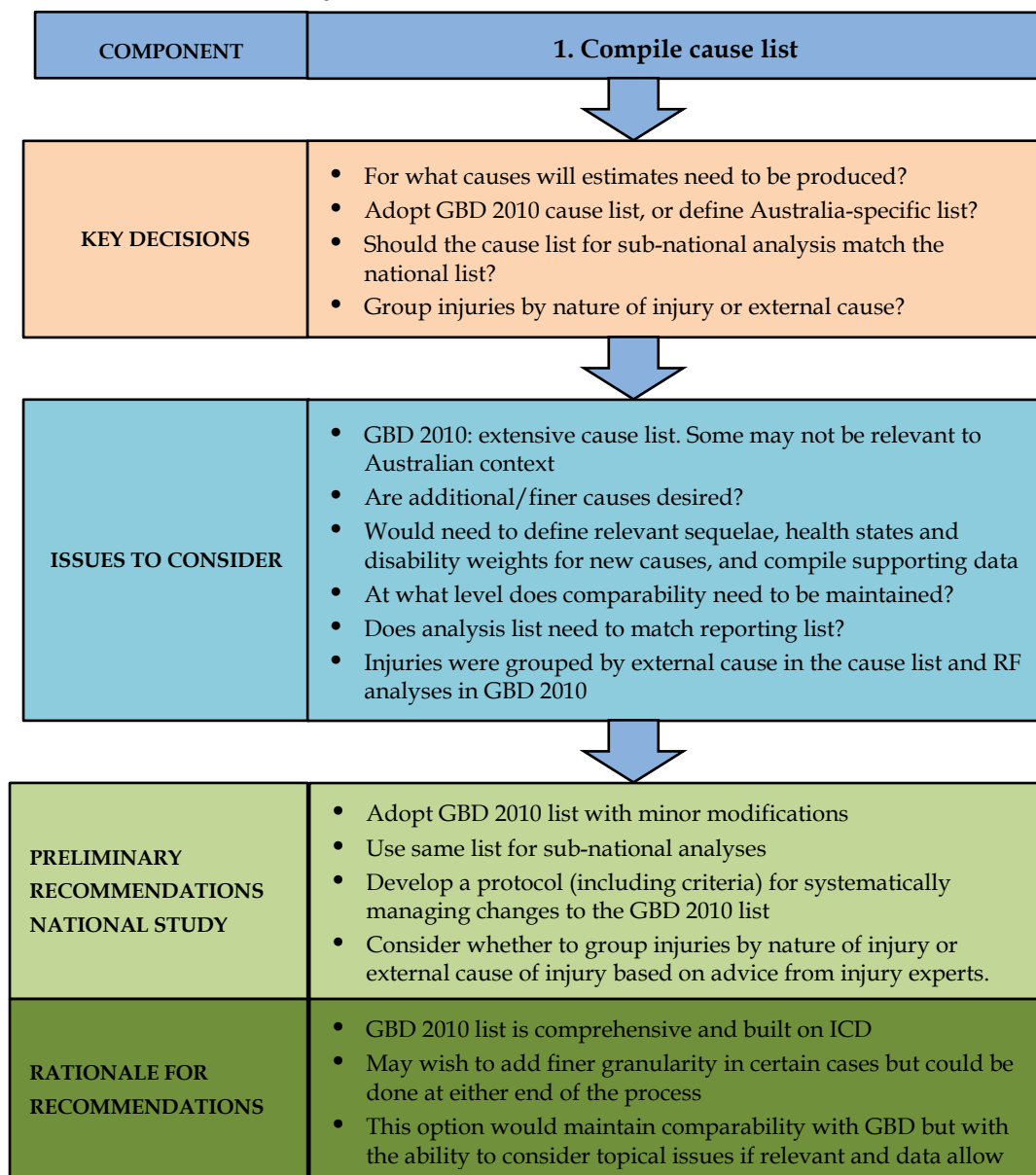
Components of the decision tree are covered in the following chapters:

- Chapter 3: Components 1 and 10
- Chapter 4: Components 2–4
- Chapter 5: Components 5–8
- Chapter 6: Components 11–12
- Chapter 7: Data sources

While the majority of this report is devoted to assessing methodological issues, there are also a number of operational and logistical issues related to applying the GBD study in the Australian context. These are outlined in Chapter 8. A full outline of our intended approach, both methodological and operational, is contained in Chapter 9. This is an amalgamation of the approaches outlined in preceding chapters with duplications removed.

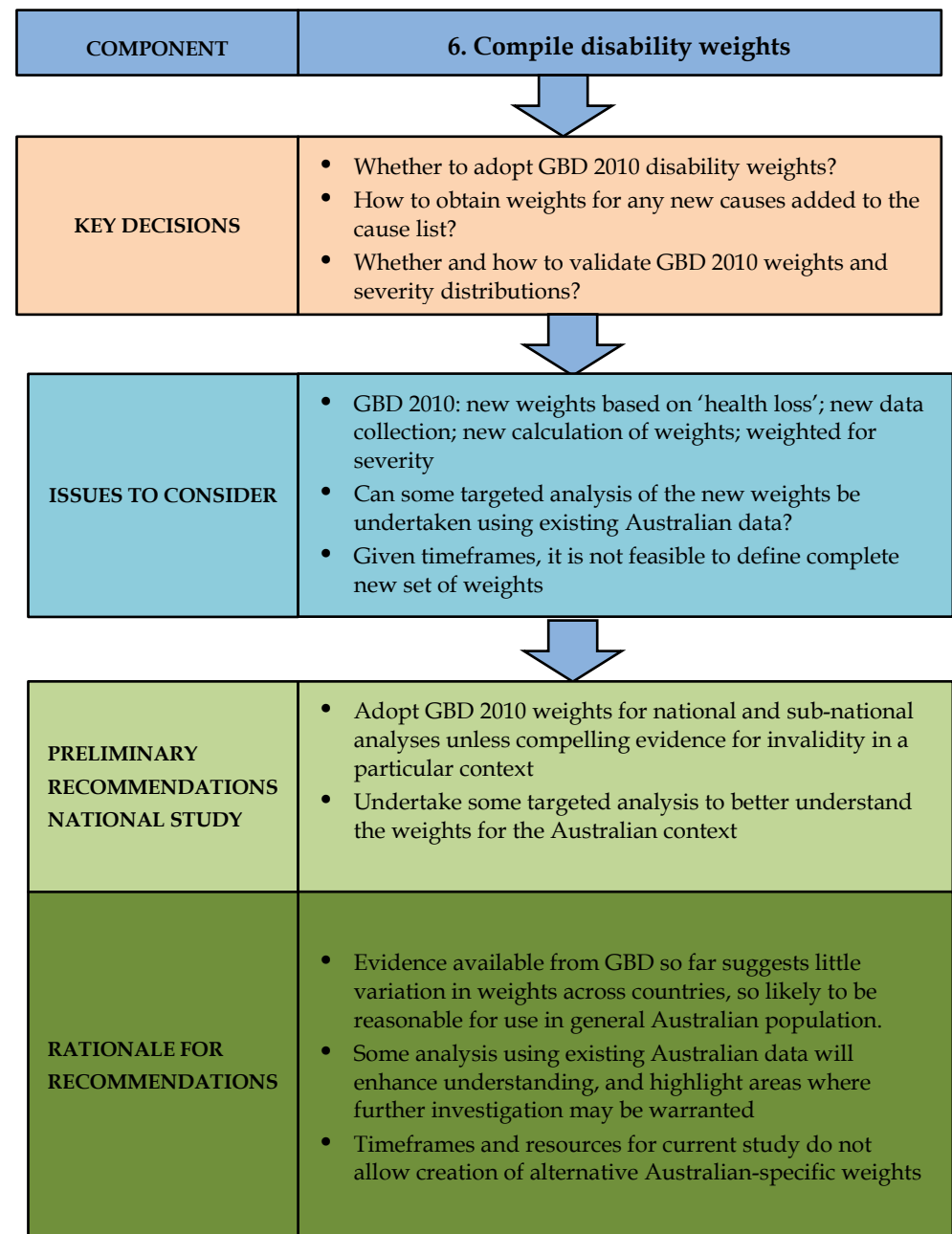
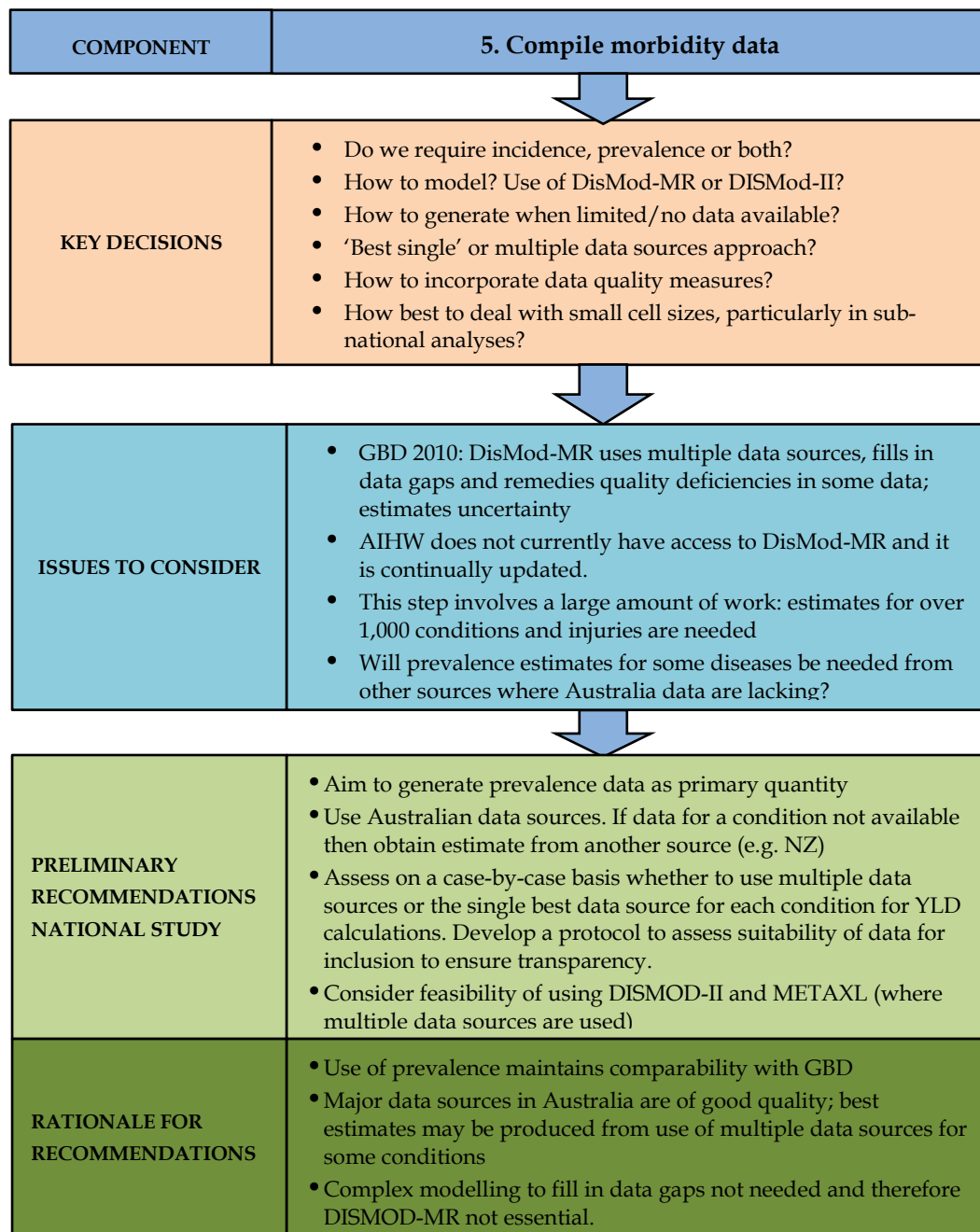
The forthcoming second phase of the work will update burden of disease data for Australia and estimates for Aboriginal and Torres Strait Islander people. This will include the provision of sub-national and population group data where valid and where capacity allows. The outputs will also include a number of publications of technical and analytical content which are expected to be finalised in 2015.

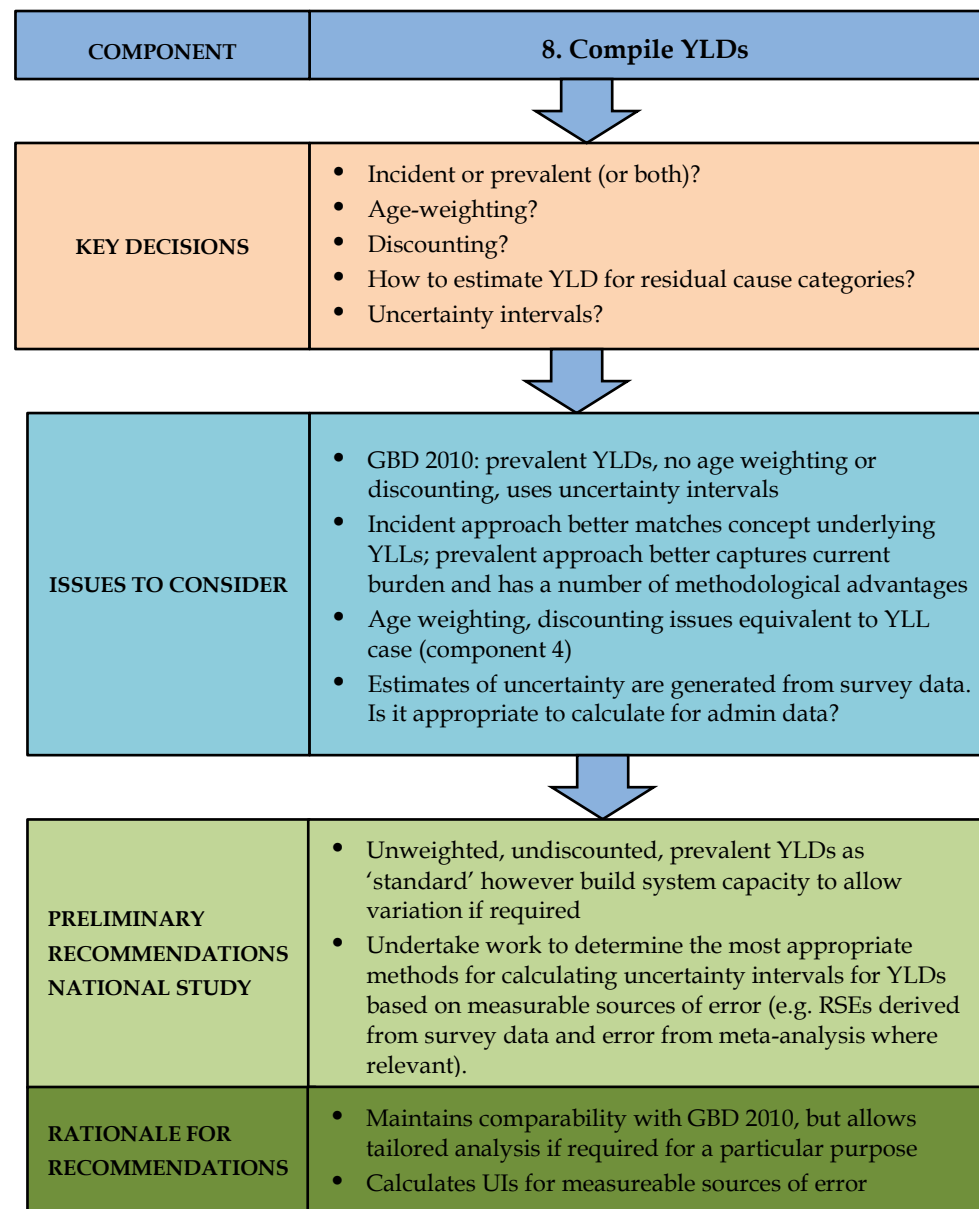
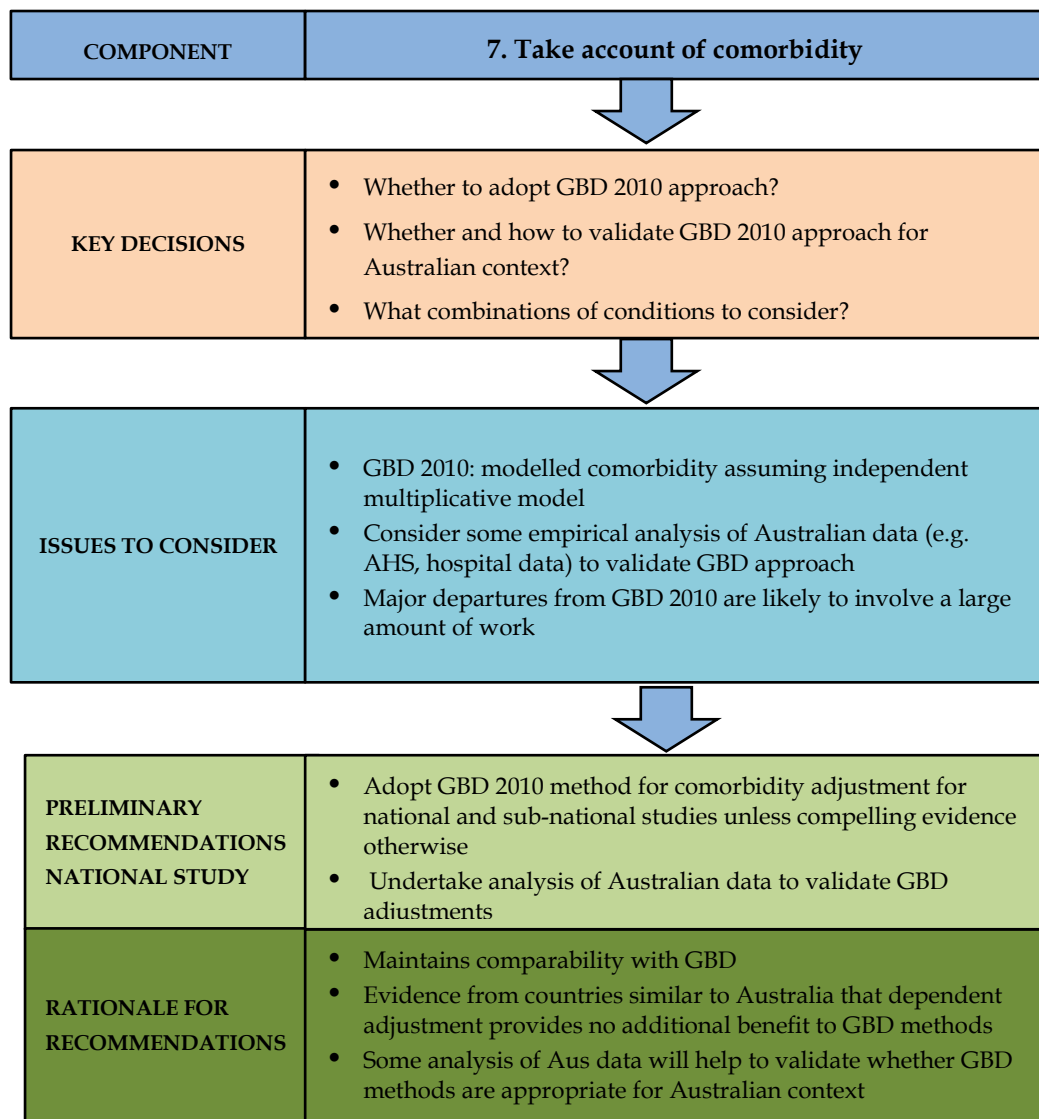
National study decision tree

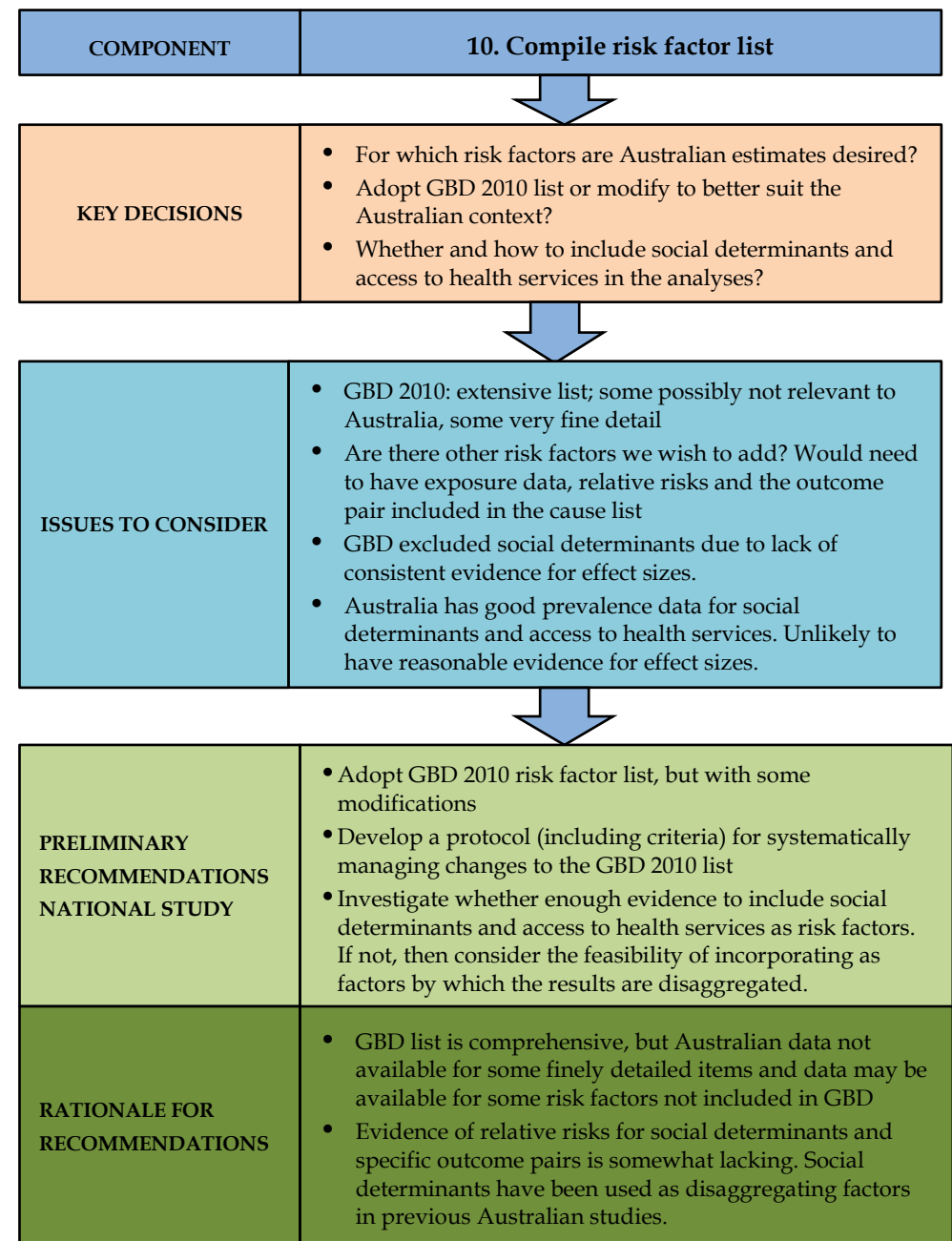
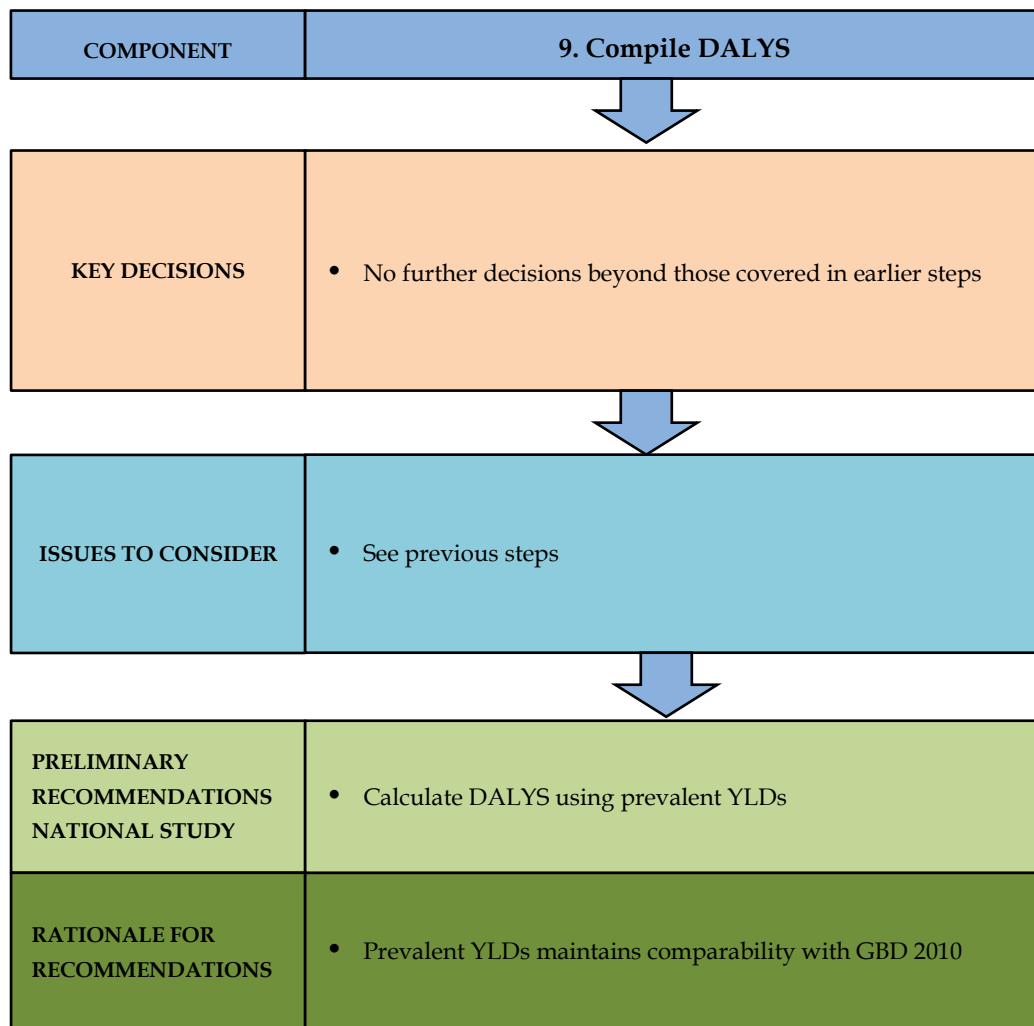


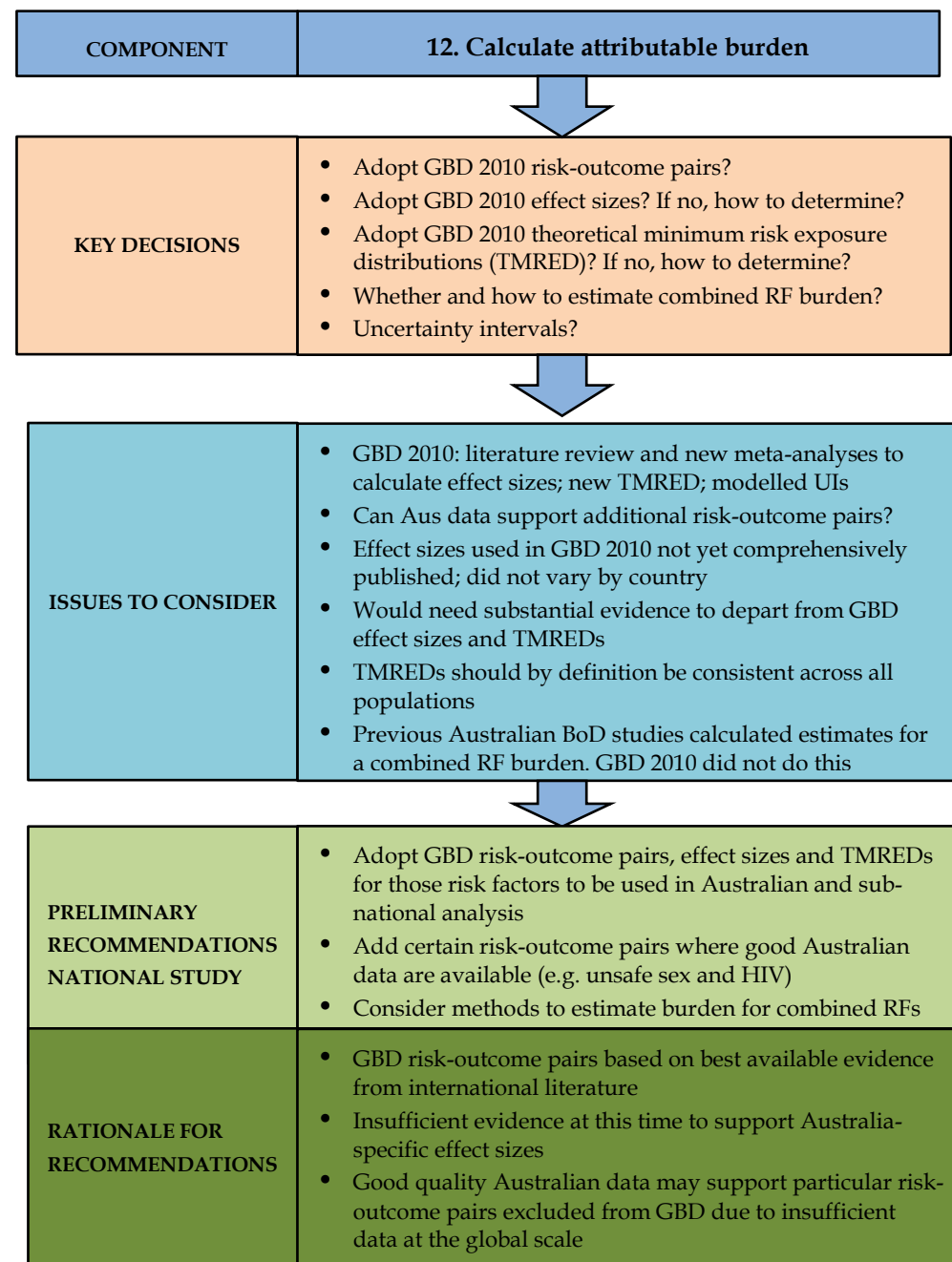
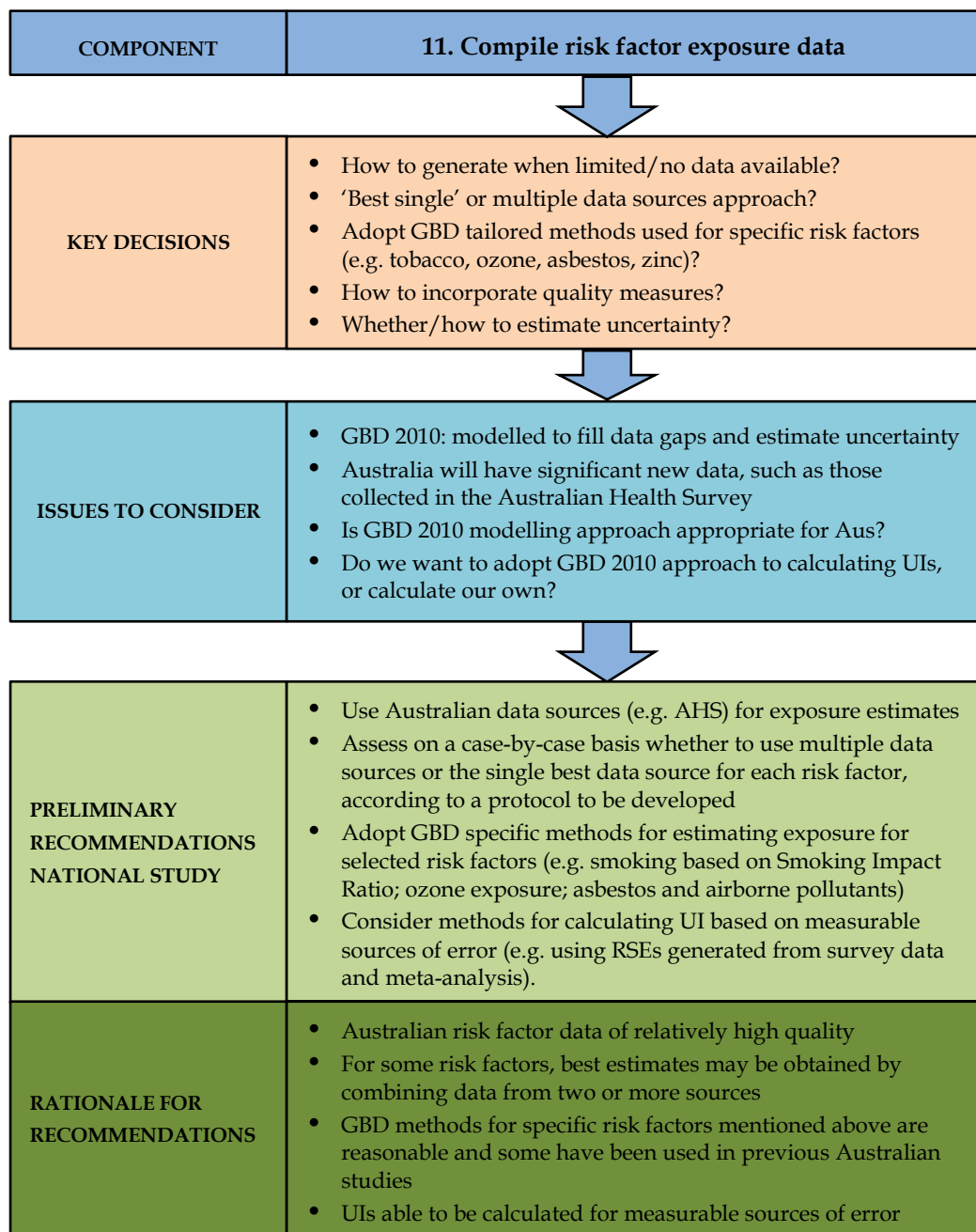
COMPONENT	3. Compile numbers of deaths
KEY DECISIONS	- Age granularity? - Year of death or year of registration (and for what year/s)? - Use multiple years of data to overcome small numbers? - Is modelling required? Of what type? In what circumstances? - How to redistribute 'garbage codes'? Use GBD method for this?
ISSUES TO CONSIDER	- GBD 2010: modelled to fill in data gaps and remedy quality deficiencies in some data - Given quality of Australian mortality data, do we need to model? If modelling not undertaken, can multiple years of data be used to increase cell sizes? - Deaths can be counted by year of occurrence of death (YoD) or year of registration (YoR). 5% of deaths in Australia are registered late. - Aus mortality data are subject to a revision process (prelim, revised and final versions are available).
PRELIMINARY RECOMMENDATIONS NATIONAL STUDY	- Use AIHW National Mortality Database as data source - Use GBD 2010 age groups as minimum, with extension to finer age groups where data supports - Investigate differences between deaths counted by YoD and YoR, and different versions of mortality data available, to inform decision on how to count deaths and for what year/s - Use GBD 2010 garbage code redistribution where appropriate. Undertake analysis of Aus data sources to inform modifications - Combine multiple years of data where numbers are small. Simple modelling strategies may be used where numbers are still too small
RATIONALE FOR RECOMMENDATIONS	- Age-groups used should be based on what the data are able to produce robust estimates for. At minimum – consistent with GBD. - Australian mortality data are of high quality - Use of multiple years data overcomes small number issues likely in sub-national analyses - GBD garbage code redistribution may be further refined using local mortality, cancer registry and other data.

COMPONENT	4. Compile YLLs
KEY DECISIONS	- Age-weighting? - Discounting? - Uncertainty intervals? If yes, what method?
ISSUES TO CONSIDER	- GBD 2010: no age-weighting or discounting; includes UIs - Do we agree with no age-weighting/no discounting approach of GBD? Does this meet needs of Australian users? Can we build in capacity to apply these if required? - Do we adopt GBD 2010 approach to calculating UIs, or calculate using our own method?
PRELIMINARY RECOMMENDATIONS NATIONAL STUDY	- No age-weighting or discounting as 'standard' - Build analysis system with capacity to allow age-weighting or discounting if desired - Undertake work to determine the most appropriate methods for calculating UIs based on measurable sources of error (e.g. error from pooling multiple years of data).
RATIONALE FOR RECOMMENDATIONS	- Maintains comparability with GBD, but allows for tailored analysis if desired for a particular purpose (e.g. discounted data for use in health expenditure analysis) - Point estimates sourced from admin data in theory are not considered to have uncertainty. A degree of uncertainty occurs when combining multiple years of data as a smoothing technique.

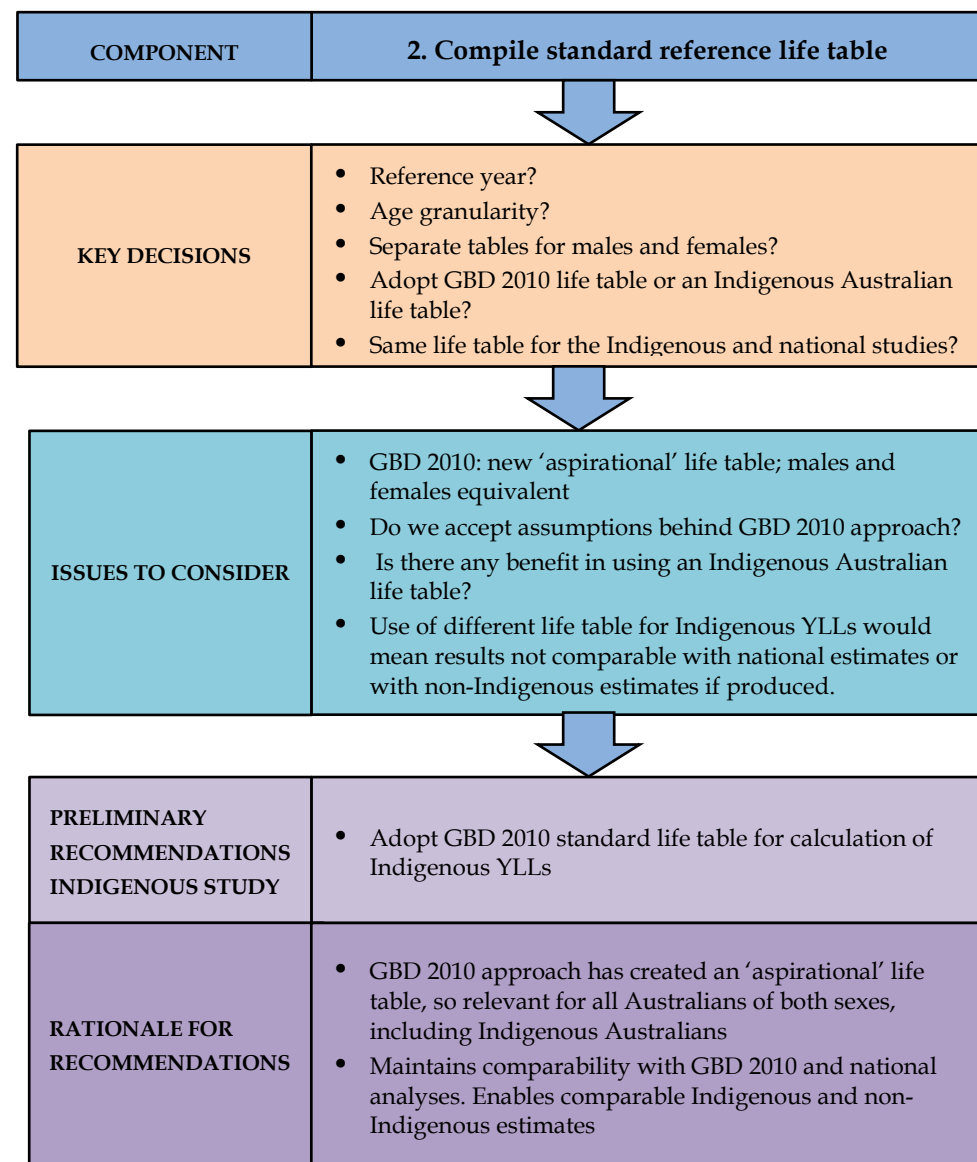
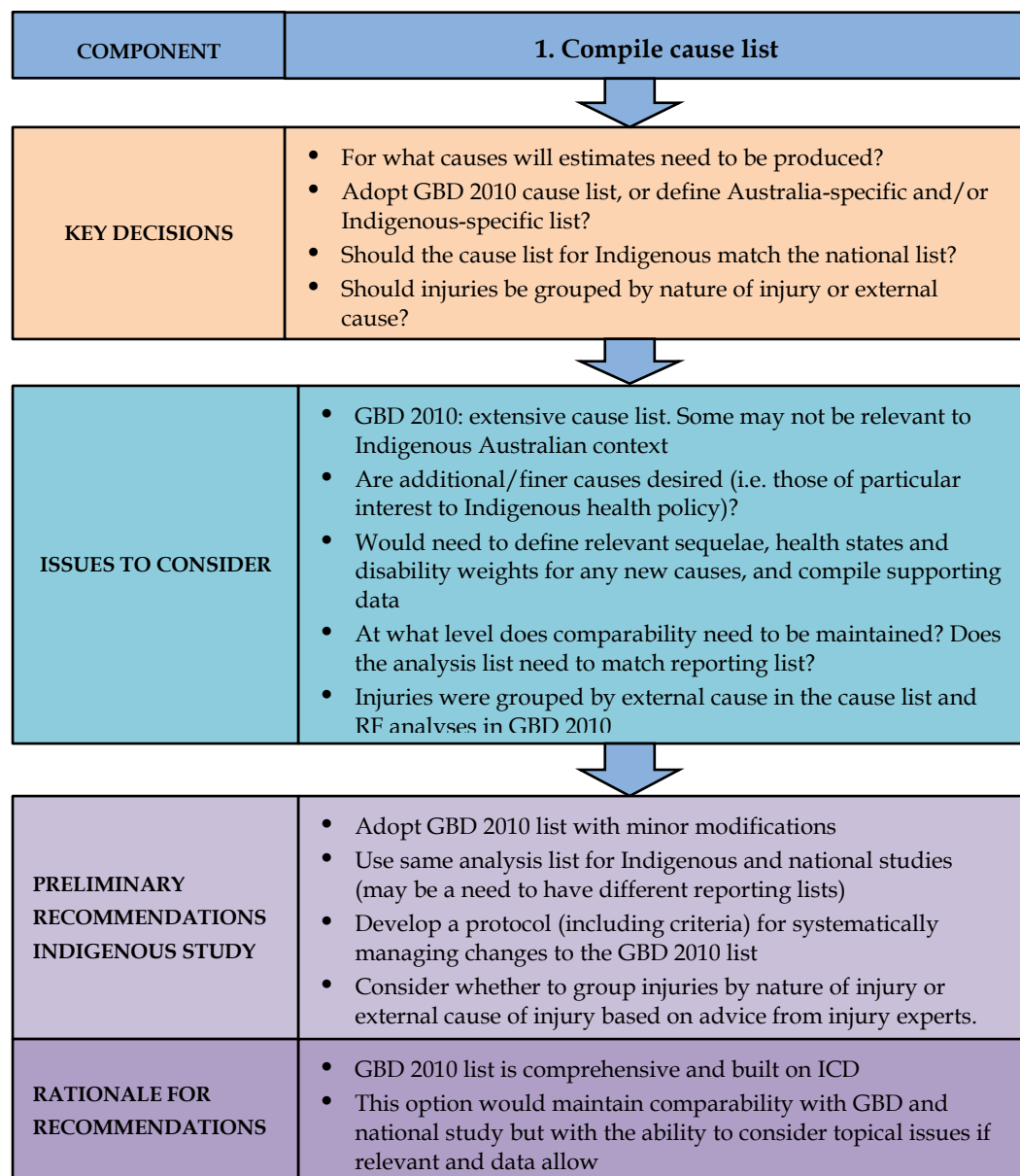








Indigenous study decision tree



COMPONENT	3. Compile numbers of deaths
KEY DECISIONS	- Age granularity? - Year of death or year of registration (and for what year/s)? - Use multiple years of data to overcome small numbers? - Is modelling required? Of what type? In what circumstances? - How to redistribute 'garbage codes'? Use GBD method for this? - How and at what level to adjust for under-identification?
ISSUES TO CONSIDER	- GBD 2010: modelled to fill in data gaps and remedy quality deficiencies in some data - Do we need to model for Indigenous estimates? - Can multiple years of data be used to increase cell sizes? - Around 15% of Indigenous deaths are subject to late registration - Do 'garbage codes' differ for Indigenous and non-Indigenous? - Indigenous under-identification adjustment factors currently available at state/territory or age/sex level. Does under-identification vary by cause of death?
PRELIMINARY RECOMMENDATIONS INDIGENOUS STUDY	- Use AIHW National Mortality database as data source - Age groups used should be based on what the data are able to support - Investigate differences between deaths counted by year of occurrence and year of registration and different versions of mortality data to inform decision on how to count deaths and for what year/s - Combine multiple years of data to overcome small number issues. Consider simple modelling strategy where numbers are still small - Adjust for under-identification at lowest level possible using available adjustment factors (e.g. ABS/AIHW data linkage studies) - Use GBD 2010 garbage code redistribution where appropriate. Undertake analysis of Australian data sources and look at differences by Indigenous status to inform modifications
RATIONALE FOR RECOMMENDATIONS	- Combining multiple years of data and using simple modelling strategies will overcome small number issues and increase stability of the estimates - Adjustments for under-identification will enable more accurate measurement of the gap

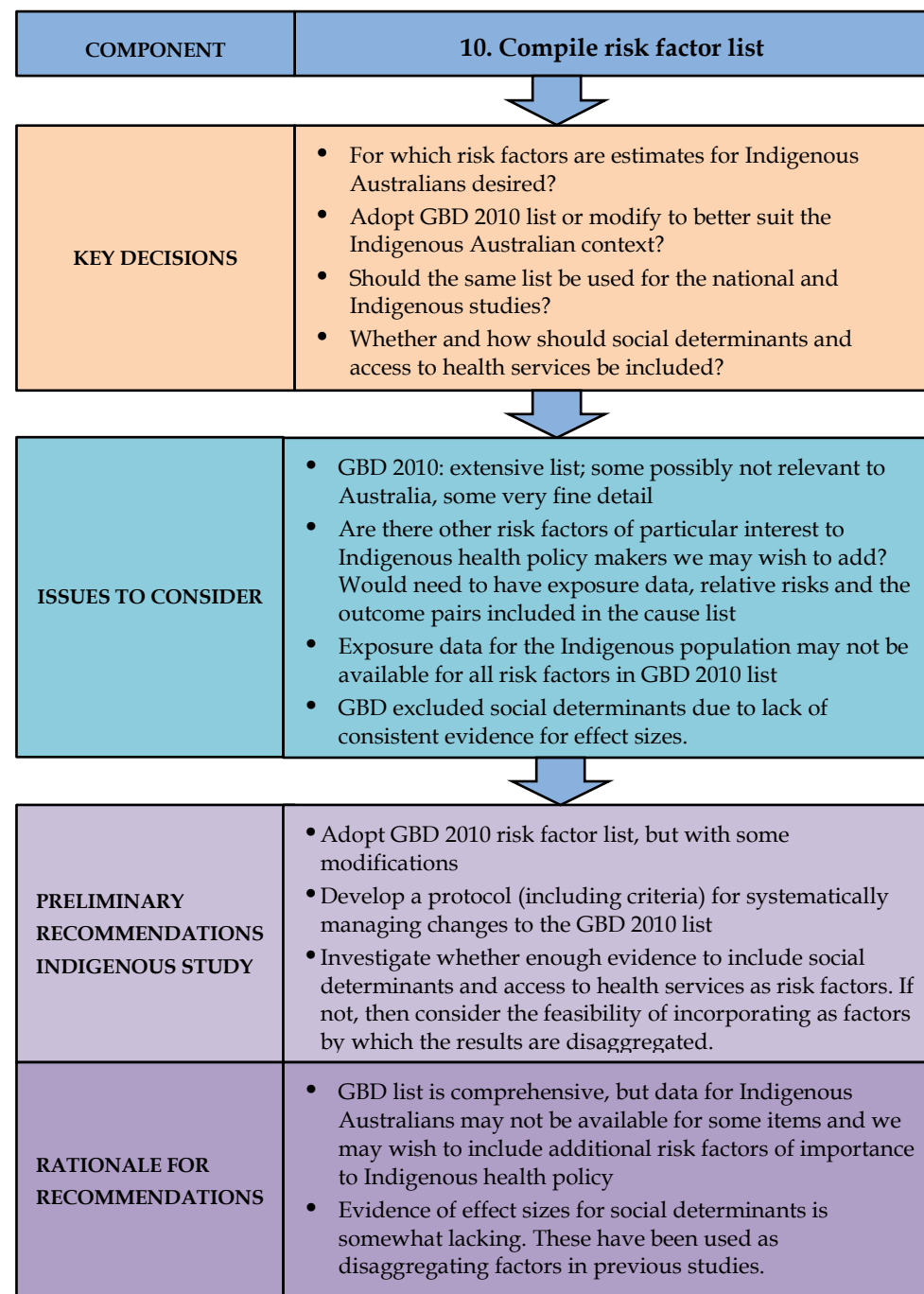
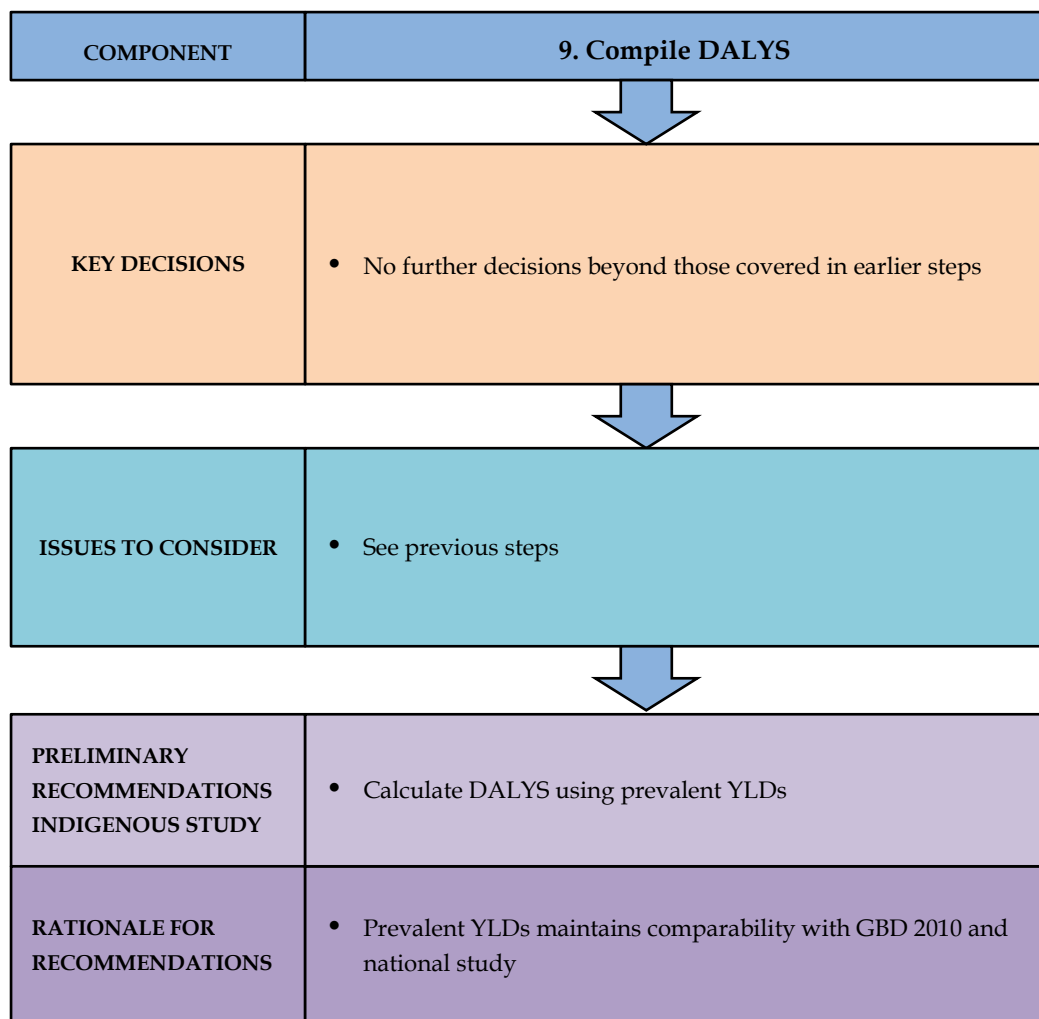
COMPONENT	4. Compile YLLs
KEY DECISIONS	- Age-weighting? - Discounting? - Uncertainty intervals? If yes, what method?
ISSUES TO CONSIDER	- GBD 2010: no age-weighting or discounting; includes UIs - Do we agree with no age-weighting/no discounting approach of GBD? Does this meet needs of Australian users? Can we build in capacity to apply these if required? - What are the sources of uncertainty that can be measured? - Do we adopt GBD 2010 approach to calculating UIs, or calculate using our own method?
PRELIMINARY RECOMMENDATIONS INDIGENOUS STUDY	- No age-weighting or discounting as 'standard' - Build analysis system with capacity to allow age-weighting or discounting if desired - Undertake work to determine the most appropriate methods for calculating UIs based on measurable sources of error (e.g. error from pooling multiple years of data and CIs for under-identification adjustment factors used).
RATIONALE FOR RECOMMENDATIONS	- Maintains comparability with GBD and national study, but allows for tailored analysis if desired for a particular purpose - UIs can be calculated for measurable sources of error. CIs may be available around under-identification adjustment factors from ABS and AIHW data linkage studies which should be incorporated

COMPONENT	5. Compile morbidity data
KEY DECISIONS	- Do we require incidence, prevalence or both? - How to model? Use of DisMod-MR or DISMOD-II? - How to generate prevalence when limited/no data available for the Indigenous population? 'Best single' or multiple data sources approach? - How best to deal with small cell sizes? - Adjust for under-identification? For what data sources and at what level?
ISSUES TO CONSIDER	- GBD 2010: DisMod-MR uses multiple data sources - AIHW does not currently have access to DisMod-MR and it is continually updated. - Simple Modelling techniques could be used to generate estimates when no Indigenous data available (e.g apply Indigenous: non-Indigenous mortality ratios; Indigenous: non-Indigenous ratios for higher level conditions) - Under-identification factors are available for hospital data at the state/territory and remoteness levels
PRELIMINARY RECOMMENDATIONS INDIGENOUS STUDY	- Aim to generate prevalence data as primary quantity - Assess on a case by case basis whether to use multiple data sources or the single best data source for each condition. Develop a protocol to assess suitability of data for inclusion. - Consider feasibility of using DISMOD-II and METAXL (where multiple data sources are used) - Consider adopting a modelling strategy to derive prevalence estimates where small numbers or no Indigenous data available - Adjust for under-identification in hospital data using adjustment factors from recent AIHW study and consider if feasible to adjust for other admin data sets.
RATIONALE FOR RECOMMENDATIONS	- Use of prevalence maintains comparability with GBD - Best estimates for some conditions may be obtained by combining data from two or more sources - Simple modelling techniques overcome small numbers and lack of data. Complex modelling (ie DISMOD-MR) to fill in data gaps not necessary.

COMPONENT	6. Compile disability weights
KEY DECISIONS	- Is GBD 2010 'health loss' concept and calculation of weights appropriate for the Indigenous population? - Whether to adopt GBD 2010 disability weights? - How to obtain weights for any new causes added to the cause list? - Whether and how to validate GBD 2010 weights and severity distributions?
ISSUES TO CONSIDER	- GBD 2010: new weights based on 'health loss'; new data collection; new calculation of weights; weighted for severity - Indigenous view of health is broader and different in concept to 'health loss' which may affect their valuation of health states for certain conditions (e.g. mental health). - Can some targeted analyses of the new weights be undertaken for the Indigenous pop using existing Aust. data?
PRELIMINARY RECOMMENDATIONS INDIGENOUSY	- Adopt GBD 2010 weights unless compelling evidence for invalidity in a particular context - Undertake some targeted analyses using Australian data to enhance understanding of the GBD 2010 weights
RATIONALE FOR RECOMMENDATIONS	- Evidence available from GBD so far suggests little cultural variation in weights. - Analyses using existing Indigenous data will enhance understanding of GBD 2010 weights , and determine whether different severity distributions could be used to account for any differences - Timeframes and resources for current study do not allow creation of a complete set of Indigenous-specific weights

COMPONENT	7. Take account of comorbidity
KEY DECISIONS	• Adopt GBD 2010 approach for comorbidity adjustment? • Whether, and how to, validate GBD 2010 methods for suitability for Indigenous Australian context? • What combinations of conditions to consider? • Should method be consistent across national and Indigenous studies?
ISSUES TO CONSIDER	• GBD 2010: modelled comorbidity assuming independent multiplicative model • Indigenous Australians have a different pattern of co-morbidity to non-Indigenous (eg more conditions, more severe) • Consider some empirical analysis of Australian data to validate GBD and also look at differences in co-morbidities between Indigenous and non-Indigenous • Major departures from GBD 2010 are likely to involve a large amount of work
PRELIMINARY RECOMMENDATIONS INDIGENOUS STUDY	• Adopt GBD 2010 method for comorbidity adjustment unless compelling evidence otherwise • Undertake analysis of Australian data (hospital, AATSIHS) to validate GBD methods and look at differences in co-morbidities between Indigenous and non-Indigenous
RATIONALE FOR RECOMMENDATIONS	• Maintains comparability with GBD and national study • Evidence from countries similar to Australia that dependent adjustment provides no additional benefit • Some analysis of Australian data will help to validate whether GBD methods are appropriate for Indigenous Australian

COMPONENT	8. Compile YLDs
KEY DECISIONS	• Incident or prevalent (or both)? • Age-weighting? • Discounting? • Uncertainty intervals?
ISSUES TO CONSIDER	• GBD 2010: prevalent YLDs, no age weighting or discounting, uses uncertainty intervals • Incident approach better matches concept underlying YLLs; prevalent approach better captures current burden and has a number of methodological advantages • Age weighting, discounting issues equivalent to YLL case (component 4) • Estimates of uncertainty are generated from survey data. Is it appropriate to calculate for admin data? What about error around adjustments for Indigenous under-identification used?
PRELIMINARY RECOMMENDATIONS INDIGENOUS STUDY	• Unweighted, undiscounted, prevalent YLDs as 'standard' however build system capacity to allow variation if required • Undertake work to determine the most appropriate methods for calculating uncertainty intervals for YLDs based on measurable sources of error (e.g. RSEs derived from survey data, error from meta-analysis where relevant, and CIs around under-identification adjustment factors used).
RATIONALE FOR RECOMMENDATIONS	• Unweighted, undiscounted, prevalent YLDs Maintains comparability with GBD 2010, but allows tailored analysis if required for a particular purpose • Calculate UIs for measurable sources of error.



COMPONENT	11. Compile risk factor exposure data
KEY DECISIONS	- How to generate when limited/no data available? 'Best single' or multiple data sources approach? - How to deal with small cell sizes for some risk factors? - Adopt GBD tailored methods used for specific risk factors (e.g. tobacco, ozone, asbestos, zinc)? - Is adjustment for under-identification required? - Whether/how to estimate uncertainty?
ISSUES TO CONSIDER	- GBD 2010: modelled to fill data gaps and estimate uncertainty - Is GBD 2010 modelling approach appropriate for Indigenous Australian estimates? - Main data source for Indigenous estimates will be ATSIHS. - Do we want to adopt GBD 2010 approach to calculating UIs, or calculate our own?
PRELIMINARY RECOMMENDATIONS INDIGENOUS STUDY	- Use national and Indigenous-specific data sources - Assess on a case by case basis whether to use multiple data sources or the single best data source for each risk factor, according to a protocol to be developed - Adopt GBD specific methods for estimating exposure for selected risk factors (e.g. smoking based on Smoking Impact Ratio; ozone exposure) - No adjustment for under-identification for AATSIHS. For other data sources used, assess whether it is appropriate and possible to adjust for under-identification - Consider methods for calculate UIs based on measurable sources of error (e.g. RSE's from survey data and meta-analysis)
RATIONALE FOR RECOMMENDATIONS	- Indigenous Australian data of relatively high quality for most risk factors - For some risk factors, best estimates may be obtained by combining data from two or more sources - AATSIHS main data source and does not require adjustment for under-identification - UIs able to be calculated for measurable sources of error.

COMPONENT	12. Calculate attributable burden
KEY DECISIONS	- Adopt GBD 2010 risk-outcome pairs, effect sizes theoretical minimum risk exposure distributions (TMRED)? If no, how to determine? - Whether and how to estimate combined RF burden? - Should risk-outcome pairs be consistent for Indigenous and national study? - Should the effect sizes used for Indigenous estimates be the same as those used in GBD 2010 and the Australian study?
ISSUES TO CONSIDER	- GBD 2010: literature review and new meta-analyses to calculate effect sizes; new TMRED; modelled UIs - Can data support additional risk-outcome pairs? - Effect sizes used in GBD 2010 are not comprehensively published - Mixed evidence on ethno variation (and Indigenous genetic predisposition) in effect sizes - Would need substantial evidence to depart from GBD effect sizes and TMREDs. TMREDs should by definition be consistent across all populations - Previous Australian studies calculated estimates for a combined RF burden. GBD 2010 did not attempt this
PRELIMINARY RECOMMENDATIONS INDIGENOUS STUDY	- Adopt GBD risk-outcome pairs, effect sizes and TMREDs for GBD risk factors to be used in the Indigenous study - Consider adding certain risk-outcome pairs where good Indigenous Australian data are available - Consider methods to estimate combined risk factor burden
RATIONALE FOR RECOMMENDATIONS	- GBD risk-outcome pairs and TMREDs based on best available evidence from international literature and should be consistent across all populations - Insufficient evidence to support Indigenous-specific effect sizes

1 Introduction

Background

What is ‘burden of disease’?

Burden of disease analysis is a way of considering the effects of different diseases and injuries in a comparable way. It uses information from a range of sources to quantify the fatal and non-fatal effects of diseases and injuries into a summary measure of health called the DALY – a disability-adjusted life year. One DALY is one lost year of ‘healthy life’ due to premature death, illness or disability, or a combination of these factors. The more DALYs associated with a disease or injury, the greater the burden, whether applied to an individual or a population. More detail on the calculation of DALYs is provided in Chapter 2.

Burden of disease analysis is widely used and extremely useful for a wide range of stakeholders as it combines the impacts of both premature death and illness or disability for a given disease or injury. Just using the fatal burden (mortality rates) for this purpose – which was all that was available in a consistent manner across all diseases prior to the use of methods such as those described here – would result in over emphasis on the predominantly fatal conditions (e.g. cancer) over the non-fatal ones (e.g. arthritis). Burden of disease analysis also allows the contribution of various risk factors to be calculated, known as the ‘attributable burden’.

Burden of disease analysis has become the global standard for bringing all useable data on various diseases together in a comparable manner to produce digestible, comparable and summarised policy-relevant evidence. Its ability to use data from a range of sources to construct a comparable measure for all diseases is a key strength of this approach and is something that cannot be done when separate studies of disease effects are conducted on a disease-by-disease basis. As such it is possible to estimate the relative contribution of each and every disease to the total burden of disease.

What are some of the uses of burden of disease studies?

Burden of disease provides information for three broad and related purposes. The first is **information on population health**, which is useful for health monitoring. This comes from the analysis quantifying the fatal and non-fatal impact of diseases, and the ability to compare the contribution from disease groups, the individual diseases, and the impact from various risk factors. As with any health monitoring, comparisons are desired by age and sex, over time, and between population groups.

The second broad group of uses is for **policy purposes and health service planning**. This has overlap with the health monitoring role outlined above, including the ability to highlight aspects such as which diseases cause the most burden, which are increasing or decreasing in burden, and which are causing most health inequalities. The attributable burden analysis provides valuable information to inform prevention policies. These types of information are also very useful for health service planning. For example, they can provide an indication of which diseases are likely to impact most on the health system, and thus which parts of the health system are likely to be affected (such as doctor visits, hospital admission or dental care). This therefore assists with future planning. The burden of disease model can be used as the basis for scenario testing of various interventions or other impacts on population

health, such as the expected change in population health from decreases in smoking rates. This type of modelling can be very beneficial both for health service planning and policy making.

The third area to highlight is that burden of disease information can be used as the foundation for **assessments of the broader impact of diseases and the cost effectiveness of interventions**. Burden of disease information can be combined with disease expenditure estimates to see how relative disease burden compares with relative expenditure, and to provide an estimate of the 'total economic burden' (covering DALYs and direct expenditure). However, care needs to be taken when analysing such information, due to various contextual issues including the cost-effectiveness of interventions, and time lags between interventions and effects. Studies on the cost-effectiveness of various interventions need a measure of disease burden, such as the DALY.

Key terms used in this chapter

Attributable burden - The disease burden attributed to a particular risk factor. It is the reduction in burden that would have occurred if exposure to the risk factor had been avoided.

DALY (Disability-adjusted life year) - A year of healthy life lost, either through premature death or equivalently through living with disability due to illness or injury.

Disability weight - A factor that reflects the severity of health loss from a condition on a scale from 0 (perfect health) to 1 (equivalent to death).

Discounting - A method used to adjust the relative value of years lived (or lost) in the future. It is based on the assumption that a year lived in the future is of less 'value' than a year lived now. 'Discounting for future benefits' is standard practice in economic analysis.

HALE (Health-adjusted life expectancy) - An estimate of the number of healthy years (free from disability or disease) that a person born in a particular year can expect to live based on current trends in deaths and disease patterns.

Incidence - The number of new cases (of an illness or event, and so on) occurring during a given period.

Prevalence - The number or proportion (of cases, instances, and so forth) in a population at a given time.

Sequelae - consequences of diseases and injuries

YLD (Years lived with disability) - measures the years of what could have been a healthy life that were instead spent in states of less than full health.

YLL (Years of life lost) - years of life lost due to premature mortality.

(Refer to glossary for a full list)

What can't burden of disease studies tell us?

Burden of disease analysis quantifies the 'health loss' from various diseases. It does not cover broader factors, such as social impacts, economic impacts, or the direct impact on the health system. It cannot provide information on workforce needs or shortages.

Since burden of disease analysis only quantifies the size of a health problem it should not be used on its own for policy and planning as it does not show what interventions will work or which are the most cost-effective. However, as outlined above, burden of disease analysis

helps to identify conditions where cost-effectiveness of interventions should be investigated to gain the maximum benefit.

Global studies

Burden of disease studies using the DALY metric began in the 1990s with the first global burden of disease study. There have been a number of these global studies:

- The first global study developed the DALY metric and quantified the disease burden in 8 regions of the world for the year 1990 and for a limited number of risk factors (Murray and Lopez 1996)
- The global study was next updated for the year 2000–2002, with a more detailed analysis and a more comprehensive risk factor comparison (Lopez et al 2006)
- WHO updated the DALY results for 2004 with projections to 2030 (WHO 2008), and the attribution to risk factors (WHO 2009)
- The latest global study (GBD 2010), conducted by the Institute for Health Metrics and Evaluation (IHME) and other academic partners, was published in December 2012 (Lancet 2012). It includes a number of revisions to the methods, which were then used to calculate DALYs for 1990, 2005 and 2010
- WHO anticipates the latest GBD results will ‘contribute to revisions for WHO global health estimates in 2013’ (WHO 2013)
- IHME has stated that it intends to produce annual updates of their global estimates.

Box 1.1: Where does Australia fit in GBD 2010?

The Global Burden of Diseases, Injuries and Risk Factors Study 2010 provides estimates for 187 countries divided into 21 regions on the basis of 2 criteria:

- Epidemiological homogeneity and
- Geographical contiguity

Australia and New Zealand form the **Australasia** region.

Regions are grouped into 7 super-regions. Australasia belongs to the **High Income** super-region, with Southern Latin America, Western Europe, High income North America and High income Asia Pacific.

Appendix 1A provides a global map showing the regions and super-regions used in GBD 2010.

Previous Australian studies

The Australian burden of disease work has been based on the principles and methods used in the Global Burden of Disease studies. There have been two major burden of disease studies at the national level in Australia:

- The first was published by the AIHW in 1999 and documented the burden of disease in 1996 including information on DALYs, the attributable burden of selected risk factors, and differences in the burden across age, sex and socioeconomic groups (Mathers et al. 1999).
- The second was published by the AIHW in 2007 and updated the data to 2003. It also included analysis by a greater range of subpopulation groups – including at the state-

level, by remoteness groups, by socioeconomic groups – as well as projections of the burden to 2020 (Begg et al. 2007).

- A separate analysis of the burden of disease in Indigenous Australians was published by the University of Queensland in 2007 (Vos et al. 2007).

Since the last national study, some states and territories have done further burden of disease work. This has included updating the jurisdiction-level estimates produced as part of the 2003 national study, and undertaking further analysis. This work is quite substantial in some cases.

Australian Burden of Disease Study

In recent years, it has become clear that the 2003 Australian burden of disease estimates are becoming increasingly out-of-date, and many stakeholders have highlighted the need for more current information. Funding from the Australian Government Department of Health and Ageing (now the Department of Health) and the Australian National Preventive Health Agency (ANPHA) became available in the first half of 2013 to update this work.

The first phase of the project is to explore a range of methodological issues and determine the best methods to use to update the Australian and Indigenous estimates – the subject of this report. The conclusions drawn will then be used to produce national estimates, Indigenous estimates, and sub-national estimates where valid.

An Expert Advisory Group (EAG) has been convened to provide expert advice to the AIHW about the methodology and content of the Australian Burden of Disease Study. This is the national advisory group to support this work, reporting to the AIHW.

The EAG includes members with expertise in relevant areas such as burden of disease methodology, health policy, epidemiology, health data analysis, and Indigenous data analysis. It includes some state and territory representatives (representation of large and small, and urban and remote), representatives of the funding agencies, and the Australian Bureau of Statistics (ABS). A list of members is included in Appendix 1B.

An Indigenous Reference Group (IRG) has been convened to provide expert advice to the AIHW on key technical issues, methodology and outputs in relation to Indigenous estimates produced as part of the Australian Burden of Disease Study. The IRG will provide advice on Indigenous estimates to the EAG, and will also take into account advice given by the EAG on national estimates to ensure alignment, where possible, between the national and Indigenous components of the Study.

The IRG includes members with expertise in relevant areas such as the demography and health of Aboriginal and Torres Strait Islander people, Indigenous mortality and life expectancy, BoD methodology, health policy, epidemiology, health data, Indigenous data issues and analysis, and data linkage. A list of members is included in Appendix 1C.

Other interest-specific working groups will be established as necessary throughout the life of the project.

Aim of this working paper

Given the timing of the release of the GBD 2010 results in late 2012, and the fact that it uses new methods in a number of areas (see Box 1.2 for summary of the main changes), it is important that these changes are evaluated for the Australian context. Therefore this first

working paper aims to assess how to best use the new GBD methods and estimates to update the Australian burden of disease estimates. The main objective is to consider the GBD methods and assess their application for the Australian context, with additional specific consideration of the Indigenous Australian context. Assessment of the actual GBD estimates for Australia is outside the scope of this working paper.

Box 1.2: Summary of main changes in Global Burden of Disease 2010 study

There were a number of methodological developments since the 2000 global study incorporated into the GBD 2010 study. These are covered in detail in the various chapters of this report. Key changes that require assessment in the context of determining the methods for the Australian study are:

- The list of **diseases and risk factors** was expanded
- A new reference **life table** was implemented
- **Discounting** of future health and **age weights** were not used
- **Uncertainty intervals** were calculated
- **Prevalent** rather than incident YLDs were calculated
- The **disability weights** were substantially revised
- **A combination of data sources** for a particular condition were used, rather than the best single data source
- An increase in the use of **modelling** to fill in data gaps
- An increase in the number of **age groups** to 20
- The definition of health has changed to only include **health loss**

Requirements and principles for Australian analysis

The assessment of how best to use GBD methods and results in the Australian Burden of Disease Study needs to take account of the broad principles that are the foundation for the Australian analysis. These are to:

1. Provide **national** estimates of fatal, non-fatal and total burden, as well as the attribution to specific risk factors, that are **up-to-date, of high-quality and meet Australia's needs**.
2. Provide **Indigenous** estimates, noting the extra challenges due to data limitations and differences in disease profiles. Because of the focus on the gap between Indigenous and non-Indigenous health, the aim will be to use comparable methods to those used for the national estimates as much as possible, though some differences are still likely due to differences in data sources, data quality and issues affecting Indigenous data.
3. Provide **sub-national** estimates (such as state/territory, regional, socioeconomic groups) where valid.
4. Maintain comparability with **GBD methods** as much as possible, with full clarity around any differences.
5. Provide **transparency** in the data sources, assumptions and methods used, with the ability to replicate the results.
6. Complete the work in an **efficient and flexible** manner, build national capacity, and set up the relevant infrastructure to enable efficient and timely **ongoing updates**.

7. Ensure **collaboration** with the various stakeholders including other burden of disease experts both nationally and internationally in order to contribute to global burden of disease work.

There is interest in assessing changes over time in the burden of disease – in particular whether the situation has changed since the previous Australian estimates (for 2003). The principle above of establishing a system that enables ongoing updates of the estimates will establish the foundation for a time series into the future; however, time series are not within the scope of the current project.

Key considerations

There are a number of contextual issues in relation to the assessment of GBD 2010 methods for the Australian context, and these are introduced below. These are overarching considerations that will be referred to throughout this working paper.

Requirements for global analysis compared with country analysis

There are different but complementary requirements for the global analysis undertaken in the GBD 2010 study and the Australian Burden of Disease Study. Both require aggregate level data, and both require disaggregations below that level.

The global work needs to estimate the overall global burden of disease – the aggregate estimates for measures such as YLLs, YLDs, DALYs and HALE – and determine which diseases and risk factors are causing the most burden at the global level. As well as these global results, the GBD study disaggregates the information into regions (groups of countries) and, in the latest release, countries. Making comparisons at the country level requires methods that can deal with large amounts of ‘missing’ data, due to many countries having limited data of the type required for burden of disease studies. The output of estimates for Australia – the first time this level of information has been available from global studies – is of obvious interest in the context of our new Australian study.

As outlined above, the highest priority for the Australian study is to obtain up-to-date, high quality estimates at the national level. The context for this aim is somewhat different from that of the global study, as Australia has relatively good health data. There are also disaggregations required for the Australian study. In particular, estimates for the Indigenous population are required. Also, sub-national breakdowns are desired.

It would be preferable if the estimates produced to meet these two sets of requirements are the same. That is, that the Australian estimates obtained from the global study are consistent with those produced as the national estimates in the Australian study. While the aim is to maintain consistency with the global methods, it is also important to fulfil country requirements, remembering that this is to be done within constrained resources and time.

The starting point for this assessment is that the Australian estimates in GBD are to be used for international comparisons. Alongside this, Australia needs to determine how to best construct national estimates using the most up-to-date and appropriate methods, with the ability to disaggregate below the national level. The resulting estimates are intended to be used for health monitoring, planning and policy-making within Australia. The reasons for any differences between the Australian estimates produced and the 2010 GBD estimates for Australia need to be explained.

Level of complexity, need for transparency

The methods used in burden of disease studies have become generally more complex over time, most notably in the latest global study. These developments have been implemented primarily to deal with previous limitations or criticisms of various aspects of the methods. Examples of developments implemented in GBD 2010 include: the quantification of uncertainty intervals; use of multiple datasets for each component of the work rather than one determined to be the best available; provision of country-level estimates; and increases in the number of conditions and risk factors included. The complexity is reflected in the large number of different types of models and techniques used. For example, the calculation of YLDs uses 6 types of models, including Bayesian meta-regression, natural history models and geospatial models among others. Micro-simulation techniques are used to make adjustments for comorbidity and to calculate uncertainty intervals. Further details on the modelling techniques are contained in later chapters.

These changes have led to an increase in the use of complex models rather than more direct calculations from data. While this may have led to improvements in estimates for regions and countries with limited data, it may have had unintended consequences such as a reduction in the transparency of the methods, and less need to obtain the latest data where available. In the Australian context where generally more complete and better quality data exist, a challenge will be to determine which of the GBD methodological advances give genuine improvements in the Australian estimates.

As well as increased complexity in the models, the scope of the analysis has broadened. There are now 241 diseases, 1,160 sequelae, and 57 risk factors included in the analysis. This has advantages, as the more complete the analysis in terms of accounting for component diseases and risk factors, the more informative the results should be. However, there may be some variation in the completeness of data and analysis for the various diseases and risk factors, which could result in inaccurate rankings. For example, there is now detailed information on various components of diet, which may artificially increase its ranking relative to other risk factors. There is also an increased risk of methodological or data errors that come with both greater breadth and detail, as it becomes more difficult to be fully across all the components of the estimation process and to ensure all of these components are handled consistently.

The result of this increased complexity is decreased clarity for stakeholders. The methods have become so complex that they need a specialist, with sufficient time, to digest the details to be able to understand them. This complexity makes it much harder to explain the methods, and risks giving the appearance of being more of a 'black box'. One of the principles for the new Australian study is for transparency of data, assumptions and methods. It is therefore important to consider whether the level of complexity in the GBD methods is appropriate in the Australian study.

Need for Indigenous estimates

A key requirement of the Australian Burden of Disease Study is to produce Indigenous estimates. Indigenous burden of disease estimates will be reported at the national level only. An assessment of the feasibility of producing Indigenous BoD estimates at the sub-national level (e.g. by state/territory) may be undertaken for future analyses; however, this is not within scope of the Indigenous component of the current study.

Indigenous Australians as a population group face a very different burden of disease to non-Indigenous Australians (both a higher burden as well as some differences in the pattern of diseases). There is therefore a need for a set of burden of disease and injury estimates specific to the Aboriginal and Torres Strait Islander population. This is particularly important in relation to the Council of Australian Governments' (COAG) commitment to closing the gap in Indigenous life expectancy within a generation (COAG 2008). Updated Indigenous burden of disease estimates will be an important source of information to monitor progress in improving Indigenous health, so as to better monitor progress against the COAG targets, and to assist Governments to develop and target interventions that reduce the main contributors to the burden of disease and injury.

The 2003 Indigenous Australian burden of disease study was the first complete assessment of the burden of disease and injury for Aboriginal and Torres Strait Islander peoples. It included estimates of more than 170 disease and injury categories which were compiled from base units at the age and sex levels. However, information on Indigenous health status for all of these categories was not available at the time. Instead ratios of the differences between Indigenous and total population rates for proxy measures of disease occurrence (e.g. mortality or hospital admissions) were used. Whether a similar approach would be used for the new Indigenous study could be considered, however, more information is now available on incidence and prevalence of certain diseases for the Indigenous population from data sources such as the 2012-13 Australian Aboriginal and Torres Strait Islander Health Survey (AATSIHS). Therefore the number of diseases for which proxy measures would need to be used to derive estimates is likely to be reduced.

A key gap in the 2003 Indigenous study was the lack of health measurement data on key risk factors such as high blood pressure, cholesterol and body mass. This information is expected to be able to be captured through the 2012-13 AATSIHS.

The 2003 Indigenous study also included a measure of the 'health gap' which was calculated as the difference between the burden of disease estimates for Indigenous Australians and what these estimates would have been if Indigenous Australians had experienced mortality and disability at the level of the total Australian population. While the primary focus of the new Indigenous study will be to produce estimates for the Indigenous population, estimates of the health gap are also relevant. The AIHW will consider how best to produce estimates of the 'health gap' for the Indigenous component of the study and whether this is feasible within the timeframes and resources allocated to the project. However at present it is the AIHW's intention to collect and input data for both Indigenous and non-Indigenous Australians where data are available for 'health gap' measurement. However, it should be noted that non-Indigenous estimates are not planned to be reported or published as a separate output for the study.

An important consideration is the accuracy of Indigenous estimates produced, given the small size of the Indigenous population, data quality issues, and lack of Indigenous status information for some diseases and risk factors. To ensure validity of the results, the level of disaggregation for diseases or risk factors may need to be broader than for the national study. The validity of the results may also be facilitated by combining several years of data.

A final consideration is whether, and how, Indigenous under-identification in administrative data collections will be adjusted for, and what sources of information could be used to derive adjustment factors. Indigenous estimates in the new Australian study will draw on the latest studies undertaken that have generated correction factors for Indigenous under-

identification for particular data sets. The benefits of using data linkage to derive estimates of Indigenous identification could also be explored.

Need for sub-national estimates

A key requirement of the Australian Burden of Disease Study is to extend the national-level estimates to sub-national estimates. This includes state/territory, socioeconomic groups, and remoteness category estimates. There may also be the possibility of constructing smaller area estimates, where valid.

Various sub-national estimates were included in the 2003 Australian burden of disease study: jurisdictional estimates, socioeconomic groups, and remoteness category estimates (at the national level). These were compiled from base units at the age/sex/statistical local area (SLA) level. As not all data sources were available at this level, assumptions were needed in order to be able to construct these small area units. For example, data that were only available at the national level would be localised to the SLA by assuming that the age/sex rates were equivalent across the country. It is expected that a similar approach could be used in the new Australian study. However, there are a number of factors to examine first, including whether the methods used in GBD can be applied in this manner, consistency between data sets and over time, advances in methods for small-area estimates that have occurred since the last Australian study, and the effect of the new Australian geographical classification (for example, whether data sources are consistently using the new classification, or whether they include sufficient information to be able to map data to the new classification).

An important consideration is the validity of sub-national results, particularly if units smaller than state/territory (i.e. sub-state) or remoteness categories are used. This will be influenced by the availability and quality of data for those aspects that vary by that disaggregation, and also by the population size in the various sub-national groups. It may be that the level of detail (for example, the detail in the disease or risk factor breakdowns) needs to be restricted to ensure validity of the results. Alternatively, some smaller groups may need to be reported together in a combined group.

Ability to update estimates on an ongoing basis

One of the requirements for the new Australian study is to design a sustainable system that will enable efficient, ongoing updates of national and Indigenous estimates. The current system of separate stand-alone 'projects', with this new study being the third Australian burden of disease project at the national level, comes with inefficiencies compared with an ongoing system of regular updates.

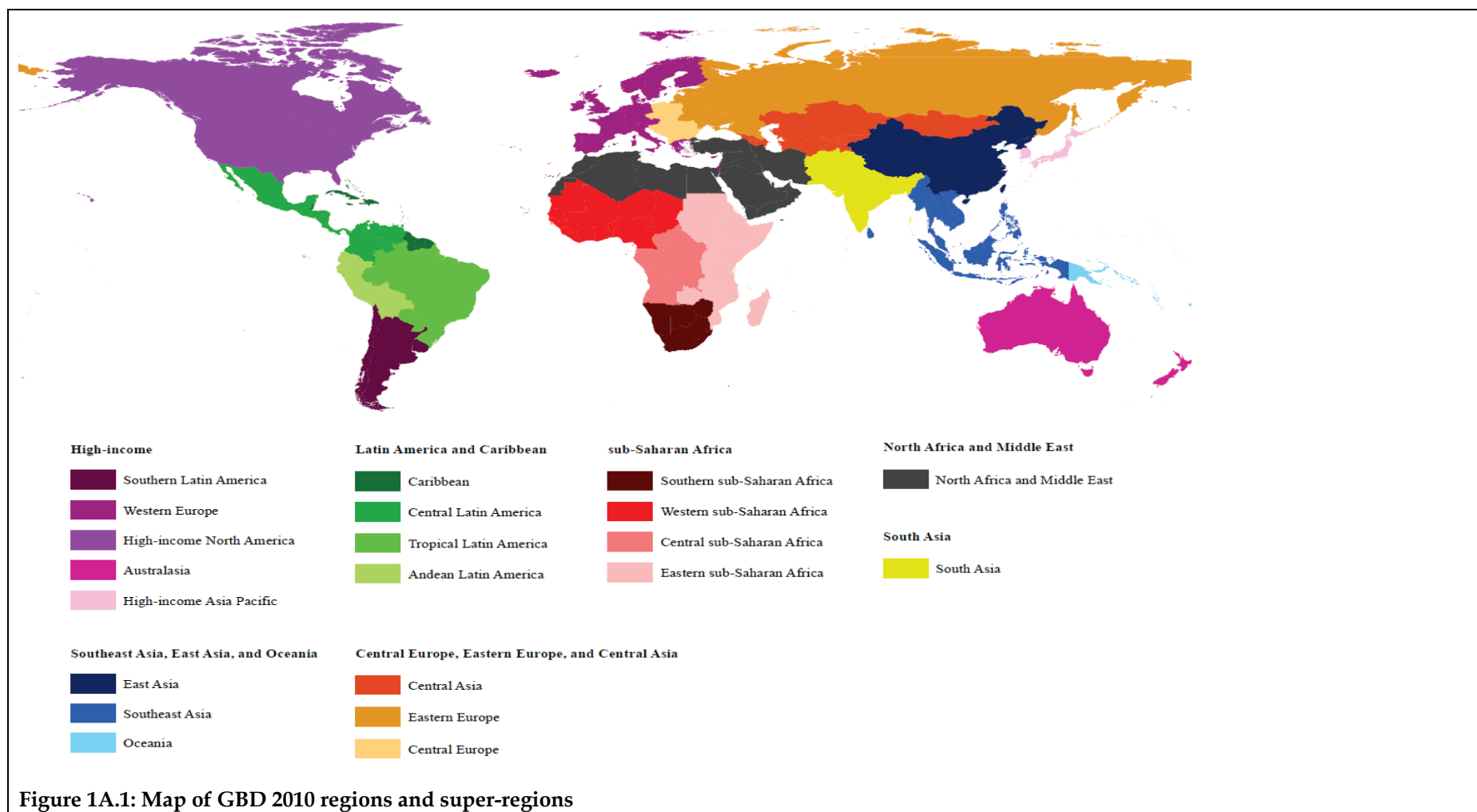
The stand-alone approach has meant that considerable work has been required each time to secure funding, train staff, update data and models, generate and validate estimates, and produce the reports. It has also meant that between these periodic studies there have been no updates to the estimates even when new input data became available. There have also been considerable inefficiencies with each jurisdiction needing to update their burden of disease estimates independently. The current system also means there is the potential for different methods to be used across jurisdictions which may make comparisons between them invalid. Considerable benefits and economies of scale would come from a national approach to this work with detailed information available at the state/territory level.

There are a number of requirements for such a sustainable system. First is the ability to input new data when available—particularly for key data sources. There is also the need for a manageable system to update estimates and include changes to input data (such as reference life tables and disability weights as IHME further develop the global studies. It is also likely that there would be potential extensions to the work that could not be done within the time and resources available for the current project, as well as the need to be able to modify certain aspects of the assumptions behind the methods depending on the needs of the user. It would therefore be preferable to design the system for running the burden of disease estimates to be flexible so that these extensions and variations could be accomplished efficiently.

There are also logistical issues that will impact on the ability to update estimates on an ongoing basis, including computing requirements. The level of complexity used in GBD 2010 has required substantial computing power. However, the quantum of data that would be collated at the single country level is substantially less than that collated at the global level, meaning that the same type of ‘supercomputing’ power is not likely to be required for a single country study.

The final aspect of building a sustainable system is to establish and maintain links with others working on burden of disease studies. This would include those within Australia and internationally, from government and academic groups, and other relevant groups.

Appendix 1A



Appendix 1B

Table 1B.1: Australian Burden of Disease Study Expert Advisory Group (as at 25 February 2014)

Name	Organisation/Jurisdiction
Ching Choi (Chair)	University of New South Wales
Emily Banks	Australian National University
Tony Barnes	Independent
Annette Dobson	University of Queensland
Sandra Eades	Baker IDI
Paul Kelly	ACT
Siew-Ean Khoo	Australian National University
John Lynch	University of Adelaide
David Roder	University of South Australia
Colin Sindall	Victoria
Peter Somerford	Western Australia
Harvey Whiteford	University of Queensland
Andrew Wilson	University of Sydney
Jeanette Young	Queensland
<i>Australian Government agencies</i>	
Michelle Marquardt	Australian Bureau of Statistics
Jack Quinane	Australian National Preventive Health Agency
David Cullen	Department of Health
Bernie Towler	Department of Health
Brendan Gibson	Department of Health

Appendix 1C

Table 1C.1: Australian Burden of Disease Study Indigenous Reference Group (as at 15 May 2014)

Name	Organisation/Jurisdiction
Len Smith (Chair)	Australian National University
Tony Barnes	Independent
Alex Brown	University of Adelaide
Debra Reid	University of Sydney
Daniel Williamson	Queensland Health
Wendy Hoy	University of Queensland
Yuejen Zhao	Northern Territory Government
Steve Guthridge	Northern Territory Government
Lisa Briggs	National Aboriginal Community Controlled Health Organisation
Daniel Christensen (TICHR)	Telethon Kids Institute
Vanessa Lee	University of Sydney
<i>Australian Government agencies</i>	
Julie Nankervis	Australian Bureau of Statistics
Shahidullah	Australian Bureau of Statistics
Kirrily Harrison	Department of Prime Minister and Cabinet
Jahan Ara	Department of Health
Fadwa Al-Yaman	Australian Institute of Health and Welfare
Michelle Gourley	Australian Institute of Health and Welfare

2 Overview of methods

This chapter provides a brief overview of the methods used to calculate disability-adjusted life years (DALYs) with a particular focus on those aspects of the Global Burden of Disease study that have changed (see Box 1.1). This chapter introduces the key inputs and underlying methodologies that apply to all burden of disease studies based on the DALY method. As such, the methodologies discussed in this chapter apply equally to Indigenous and the national components of the Australian Burden of Disease Study. Note that more detail on specific components of the methods is included in later chapters.

The burden of disease, as measured by the DALY, quantifies the gap between a population's actual health and an ideal or reference level of health. The DALY extends the concept of potential years of life lost due to premature death (YLL) to include equivalent years of 'healthy' life lost by virtue of ill health (quantified as years lived with disability, YLD). Box 2.1 provides an illustration of the concepts behind the DALY.

Box 2.1: Calculating disability-adjusted life years

Consider a single individual, Jim. His is a typical case of a person with a serious disease. He is fully healthy until he gets the disease, aged 50. Evidence about the disease shows he is likely to live with it until he dies aged 60. Based on further evidence, those 10 years suffering his particular disease will be equal to only 3 years of full health. This means he will 'lose' 7 healthy years (or YLDs) even though he is alive. (In technical terms, his disease has a 'severity weight' – often known as a 'disability weight' – of 0.7. As other examples, if the severity weight had been 0.55 he would have 'lost' 5.5 of those 10 years; and 1.8 years if it had been 0.18.)

As well as the 7 healthy years lost through ill-health, Jim will lose many years through dying too early. At the age he is likely to die, 60, a male in the baseline year would normally go on to live until he is 81. This means Jim will lose a further 21 years (or YLLs).

Jim's total disability-adjusted life years (DALYs) are therefore 7 plus 21, making 28.

Using this reasoning, we can take all the people getting Jim's disease and start to build towards the total burden of disease. This means drawing on surveys and other research that shows:

- how many people are newly diagnosed with the disease in the baseline year
- what sex and age groups they are in
- how long people in each group will typically have their disease for
- the average age at which the people in each group are likely to die.

Next we can calculate subtotals for each group and then total them to get the diseases DALYs for the whole population. The steps taken for this disease can then be applied to all other forms of disease and injury that arose in the baseline year.

If desired, projected trends in disease patterns and changes in population can be overlaid on the methods described above to estimate the burden in future years.

Note: This example uses an incidence based approach to calculating YLDs and does not discount for time (see later sections on discounting and incidence versus prevalence)

Key terms used in this chapter

age weighting – method used to adjust the relative value of years lived at different ages, for example, to value a year lived by a young adult more highly than a year lived at older ages. Age weighting means that some age groups will have greater influence on the results than others.

comorbidity – having two or more health problems at the same time

discounting – a method used to adjust the relative value of years lived (or lost) in the future

life table – a table that shows, for each age, the probability that a person of that age will die before their next birthday

uncertainty interval – a statistical term describing a range (interval) of values that are likely to include the 'correct' estimate

(Refer to glossary for a full list)

Outline of inputs and key decisions

The burden of disease framework relies on a comprehensive assessment of health in a community and uses time in the form of lost years of healthy life as the basic unit of measurement. Underpinning the analysis is a series of social value choices and other inputs that determine the methods used to calculate DALYs. These choices include the aspirational life table to be used, whether and how to adjust for comorbidity, the use of incidence or prevalence estimates for calculating YLD, and whether future health loss (or gain) is to be valued differently to current health loss (or gain). The major inputs and choices to be made are outlined below and, in some cases, detailed further in other chapters.

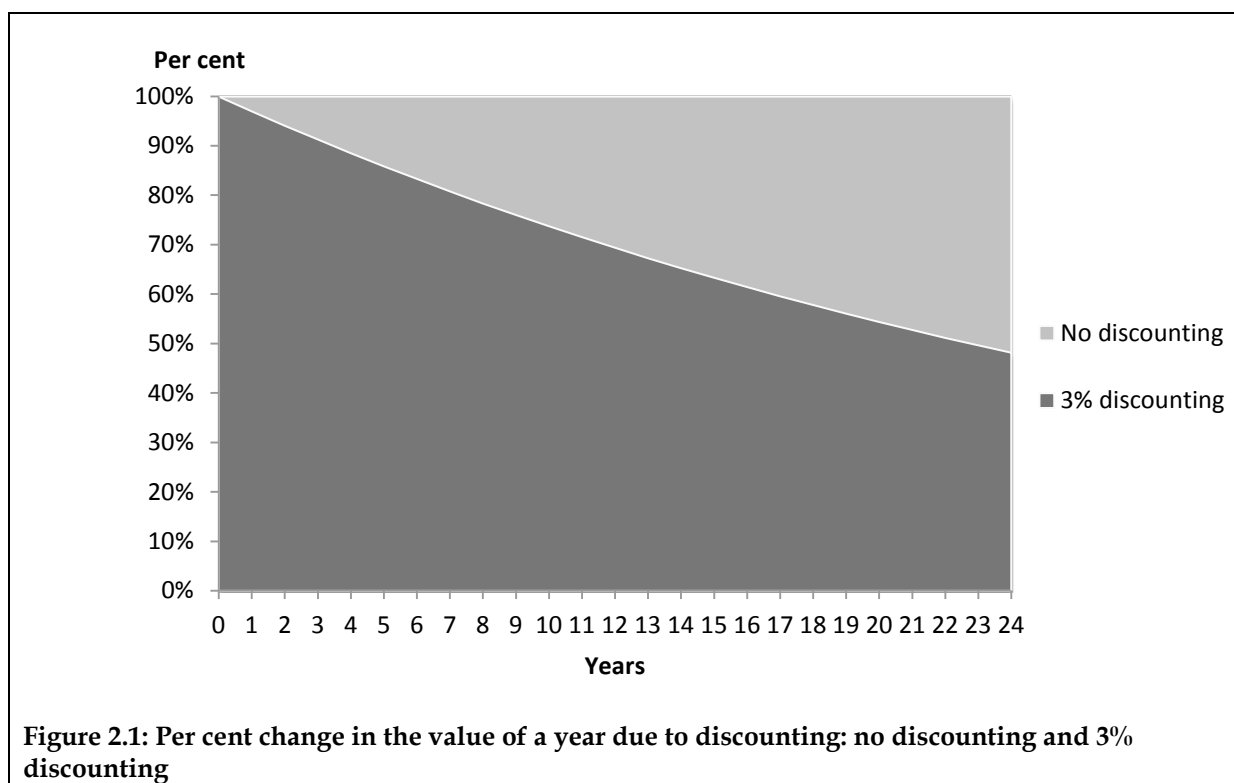
Cause (or condition) list

Underpinning any burden of disease analysis is a cause (or condition) list. This is an exhaustive list of diseases based on International Classification of Diseases (ICD) codes for which analysis is meaningful and possible. The cause list also forms the basis for risk factor analysis, where the burden attributed to a condition can be related to a particular risk factor – for example, a portion of all lung cancer is caused by smoking (the risk factor). The 2003 Australian burden of disease study included estimates for 186 diseases and 14 risk factors. In contrast, GBD 2010 reported estimates for 241 individual diseases and causes of injury (described by 1,160 sequelae), and 57 individual risk factors. The cause and risk factor lists are discussed in detail in Chapter 3.

Discounting

'Discounting for future benefit' is standard economic practice and was used in the first GBD study (for the year 1990) to estimate the net present value of years of life lost (AIHW 1999). That is, healthy years lived in the present are valued more than those lived in the future. The incorporation of discounting is in line with methods used in cost-effectiveness analysis that suggest that, without discounting future health, decision makers would be led to make choices that call for a disproportionate sacrifice from the current population. For example, the chance of disease eradication through research would lead to unlimited health gains and provide reason to use all health resources for research purposes to benefit future populations.

The effect of a 3% discount rate (standard discount) is that a year of healthy life gained in 25 years' time is worth less than half of a year gained now (Figure 2.1).



The arguments for and against discounting for time have been well documented in recent reports. Both the 1996 and 2003 Australian studies discounted at 3% per year. The reasons for this decision as recorded in the 1996 study report (AIHW: Mathers et al. 1999, p12) included:

- to be consistent with methods employed in cost effectiveness analyses, and
- to prevent giving excessive weight to deaths at younger ages.

The GBD 2010 project consulted with a group of philosophers, ethicists and economists who reached consensus that a year of health should be of equal value regardless of when it is lived. This choice not to discount for time is a simpler approach that is likely to be interpreted more easily. The GBD team cites arguments in a number of journals as being behind the decision (supplement to Murray et al. 2012, p12).

While no discounting enables easier interpretation of the estimates from GBD 2010, it has implications for using these estimates in cost-effectiveness analyses.

Age weighting

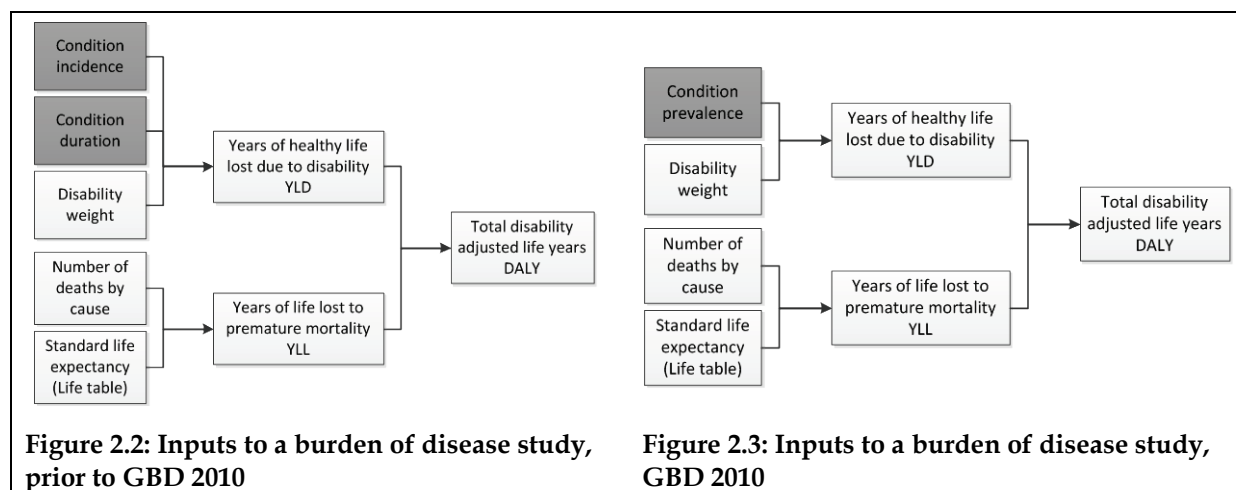
Age-weighting is a method used to assign larger importance to certain age groups compared with others. Previous global studies weighted health loss in young adults more highly than health loss at younger and older ages; whereas GBD 2010 chose to treat each age group as having the same importance. This latter approach is consistent with previous Australian studies which applied uniform age weights.

Inputs for calculating DALYs

DALYs for a disease or health condition are calculated as the sum of the years of life lost due to premature mortality (YLL) in the population and the equivalent ‘healthy’ years lost due to disability (YLD) relating to the health condition.

$$\text{DALY} = \text{YLL} + \text{YLD}$$

The key inputs for calculating DALYs are shown in Figures 2.2 and 2.3. These figures show 2 different options for calculating the YLD component – using an incidence or prevalence approach. These inputs are described briefly below, and discussed in more detail in chapters 4 (YLL) and 5 (YLD).



Inputs for calculating YLLs

Years of life lost (YLL) for an individual is the difference between their age at death and the aspirational life expectancy for a person of that age. The aspirational life expectancy for males and females at every age is defined by a chosen reference life table.

Mortality data

The main input for measuring the fatal burden of a condition is the number of deaths for that condition in the relevant population. The source of this information in Australia is coded causes of death data based on vital registration, where the cause of death is coded using ICD-10. More information about Australian mortality data is in Chapter 4.

Life table

As noted above, YLL is calculated as the difference between the age at death and the aspirational life expectancy at that age, as defined by a chosen reference life table. The choice of life table will therefore affect the quantum of burden due to premature mortality.

The standard reference life table for both Australia’s 2003 study and the GBD 1990 was based on the highest observed life expectancy at the time – that of Japanese females – and also used the observed male-female gap in life expectancy in comparable countries of 2.5 years. This resulted in the standard reference life expectancy at birth being 80.0 years for males and 82.5 years for females.

In the twenty years since the 1990 analysis, life expectancy has improved in low mortality countries. In Japan, for example, it increased for females to 85.9 years at birth. In addition, rather than relying solely on one country’s mortality rates, an approach using multiple

countries was adopted. This was done by constructing a hypothetical life table for GBD 2010 using the lowest age-specific death rates from around the world, and applying these rates to both sexes. The new reference life table (supplement to Murray et al. 2012, p13–14) has a life expectancy at birth of 86.0 years for both males and females.

The implications of using the GBD 2010 reference life table are discussed further in Chapter 4.

Inputs for calculating YLDs

The concepts underlying the calculation of YLDs are more complex than for YLLs. They attempt to capture the natural history, severity, prognosis and consequences of each condition in the cause list and to quantify their impact on a population in terms of the difference between a life lived in full health and a life lived with one or more health problems. This can be illustrated by generic disease models.

Disease models

A (conceptual) disease model is a description of the experience of a disease in a population. It describes the starting point of the disease (incidence), how many people have the disease (prevalence) and for how long (duration) and consequently how many people are either cured (remission) or die (case fatality) (Figure 2.4). See Box 2.2 for a description of these terms.

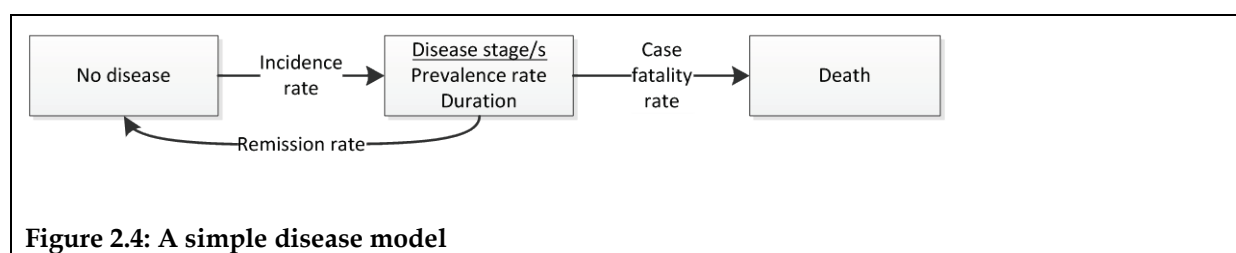


Figure 2.4: A simple disease model

In a more complex model, the disease stage as defined by prevalence and duration may be one of a number of stages of the disease for which there is an input path (incidence or progression) and up to three output paths (remission, fatality, progression to the next stage) (Figure 2.5).

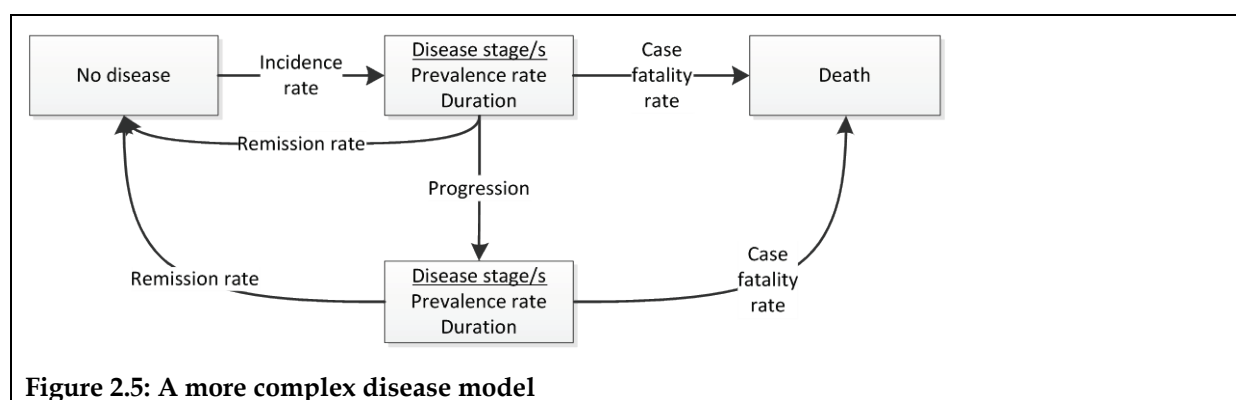


Figure 2.5: A more complex disease model

A disease model is created for each stage of each condition in the cause list that caused disability. The disease models in GBD 2010 were created at the sequelae level, requiring 1,160 models to make up the morbidity database for that study.

Box 2.2: Key elements of a disease model

Incidence rate

The proportion of the population who contract a condition or disease during a specified time period. This reflects the number of new cases of disease over a given time period – usually 1 year.

Prevalence rate

The proportion of the population who have a condition or disease at a point in time or during a specified time period. This reflects the number of new and existing cases of a disease at a point in time or over a given time period – usually 1 year.

Duration

The length of time for which the effects of the condition are experienced. This may include asymptomatic periods.

Remission rate

The proportion of people with a condition for whom disease activity ceases, either through remission or cure, within a specified time period.

Case fatality rate

The proportion of people with a condition who die as a result of that condition within a specified time period.

Choice of prevalence or incidence approach

Until the GBD 2010 study, incident YLD was usually used in the calculation of DALYs. This was calculated as the product of the incidence rate, disability weight and duration for each stage of a disease. That is, only new cases of a disease in the reference time period were included.

In GBD 2010, YLD was calculated using a prevalence approach. This requires estimation of the prevalence of the condition (disease or injury) in the specified time period. The number of years of healthy life lost is obtained by multiplying the prevalence rate by the disability weight. Therefore all cases of a disease in the reference time period are included in the analysis.

There are a number of reasons behind the decision to use the prevalence approach for calculating YLDs. Primarily it is easier to adjust for comorbidity using the prevalence approach. In addition, the arguments for prevalent YLD (which include the usefulness of having burden assigned to the age group in which the health loss is occurring, and the impact of arbitrarily assigning a point of incidence for a disease) outweigh the arguments for incident YLD (which are that both YLL and YLD are incidence measures, and that recent trend information is important) (supplement to Murray et al. 2012).

This issue is discussed in more detail in Chapter 5.

Disability weights

Disability weights attempt to capture the severity of the effects of a disease or injury on a scale from 0 (perfect health) to 1 (equivalent to death). They aim to quantify societal preferences for different health states. The weights do not represent the lived experience of any health state, or imply any societal value of the person in a particular health state. Rather, they quantify societal preference for health states in relation to the 'ideal' of good health.

Disability weights are used as a multiplier in the calculation of YLD and, as such, any change has a direct influence on the quantum of YLD (that is, a larger disability weight will result in an increase in YLD). Previous global and other burden of disease studies, including in Australia (1996 and 2003), used the same set of weights (with minor amendments), which were derived from international burden studies that quantified societal preferences.

The GBD 2010 project has explicitly attempted to separate the health effect from other factors and has employed new methods to estimate the disability weights for 'health states' rather than conditions. It aimed to respond to suggestions for a more inclusive measurement exercise which represented the broader perspective of different cultures, communities and societies. This has resulted in a new and significantly different set of disability weights to previous studies.

Such fundamental changes to the disability weights are likely to cause changes in both the quantum of YLD for a condition and the relative ranking of conditions. Adoption of these new weights for Australian burden of disease estimates assumes they adequately reflect Australian preferences. This may or may not be the case, particularly for Aboriginal and Torres Strait Islander people. The issue of disability weights is discussed in more detail in Chapter 5.

Comorbidity adjustments

It is not uncommon that someone with a chronic condition like diabetes will also have heart disease, or that a patient requires treatment for more than one condition during a hospital visit. These are examples of **comorbidity**, which is when a person has two or more health problems at the same time. Comorbidity is problematic for a burden of disease study because the disability weights are derived for a condition in isolation. In most cases, the collective disability caused by more than one condition is not likely to be as great as the sum of the individual condition weights, though in some cases it may be greater. For example, the effect of deafness and blindness together may be less 'disabling' than might be expected based on adding together the disability weights for the two conditions.

For the 2003 Australian study a multiplicative model was used to adjust the disability weights for the 21 most common non-fatal conditions of old age. Comorbidity between the 21 conditions was estimated using National Health Surveys and Hospital Morbidity data. In allocating the new combined weight, the change in total weight was taken from the milder of the conditions (Begg et al. 2007, p25). For GBD 2010, the analysis of comorbidity involved micro-simulation of each age-sex group for combinations of diseases and a multiplicative model for combining the disability weights (supplement to Murray et al. 2012). The individual health loss represented by the combined disability weight is allocated to each condition proportionally using the disability weight of the condition on its own. More detail about comorbidity adjustments is provided in Chapter 5.

Uncertainty intervals

Similar to confidence intervals, 'uncertainty intervals' provide an indication of the accuracy of the estimates and the information on which they are based. There are a number of stages in the GBD 2010 methodology where uncertainty is estimated. There is uncertainty in:

- Cause-of-death model predictions for each age, sex, country, year, and cause. It has been stated that both uncertainty in the parameter estimation and in model specification are accounted for (Lozano et al. 2012).

- Bayesian meta-regression models (and, for some sequelae, alternate approaches) used to calculate prevalence estimates by age, sex, country and year for each sequelae.
- The construction of the disability weights.
- The estimation of prevalence of sequelae and risk factor exposures.

Simulation methods were used to calculate the uncertainty intervals in GBD 2010. Further details on the methods used to calculate the uncertainty intervals for the various components of GBD 2010 are provided in the relevant chapters of this working paper.

The calculation of uncertainty has the potential to add significant value to the outputs from burden of disease analysis. However, it has added complexity to the methodology and, as noted below, may require different approaches for the Indigenous and the national components of this study.

Key issues in the Indigenous context

In addition to the impact of the methodology changes highlighted above, several additional factors need to be considered when calculating burden of disease estimates for Aboriginal and Torres Strait Islander people.

1. Health statistics for Indigenous Australians are subject to a number of issues that do not affect data for the total Australian population, for example:
 - not every data collection has an Indigenous status flag to enable analysis for the Indigenous population. Similarly, many epidemiological studies do not provide separate results for the Indigenous and non-Indigenous populations, making it hard to apply their results to Indigenous burden of disease calculations.
 - for administrative data collections containing an Indigenous flag, studies have found an under-identification of Indigenous people. These include some important data sources for the Australian burden of disease study, such as hospital data, mortality data and Medicare data.

For this reason, the estimates of uncertainty intervals and methods for estimating inputs for the Indigenous burden of disease study could be different from the national study, to the degree the particular data source allows.

2. Indigenous Australians make up a very small proportion (around 2.5-3%) of the total Australian population, and around a quarter live in remote areas. The small size and geographic distribution of the Indigenous population may require specifically-designed methods to produce prevalence estimates and produce uncertainty intervals. Testing could be done to assess whether this is required.
3. In addition, Aboriginal and Torres Strait Islander peoples hold a 'whole-of-life' view on health that recognises that achieving optimal conditions for health and wellbeing requires a holistic view that encompasses the social, emotional and cultural wellbeing of the whole community. This differs in concept from GBD 2010's disability weights, which are focussed on health (rather than welfare) loss. Indigenous health state preferences may therefore differ somewhat from those used in GBD 2010. Whether the Indigenous burden of disease study will require different disability weights from the national study could be investigated.

These issues, as they relate to various components of the burden of disease analysis, are explored in greater depth in the following chapters.

3 Cause and risk factor lists

The cause and risk factor lists are the primary organisational framework for producing burden of disease estimates. The two lists are related and decisions made regarding one may result in changes to the other.

The **cause list** (also referred to as condition list) details the specific diseases and causes of injury (or groups thereof) for which estimates of numbers of deaths, years of life lost (YLL), years lived with disability (YLD) and disability-adjusted life years (DALY) will be made. It is, in principle, a set of mutually exclusive and collectively exhaustive categories, with all fatal and non-fatal health outcomes being attributed to an item on the list.

The **risk factor list** details the specific risk factors to be investigated as underlying causes of the estimated burden through their causal association with particular diseases and causes of injury. In contrast to the cause list, which is exhaustive and where an established classification system (the ICD) exists, the list of potential risk factors is near limitless and there is often no consensus in the literature on what level(s) of exposure constitute 'risk'. A predetermined set of criteria is used to guide the development of the list, taking into account the potential for modification of exposure in the population, the availability of data on exposure, and the quality of evidence about causal effects and their magnitude.

This chapter provides an overview of the cause and risk factor lists used in GBD 2010, outlines the main differences from those used in the previous Australian and Indigenous burden of disease studies, highlights some of the key issues for consideration, and outlines the intended approach for the national and Indigenous components of the Australian Burden of Disease Study.

Key terms used in this chapter

attributable burden - the disease burden attributed to a particular risk factor. It is the reduction in burden that would have occurred if exposure to the risk factor had been avoided

effect size - a statistical measure of the strength of the relationship between two variables, which is relatively independent of sample size, for example, the relative risk or the odds ratio

health states - groupings of sequelae that reflect key differences in symptoms and functioning

relative risk - the risk of an event relative to exposure, calculated as the ratio of the probability of the event occurring in the exposed group to the probability of it occurring in the non-exposed group

risk factor - any factor which represents a greater risk of a health disorder or other unwanted condition or event

social determinants of health - the economic and social conditions (such as income, level of education and employment status) that influence health status

(Refer to glossary for a full list)

Disease and injury causes

The GBD 2010 cause list

The GBD 2010 project uses a hierarchical list of disease and injury causes and their sequelae, which can be considered at five different levels (see Table 3.1). Levels 1 to 4 represent diseases and causes of injury that can be directly related to the International Statistical Classification of Diseases and Related Health Problems, version 10 (ICD-10). Level 5 consists of sequelae, that is, conditions or injuries resulting from events/diseases at the levels above. At each level, the list is mutually exclusive and exhaustive. Although each of the levels 2 to 4 are a disaggregation of the level above, level 5 is different in that a single sequela may relate to more than one cause (see section on sequelae and health states below). Many of the level 3 causes have no level 4 causes below them in the hierarchy, leading to a total of 241 mutually exclusive and exhaustive categories at the lowest level of disaggregation of causes.

The GBD 2010 cause list was developed by taking the previous GBD cause list (GBD 2000) and adding causes on the basis of three criteria:

- potentially large causes of burden
- causes for which there is substantial health policy interest
- feasibility of measurement.

In GBD 2010 published material, results are provided for all causes at levels 1 to 4, that is, for a total of 291 cause categories (including both individual causes and groups of causes).

Table 3.1: GBD 2010 cause-sequelae list: examples of hierarchical structure

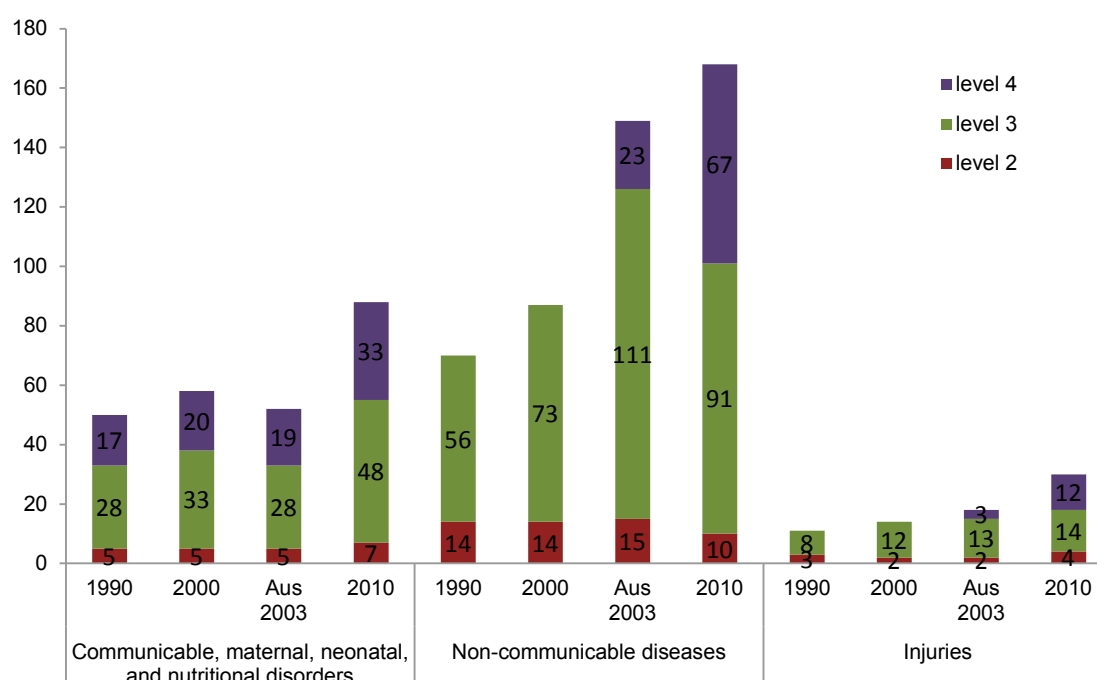
Level	Number of entries	Example 1	Example 2	Example 3
1	3	Communicable, maternal, neonatal and nutritional disorders	Non-communicable diseases	Injuries
2	21	Maternal disorders	Cardiovascular and circulatory diseases	Transport injuries
3	155	Maternal haemorrhage	Ischaemic heart disease	Road injury
4	112	[no further disaggregation]	[no further disaggregation]	Pedestrian injury by road vehicle
5 (sequelae)	1,160	Moderate anaemia due to maternal haemorrhage	Myocardial infarction due to ischaemic heart disease	Fracture of radius or ulna due to pedestrian injury by road vehicle

Source: Murray et al. 2012, supplementary appendix.

It should be noted that the GBD cause list is subject to annual revision by the GBD project team at the Institute for Health Metrics and Evaluation (IHME). Minor changes have been flagged for 2013 with a major revision for 2015 that will result in 400+ causes. Changes to the cause list are being managed by IHME through a defined protocol.

Sequelae and health states

The GBD 2010 list contains substantially more causes than in previous studies (Figure 3.1). In particular, the number of sequelae has more than doubled – 1,160 in GBD 2010 compared with 474 in GBD 2000.



Note: The 2003 Indigenous burden of disease study used the same cause list as the Australian study.

Sources: Begg et al. 2007; Web Figure 1, supplementary appendix, Murray et al. 2012.

Figure 3.1: Number of causes in the various revisions of the GBD cause list and 2003 Australian burden of disease study cause list, disaggregated by different levels in the cause hierarchy

Sequelae for each condition were developed by disease and injury experts, and included 'main outcomes from a disease that could potentially make an important contribution to the burden of a given disease or injury and which could in principle be measured' (Murray et al. 2012). In some cases, an outcome could reasonably be considered a sequela of one disease or a disease in its own right—for example, chronic kidney disease secondary to diabetes. In other cases, the same outcome appears several times because more than one disease contributes to a clinical outcome—for example, heart failure is treated as a sequela of a large number of diseases rather than a disease entity of its own. In order to avoid double counting of the associated burden, each outcome appears only once in the final cause-sequelae list.

A major innovation for GBD 2010 has been to identify which sequelae are common across conditions; these common sequelae are called **health states**. The GBD 2010 study identified 220 health states, to which each of the 1,160 sequelae can be mapped. This shorter list of health states facilitates the disability weight measurement component of the study (see Chapter 5) and allows estimation of the total burden of specific health states that may be caused by different conditions. For example, the sequelae 'heart failure due to iron deficiency anaemia', 'heart failure due to ischemic heart disease' and 'heart failure due to obstructive pulmonary disease' all map to the same health states: 'Heart failure: mild/moderate/severe'.

How the cause list is used

Mortality

Calculation of mortality and years of life lost (YLL) due to premature mortality are based on the causes at the lowest level of disaggregation—that is, level 3 or (where existing) 4 in the hierarchy. The sequelae (level 5) are not used for mortality analysis. For each cause that is a

‘valid’ cause of death (according to the ICD coding rules and GBD criteria for specificity), the number of deaths in each age and sex group are estimated and used to calculate death rates and YLL. In the GBD 2010 list, 191 of the 241 individual causes were considered valid causes of death, with mortality results being published for a total of 235 of the 291 cause categories (that is, individual causes and groups of causes).

More detail relating to mortality, causes of death and YLL calculations is in Chapter 4.

Morbidity

The sequelae are the key component of the list for the GBD 2010 morbidity analyses – that is, calculation of healthy life expectancy (HALE) and years lost due to disability (YLD). Both of these calculations are based on estimating the prevalence of each of the 1,160 sequelae and then building YLDs and HALE for each cause at levels 1 to 4 by combining the results for the relevant sequelae.

More detail relating to morbidity calculations is included in Chapter 5.

Risk factor attribution

Attribution of burden of disease to various risk factors requires these risks to be causally linked to one or more conditions in the cause list (risk-outcome pairs). More detail on the comparative risk assessment methodology is included in Chapter 6.

Strengths and weaknesses of the GBD 2010 approach

The large number of causes included in GBD 2010 provides information about a wider range of individual diseases, causes of injury and sequelae than was previously available. This is important for the prioritisation of existing treatments or prevention efforts and the development of new ones. For example, separating out various diarrhoeal diseases shows where rotavirus vaccinations may be most effective, and more detail in causes of road injury highlights the split between pedestrian and vehicle occupant deaths and the influence of road safety measures.

Further, the increased granularity in some disease groups, such as mental disorders, means that results are based on prevalence data for a larger number of individual disorders, rather than prevalence data for a small number of major disorders and crude estimates for a ‘residual’ group. This should provide more reliable estimates of the YLDs and DALYs for these types of conditions.

However, the need to estimate mortality and prevalence at such a low level leads to challenges in terms of data availability, particularly at the global level, with adjustments and modelling needing to be used to overcome missing or inadequate data. These models may be extremely difficult to validate. In particular, allocation of cause of death may be difficult or unreliable at lower levels.

Differences from the 2003 Australian studies

As shown in Figure 3.1, the number of causes in GBD 2010 was considerably greater than in the 2003 Australian and Indigenous studies. The latter studies included 186 causes at the lowest level of disaggregation, compared with 241 in GBD 2010.

Applicability of GBD 2010 cause list to the Australian context

Prior to being used in any Australian Burden of Disease Study, the GBD 2010 cause list needs to be assessed for its applicability to the Australian and Indigenous contexts. The basis for this assessment is outlined below, the result of which may indicate changes needed to the existing list of diseases and injuries.

To systematically and transparently manage any changes to the GBD 2010 cause list for use in the Australian Burden of Disease Study, a defined protocol would be useful. This would be developed in consultation with the Expert Advisory Group, and other reference groups established for the project and can refer to criteria used in GBD 2010 and the recent New Zealand BoD study as a starting point.

Key issues at the national level

Relevance

Are all of the causes relevant for Australia?

The GBD 2010 cause list is comprehensive and in all likelihood covers the vast majority of conditions and injuries of relevance for Australia. It does, however, include as individual causes (that is, a category at level 3 or 4) a number of conditions that are extremely rare in Australia – for example, cholera, diphtheria, tetanus, rabies, yellow fever and leprosy. While these causes may have little relevance to Australia, deletion of individual causes will affect estimates for other causes disrupting comparability with the GBD 2010 estimates, unless they are combined with causes within the same hierarchical grouping.

YLDs for causes such as these, which are considered absent in Australia, would be considered to be zero if an incident model was used. This is consistent with previous Australian and international studies. Using a prevalence-based YLD model, however, (as GBD 2010 does), means that DALY estimates include YLDs for diseases for which there were no incident cases in the reference year but which were the cause of lost health (that is, ongoing disability) experienced in that year. This would be the case now in Australia for polio, for example. This means that prevalent YLDs for conditions where no new cases are now occurring will not necessarily be zero, as there may be ongoing disability from cases that occurred in previous years. More detail about the difference between incidence- and prevalence-based models and the impact on YLD calculation is included in Chapters 2 and 5.

Are there additional causes relevant to Australia that meet the inclusion criteria?

Conversely, there may be specific conditions that are considered important in the Australian context (such as Indigenous health issues or topics of public health policy interest) but which in GBD 2010 were part of a grouped cause category. For example, the 2003 Australian studies produced separate estimates for Type 1 and Type 2 diabetes, whereas GBD 2010 did not. In addition, mesothelioma and cystic fibrosis (two conditions of public health interest in Australia) are included in their respective ‘other’ categories (*Other malignant neoplasms* and *Other endocrine, nutritional, blood and immune disorders*) in the GBD cause list.

Amount of granularity

What is the effect of the increased number of causes on results at higher levels?

There are flow-on effects from adding a cause to the cause list. Because the burden of disease aims to capture all ill-health, the addition of a cause to the list is actually splitting an existing cause creating two new causes: the specific cause and the residual of the original cause. Data must therefore exist for the specific cause and also for the residual cause category.

The addition (or rather, separation out) of new causes at the lower levels has the effect of reducing the size of the residual categories that are used to hold conditions that are rare or for which data are limited. The obvious outcome is that results for these residual categories will not be comparable to previous studies even if the same methodology is used for the DALY calculations. What is less clear is the effect this increased granularity has on the results at higher levels of aggregation.

For example, if a condition that would otherwise be in the residual category 'Other mental and behavioural disorders' is added to the list at level 3 or 4, then the results for the level 2 cause 'Mental and behavioural disorders' include specific prevalence estimates, YLD and (if relevant) YLL calculated for that condition. These estimates are (presumably) more accurate than the estimates made for the residual categories, which are generally made using YLD to YLL ratios for similar or related conditions (often a group of others aggregating to the same level 2 cause). This ratio estimation method means that in many cases the change to using specific prevalence estimates will lead to an increase in the DALYs at higher levels of aggregation, but in some to a decrease, depending on the distribution of deaths across categories and the relative 'importance' of the new condition within the group as a cause of death and disability.

Data availability

What if data for Australia are not available for a particular cause included on the list?

There may be causes included in the GBD 2010 list which are considered relevant to the Australian context but for which good quality prevalence data may not be available. In such cases, an option is to use the Australian estimate from GBD 2010 as the default prevalence estimate.

More information on potential data sources for YLL and YLD calculation is in Chapter 7.

What will be the impact of Australian modifications to the GBD 2010 list?

Australian modifications to the analysis cause list will need to take into account the following consequences that may result in differences between the Australian estimate and the GBD estimate:

1. In mortality analysis, the addition of a cause may affect the redistribution of causes of death not relevant to burden of disease studies, thereby potentially changing the YLL estimate for other causes in the list. In morbidity analysis, it is necessary to determine the cause specific sequelae that will allow the cause to map to an existing health state/disability weight combination.
2. Additions of new causes will impact the estimates of other causes within the same hierarchical grouping. Decisions as to whether to separate out such conditions as individual causes, most likely at level 4 in the hierarchy, would need to be made on a case-by-case basis, taking into account data availability (including disability weights) and the likely quality and usefulness of the resulting estimates. The associated tasks of defining relevant sequelae, mapping these to appropriate health states and identifying or developing associated disability weights also need to be considered.

Use of ICD-10-AM

Can the cause list (developed using ICD-10) be mapped easily to/from ICD-10-AM? How should residual ICD-10-AM codes be allocated? How can any ICD-10-AM codes mapping to more than one cause be reconciled?

ICD-10-AM is the coding system used to classify diseases and problems in Australian inpatient episodes. It is the Australian modification of the WHO ICD-10 system, and has been developed to serve particular needs. An ICD-10 cause list for GBD 2010 is available (see Lozano et al. 2012 supplementary appendix; Web Table 3) and can form the basis of an ICD-10-AM mapping. Standard concordances between ICD-10 and the various editions of ICD-10-AM exist to facilitate this. As was done for previous Australian studies, residual ICD-10-AM codes can be manually allocated to appropriate cause categories. The advice of coding experts may be required to determine the most appropriate allocation of codes that do not map directly to a single ICD-10 code or which appear to map to multiple cause categories.

Injury listed by nature of injury or external cause of injury

GBD 2010 used the external cause of injury (i.e. transport accident, self-harm, falls) for injuries in the cause list and risk factor list, and the nature of injury (i.e. fractured femur) as the sequelae. The 2006-2010 New Zealand BoD study (see Appendix A) used the nature of injury for injuries in the cause list and external causes in the risk factor list. The most appropriate approach for the Australian BoD studies would need to be determined, in consultation with the Expert Advisory Group and relevant experts in the injury field.

Specific issues relating to Aboriginal and Torres Strait Islander people

Relevance

Are there some causes which are not relevant to Indigenous Australians?

In the 2003 Aboriginal and Torres Strait Islander Australian burden of disease study (Vos et al. 2007), several outcomes were excluded as they were assumed to be absent in the Indigenous Australian population. These were:

- diphtheria (also assumed absent in the non-Indigenous population)
- poliomyelitis (also assumed absent in the non-Indigenous population)
- rubella
- occupational overuse syndrome, and
- other conditions arising in the perinatal period.

In addition, trachoma was assumed present only in the remote population.

The status of each of these conditions within the Indigenous population would need to be reviewed and, if required, consultation with the Expert Advisory Group and Indigenous Reference Group undertaken so that informed decisions as to the inclusion or exclusion of these and other rare or potentially absent causes can be made.

As noted above, an incident-YLD model was used in the 2003 study, which did not take into account the effects of past cases on current and future disability. This will need to be kept in mind for a prevalent-YLD model (as used in GBD 2010).

Are there additional causes that are particularly relevant to Indigenous Australians and that meet the inclusion criteria?

As noted above, there may be some causes of relevance to Indigenous health that were included in GBD 2010 as part of a grouped cause category – for example, Type 2 diabetes. Consultation with the Expert Advisory Group and Indigenous Reference Group will help to identify such causes. If data are available to produce prevalence estimates of reasonable quality, or to model them, it may be possible to separate these causes out and produce individual estimates of disease burden. As noted above, however, adding a cause to the list also requires defining relevant sequelae, mapping these to existing health states and defining an appropriate disability weight if an appropriate health state does not exist. This will need to be considered on a case-by-case basis. Alternatively, as noted above, disaggregation of results may be possible.

Any new causes added to the list should also feature in the national estimates to ensure comparability of estimates for causes at higher levels.

Data availability

While it is expected that data for the Indigenous population will be available for many causes included in the GBD 2010 list, there will be some causes for which data are not available. In such cases, simple modelling techniques may need to be used to generate Indigenous estimates (see Chapter 7 for examples of such techniques). The methods used in the 2003 Indigenous study and the recent New Zealand BoD study will provide a good starting point for this.

What about small numbers for some conditions?

There may be a need to group some causes together to report at a higher level than used in GBD 2010 for the Indigenous study to overcome small number issues. As well as the possibility of making changes to the analysis cause list, there is also the option of making changes to the reporting list. It may therefore be the case that, while the same analysis list is used for the Australian and Indigenous studies, the reporting lists may differ.

Specific issues relating to sub-national estimates

Capacity for disaggregation

Do the data sources allow disaggregation? If not, is sub-national estimation based on population distribution valid?

Many of the national Australian data sources include variables that enable disaggregation at various levels (for example, by geographic region or socioeconomic status). However, depending on the data source and condition, estimates at sub-national levels may be subject to considerable uncertainty. Decisions as to whether valid estimates are able to be produced at all the desired levels of disaggregation would need to be made on a case-by-case basis as the individual data sources and causes are investigated. In some cases it may be necessary to produce estimates at higher levels of aggregation of causes or larger geographic areas rather than at the lowest levels.

Summary

The GBD 2010 cause list is substantially larger than in previous studies, and appears comprehensive. However, there are some cases where finer granularity would better meet the needs of Australian policy-makers and planners, and for which Australian data are

available for analysis—for example, producing separate estimates for Type 1 and Type 2 diabetes. This could be achieved either by modifying the cause list or by post-hoc adjustment or disaggregation of the results. Conversely, there are causes in the GBD 2010 list which are rare in Australia or for which good quality prevalence data may not be available. In such cases, an option is to use the Australian estimate from GBD 2010 as the default prevalence estimate.

Risk factors

The GBD 2010 risk factor list

The risk factor list for GBD 2010 is also hierarchical, and includes a total of 69 risks (or clusters of risks) (see appendix Table 3A.1). The list was compiled by taking the risk factors used in the Comparative Risk Assessment 2000 study (Ezzati et al. 2002) and adding or subtracting items based on four criteria:

- convincing or probable evidence on relative risks by cause
- sufficient data to estimate exposure by region
- potential size of the attributable burden of disease
- the extent to which the burden is modifiable.

Key additions in the GBD 2010 list include breastfeeding (for its effects on the child), bone mineral density, intimate partner violence and a number of specific occupational exposures. Both bone mineral density and intimate partner violence were included in the 2003 Australian burden of disease study. The category of dietary exposures has been considerably expanded to cover 14 components of diet, compared with a single dietary risk factor of 'low fruit and vegetable intake' in previous work.

A number of risk factors used in the Comparative Risk Assessment 2000 study have been dropped from the GBD 2010 list due to insufficient direct measurement of exposure at the regional level, or inadequate information on relative risks or effect sizes. These include 'unsafe sex', 'unsafe health-care injections', 'global climate change' and 'lack of contraception'. Of these, only 'unsafe sex' was included in the 2003 Australian study.

At level 1 in the GBD 2010 risk factor hierarchy are groups of risk factors that are related by mechanism, biology, or potential policy intervention—for example, 'Tobacco smoking, including second-hand smoke'. Level 2 is the level at which most risk factors are presented, but in two cases (suboptimal breastfeeding and occupational carcinogens) a third level of disaggregation is provided to enable calculation of detailed, but relatively small, risk estimates (see Table 3.2). In total there are 57 individual risk factors at the lowest level of disaggregation. In practice, GBD 2010 results are presented for 67 risks or clusters of risks. Combined estimates are not provided for the level 1 groups 'air pollution' and 'physiological risk factors', since the risk factors in these groups do not operate independently of each other—for example, the effects of high body mass are partly mediated through high fasting plasma glucose, and solid fuels used for cooking contribute to both household air pollution and ambient particulate matter pollution.

Table 3.2: GBD 2010 risk factor list: examples of hierarchical structure

Level	Number of entries	Example 1	Example 2	Example 3
1	10	Physiological risk factors	Dietary risk factors and physical inactivity	Occupational risk factors
2	44	High blood pressure	Diet low in calcium	Occupational carcinogens
3	15	[no further disaggregation]	[no further disaggregation]	Occupational exposure to asbestos

Source: Murray et al. 2012, supplementary appendix.

Results for the GBD 2010 risk factors analysis (Lim et al. 2012) represent the effects of each individual risk factor or cluster of risk factors, holding all other independent factors constant. The effect of multiple risk factors is not accounted for, meaning that the sum of the effects of individual factors on a particular outcome may be considerably greater than 100% of total deaths and DALYs. For example, adding the proportion of ischaemic heart disease DALYs attributed to each individual risk factor associated with that condition gives a total of 442% of DALYs accounted for. The GBD 2010 study did not produce any estimates of the overall (joint) burden of all of the risk factors included in the study.

How the risk factor list is used

Calculations for attributable burden are based on estimates of exposure to each of the risk factors at the lowest level of disaggregation – that is, levels 2 or (where existing) 3. More detail on the comparative risk assessment methodology is included in Chapter 6.

Strengths and weaknesses of the GBD 2010 approach

As with the cause list, the expanded number of risk factors provides information on a wider range of issues than was previously available. However, this brings challenges relating to data availability, and the need for imputation and estimation methods to deal with missing or incomplete data. Further, data on direct exposure is often not available, and indirect measures need to be used, potentially leading to considerable uncertainty around the estimates.

The strict criteria for inclusion of risk-outcome pairs in GBD 2010 (see Chapter 6) increases robustness of the estimates, but means that some pairs included in previous studies were excluded from GBD 2010. In some cases this meant omitting factors which probably account for a considerable amount of the global health burden – for example, unsafe sex as a cause of HIV.

More information about the strengths and weaknesses of the GBD 2010 approach to the calculation of attributable burden are included in Chapter 6.

Differences from the Australian 2003 studies

The 2003 Australian and Indigenous burden of disease studies considered 14 individual risk factors, compared with 57 in GBD 2010. The major differences are the expansion of the dietary risks, from a single category of ‘low fruit and vegetable consumption’ in previous Australian and global studies to a set of 14 individual dietary components in GBD 2010, and the publication of estimates for 18 specific occupational risks, rather than a combined estimate for ‘occupational exposures and hazards’.

Applicability of GBD 2010 risk factor list to the Australian context

Prior to being used in any Australian Burden of Disease Study, the GBD 2010 risk factor list needs to be assessed for its applicability to the Australian and Indigenous contexts. The basis for this assessment is outlined below, the result of which may indicate changes needed to the existing list of risk factors.

To systematically and transparently manage any proposed changes to the GBD 2010 risk factor list for use in the Australian Burden of Disease Study, a defined protocol is required. This would be developed in consultation with the Expert Advisory Group and other relevant reference groups established for the project, and can refer to criteria used in GBD 2010 and the recent New Zealand BoD study as a starting point.

Key issues at the national level

Relevance

Are each of the GBD 2010 risk factors relevant for Australia?

The majority of risk factors in the GBD 2010 list are relevant for Australia, however there are some for which exposure is low at the national level. These include:

- unimproved water source
- unimproved sanitation
- household air pollution from solid fuels
- residential radon.

These risk factors could potentially be dropped from the national list, however, their applicability in the Indigenous context needs to be considered.

Are there additional risk factors relevant to Australia that meet the inclusion criteria?

The 2003 Australian burden of disease study estimated the burden attributable to unsafe sex by calculating the proportion of various outcomes (such as HIV and hepatitis) attributed to specific exposures in relevant national data collections. As noted above, this risk factor was excluded from GBD 2010 due to insufficient exposure data available at the global level or a way of calculating the proportion of HIV cases attributed to unsafe sex by country. As the data collections used in the 2003 Australian studies still exist, this analysis could be replicated for the current Australian study. In addition, the Second Australian Study of Health and Relationships currently underway should provide representative information about unsafe sex. These data are expected to become available in late 2014.

Access to health services (or lack of) is of particular relevance in Australia and could be assessed against the inclusion criteria developed.

What about social determinants of health

Neither the GBD nor previous Australian studies included social determinants of health, such as income, level of education or socioeconomic status, although the Comparative Risk Assessment study 2000 did examine the distribution of risk by poverty (Blakely et al. 2004). Some of these factors have a strong correlation with health outcomes, and are particularly relevant for Australia's Indigenous population.

Social determinants were not included in GBD 2010 'largely because of the study requirement that evidence meet the World Cancer Research Fund criteria of convincing or probable evidence and that epidemiological data be available to estimate effect sizes for a

risk factor on specific causes of death or disability' (US Burden of Disease Collaborators 2013, p.E15). Much of the available evidence relates social determinants of health to all-cause rather than specific outcomes. Further, due to the complex way these factors affect health, results are often inconsistent with widely varying effect sizes.

Although good quality, recent data on at least some of the social determinants of health are available in Australia, the same issues as faced in GBD 2010 regarding limited evidence of outcomes and effect sizes are likely to apply in the Australian context. A more detailed investigation of this issue is planned, including discussion with the Expert Advisory Group and Indigenous Reference Group. If there are not sufficient data to enable inclusion of social determinants as a risk factor, an alternative option is for social determinants to be incorporated in the study as disaggregating variables in the reporting of DALYs. For example, differentials in DALY rates for selected relevant social determinants of health (e.g. socioeconomic group, employment status, highest year of school completed and housing tenure) could be reported to show differences in the ranking and age-standardised rates of DALYs by disease groups and/or risk factors. The 2003 Australian BoD study reported differentials in the ranking and rates of DALYs by socioeconomic quintile.

Data availability

Are Australian exposure data available for each of the GBD2010 risk factors?

The 2011–12 Australian Health Survey should provide national information for many of the GBD 2010 listed risk factors, excluding the occupational and environmental risks. Other large-scale surveys such as the National Infant Feeding Survey, the Longitudinal Study of Australian Children, the 45 and Up study and the Australian Longitudinal Study on Women's Health will also provide good information for several risk factors.

Data on exposure for at least some of the remaining risks may be obtained from other sources such as the state and territory Environmental Protection Agencies, the National Hazard Exposure Worker Surveillance Survey, the Personal Safety Survey and national child protection data.

Those risk factors included on the GBD 2010 list which do not have adequate exposure data but are deemed relevant to the Australian context could either be excluded from the Australian modified risk factor list, or modelling techniques employed (method to be determined) to derive national exposure estimates. Advice from the Expert Advisory Group and/or relevant reference groups would be sought to inform these decisions.

More information on potential data sources for risk factors is in Chapter 7.

Specific issues relating to Aboriginal and Torres Strait Islander people

Relevance

Are each of the GBD 2010 risk factors relevant to Indigenous Australians?

It is likely that all of the risk factors deemed relevant to the national population will also be relevant to the Indigenous population, though the level of exposure and the scale of the impact of each risk factor will vary between the two groups. However, there may be some factors that are relevant to the Indigenous population but which are not relevant at the national level and may be excluded in the national study – for example, unimproved water supply and unimproved sanitation. The extent to which Aboriginal and Torres Strait Islander people are affected by these risk factors would need to be determined to decide whether they should be included in the Indigenous component of the current study.

Are there additional risk factors relevant to Indigenous Australians that meet the inclusion criteria?

Smoking during pregnancy is considerably more common among Indigenous women compared with other Australian women. It was included within the 'tobacco' risk factor in the 2003 Australian studies, and associated with the outcome of low birthweight. However, this was not included as an outcome in GBD 2010, with the relevant ICD-10 codes being split across the cause categories 'Preterm birth complications' and 'Other neonatal disorders'. Exposure data on smoking during pregnancy are available from the National Perinatal Data Collection, and relative risk estimates have been published. The creation of 'low birthweight' as a cause, with the resulting change to the two GBD 2010 cause categories noted above, would need to be considered in consultation with the Expert Advisory Group.

Petrol sniffing has been, and is still, a considerable problem in some Indigenous communities. The 2003 Aboriginal and Torres Strait Islander burden of disease study did not include this as a risk factor 'because of the problems estimating the prevalence of this condition and the uncertainty about the exact nature of long-term harms' (Vos et al. 2007, p.25). Investigation would be needed as to whether adequate exposure data and robust relative risk estimates can be obtained for inclusion of 'petrol sniffing' or the broader category of 'inhalant use' as a risk factor in the Indigenous component of the current study.

As for the total population (and discussed under 'issues at the national level'), social determinants of health and access to health services are risk factors that are also relevant for the Indigenous population and could be investigated for their potential inclusion in the Indigenous BoD study (either as risk factors if there is adequate evidence for their inclusion or as disaggregating factors for the results). Aboriginal and Torres Strait Islander people often experience considerably greater social and economic disadvantage compared with other Australians. Access to health services is also a major issue for Indigenous people, both for those in remote areas who may be physically distant from relevant services and for those in urban locations who may have trouble accessing appropriate services.

Data availability

Are adequate exposure data for Indigenous Australians available?

The 2012–13 Australian Aboriginal and Torres Strait Islander Health Survey should provide reasonable information on exposure for many of the risk factors included in GBD 2010. Previous national surveys, such as the National Aboriginal and Torres Strait Islander Social Survey, will provide additional data, which can also be supplemented by sub-national surveys or Indigenous-specific studies such as the Longitudinal Survey of Indigenous Children.

Information on certain environmental exposures (air pollution, lead exposure) should be able to be estimated in the same manner as for the national study, since these estimates are made on a geographical basis and not at the person level. Specific sources for data on exposure to unimproved water and sanitation would need to be identified should these factors be considered relevant for inclusion in the Indigenous study (see above).

Data on occupational exposures may be difficult to obtain as the major data sources do not have an Indigenous identifier.

Specific issues relating to sub-national estimates

Capacity for disaggregation

Do the exposure data allow disaggregation? If not, is sub-national estimation based on population distribution valid?

As for the cause data, many of the national Australian risk factor data sources include variables that enable disaggregation at various levels (for example, by geographic region or socioeconomic status). However, depending on the data source and risk factor in question, estimates of exposure at sub-national levels may be subject to considerable uncertainty. As was done for GBD 2010 and as suggested above for the morbidity data for calculating YLDs, modelling techniques may be able to be used to impute missing or incomplete data if sufficient information is available to support this (for example, proxy measures, good-quality epidemiological studies and/or relevant covariates to inform estimation). Decisions as to whether valid estimates are able to be produced at all the desired levels of disaggregation would need to be made on a case-by-case basis, in consultation with the Expert Advisory Group.

Summary

A substantially expanded list of risk factors was included in GBD 2010 compared with previous global and Australian studies. This provides more detailed information but brings challenges relating to data availability. As a result, it may not be possible to include all of the GBD 2010 risk factors in the national and Indigenous components of the study. Conversely, it may be possible for the Australian and Indigenous studies to consider risk factors (and therefore outcomes) that were not included in GBD 2010 because of a lack of adequate data at the global level.

Intended approach

The following high-level decisions have been identified which are relevant to both the national and Indigenous components of the Australian Burden of Disease Study, to be discussed with the Expert Advisory Group and the Indigenous Reference Group where relevant.

Cause list

- Adopt the GBD 2010 cause list as a starting point for an Australian-modified cause list that will be used for both the Indigenous and national studies. This will maintain comparability with GBD 2010 as far as practicable, while ensuring the cause list is appropriate to the Australian and Indigenous contexts, and meets needs of the Australian policy and research communities.
- Modifications to the GBD 2010 list to be considered on a case-by-case basis. A protocol for systematically and transparently managing proposed changes to the GBD 2010 cause list for the Australian and Indigenous context should be developed and implemented in consultation with the Expert Advisory Group and Indigenous Reference Group established for the project.
- Consider whether to group injuries by nature of injury or external cause of injury in the Australian modified cause list, in consultation with the Expert Advisory Group and relevant experts in the injury field.

Risk factor list

- Adopt the GBD 2010 risk factor list as a starting point for an Australian-modified risk factor list, and an Indigenous-modified risk factor list.
- Modifications to the GBD 2010 risk factor list to be considered on a case-by-case basis for the Australian and Indigenous studies. A protocol for systematically and transparently managing proposed changes to the GBD 2010 risk factor list for the Australian and Indigenous contexts should be developed and implemented in consultation with the Expert Advisory Group and Indigenous Reference Group established for the project.
- Undertake work to investigate whether there is enough evidence to include social determinants and access to health services as risk factors in the Australian study. If not, then consider the feasibility of incorporating these risk factors in the analysis as factors by which the results are disaggregated.

Additional considerations for the Indigenous component:

- Undertake further investigation of the causes of disease and risk factors included in GBD 2010 that are not relevant to Indigenous Australians, and those that may be relevant which were excluded from GBD 2010 to inform modifications to the cause and risk factor lists for the Australian BoD study.

Appendix 3A

Table 3A.1: GBD 2010 hierarchical risk factor list

Level	Risk factor	Level	Risk factor
1	Unimproved water and sanitation	2	Diet low in milk
2	Unimproved water source	2	Diet high in red meat
2	Unimproved sanitation	2	Diet high in processed meat
1	Air pollution	2	Diet high in sugar-sweetened beverages
2	Ambient particulate matter pollution	2	Diet low in fibre
2	Household air pollution from solid fuels	2	Diet low in calcium
2	Ambient ozone pollution	2	Diet low in seafood omega–3 fatty acids
1	Other environmental risks	2	Diet low in polyunsaturated fatty acids
2	Residential radon	2	Diet high in trans fatty acids
2	Lead exposure	2	Diet high in sodium
1	Child and maternal under-nutrition	2	Physical inactivity and low physical activity
2	Suboptimal breastfeeding	1	Occupational risk factors
3	Non-exclusive breastfeeding	2	Occupational carcinogens
3	Discontinued breastfeeding	3	Occupational exposure to asbestos
2	Childhood underweight	3	Occupational exposure to arsenic
2	Iron deficiency	3	Occupational exposure to benzene
2	Vitamin A deficiency	3	Occupational exposure to beryllium
2	Zinc deficiency	3	Occupational exposure to cadmium
1	Tobacco smoking, including second-hand smoke	3	Occupational exposure to chromium
2	Tobacco smoking	3	Occupational exposure to diesel engine exhaust
2	Second-hand smoke	3	Occupational exposure to second-hand smoke
1	Alcohol and drug use	3	Occupational exposure to formaldehyde
2	Alcohol use	3	Occupational exposure to nickel
2	Drug use	3	Occupational exposure to polycyclic aromatic hydrocarbons
1	Physiological risk factors	3	Occupational exposure to silica
2	High fasting plasma glucose	3	Occupational exposure to sulphuric acid
2	High total cholesterol	2	Occupational asthmagens
2	High blood pressure	2	Occupational particulate matter, gases and fumes
2	High body mass index	2	Occupational noise
2	Low bone mineral density	2	Occupational risk factors for injuries
1	Dietary risk factors and physical inactivity	2	Occupational low back pain
2	Diet low in fruits	1	Sexual abuse and violence
2	Diet low in vegetables	2	Childhood sexual abuse
2	Diet low in whole grains	2	Intimate partner violence
2	Diet low in nuts and seeds		

4 Mortality—Years of life lost

What does it measure and why is it important?

YLL measures the years of life lost due to premature deaths; that is, the fatal component of burden of disease. It is calculated by counting the number of deaths at each age, multiplied by the remaining life expectancy at this age according to a reference life table.

In order to estimate YLL, a critical step is to compile the numbers of deaths by age and sex. The total number of deaths from all-causes combined is known as the ‘mortality envelope’ (see also Box 4.1). This total is then broken down by cause of death for each age group and sex to determine the burden of premature mortality due to specific causes.

This chapter reviews the data sources and methods used in estimating YLL in the GBD 2010 study in comparison with past burden of disease studies in Australia. The GBD methods are assessed for their application in the Australian context, and in regard to estimating YLL for Indigenous Australians. It then outlines the intended approach for the calculation of YLL estimates in the national and Indigenous components of the Australian Burden of Disease Study.

Key terms used in this chapter

direct methods – direct methods measure directly the degree of under-coverage in death registration data through, for example, data linkage processes. In contrast to indirect methods, direct methods do not require assumptions.

‘garbage codes’ – in burden of disease studies, a cause of death which cannot or should not be considered as an underlying cause of death, or which relates to an intermediate cause in the chain of events leading to death, or which provides insufficient detail to ascertain an appropriate cause.

indirect methods – indirect methods measure the degree of under-coverage in death registration data by reconciling deaths over a time period to population estimates in the two end time points.

mortality envelope – the total number of deaths from all causes.
(Refer to glossary for a full list)

Overview of GBD 2010 data sources and methods

Compiling numbers of deaths

GBD 2010 estimated YLLs for 187 countries with varied data sources, collection methods and data quality issues. These are discussed in detail below.

Cause list

Calculation of mortality and years of life lost (YLL) due to premature mortality in GBD 2010 was based on causes at the lowest level of disaggregation—that is, level 3 or (where existing) 4 in the GBD 2010 hierarchy. The sequelae (level 5) are not used for mortality analysis (see Chapter 3 for a description of cause levels).

For each cause in the GBD 2010 cause list considered to have a fatal burden, the number of deaths in each age and sex group were estimated and used to calculate death rates and YLL. In the GBD 2010 list, 191 of the 241 individual causes were considered causes that have a fatal burden, with mortality results published for a total of 235 of the 291 cause categories (that is, individual causes plus groups of causes).

Age groups

YLL estimates were produced for 20 age groups for each sex separately: early neonatal (0–6 days), late neonatal (7–27 days), post neonatal (28–364 days), 1–4 years, then 5 year age groups from 5–9 years to 75–79 years, and 80 years and older.

Data sources and methods

Mortality estimates in GBD 2010 were heavily reliant on complex modelling strategies using multiple data sources (e.g. vital registration data, sample registration data, verbal autopsy, and survey data on birth history). This facilitated filling data gaps for countries with little or no data or incomplete vital registration systems, as well as completing time series for countries with high quality vital registration such as Australia.

Indirect demographic methods were used to assess and adjust the death registration data. The adjusted numbers and data from all sources were then used as input to a range of modelling processes to yield mortality estimates for the 187 countries, including those without a complete vital registration system (see Box 4.1).

Australian vital registration data were used as an input in modelling death numbers in GBD 2010. These data were obtained from the World Health Organisation (WHO) Mortality Database and based on year of registration. According to the GBD 2010 data visualisation tools on IHME's website, it appears that cancer registry data was also used in GBD 2010 for ascertaining number of cancer deaths. Mortality data were only available up to 2006 and GBD 2010 made estimates for the remaining years to 2010 using projection models.

In this modelling approach, it is the AIHW's understanding that although mortality estimates for Australia were largely determined by its own data, they were also impacted to a small degree (around 10%) by data from the region (New Zealand) and the super-region (similar high income countries).

Box 4.1 Mortality estimates in GBD 2010

Mortality envelope

To estimate the all-cause mortality envelope, GBD 2010 estimated trends in under-5 mortality and adult mortality for each country using both vital registration data and survey data on complete and summary birth histories. The incompleteness of the vital registration data was assessed and corrected for with improved death distribution methods (DDM). A stream of indirect demographic methods, DDM works by reconciling the age distribution in deaths over a time period and the age distribution in population estimates in the two end time points. Refined methods were applied to survey data on sibling survival to correct for survivor, zero-sibling, and recall bias. The corrected data were used to generate the estimates of under-5 mortality and adult mortality with a method referred to as 'Gaussian process regression (GPR)'. The results were then used as inputs to a relational model life table system to develop a set of age- and sex-specific mortality rates. This system, namely model life table system with flexible standards (MLTFS), was designed in GBD 2010 to choose the appropriate standard age pattern of mortality for a specific country in the modelling, based on how much is known about the country from reliable sources.

Cause-specific mortality

To estimate mortality by cause, data from various sources were assessed and adjusted in terms of completeness, diagnostic accuracy, missing data, stochastic variations, and probable causes of death.

The preferred and most commonly applied modelling technique was CODEm (Cause of Death Ensemble model); this was applied when the database contained adequate numbers of deaths.

Other models used included:

- negative binomial models (where the number of deaths in the database were too low to generate stable estimates)
- fixed proportion models (where death was a rare event)
- natural history models (where it was suspected that data systems did not capture sufficient information)
- mortality shock regressions (where significant numbers of deaths occurred as a result of natural disasters or collective violence).

Special treatment was also given to certain specific diseases, including diarrhoea, lower respiratory infection, cirrhosis, maternal disorders, liver cancer and chronic kidney disease.

Reference life table

As introduced in Chapter 2, the aim of the reference life table is to represent the 'ideal' or aspirational healthy lifespan for all people. Earlier GBD studies used separate reference life tables for males and females. For GBD 2010, a new, unisex reference life table was developed using the lowest observed death rate for any age group in any country with a population exceeding 5 million. The result is a life table with the life expectancy at birth of 86.0 years for both males and females.

Uncertainty interval (UI) estimates

In contrast to previous GBD studies, GBD 2010 provided UIs for the mortality estimates, quantifying the uncertainty due to sampling error, known non-sampling error, missing data,

and model parameters. In some cases these UIs are relatively wide, as they are influenced by global or regional noise and model-uncertainty, rather than being informed closely by the quality of any target country's own mortality data.

The UI calculations in GBD were made possible by the modelling approach where UIs can be included in each step of the estimation process, as well as simulations which generate 1,000 draws to calculate UIs.

Identifying and redistributing deaths not suitable for BoD analysis ('Garbage codes')

Some deaths may be coded to a cause of death that is not appropriate for the purpose of burden of disease analyses. This includes deaths that have been coded with inaccurate ICD-10 codes (not an issue in Australia), as well as accurately coded deaths that are not considered a 'valid' cause of death for burden of disease purposes.

GBD 2010 used the term 'Garbage codes' for such deaths. The Australian Burden of Disease Study Expert Advisory Group and the AIHW consider this term too strong as it implies that the deaths are incorrectly coded and are discarded. Rather, for many of these deaths the level of detail provided in the death certification process was inadequate for assigning a specific GBD cause; however they are still suitable for other purposes. For example, deaths coded as 'stroke unspecified' were redistributed to a more specific cause of 'ischaemic stroke' or 'haemorrhagic/other stroke' for GBD. It should also be noted that some cause codes considered 'garbage codes' by GBD are used for specific purposes in Australia. For example, 'event of undetermined intent' (ICD-10 Y10-Y34) has been used in Australia for specifically identifying deaths that are under coronial investigation (ABS 2010).

As a result, these codes are being described as codes requiring redistribution in the Australian Burden of Disease Study.

GBD 2010 used the following groupings of redistribution codes:

- ill-defined and impossible causes of death
- symptoms, signs and abnormal findings
- intermediate causes
- unspecified cause or sequelae in each chapter
- immediate causes that are the final steps in a disease pathway leading to death
- hypertension and atherosclerosis
- unknown ICD cause of death, and
- nature of injuries (S, T code), provisional assignment of new diseases of uncertain aetiology (U code), and service codes (Z code).

GBD 2010 identified redistribution codes at a much more detailed level than previous GBD studies, with 2,759 codes identified in available death data recorded in ICD 10, at the fourth digit level. AIHW understands that GBD 2010 classified 18% of the recorded deaths in Australia as having codes requiring redistribution.

For each redistribution code identified in GBD 2010 a more suitable (target) code was determined on the basis of pathophysiology. The redistribution algorithms in GBD 2010 were developed on the basis of published scientific literature available, expert judgment, statistical analysis, and in the majority of cases, proportionate allocation across target causes.

The final redistribution proportions used in GBD 2010 are presented by category in the appendix to Lozano et al. 2012 (pages 71 to 103); however this table does not include a full inventory of the ICD codes defined in each category.

The GBD methodology to redistribute inappropriately assigned codes continues to be developed. Future GBD studies are expected to include a method to redistribute intermediate causes according to aetiologic fraction analysis, which is based on the observed negative correlations between the numbers of intermediate causes and underlying causes across demographically similar regions. The impact of this change is currently unknown.

Methodological differences from the 2003 Australian and Indigenous studies

A summary table comparing the methodologies used for YLL calculation between GBD 2010 the 2003 Australian BoD study, and the 2003 Indigenous BoD study is included in the appendix to this chapter (Appendix 4A). The main differences are discussed below.

2003 Australian study

Compiling numbers of deaths

Like most developed countries, Australia has a well-developed death registration system and causes of death are coded to an international standard. Previous Australian BoD studies used official Australian death registration data rather than statistical modelling for compiling numbers of deaths.

The age groups used in the 2003 Australian study were slightly different to those used in GBD 2010. The Australian study used 22 age groups (0–1, 1–4, 5–9,...95–99, 100+ years), compared to 20 age groups in GBD 2010 (0–6 days, 7–27 days, 28–364 days, 1–4 years, 5–9... 80+).

Reference life table

In the 2003 Australian BoD study, YLL calculation was based on a reference life table which was internationally recognised and used in most other burden of disease studies at that time. In that reference life table, the life expectancy at birth was 80.0 years for males and 82.5 years for females.

The development of the new reference life table in GBD 2010 which increased life expectancy at birth to 86 years for both males and females will alter the YLL:YLD ratio by increasing the relative significance of fatal conditions.

Discounting and age weighting

Discounting for time and age-weighting are described in Chapter 2. In the 2003 Australian BoD study, YLL was calculated using a 3% discount rate to be consistent with previous studies. Uniform age weighting was applied across the age of death.

UI estimates

UI estimates were not estimated in past Australian BoD studies.

Redistribution of deaths

The 1996 Australian burden of disease study made a number of decisions based on local considerations about what to do with the deaths that needed to be reassigned to another cause. The 2003 Australian burden of disease study largely retained these redistribution decisions, with some additional algorithms developed mainly in response to assessments of the death coding software system introduced in 1997 (see Box 4.2).

2003 Indigenous study

The 2003 Indigenous BoD study calculated YLL using the same methodological elements as the 2003 Australian BoD study, including the same reference life table, discounting, age-weighting and redistribution of inappropriately assigned codes. Uncertainty intervals were not calculated for Indigenous YLLs.

Box: 4.2 Example of death code redistribution rules in 2003 Australian burden of disease study

Compared with the previous system, the new death coding software introduced in 1997 had a greater preponderance to code deaths to codes for 'unspecified' categories (Begg et al. 2007). To minimise non-specific cause codes in the 2003 study, deaths coded to 'exposure to unspecified factor' (ICD-10 code X59) were reallocated to 'falls' if they also had a 'fracture' code in the multiple cause of death data. Likewise, deaths coded to 'unspecified septicaemia' (ICD-10 code A419) were reallocated to 'nephritis & nephrosis' if they also had an 'acute renal failure' code. This system is believed to be different to that used in GBD at the time.

For some deaths the new system assigned inappropriate underlying causes where another code would have been more informative (Begg et al. 2007). The assigned codes are most appropriately regarded as risk factors for more specific underlying disease processes and preferably should not be used in primary underlying cause of death tabulations. Following the assessment, deaths previously coded to tobacco dependence, obesity, hypercholesterolaemia and dyslipidaemia were reallocated to related categories based on probability analysis of multiple-cause information.

Adjusting for under-identification

Despite high quality and completeness of Australia's vital registration system, Indigenous deaths in Australia are considered to be underestimated due to incomplete recording of Indigenous status. The resulting undercount varies by jurisdiction and over time. When calculating Indigenous mortality rates, this is further complicated by potentially inaccurate Indigenous population estimates. This arises from an undercount of Indigenous people in the population censuses, non-response to the Indigenous status question in Census forms, and unexplained Indigenous population growth observed between Censuses (ABS 2008).

Australia's 2003 Indigenous BoD study took note of the under-identification in both Indigenous death and population statistics. The researchers chose to anchor the results to the Census population estimates in 2001. Under-counting in Indigenous deaths was then quantified and adjusted for using an indirect demographic method and results from a Western Australia linked study. Adjusted mortality rates were then used to estimate the YLL for Indigenous Australians (see Box 4.3).

The 2003 Indigenous BoD study assessed the recording of mortality by age, sex and cause for Indigenous Australians as of good quality. The cause of death structure by age and sex from recorded Indigenous deaths were then applied to adjusted mortality estimates to derive counts of death by age, sex and cause (Vos et al. 2009).

Box 4.3 Adjustment to Indigenous deaths in 2003 Indigenous BoD study

Indigenous population counts of two successive censuses, a generalized growth balance (GGB) method was applied to adjust for under-identification in Indigenous adult mortality rates. GGB is an indirect demographic method to adjust for incompleteness in the vital registration system. It works by reconciling deaths within a specified time period to population estimates in the two end time points. The infant mortality rates were adjusted according to infant mortality estimates in Western Australia (WA) using linked birth and death data. This yielded a set of age- and sex- specific mortality rates for the 1996-2001 period. Under the assumption of no changes in mortality from the 1996 to 2001 period to 2003, the 2003 Indigenous mortality envelope was calculated by multiplying these mortality rates by the 2003 Indigenous population estimates.

Other past BoD studies that included an Indigenous population

There have been a small number of other past BoD studies that included an Indigenous population, e.g., the 1999-2003 Northern Territory BoD study, the 2001 New Zealand BoD study and the 2006-2016 New Zealand BoD study. In the first two studies, YLLs for the Indigenous population were calculated using the same methodology as YLLs for the non-Indigenous population. The 2006-2016 New Zealand BoD study also applied largely the same methodology to Indigenous and non-Indigenous YLLs, with inappropriately assigned cause of death code redistribution scaled to maintain age-sex-ethnicity mortality envelopes. More information on the 2006-16 New Zealand BoD study can be found in Appendix A of this working paper.

Applicability of GBD 2010 methods to the Australian and Indigenous contexts

Compiling numbers of deaths

Age groups

As mentioned earlier in the chapter, GBD 2010 used 5 year age groups up to 80 years and over, with a finer split for infants. Analysis of Australian deaths data show that the majority of deaths in Australia occur in people over the age of 80 years (see Figure 4.2). Similarly, deaths in neonates (that is, babies less than 1 year old) mostly occur in the first 6 days.

As Australian deaths data are available by single year of age at death, death numbers could be compiled at this level for the calculation of YLLs, however these would then need to be combined into specific age-groups for reporting purposes to overcome small numbers.

Under the broad principle of comparability and consistency with GBD 2010, the Australian Burden of Disease Study aims to report YLLs for at least the same age groupings as used in GBD 2010, with disaggregation of the 80 years and over age group where this is supported by the available data. Ultimately, the final age groups used for reporting YLLs should be based on what the data are able to produce robust estimates for.

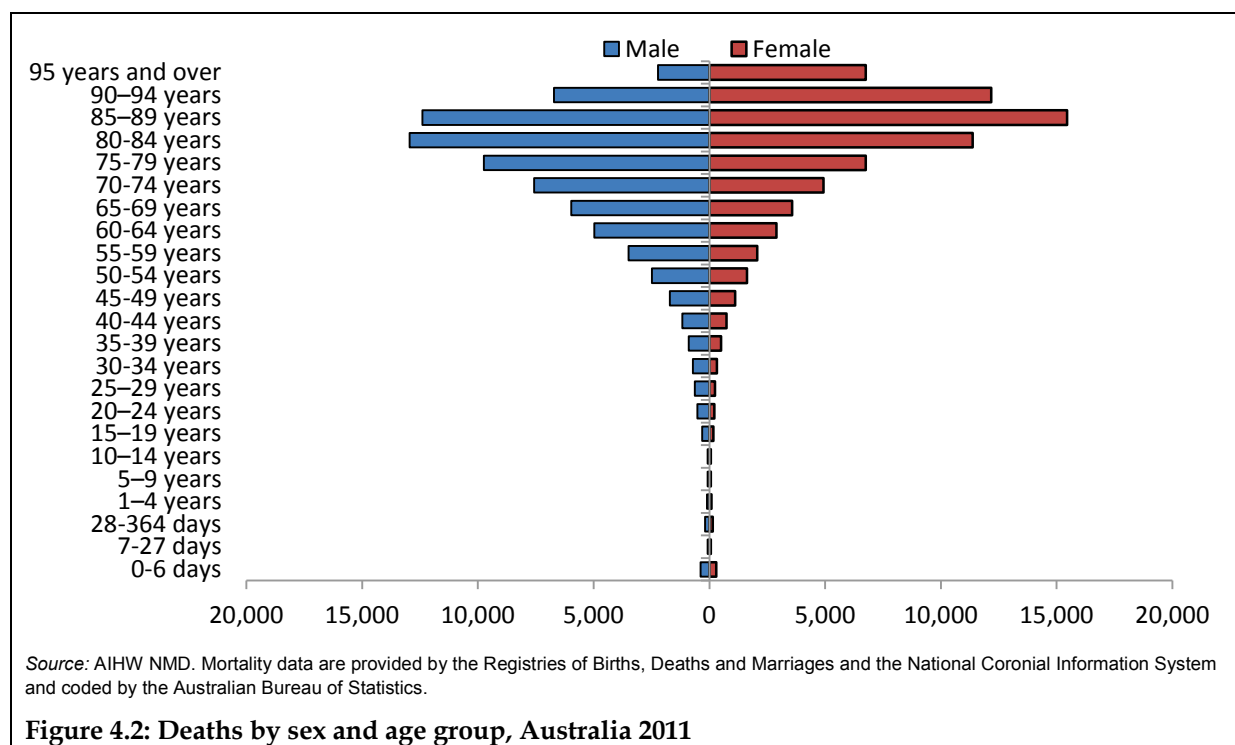


Figure 4.2: Deaths by sex and age group, Australia 2011

Data source and methods

The multi-data source model-based approach used in GBD 2010 was designed to fill in data gaps. While GBD offers national level YLLs with high global comparability, it does not currently include estimates for Indigenous and non-Indigenous Australians or sub-national estimates. It is also not clear the degree to which Australian estimates are influenced by other countries' data.

Using multi-source and multi-modelling methods as per GBD is not considered necessary for producing Australian YLL estimates due to the high quality of deaths data from the Australian vital registration system. The AIHW maintains an up-to-date compilation of cause of death unit record files in the National Mortality Database (NMD). This is the single, preferred data source for compiling number of deaths in Australia.

For some causes of death, there may be insufficient numbers of deaths to obtain reliable and robust estimates. This is particularly the case for rare conditions, sub-national estimates and Indigenous estimates. This problem may be avoided by combining multiple years of data.

Reference life table

The broad principle of consistency and comparability with GBD suggests that the standard reference life table for GBD 2010 be adopted in the calculation of YLLs for the Australian study. This life table is also suggested for sub-national estimates to maintain consistency and comparability with national estimates. Using this reference life table for the Australian BoD study is further supported by the following:

- The reference life table for GBD 2010 is an expression of aspirational life expectancies, based on lowest observed mortality rates worldwide. Therefore it is relevant for all Australians of both sexes.

- The reference year for GBD 2010 is close to the reference year likely to be used for the current Australian study (2011). In the period of one year, the change to life expectancy is expected to be minimal.
- Using the life table for GBD 2010 will likely result in a higher number of YLLs in the Australian studies compared to estimates using Australian life tables. However, the absolute estimates of YLL have limited policy importance. It is the patterns of YLLs across diseases and changes over time that are most relevant.
- The approach is consistent with GBD 2010 and therefore enhances comparability for Australian estimates of YLLs.
- It promotes some level of transparency in the attribution of differences between Australian and global estimates as any differences would likely be due to other factors.

Discounting and age weighting

As outlined in Chapter 2, GBD 2010 did not discount for time, and did not apply age-weighting.

For the Australian BoD Study, changes to the discounting or age-weighting method used in GBD 2010 will lead to inconsistency and incomparability to GBD for YLLs. However, for use in some forms of economic analysis, discounting is the norm (e.g. discounted data for use in health expenditure analysis). Therefore, for some end-users, discounted output may be more desirable. Likewise, differential impacts of deaths occurring at different ages may be of interest to some users, which makes age-weighting relevant.

Therefore an option to be considered for the Australian BoD Study is to produce the estimates without age-weighting or discounting for time (to be consistent with GBD), but also build system capacity to enable different options for discounting and age weighting to be included in analysis for specific purposes.

Uncertainty interval (UI) estimates

In theory, point estimates sourced from (high-quality) administrative data are not considered to have uncertainty. However, a degree of uncertainty can occur when multiple years of data are combined to overcome small number issues. An appropriate method for estimating uncertainty intervals for YLL estimates based on measureable sources of error needs further consideration. At a minimum, they could capture error derived from pooling multiple years of data.

Redistribution of deaths

As mentioned earlier in the chapter, GBD 2010 redistributed approximately 18% of the recorded deaths in Australia. This is a relatively large proportion compared with the 2003 Australian BoD study, where 5.1% of deaths were redistributed.

The AIHW is conducting an assessment of the extent and impact of GBD 'garbage codes' and redistribution rules, using data from the AIHW NMD. Alternative redistribution methods which reflect the Australian context may be explored. Expert advice around Australian modifications to cause of death coding could be sought to inform these data assessments.

The AIHW is also assessing the level and extent of redistribution codes in the Australian data by Indigenous status to determine if there are any key differences between Indigenous and non-Indigenous Australians. This will inform overall data quality and may inform the final redistribution algorithms.

Using Australian mortality data in BoD analyses

As stated earlier in this chapter, GBD 2010 reported YLLs using Australian death data up to 2006 (based on year of registration) and cancer registry data. The data used were relatively old for the 2010 estimates. Australian cause of death data are produced annually and revised 12 and 24 months after the initial release. The revisions apply only to causes of death and reduce the number of ill-defined and un-specified codes.

Year of death versus year of registration

The Australian cause of death data can be counted by year of registration of the death or by year of occurrence of the death. The registration year refers to the time period in which deaths were registered and processed. There are approximately 5% of deaths that occur in a calendar year that are not registered in that year; these deaths are usually captured in the following processing year.

Counting deaths based on year of occurrence matches the 'year of burden' concept for morbidity data. However, this will present a point of difference with GBD 2010 which used year of registration and also has the issue of late registrations if the latest year of mortality data available is used. This issue is planned to be further explored and considered by the Expert Advisory Group and Indigenous Reference Group established for the project.

Reference year/s for analysis

The choice of a reference year for the new Australian Study requires consideration of the data revisions process. Since 2006, Australian cause of death has undergone revisions to enhance the specificity of the cause of death coding.

In the first release of cause of death data (the preliminary version) a non-specific cause code (unknown or intent undetermined) is applied to deaths that are certified by a coroner where the coroner has not finalised cause within the processing period.

During the course of the following processing period (i.e. for the subsequent year of deaths data), cause of death in the current preliminary version will be revised where the coroner case has become closed. The resulting data are referred to as a 'revised' version and are available 12 months after the release of the preliminary version. The same cycle occurs one more time, resulting in a 'final' version 24 months after the release of the preliminary version.

The revisions process results in a shift in the distribution of causes of death only – not the number of deaths registered. An example of the reduction in the number of non-specific causes and a corresponding increase in the number of specific causes over subsequent revisions is shown in Table 4.2. Here deaths coded to a non-specific cause R99 decrease with subsequent revision while the number of deaths coded with specific causes increases. It should also be noted that the revisions to data occur only to deaths that are reported to a coroner. The criteria for reporting a death to the coroner varies by jurisdiction, but in general, in Australia, this occurs where the cause is unknown; the person died in an unnatural or violent manner; the death occurred during or as a result of an anaesthetic; the death occurred to a person being 'held in care' or custody immediately before they died; the doctor was unable to sign or certify the cause of death; or the identity of the person is unknown.

Table 4.2 Distribution of deaths for selected causes, in the preliminary, revised, final mortality data collection for 2009

Cause of death and ICD-10 code	Number of deaths			Changes (%)	
	Preliminary	Revised	Final	Preliminary to revised	Revised to final
Other ill-defined and unspecified causes of mortality (R99)	649	337	316	-48.1%	-6.2%
Car occupant injured in transport accident (V40-V49)	743	819	853	10.2%	4.2%
Other land transport accidents (V80-V89)	127	84	83	-33.9%	-1.2%
Falls (W00-W19)	1,370	1,435	1,450	4.7%	1.0%
Accidental poisoning by and exposure to noxious substances (X40-X49)	799	964	1,004	20.7%	4.1%
Intentional self-harm (X60-X84) ^(a)	2,130	2,284	2,335	7.2%	2.2%
Assault (X85-Y09)	210	261	272	24.3%	4.2%
Event of undetermined intent (Y10-Y34)	993	547	390	-44.9%	-28.7%

Source: ABS 2013.

The latest unit-record deaths data currently available in Australia are the preliminary version for 2011, revised data for 2010 and final data for earlier years. It is expected that final data for 2010, revised data for 2011 and preliminary data for 2012 will be available by mid-2014.

While most of the revisions in the cause of death coding in Australia occurs between the preliminary and revised versions, this may vary by state/territory and Indigenous status.

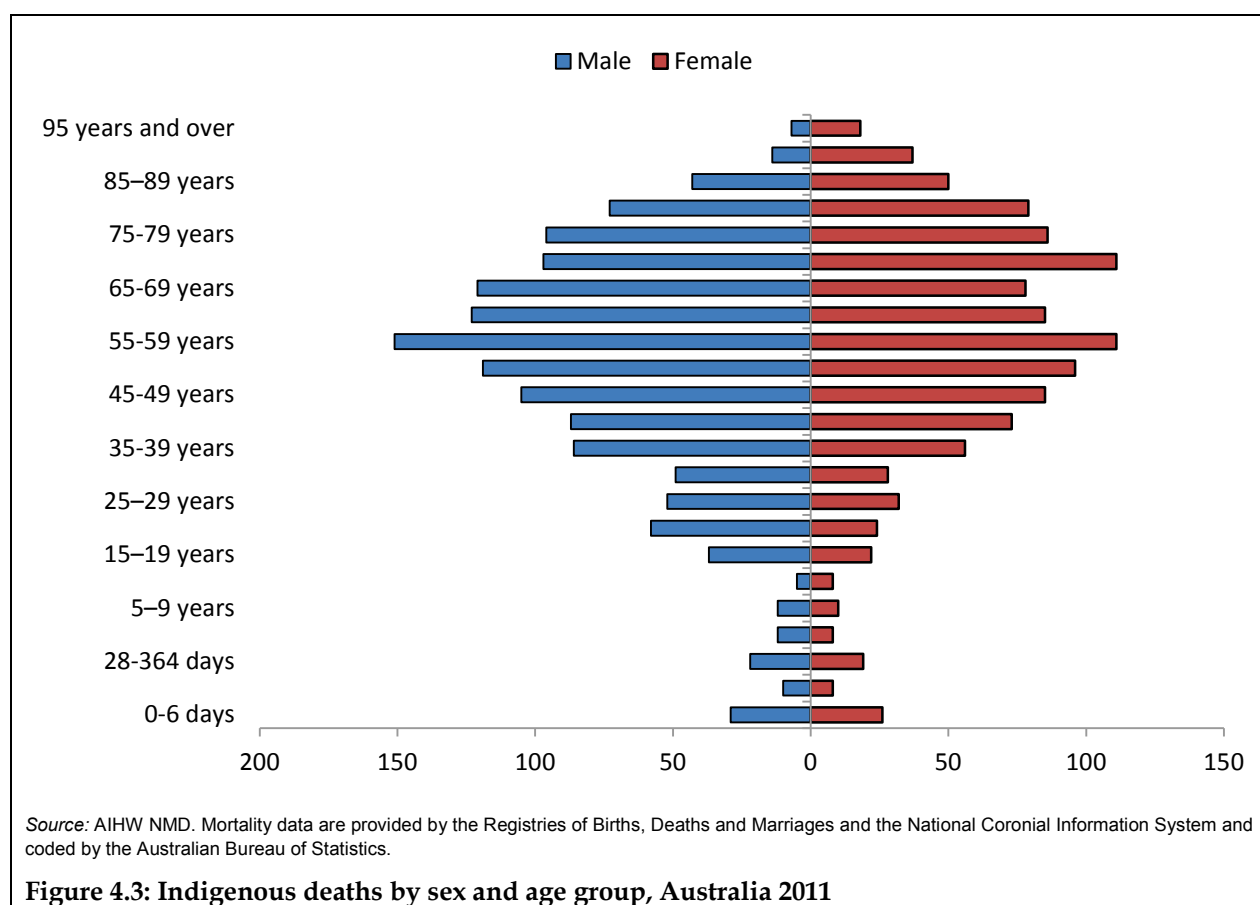
Analyses to investigate differences in death counts reported by year of registration versus year of occurrence, and between preliminary, revised and final versions of the mortality data over a number of years and by state/territory and Indigenous status will be useful to determine the most appropriate method to count deaths for both the national and Indigenous components of the Australian BoD Study.

Issues to consider for the Indigenous estimates

Age groups

As for the national estimates, the age groups used to report YLLs for the Indigenous study should reflect the finest level for which the data are able to produce robust estimates.

Indigenous deaths occur at younger ages than nationally, with a small proportion at the age of 80 or over (Figure 4.3). Due to small numbers in the older age groups, the highest age group reported for the Indigenous estimates is expected to be lower than for the national estimates.



Small number issues

Small numbers are likely to arise in the estimation of Indigenous death numbers, particularly for certain causes of death in certain age groups. These may be an issue for robustness of estimates. While the modelling approach of GBD 2010 might be used to solve this problem, a more practical alternative is to base Indigenous death estimates on multiple years' data, or to combine deaths of multiple causes or age groups. The standard reporting of Indigenous deaths data by the AIHW has been 5 years combined to overcome small numbers.

Where the combined numbers are still too small to produce robust estimates, this approach may be supplemented with simple modelling techniques. Historically, Indigenous: non-Indigenous rate ratios were used as the basis for addressing small number issues for estimates of Indigenous deaths; however, alternative methods such as the use of rate differences to model the data may be explored.

Indigenous under-identification in death registration data

GBD 2010 did not provide Indigenous estimates. However, indirect demographic methods were used to adjust for incomplete death records in many countries. The 2003 Indigenous study also used indirect methods to correct for under-identification in Indigenous deaths. However, these methods are no longer considered appropriate in estimating Indigenous deaths in Australia, due to their dependence on demographic assumptions and the requirement for high quality population estimates (See box 4.4). Furthermore, the ABS recommended using a direct method involving linkage studies to estimate Indigenous life tables and using these to assess and adjust for undercount.

Box: 4.4 Indirect vs. direct demographic methods to measure Indigenous undercount

There are two main streams of demographic methods to measure and correct for undercounting of number of deaths: indirect and direct.

Indirect methods

The rationale of indirect methods is to reconcile deaths over a time period to population estimates in the two end time points. The GBD 2010 study adjusted for incomplete vital registrations using indirect methods. In Australia, indirect adjustment methods were used to measure the under-identification in Indigenous death data and estimate Indigenous life expectancy before the 2006 Census. Indirect methods were also used in the 2003 Indigenous BoD study to adjust for undercount in Indigenous death.

Despite their wide application, indirect methods have their weak points in estimating under-coverage in death registrations. They are dependent on a range of demographic assumptions, and the quality of the population estimates at the end points. After the 2006 Census, the ABS conducted an assessment of three indirect methods and one direct method to investigate their impact on Indigenous mortality under-coverage and life expectancy estimates. All of the assessed indirect methods concluded that coverage of Indigenous deaths was considerably lower in 2001-06 than in 1996-2001 – a result which is implausible when compared to observed data. The assessment also found Indigenous life expectancy estimates derived from indirect methods are very sensitive to the accuracy of the population estimates (ABS 2008).

Direct methods

Direct methods measure directly the degree of under-coverage in death registration data through, for example, data linkage processes. In contrast to indirect methods, direct methods do not require assumptions. The ABS has developed a direct measure of the Indigenous death under-coverage in their Indigenous Mortality Data Quality study (see ABS 2008 for details of this measure). This method was found to provide more plausible results than the indirect methods. The ABS indicated it was their preferred way to calculate under-coverage of Indigenous death data, and hence Indigenous life expectancy estimates (ABS 2008).

The ABS's Indigenous Mortality Data Quality study (which links Census data with death registration data) and the AIHW's Enhanced Mortality Database study (which links death registration data with a number of administrative data), are two such linkage studies and are currently being repeated with updated data. A summary comparison of these two studies is listed in Table 4.1; and further information on these two studies can be found in Chapter 7 (Data sources).

Adjustment factors from the ABS study are available by state/territory, and adjustment factors from the AIHW study are available by state/territory as well as age and sex. Adjustment factors from either of these studies can be used to correct for undercounted Indigenous deaths in the Indigenous BoD study.

Table 4.1: Comparison between the ABS and AIHW linkage studies on the quality of Indigenous mortality data and life expectancy

	ABS study	AIHW study
Linked datasets	Death registration to Census data	Death registration to perinatal data, aged care data, hospital morbidity data
Coverage time	Around the Census year (i.e., 2005-07 in the last study) 2010-2012 in the current study.	Intercensal period (i.e., 2001, 2002, 2003, 2004, 2005, and 2006 in the last study) 2007-2010 in the current study, can be disaggregated down to yearly estimates.
Disaggregation level for under-identification adjustment factors	Australia NSW Qld WA NT Vic, SA, Tas, ACT and other territories	State/territory, age and sex (except for Tas and ACT).
Disaggregation level for life expectancy	Sex and geographical areas, including Australia NSW Qld WA NT	State/territory and sex (except for Tas and ACT).

It will be useful to investigate potential ways in which adjustments factors could be derived and/or applied at the cause level.

Late registrations

A higher proportion of Indigenous deaths than non-Indigenous deaths are registered late (i.e. a year or more later than the year in which they occur) (15% compared to 5%). This issue needs to be considered when deciding whether to analyse deaths by year of occurrence or year of registration.

Reference life table

Using a reference life table that does not highlight differentials in mortality at younger ages may underestimate the gap between Indigenous and non-Indigenous Australians. Therefore, to preserve comparability between Indigenous and non-Indigenous estimates, the same reference life table should be used to derive YLLs for both groups, as well as for sub-national estimates.

The GBD 2010 reference life table represents an aspirational life span for all human populations. It has remarkably higher life expectancies than estimates based on recent Indigenous life tables (that is, 69.1 and 73.7 years at birth for Indigenous Australian males and females, respectively, in 2010–2012; ABS 2013). Using the GBD 2010 reference life table will emphasise deaths at younger ages and result in substantially greater YLLs for Indigenous Australians compared to the 2003 Indigenous study.

Uncertainty intervals

In addition to error derived from pooling multiple years of data, confidence intervals produced around under-identification adjustment factors applied in Indigenous analysis could be incorporated into the calculation of uncertainty intervals for YLLs for Indigenous estimates.

Intended approach

The following outlines our intended approach for undertaking mortality estimates for the Australian Burden of Disease Study, subject to further consultation with the Expert Advisory Group and the Indigenous Reference Group:

- Adopt the standard reference life table defined in GBD 2010 for calculation of Australian and Indigenous YLLs.
- Use Australian cause of death data held in the AIHW National Mortality Database to compile numbers of deaths for Australian, sub-national and Indigenous YLL estimates.
- Reporting age groups for YLLs should be based on what the data are able to support. For the national estimates, at a minimum, age-groups should reflect those used in GBD 2010 to enable comparison, and extended to older age-groups where possible. For the Indigenous estimates, age-groups should reflect what the data are able to produce robust estimates for (for example, the highest age-group is expected to be lower than for the national study due to small number issues in the older age groups).
- Investigate differences in death counts reported by year of registration versus year of occurrence, and between preliminary, revised and final versions of the mortality data over a number of years.
- To avoid problems with small numbers and provide greater stability in YLL estimates, combine multiple years of data to derive counts for Indigenous deaths, deaths due to rare conditions, and for sub-national estimates. Simple modelling strategies may be used where numbers are still too small for stability.
- Follow GBD methods for identifying and redistributing inappropriately assigned cause of death (garbage) codes where appropriate. Analyses should be undertaken on Australian data sources to inform modifications where the GBD rules may not be suitable in the Australian or Indigenous contexts.
- No discounting for time or age-weighting (as per GBD 2010) for the standard analysis; however, system capacity may be built in to estimate YLLs using discounting and age-weighting to provide further functionality and usefulness.
- Determine the most appropriate methods for calculating uncertainty intervals for YLLs based on measurable sources of error.

Additional considerations for the Indigenous estimates

- Apply adjustment factors for Indigenous under-identification in mortality data at the lowest possible level. These could be derived from either the ABS's Indigenous Mortality Data Quality study or AIHW's Enhanced Mortality Database study.

Appendix 4A

Comparison of YLL estimates in GBD 2010 and the 2003 Australian BoD study

		GBD 2010	2003 Australian BoD	Notes
Quality of vital registration system		Lack of complete vital registration systems in most developing countries	High level death registration system and high quality information on causes of death.	The YLL estimates in GBD 2010 were heavily model-based and reliant on various data sources to address the lack of complete vital registration system in most developing countries. The case in Australia is different since there is a high level death registration system with high quality information on causes of death.
Standard reference life table		GBD-derived life table: Life expectancy at birth of 86 years for both males and females	Internationally recognised life table used at most other BoD studies, where life expectancy at birth were 80.0 years for males and 82.5 for females.	
Mortality envelope	Data source	Vital registration data Survey data on complete and summary birth histories. Australian data to 2006 appear to be sourced from WHO, in aggregated form. 2007–2010 data derived from modelling, with potentially some influence from regional trends.	Cause of death unit record files (CODURFs) sourced from the AIHW NMD	
	Method	Indirect adjustment (DDM) Gaussian process regression Relational model life table system	Straightforward mortality calculation	
	Corresponding life expectancy estimates at birth	79.2 years for Australian males and 83.8 for Australian females	78.3 years for Australian males and 83.2 for Australian females	
Cause-specific mortality	Data source	Vital registration, verbal autopsy, mortality surveillance, censuses, surveys, hospitals, police records, and mortuaries.	Cause of death unit record files (CODURFs) sourced from the AIHW NMD	
	Method	Six different modelling strategies were applied depending on the strength of the data.	Directly from data	

Comparison of YLL estimates in GBD 2010 and recent BoD studies which included an Indigenous population

		GBD 2010	ATSI BoD 2003	ATSI YLL in NT BoD 1999-2003	Māori YLL in NZ BoDs 2001 & 2006-16	Notes
Quality of vital registration system		Lack of complete vital registration systems in most developing countries	Under-identification of Indigenous deaths, but relatively good information on causes of death.			The ABS's investigation has revealed some issues with indirect adjustments in the context of Australia's Indigenous mortality. The ongoing linkage projects by ABS and AIHW respectively, based on direct adjustment methods, can provide insights for the YLL estimates in the Indigenous BoD.
Mortality envelope	Data source	Vital registration data Survey data on complete and summary birth histories	Official country-level cause of death data	Official country-level cause of death data		
	Method	Indirect adjustment (DDM) Gaussian process regression Relational model life table system	Indirect adjustment (GGB) to address under-identification.		Same as non-Māori YLL	
	Corresponding life expectancy estimates at birth	79.2 years for Australian males and 83.8 for Australian females	64 years for Indigenous Australian males and 69 years for Indigenous Australian females in 1996-2001.	59 years for NT Aboriginal males, and 65 years for NT Aboriginal females.		
Cause-specific mortality	Data source	Vital registration, verbal autopsy, mortality surveillance, censuses, surveys, hospitals, etc.	Cause of death unit record files (CODURFs)	Cause of death unit record files (CODURFs)		
	Method	Six different modelling strategies were applied depending on the strength of the data.	Cause of death structure by age and sex from recorded Indigenous deaths were applied onto adjusted mortality estimates.		Same as non-Māori YLL garbage code redistribution scaled to maintain age-sex-ethnicity mortality envelopes in 2006-16 BoD	

5 Morbidity—Years lived with disability

What does it measure and why is it important?

It is important to emphasise that health is more than avoiding death. Individuals, households and health systems devote enormous resources to the prevention, cure and treatment of non-fatal consequences (sequelae) of diseases and injuries. As we live longer, the time spent living with disease may also likely to increase. As such, it is equally important to measure non-fatal outcomes when estimating the burden of disease for policy making and planning.

YLD—‘years lived with disability’—is a measure of the number of healthy years lost due to ill-health—that is, the non-fatal component of burden of disease. YLD estimation attempts to capture the frequency, severity, comorbidity and consequences of each condition in the cause list and to quantify their impact on a population in terms of the difference between time lived in full health and time lived with one or more health problems.

YLD estimation is more complex than for YLLs, and is further complicated by a choice of the underlying method. This chapter reviews the method for compiling YLDs used in the GBD 2010 study, and compares it to the methods used in previous Australian burden of disease studies. The GBD methods are assessed for their application in the Australian context, and the Indigenous Australian context.

Key terms introduced in this chapter

disease model – representation of clinical conditions designed to summarise what is known about the disease epidemiology, prevention and treatment.

incidence - in epidemiological terms ‘incidence’ refers to the occurrence of a disease or event; the ‘incidence rate’ is the number of new cases occurring during a specified time period.

prevalence - refers to the existence of a disease or event, whether or not it is newly occurring; the ‘prevalence rate’ is the number of cases existing at a point in time or over a specified time period.

disability weight - a factor that reflects the severity of health loss from a particular **health state** on a scale from 0 (perfect health) to 1 (equivalent to death).

sequelae - consequences of diseases and injuries.

health state - groups of sequelae reflecting key differences in symptoms and functioning.

comorbidity – a health problem/disease that exists at the same time as another health problem/s.

cross-walking – technique to standardise inputs for YLD calculation.

(Refer to glossary for a full list)

Overview of GBD 2010 methods

Two main components lie behind the calculation of YLD for a particular disease or injury: a measure of the frequency of a disease in a population (based on a disease 'model'), and an adjustment (the 'disability weight') for the severity of its effects. Both components are measured for each stage or consequence of disease (sequelae).

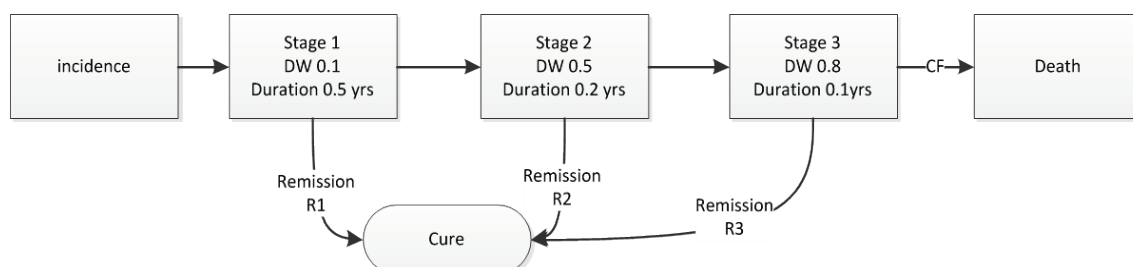
Expert groups comprised of leading experts on diseases and injuries were engaged to assist with both these components. These expert groups assisted with lay descriptions for health states for the revised disability weights, and provided input on epidemiological models for diseases and injuries, reviewed the estimates and helped guide the interpretation of results in light of broader epidemiological evidence. As a result of consultation with these expert groups, GBD 2010 made major methodological changes to both components – firstly, by moving from calculating disease frequency using incidence to using prevalence and, secondly, by revising the disability weights. These changes are outlined in detail below.

What is the difference between incident and prevalent YLD?

'Incidence' in epidemiological terms refers to the occurrence of a disease or event; the 'incidence rate' is the number of new cases occurring during a specified time period in a population. 'Prevalence' refers to the existence of a disease or event, whether or not it is newly occurring; the 'prevalence rate' is the number of cases existing in a population at a point in time or over a specified time period. At a basic level, therefore, incident YLD for a particular year reflects the burden of disease due to conditions and injuries newly occurring in that year, whereas prevalent YLD reflects the burden of disease due to all conditions and injuries existing in that year, regardless of when they originally occurred. The mechanics of calculating the two types of YLD are outlined in Box 5.1.

Box 5.1: Calculating incident and prevalent YLD

In the example below, a generic disease model shows the progression of a disease from point of incidence to cure or death. The proportion moving from stage to stage (or through levels of severity) is affected by remission and case fatality rates.



In this model, the incident YLD is calculated by multiplying the incidence by the disability weight (DW) by the duration for each stage of disease, as follows:

$$YLD_{inc} = INC_1 * 0.1 * 0.5 + INC_2 * 0.5 * 0.2 + INC_3 * 0.8 * 0.1$$

where

INC_1 = the incidence of the disease

INC_2 = incidence of stage 2 = $INC_1 - R_1$ (i.e. incidence of stage 1 – remission from stage 1)

INC_3 = incidence of stage 3 = $INC_2 - R_2$ (i.e. incidence of stage 2 – remission from stage 2).

The GBD 2010 methodology used prevalent YLD for the calculation of DALYs. This method multiplies the prevalence by the disability weight for each stage of disease. Thus the YLD for the disease modelled above would be:

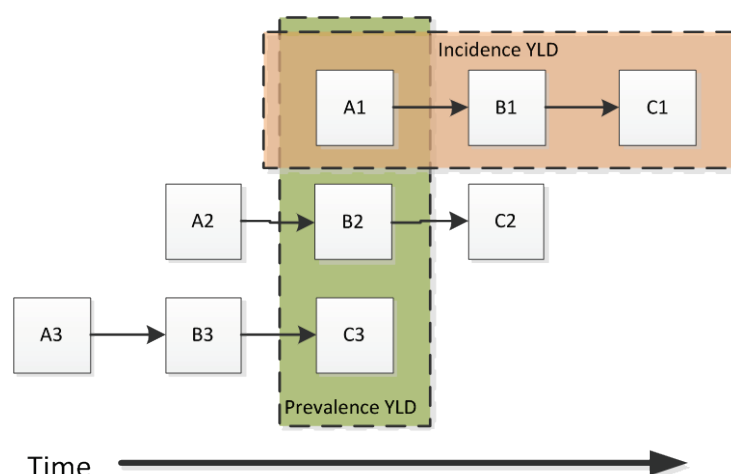
$$YLD_{prev} = PREV_1 * 0.1 + PREV_2 * 0.5 + PREV_3 * 0.8$$

where $PREV_1$, $PREV_2$ and $PREV_3$ is the prevalence of disease at stages 1, 2 and 3, respectively.

It is important to recognise that the different methods for calculating YLDs are measurements of different quantities – they are not alternative approaches to quantifying the same thing.

The major difference between incident and prevalent YLD is that each reflects populations at different points **in time**. The incidence model uses information about remission and case fatality to project the disease experience of new (incident) cases forward in time, quantifying the **current and future** health effects of that condition throughout different disease stages. In contrast, the prevalence model quantifies the current health effects of **existing** cases in the reference period taking into account the stages of the disease. This is illustrated in Box 5.2. In effect, incident YLDs could be thought of as a cohort approach whereas prevalent YLDs could be thought of as a cross-sectional (or period) approach. The latter has the effect of attaching the burden of disease to the age group at that point in time: the incidence method attaches all current and future burden to the age group in which the disease first occurred.

Box 5.2: The relationship between incident and prevalent YLD



This figure illustrates the relationship between incident YLD and prevalent YLD for the same disease, where A1, A2 and A3 represent the incidence of the disease in successive years, and B and C represent later stages of the disease.

Incident YLD is derived from the disease model based on incidence A1 and is calculated using A1, B1 and C1 (horizontal/brown box). Prevalent YLD is derived from incident cases of disease at all three points in time, A1, A2 and A3, but only counts the burden occurring at the reference time point so is calculated using A1, B2 and C3 (vertical/green box). The two values are most different for diseases of long duration.

Assuming the other parameters of the disease (progression, remission and case fatality rates) are constant then YLD_{prev} will be greater than YLD_{inc} when $B2 > B1$ and/or $C3 > C1$, which will occur when incidence at A3 and incidence at A2 are greater than incidence at A1. Thus for diseases with decreasing incidence, prevalent YLD will be larger than incident YLD.

These differences between incident and prevalent YLD influence the interpretation of the resulting DALYs. For example, DALYs calculated using incident YLD for time period T measure 'the health loss connected with events (of whatever type) occurring in T, or [...] the total health loss tied to new health problems in T', whereas DALYs calculated using prevalent YLD for time period T might be best described as 'representing overall population health for the period' (Schroeder 2012). Both of these measures are useful in different ways: incident YLD may provide information relevant to prevention and evaluation of preventive interventions, whereas prevalent YLD may be more relevant to questions about health service use, evaluations of interventions aimed at existing cases, and health expenditure.

Table 5.1: Summary of the differences between incident and prevalent YLD

Incidence	Prevalence
Cohort approach	Cross-sectional approach
All future expected health loss counted in year of incidence; does not include burden from existing cases.	Health loss is spread over the duration of the illness; includes burden from new and existing cases occurring in the time point of interest
All health loss (current and future) counted in age group of incidence	Health loss spread across the age groups in which it is prevalent
Similar conceptually to YLLs with the consequence of the event being counted only in the time period in which it commences	Different to YLL approach with the consequence of the event being counted in the time period in which it is experienced
Useful for questions about prevention and evaluations of preventive health	Better 'match' to disease expenditure figures. Useful for questions about health service use and evaluations of interventions aimed at existing cases.
More sensitive to current trends	Used in HALE (Healthy Life Expectancy) calculation
Incidence data usually sourced from registers or surveillance where information about comorbidities is usually not available	Easier to adjust for comorbidity (where survey and other data captures information about other conditions)

Generating disease and injury models

Across the 291 disease and injury causes and cause groups listed in the GBD 2010 study, 289 were identified as causing ill-health. The major outcomes of these diseases and injuries were captured in 1,160 defined sequelae. Sequelae for each condition were developed by disease and injury experts, and include the 'main outcomes from a disease that could potentially make an important contribution to the burden of a given disease or injury and which could in principle be measured' (Murray 2012). Sequelae could include the disease itself, stages of the disease, or the outcomes associated with that disease. Some clinical disorders could be classified as both a disease, and as a consequence of another disease (for example, chronic kidney disease secondary to diabetes).

Complexity of models

The GBD 2010 project modelled each of the 1,160 disease sequelae. Each sequela required prevalence estimates for each sex, age group and geographic region. These estimates were derived from prevalence data as well as incidence, mortality and duration data. These disease models are not yet available and so cannot be assessed for validity in the Australian and Indigenous Australian contexts.

Models were developed after an extensive, systematic literature search for sources of data for incidence, prevalence, remission, duration and excess mortality due to the disease, including published studies, case notification data, population-based registry data, clinic and hospital data, outpatient data and household surveys supplemented where necessary by treatment data, police records and survey data.

Models were developed using meta-analysis techniques to take advantage of all available data. A new disease modelling tool, DISMOD-MR, was developed based on a generalised negative binomial model to run Bayesian meta-regression analysis across all data for all diseases and countries. This tool currently requires substantial computing resources. A more user-friendly, easy-to-navigate version of DISMOD-MR, is planned for development, but is not yet available.

In order to utilise as much data as possible, a number of techniques were used to estimate the required inputs for YLD estimation:

- To account for data inconsistencies (e.g. different age groups) and methodological differences (e.g. varying case definitions or different diagnostic techniques) across data sources, techniques were developed to standardise inputs. This is referred to as ‘cross-walking’.
- To account for missing data, prediction models (mostly Bayesian meta-regression models) using data on covariates were constructed. Alternatively, country-level estimates were modelled from those for the region (e.g. Australasia), or the super-region (e.g. High Income).
- To account for specific disease or data-related issues, other techniques used included:
 - Natural history models
 - Geospatial models
 - Back-calculation models (calculates incidence from mortality and case-fatality)
 - Registration completeness models (to adjust for incomplete data due to health-system access and other covariates).

Disability weights

Since time is the common currency for both YLL and YLD, the amount of time lived in non-fatal health states needs to be defined, measured and numerically valued. The valuation process results in a set of weights “based on individuals’ perceptions of the impact on people’s lives from a particular disability” (IHME 2013a). Commonly known as disability weights, they reflect the severity of a disease or injury on a scale from 0 (perfect health) to 1 (equivalent to death).

In earlier GBD studies, disability weights were derived mainly based on the views of healthcare professionals (see Box 5.3).

Box 5.3: Disability weight estimation in previous GBD studies

‘The previous comprehensive estimation of the global burden of disease (undertaken in the final revision of GBD 1990, which was published in 1996) used the judgments of a small group of health-care professionals to establish disability weights for 483 sequelae of 131 diseases and injuries. These disability weights were used widely in WHO’s revisions of the GBD for 1999–2002, and 2004, and in several national and subnational burden of disease studies. Additions and amendments to the 1996 GBD weights have been assimilated selectively, largely on the basis of the Dutch Disability Weights study, which adapted the GBD measurement approach from the 1996 study, with specific modifications to the descriptions of health states and addition of several states’ (Salomon J. A. et al. 2012).

The GBD 2010 study explicitly attempted to separate the health effect from other societal factors (in particular, welfare effects) and employed new methods to estimate these factors (see Box 5.4). It aimed to respond to suggestions for a more inclusive measurement exercise which represents the broader perspective of different cultures, communities and societies using a transparent, standardised and replicable approach.

Unlike sequela (which are the direct consequence of a particular disease or injury), **health states** are the functional consequences or symptoms experienced by people with that

sequela. For example, people in the ‘diagnosis and primary therapy phase of breast cancer’ or ‘diagnosis and primary therapy phase of bowel cancer’ (sequelae of breast cancer and bowel cancer respectively) both experience the health states ‘cancer diagnosis and primary therapy’ as well as ‘generic uncomplicated disease with worry and daily medication’. For many other diseases, health states are groups of very similar sequelae with similar severity. For example, ear pain (the health state) is associated with some infections and some injuries, but the health loss associated would be expected to be similar in both cases.

Across the 1,160 disease and injury sequelae, 220 unique health states were defined, each accompanied by a simple, non-clinical description of functional consequences and symptoms that lay persons are able to understand and compare. The individual health states were defined to include levels of severity – for example, heart failure: mild, heart failure: moderate, heart failure: severe.

For GBD 2010, disability weights were derived based on the view of the general public rather than of health professionals. The survey instrument allowed respondents to make pairwise comparisons between two health states. Respondents were surveyed in two ways: 13,902 adults participated in household surveys (face-to-face interviews in Bangladesh, Indonesia, Peru and Tanzania, and telephone interviews in the United States), and 16,328 individuals in an open-access web-based survey. At least 500 of the web-based survey participants were based in Australia (Salomon 2012). The result is a set of weights which is claimed to reflect consistent results across different cultural environments (Salomon 2010; Salomon et al. 2012).

Box 5.4: Disability weight estimation for 2010 burden of disease study

“Responding to critiques of previous efforts to measure disability weights for the GBD, we focused on eliciting judgments about health loss rather than welfare loss; used a new, standardised approach to measurement with simple paired comparison questions; included a major emphasis on surveying respondents from the general public; and used primary data collection in diverse communities to examine hypotheses about cultural variation in assessments of disability” (Salomon J. A. et al. 2012).

Data modelling and transformation

Both probit regression modelling and rescaling were applied to pairwise survey responses. This generated a disability weight for each of the 220 health states. These weights were then mapped back onto the 1,160 sequelae.

Data validation and uncertainty intervals

GBD 2010 also analysed the results between the surveys, which ascertained a high degree of consistency between the six settings of cultural and educational difference – and hence the derived disability weights were considered applicable across the global study.

Further, the regression process in GBD 2010 generated uncertainty intervals around the disability weights.

The disability weights in GBD 2010 broadly follow those used in the previous GBD study published in 2004, particularly in the moderate to severe range. However, in the mild end with disability weights below 0.2 – where most weights reside in GBD 2010 – many states have lower weights than previously was the case. IHME has partly attributed this change to a more detailed disaggregation of mild conditions, as well as the clearer lay descriptions in the new study. The disability weights in GBD 2010 for several states are higher than

previously. These states include, for example, untreated epilepsy, illicit drug use disorder, and acute lower back pain (Salomon J. A. et al. 2012). Because the disability weight is a multiplier of the prevalence estimate, the impact of these changes is reflected both in the ranking of conditions, and in the final YLD calculation.

Severity distribution

While disability weights defined for individual health states are considered constant across countries and cultures, the severity levels for particular sequelae can be distributed differently in different countries or sub-populations. For example, 'dementia', a sequela of 'Alzheimer's disease and other dementias' includes the health states dementia: mild (DW=0.082), dementia: moderate (DW= 0.346) and dementia: severe (DW=0.438). However, the distribution of mild, moderate and severe dementia cases may vary between countries. Consequently, the aggregation to the sequelae level following application of the disability weights will reflect these different distributions.

Global information of the distribution of severity is limited. In GBD 2010 the severity distribution was determined using both published literature and by creating severity distributions using survey data:

- where published literature was available, severity distributions were obtained using meta-analysis of pooled data, rescaled to preserve internal consistency.
- where no published literature was available, severity distributions were calculated using World Mental Health surveys, plus three SF-12 surveys from the USA^{1,2} and Australia³, mapped to the disability weights (at the health state level) and adjusted for comorbidity. It should be noted that the mapping of SF-12 outcomes to health states for GBD 2010 was based on a 'convenience sample of IHME staff who had not worked on GBD' and is being further refined for future GBD iterations.

In GBD 2010, it appears that the calculated severity distributions were applied globally; however, this method may provide a mechanism for obtaining country or population-specific (or country-weighted) disability weights with sufficient data.

Comorbidity adjustment

As introduced in Chapter 2, burden of disease estimates must also account for the fact that people may experience more than one condition at a time (comorbidity).

Comorbidity can occur by chance (**independent comorbidity**) or because the conditions are related to each other either because of common risk factors, or because one disease (e.g. diabetes) is itself a risk factor for other diseases (e.g. heart disease) (**dependent comorbidity**). Both types of comorbidity are problematic for burden of disease estimation because the disability weights for health states are derived in isolation of other conditions, and so adjustments must be made to avoid over-estimating the disease burden.

¹ Medical Expenditure Panel Surveys (MEPS)

² US National Epidemiological Survey on Alcohol and Related Conditions (NESARC)

³ Australian Mental Health Survey

Comorbidity tends to increase with age. As such, comorbidity adjustment tends to be very small at younger ages, but larger at older ages.

For GBD 2010, the analysis of independent comorbidity involved micro-simulation of each country-age-sex group for combinations of diseases and a multiplicative model for combining the disability weights (supplement to Murray et al. 2012). The individual health loss represented by the combined disability weight is allocated to each condition proportionally using the disability weight of the condition on its own.

Analysis of Medical Expenditure Panel Surveys (MEPS) data in the USA, and a subsequent analysis of Canadian and New Zealand data (NZ MoH 2013) confirmed the assumption that independent comorbidity accounted for the vast majority of comorbidity.

As GDB 2010 used a prevalence approach to YLD estimation, adjustment for comorbidity was more straightforward as only disease combinations present during the reference period needed to be measured or estimated. This is computationally easier than including a temporal component to estimate future combinations as in the incidence-based model.

Methodological differences from the 2003 Australian and Indigenous studies

Incidence versus prevalence methodology

The 2003 Australian and Indigenous studies derived DALYs from incident YLD rather than prevalent YLD as in the GBD 2010 study. However, prevalent YLDs were also calculated for the 2003 studies. The top five contributors to total incident and prevalent YLD for the Australian population are shown in Table 5.2. This demonstrates the effect of this different methodology on the top 5 causes. For some other conditions there may be more extensive changes. For example, polio would not have YLDs under the incidence approach (as there are no new cases in Australia), but would have YLDs using the prevalence approach (as some people are still living with the consequences of polio contracted many years ago).

Table 5.2: Comparison of top five causes of YLD, by incidence and prevalence, Australia, 2003

Rank	Incident YLD		Prevalent YLD	
	Cause	% of YLD	Cause	% of YLD
1	Anxiety and depression	14.1	Anxiety and depression	14.3
2	Sense organ disorders	8.3	Sense organ disorders	7.3
3	Type 2 diabetes	7.8	Type 2 diabetes	5.5
4	Dementia	5.2	Asthma	4.4
5	Asthma	4.4	Dementia	4.2

Sources: Annex tables 12 & 13, Begg et al. (all Australian data)

Cause list and disability weights

The list of the five conditions contributing the most to total prevalent YLD in the 2003 Australian study is quite different to that for Australia in GBD 2010 (low back pain, major depressive disorder, neck pain, falls and anxiety disorders). Key reasons for this difference are the changes to the cause list in GBD 2010 (for example, the addition of neck pain and the

separation of depression and anxiety) and the change in disability weights applied to particular conditions, particularly musculoskeletal conditions and neurological and sense disorders (Table 5.3).

Because disability weights are used as a multiplier in the calculation of YLD, any change in the value of these weights also has a direct influence on the estimate of YLD. The 2003 Australian studies used a combination of disability weights from previous GBD studies and the Dutch Disability Weights studies, reflecting societal preferences mainly based on views of healthcare professionals. These weights incorporated elements of welfare loss, as well as health loss. In contrast, disability weights in GBD 2010 were derived from large-scale surveys on the general public, and aimed to measure only 'health loss'.

Table 5.3: Disability weights for selected specific conditions, 2003 Australian and Indigenous studies and GBD 2010

Condition	2003 Australian weight	GBD 2010 weight
Chronic back pain	0.103	0.366 without leg pain 0.374 with leg pain
Osteoarthritis	0.127 all cases, males 0.117 all cases, females	0.023 mild 0.079 moderate 0.171 severe
Dementia	0.41-0.42 age <75 years (mild) 0.48 age 75-79 years (moderate) 0.54 age 80+ years (severe)	0.082 mild 0.346 moderate 0.438 severe
Hearing loss, severe to complete	0.145 all cases	0.031-0.033 without tinnitus 0.065-0.092 with tinnitus
Vision loss	0.266 severe, glaucoma-related 0.228-0.246 macular degeneration (varied by age) 0.103-0.136 severe, cataract-related (varied by age) 0.09-0.24 other causes (varied by age)	0.004 mild 0.033 moderate 0.191 severe 0.195 blindness
Asthma	0.054	0.009 controlled 0.132 uncontrolled
Chronic obstructive pulmonary disease	0.168 all cases, males 0.159 all cases, females	0.015 mild 0.192 moderate 0.383 severe

Sources: AIHW unpublished data (2003 Australian study); Salomon J. A. et al. 2012 (GBD 2010).

Comorbidity

The 2003 Australian estimates included an adjustment for independent comorbidity derived using empirical estimates from national health surveys and admitted-patient hospital data. A multiplicative model was used to adjust the disability weights for the 21 most common non-fatal conditions of old age. Apart from dementia, no attempt was made to adjust for mental-physical comorbidities. Comorbidity between the 21 conditions was estimated using National Health Surveys and Hospital Morbidity data. In allocating the new combined weight, the change in total weight was taken from the milder of the conditions (Begg et al. 2007, p25).

Disease modelling software

The 2003 Australian and Indigenous studies used a disease modelling tool known as DISMOD II. This tool was developed by the WHO for use in burden of disease analyses to check the consistency of estimates of incidence, prevalence, duration and case fatality for diseases, and to estimate missing parameters. It uses mathematical modelling techniques which exploit the dependency between disease incidence, prevalence and mortality. It is based on single (best) estimates as inputs to the modelling process. DISMOD II is freely available from the WHO website, and can be run locally using modest computational resources.

GBD 2010 extended the DISMOD methodology to use Bayesian meta-regression techniques on pooled data, rather than a single data source, and to borrow strength based on regional patterns and country-level covariates to estimate missing data points. This extended version, known currently as DISMOD-MR, is computationally intensive and is not currently available outside the University of Washington.

Applicability of GBD 2010 methods to the Australian and Indigenous contexts

Prevalence approach

YLDs calculated from prevalence are assigned to age-groups according to the ages at which the loss of health is experienced, rather than the age when the condition first occurred. This analysis method is useful for most planning and decision-making purposes as it matches the current burden.

Computationally, prevalent YLD is simpler to implement as duration data is not required; however it does require the estimation of point prevalence data. Additionally, the incorporation of comorbidity data is computationally more straightforward using the prevalence method.

Previous concerns of prevalence YLD being fundamentally different from YLLs (which are an incident measure) are no longer considered a problem (Schroeder 2012) and comparison of the top 5 causes of YLD by both the incident and prevalence methods in the 2003 Australian study showed little difference in both the conditions in the top 5 and the estimate.

The prevalence approach to YLD calculation is therefore considered appropriate for use in both the national and Indigenous YLD estimates.

Multiple data sources

GBD 2010 methodology uses multiple data sources to estimate prevalence of a condition, thus taking into account more information than when only relying on one single data source. It seems reasonable to do this when there are multiple high-quality data sources available for a particular condition; however, Australia may not need to do this for all conditions.

Judgement will be required about when it is likely to be useful to use more than one data set to estimate prevalence. A clearly defined protocol describing the process for inclusion of data sets could be developed in consultation with the Expert Advisory Group and used to ensure transparency of all data sources.

GBD 2010 used software (DISMOD-MR) specifically designed to take as inputs multiple data sets. A reasonable alternative if this software is not available for use in Australia may be to use meta-analysis to derive a pooled estimate from multiple data sources that can then be used as inputs into DISMOD II. The software options for undertaking the YLD analysis are described in more detail in Chapter 8.

Validity of disability weights, comorbidity assumption and severity distributions

National YLD estimates

The set of disability weights in GBD 2010 used a new survey instrument to measure health loss only (as opposed to welfare loss), based on the views of the general public.

To follow the broad principle of consistency and comparability with GBD 2010, it is likely that the Australian BoD Study will need to generate at least one set of YLDs based on the same disability weights as are used in GBD 2010.

To adequately reflect Australian judgements of health state preferences, the pairwise comparison survey would ideally be replicated comprehensively in Australia. However, this would be costly and time consuming. In view of time constraints and the need for comparability of methods with the 2010 GBD study, adopting disability weights from GBD 2010 is desirable. This is supported by consistency of GBD survey results between countries and across population groups with different cultural and educational backgrounds. Furthermore, as a large proportion of the survey population used to estimate disability weights was from the United States, and more than 500 of the 16,328 internet-based responses were from Australia, these weights are likely to be generalisable to the Australian context.

As part of the Australian study, the suitability of GBD 2010 weights in Australia may be assessed to some degree using comparisons with SF-12 measures that are available in some Australian datasets. Unless the assessment suggests major anomalies, it is reasonable to adopt the disability weights used in GBD 2010 for the Australian study.

Similarly, comorbidity and severity distributions could be assessed against Australian data sources to measure the applicability of GBD methods.

Indigenous YLD estimates

The formulation of new disability weights for GBD 2010 has focused on judgements of health—rather than welfare—loss, but for Indigenous people, a holistic view of health necessarily includes both. Indigenous health state preferences may well differ from those used in GBD 2010.

While the reductionist, biomedical view of health adopted by GBD can only be defined, measured and valued by Indigenous people through a culturally-appropriate survey, any generation of separate Indigenous disability weights will, in turn, invalidate comparisons with national, and GBD results.

A suggested approach, despite these misgivings, is to recognise the applicability of the GBD 2010 disease weights to a wide variety of culturally diverse settings, but to note that the universality of these measures still requires further systematic scientific investigation.

For the Indigenous estimates, we are conscious that Indigenous people:

- may take a broader view of the disability than entailed by diseases
- may place different relative valuations (assessments of health losses).

While it is not considered feasible to undertake a full-scale survey to derive an alternative set of disability weights for the Indigenous population in the timeframes of this project, some further data analysis using Australian data sets (e.g. using SF-12 measures from the AATSIHS) may be useful to support interpretation of results.

Comorbidity and severity distributions for the Indigenous populations could be assessed or validated against Australian data sources (e.g. the AATSIHS and hospital data) to measure the applicability of GBD methods.

Intended approach

The following outlines our intended approach for the morbidity estimates for the Australian Burden of Disease Study, subject to further consultation with the Expert Advisory Group and the Indigenous Reference Group:

- Adopt the GBD prevalence-based methodology for primary estimation of YLD.
- YLDs to be unweighted and undiscounted. Options to parameterise this choice to be included in the system design.
- Use Australian and Indigenous-specific data sources for compiling morbidity data. If data for a condition are not available for Australian estimates, then consider using the Australian prevalence rate from GBD 2010 which have been modelled from similar countries (from the region, and super-region where possible). If data for a condition are not available for the Indigenous population, or numbers are small, then consider adopting modelling strategies to derive prevalence estimates for the Indigenous population.
- Assess on a case-by-case basis whether to use multiple data sources or the single best data source for each condition for YLD calculations. Develop a protocol to assess suitability of data for inclusion to ensure transparency.
- Adopt the disability weights developed by GBD 2010 unless there is compelling evidence for invalidity in a particular context. Consider some targeted analysis using Australian data sources (e.g. AHS/ AATSIHS SF-12 component) to support interpretation in the Australian and Indigenous contexts.
- Adopt GBD method to adjust for comorbidity, subject to validation against Australian data sources for both the Australian and Indigenous populations.
- Undertake work to determine the most appropriate methods for calculating uncertainty intervals for YLDs based on measurable sources of error (e.g. RSEs derived from survey data and error from meta-analysis where relevant).

Additional considerations for the Indigenous estimates:

- Adjust for under-identification in hospital data using adjustment factors from the most recent AIHW hospital data quality study. Assess whether it is feasible to adjust for under-identification in other administrative data sources.

6 Burden of risk factors

A major component of previous global and Australian burden of disease analyses has been the estimation of the burden attributable to key risk factors. Quantification of the impact of risk factors in this way assists in making evidence-based decisions about where to direct efforts to improve population health and prevent disease and injury.

This chapter provides an overview of the methods used to calculate the burden of risk factors in GBD 2010, including the strengths and weaknesses of this approach, outlines differences from the methods used in the 2003 Australian and Indigenous studies, discusses the applicability of the GBD 2010 methods in the Australian context and outlines the most appropriate way of estimating the burden of risk factors for Australia. The risk factor list is discussed in Chapter 3 and is therefore not discussed in this chapter.

Key terms used in this chapter

attributable burden - the burden associated with a particular risk factor. It is the reduction in burden that would have occurred if exposure to the risk factor had been avoided

counterfactual - the alternative risk factor exposure distribution chosen for comparison with the observed distribution, in order to estimate the contribution of that risk factor to the burden of disease.

effect modification - a change in the observed magnitude or direction of an association between exposure and outcome with the addition of a third variable (such as age or sex).

effect size - a statistical measure of the strength of the relationship between two variables, which is relatively independent of sample size. Many statistical measures are effect sizes, for example, the standardised mean difference, the odds ratio, the relative risk, and measures of 'variance accounted for' by a model, such as R-squared or Pearson's correlation.

meta-analysis - the use of statistical methods to combine the results of individual studies. A pooled effect size is calculated using a weighted average method which takes into account differences in sample sizes and event rates between studies.

meta-regression - in meta-analysis, a way of investigating how study characteristics are associated with the effect size

population attributable fraction (PAF) - for a particular risk factor, the percentage reduction in disease, illness, disability or death in a population that would occur if exposure to the risk factor was avoided.

relative risk (RR) - the risk of an event relative to exposure, calculated as the ratio of the probability of the event occurring in the exposed group to the probability of it occurring in the non-exposed group. A relative risk of 1 implies no difference in risk; $RR < 1$ implies the event is less likely to occur in the exposed group; $RR > 1$ implies the event is more likely to occur in the exposed group.

theoretical-minimum-risk exposure distribution (TMRED) - the risk factor exposure distribution that will lead to the lowest conceivable disease burden.

(Refer to glossary for a full list)

Overview of GBD 2010 methods

The process for estimating the burden of disease attributable to selected risk factors involves five key steps: selection of risk-outcome pairs; estimation of exposure distribution; estimation of effect sizes; choice of counterfactual; and finally the calculation of burden (Figure 6.1). Each of these steps is described in more detail below, with particular attention given to new approaches adopted in GBD 2010 and a comparison with previous studies. The final choice of risk-outcome pairs used in GBD 2010 is outlined in Appendix 6A.



Figure 6.1: The basic process for calculating the burden of disease attributable to risk factors

Selection of risk-outcome pairs

The first step in the process involves determining which risk factors and outcomes should be included in the analysis. The choices are based on criteria defining policy relevance, data availability, causal associations and generalisability. The criteria used in GBD 2010 were:

- the likely importance of a risk factor to disease burden or policy
- availability of sufficient data and methods to enable estimation of exposure distributions
- sufficient evidence for causal effects based on high-quality epidemiological studies, and sufficient data to estimate outcome-specific aetiological effects sizes per unit of exposure
- evidence to support the generalisability of effect sizes to populations other than those included in the available studies, or satisfactory models for extrapolating them (Lim et al. 2012).

The application of these criteria resulted in several changes to the risk-outcome pairs used in previous studies (GBD 2000 and 2004, and Australian 2003). For example, ‘unsafe sexual practices,’ included as a risk factor in previous Australian burden of disease studies, was excluded in GBD 2010. Despite being both an important contributor to disease burden and of policy relevance, robust estimates of exposure and consistent definitions of ‘safe sex’ were lacking. Further, evidence suggests that effect sizes differ markedly across populations globally. For example, the generalised HIV epidemic in sub-Saharan Africa results in different relative risks than those observed in other regions, such as Australasia, which experience an epidemic largely concentrated in particular subpopulations. This lack of generalisability prevented an accurate estimate of the proportion of sexually transmitted infections that are attributable to unsafe sex by country on a global scale.

Estimation of exposure distribution

This step involves collating estimates of population exposure for the chosen risk factors for each age-sex group in the analysis. For GBD 2010, a systematic search was undertaken to identify published and unpublished data sources. Research strategies employed included database searches, manual reference checking and expert advice, resulting in data drawn

from both published and unpublished works, including censuses, health and nutrition surveys, community-based studies, and environmental measurements.

Estimating risk-factor exposure distribution at a global level presented difficulties, with data regularly incomplete or missing for entire regions or populations. In these circumstances, estimates were imputed via modelling using a wide array of relevant covariates.

For some risk factors, specific methods were used to generate exposure estimates. For 'tobacco smoking and second-hand smoke', the literature suggests that present prevalence data do not take into account likely variations in the intensity and duration of exposure, and so methods have been developed to estimate cumulative exposure. The GBD 2010 (like previous global and Australian studies) used lung cancer mortality (the smoking impact ratio, as described in Peto et al. 1992) as a marker of cumulative population exposure to smoking for the outcomes of cancers and chronic respiratory disease. For all other outcomes, 10-year lagged smoking prevalence was used. For 'occupational exposure to asbestos', mesothelioma mortality was used as a marker of asbestos exposure, similarly to the smoking impact ratio method. Exposure to 'airborne pollutants' was modelled based on two estimates of ambient particulate matter concentration (a global atmospheric simulation model and satellite observations) over a global grid. Ozone exposure was estimated using the global atmospheric simulation model. Zinc deficiency was estimated on the basis of dietary sources of zinc in the food supply.

Estimation of effect sizes

Some effect sizes were drawn directly from published studies or meta-analyses of epidemiological studies. In cases where no recent reviews or syntheses were available, new meta-analyses were undertaken as part of GBD 2010. These analyses were conducted to determine new effect sizes for some dietary risk factors, air pollution and unimproved water and sanitation.

Effect sizes used in GBD 2010 were adjusted for confounders but not for factors along the causal pathway. For example, the relative risk of ischaemic heart disease due to physical inactivity was not adjusted for high blood pressure.

GBD 2010 endeavoured to incorporate a broad evidence base from multiple data sources. Effect sizes were included from studies if statistically significant results were observed in either a) at least one age group or b) the pooled results across all age groups.

For some risk-outcome pairs, effect sizes are only available for either morbidity or mortality. In these cases, the relative risk was assumed to apply to both, unless there was evidence to suggest this was not the case. For example, for all dietary risk factors the same relative risks were applied to mortality and morbidity associated with ischaemic heart disease, except for diets low in seafood omega-3 fatty acids. A review of literature found dietary intake of omega-3 acids reduced mortality from IHD, but not morbidity. Therefore, only death from IHD was included as a risk-outcome pair with omega-3 intake.

For some conditions, evidence from epidemiological studies shows that effect sizes change with age. In the case of cardiovascular disease risks, GBD 2010 estimated age-specific relative risks for all cardiovascular disease risk factors using meta-regression. Previous burden of disease studies have incorporated reductions in relative risks with increasing age for cardiovascular outcomes, though often not in a systematic fashion.

Choice of counterfactual distribution

Counterfactual analysis involves estimating the contribution of a particular cause or risk factor to health loss by comparing the observed levels of health in a population to those that would be expected under an alternative, hypothetical scenario (the counterfactual). This scenario could be an increase or decrease in levels of exposure or changes in behaviour compared to the observed situation, including zero exposure, depending on the factor being examined. This optimal exposure is chosen to provide benefits in population health terms.

For GBD 2010, as in previous studies, a 'theoretical-minimum-risk exposure distribution' (TMRED) scenario was adopted. This involves determining the exposure distribution that will lead to the lowest conceivable disease burden. For some risk factors the choice of TMRED is obvious and involves zero exposure to risk: all persons are lifelong non-smokers; all babies are fully breastfed; all persons are highly active. However, for many risk factors zero exposure is not appropriate, either because it is physiologically impossible (for example, blood pressure or body mass index) or because there are lower limits beyond which exposure cannot feasibly be reduced (for example, particulate matter air pollution). In these cases, epidemiological evidence is used to determine the optimal level of exposure. The agreed level may reflect the lowest level at which a dose-response relationship can be observed within a meta-analysis of cohort studies, or the lowest risk factor exposure distribution observed globally.

GBD 2010 made adjustments to several of the counterfactual distributions used in previous studies, based on the availability of recently published information. For example, GBD 2010 revised the TMRED for systolic blood pressure from 115mm Hg down to 110 mm Hg due to recently published randomised trials. Similarly, the counterfactual ideal BMI was changed to a range of 21-23kg/m², compared to 21kg/m² used previously, including in the 2003 Australian studies. The definition of exposure and TMRED used for each risk factor in both the GBD 2010 and the 2003 Australian studies are included in Appendix 6A.

Significantly, it appears that no risk factor distributions used in GBD 2010 differed by sex. Previous burden of disease studies used a TMRED for low bone mineral density, measured as bone density at the femoral neck, which differed between males and females. However, GBD 2010 applied the same distribution across sexes, reflecting the principle of using estimates that are generalisable across populations. Differences by sex did occur at the risk factor level with the inclusion of intimate partner violence only for females, as data availability limited the ability to calculate estimates for males.

Calculation of attributable burden

The burden attributable to a particular risk factor is obtained by calculating the population attributable fraction (PAF) for each outcome and age-sex group separately for mortality and morbidity (where the effects of exposure are different). These PAFs are then multiplied by the relevant YLL or YLD estimates to obtain estimates of attributable fatal and non-fatal burden. These estimates are combined to calculate the total attributable burden for that risk factor (see Box 6.1). Simulations, based on 1,000 draws from the posterior exposure distribution, effect size, and outcomes by age, sex, country and year, were used to estimate uncertainty.

Box 6.1: Calculating attributable burden

For risk factors with a continuous distribution (e.g. blood pressure), the population attributable fraction (PAF) for a particular cause is calculated as:

$$PAF = \frac{\int_{x=0}^m RR(x)P1(x)dx - \int_{x=0}^m RR(x)P2(x)dx}{\int_{x=0}^m RR(x)P1(x)dx}$$

where RR(x) is the relative risk of that cause at exposure level x
P1(x) is the population exposure distribution
P2(x) is the counterfactual exposure distribution
m is the maximum exposure level.

For risk factors with a categorical distribution (e.g. physical activity), the PAF for a particular cause is calculated as:

$$PAF = \frac{\sum_{i=1}^n P_i(RR_i - 1)}{\sum_{i=1}^n P_i(RR_i - 1) + 1}$$

where RR_i is the relative risk of that cause for exposure category i
P_i is the fraction of the population in exposure category i
n is the number of exposure categories.

In each case, the attributable burden (AB) for a particular cause due to the risk factor is calculated as:

$$AB = PAF \times B$$

where B is the burden of disease for the cause in question.

Methodological differences from the 2003 Australian and Indigenous studies

The choice of risk factors and outcome pairs

As noted above, GBD 2010 included substantially more risk factors than the 2003 Australian and Indigenous studies, including 15 individual components of diet, a number of additional occupational exposures, and new risk factors such as high fasting plasma glucose and suboptimal breastfeeding. Calculations of disease burden across these risk factors required the input of considerably more data sources than previous studies and significant efforts to determine risk factor exposure distributions globally.

Despite the expansion of risk factors since 1990, some risk factors included in previous Australian studies were excluded due to methodological and data limitations (see Chapter 3 for information on which risk factors were excluded). These limitations were often a result of a lack of global data availability and/or data quality, the demand for generalisability of estimates across populations, and the potential for confounding. Within an Australian context, good quality national data may mean these concerns are not apparent and allowing the inclusion of additional risk-outcome pairs. Conversely, some risk factors, particularly occupational exposures to pollutants, may not warrant inclusion in Australia, due to the

absence of relevant industries or the presence of strict environmental controls that limit exposure.

Theoretical minimum risk exposures

Like GBD 2010, previous Australian burden of disease studies have drawn on systematic reviews and meta-analyses of epidemiological studies to obtain reliable estimates of risk factor exposure distributions. Additionally, some estimates were drawn from Australian sources, such as national data collections (such as the National Hospital Morbidity Database), surveys (the National Health Survey) and longitudinal studies (the Australian Longitudinal Study of Women's Health).

At times, TMREDs used in GBD 2010 have differed from those used in previous Australian studies. For example, the optimal alcohol exposure in GBD 2010 was 'no alcohol consumption,' whereas the 2003 Australian study used 'low alcohol consumption' as measured in the National Health Survey 2001. Low alcohol consumption was defined by alcohol standards outlined by the National Health and Medical Research Council and, significantly, these guidelines have since changed. The choice of a theoretical minimum risk exposure for alcohol is further complicated by the potential protective cardiovascular health effects observed with low levels of consumption. Despite using a TMRED of no alcohol consumption, some adjustment appears to have been made for the protective effects of low-level consumption in GBD 2010. In addition to the range of conditions attributed to alcohol, including diabetes mellitus, cirrhosis of the liver, and a number of cancers, alcohol is estimated to reduce the attributable burden of DALYs from cardiovascular and circulatory diseases by 2.8%. A review of current evidence will be required to determine an appropriate TMRED for use in the current Australian study.

Treatment of joint and multiple risk factors

Attempts were made in GBD 2010 to calculate the attributable burden due to the joint effect of multiple risk factors. Although the attributable burden for various causes can be added together to produce a total for a particular risk factor, it is not valid to add together the burden attributable to different risk factors to estimate the total burden for a specific cause. A multiplicative, rather than additive, approach (outlined in Box 6.2) is needed to allow for the measurement of the combined effect of joint risk (Ezzati et al. 2004).

Box 6.2: Calculating joint risk factor effect

Following the multiplicative approach employed in previous Australian burden of disease studies, the joint PAF of multiple risk factors are calculated as:

$$\text{Joint PAF} = 1 - \prod_{i=1}^n (1 - \text{PAF}_i)$$

where PAF_i is for individual risk factor i .

For example, high blood pressure, physical inactivity and tobacco use are three important risk factors for ischaemic heart disease globally. Their population attributable fractions are 0.53, 0.31 and 0.31 respectively.

To calculate the joint effect of these three risk factors on ischaemic heart disease:

$$= 1 - (1 - 0.53) \times (1 - 0.31) \times (1 - 0.31) = 0.78$$

Therefore, 78% of ischaemic heart disease globally can be attributed to high blood pressure, physical inactivity and tobacco use.

A limitation of this methodology is the assumption that risk factors are both biologically independent and uncorrelated. According to Ezzati and others (2004), the assumption of independence of risk factors may not hold when:

- the effects of distal factors (such as physical inactivity) are mediated through proximal factors (such as high blood pressure), particularly when these factors share a common causal pathway
- effect modification occurs. For example, the interaction between tobacco smoking and lung cancer is altered by exposure to asbestos. It is unclear how, or if, the significantly larger relative risks associated with the combined exposure to these two risk factors has been taken into account in GBD 2010.
- there is correlation between exposure to particular groups of risk factors.

The combined effects of clusters of risk factor exposures were calculated in GBD 2010 using the above approach, with risk factor clusters presented in Table 6A.1 (Appendix 6A).

However, due to the high degree of correlation and mediation within the physiological risk factors and air pollution clusters, the joint effect of these risk factors was not calculated.

Previous Australian and Indigenous studies also adopted the method proposed by Ezzati and others (2004) to calculate the combined effect of 14 and 11 risk factors respectively. The joint effect of these risk factors was disaggregated to the level of seven broad cause groups (cancer, cardiovascular disease, mental, neurological, injury, diabetes and other), by sex, and by age-group (0-44, 45-64, and 65+ in the national study, and 0-14, 15-34, 35-54, and 55+ in the Indigenous study).

Uncertainty intervals

In GBD 2010, uncertainty analysis occurs throughout the study including for estimates of risk factor attribution. Estimates of uncertainty around PAFs relied on a simulation method based on 1,000 draws from the exposure distribution, effect size, and outcome by age, sex, country and year..

Applicability of GBD 2010 methods to the Australian context

Exposure estimates and multiple data source approach

As mentioned previously, GBD 2010 used all available data sources for the estimation of risk factor exposure estimates. For the national and Indigenous estimates, it is expected that some risk factors will benefit from the use of multiple data sources (i.e. a mix of self-reported data from the Australian Health Survey, and epidemiological studies), while other risk factors will be most accurately estimated using a single good quality national data source. Therefore whether a multiple or single best data source approach will be taken may best be assessed on a case-by-case basis for each risk factor.

Choice of effect size, risk-outcome pairs and exposure distributions

While effect sizes based solely on Australian epidemiological studies could be sourced for some of the risk-outcome pairs included in GBD 2010, complete and timely data are not available to estimate relative risks for all risk factors. The availability of effect sizes for subnational populations will be even further limited.

One criterion used to decide upon the inclusion of a risk-outcome pair in the GBD 2010 analysis was the availability of 'evidence to support generalizability of effect sizes to populations other than those included in the available epidemiological studies or satisfactory models for extrapolating them' (Lim et al. 2012, 2226). The GBD authors maintain that there is increasing evidence to support the view that relative risks reflect intrinsic biological relationships which are common across all humanity (Harvard Initiative for Global Health 2009). Reflecting this principle of generalisability, past Australian BoD studies have used relative risk estimates which were based on meta-analyses of both national and international epidemiological studies (Begg et al. 2007; Vos et al. 2007). If the relative risks used in GBD studies are generalisable, then there may be no compelling requirement for Australian-specific estimates.

Limitations surrounding generalisability and data quality prevented the inclusion of some risk factors in GBD 2010. However, good quality Australian data may support the inclusion of particular risk-outcome pairs that were excluded such as unsafe sex and HIV/AIDS. Chapter 7 of this report describes potential sources of good quality data nationally. The 2011-12 Australian Health Survey and Australian Aboriginal and Torres Strait Islander Health Survey will be important data sources for the Australian BoD study, along with other surveys and epidemiological studies.

Theoretical minimum risk exposures draw on a variety of epidemiological sources to determine the exposure that would result in the lowest conceivable disease burden. This may be determined via meta-analyses or as the lowest exposure distribution observed globally. Therefore, they should, by definition, remain consistent across all populations and it would seem appropriate to adopt the TMREDs used by GBD 2010 for the current Australian study.

Uncertainty analysis

Uncertainty intervals could be calculated using error estimates generated from survey data. For example, the Aboriginal and Torres Strait Islander Health Survey will be an important

data source for the Indigenous estimates. Relative standard errors from the survey can be used to calculate uncertainty intervals. It would be appropriate for this issue to be included in a review of uncertainty interval methods for the whole study.

Issues relating to Aboriginal and Torres Strait Islander people

Are the GBD 2010 effect sizes appropriate for the Indigenous population?

There is a small amount of evidence in the literature which suggests that there may be some ethno-racial variation in effect sizes for certain risk-outcome pairs (Shai et al. 2006, Lee et al. 2001, Hunt et al. 2005, Vasilevska et al., 2012). Researchers have proposed that Indigenous Australians have an increased genetic susceptibility to certain diseases. The 'thrifty genotype' hypothesis, for example, maintains that Aboriginal and Torres Strait Islander persons have a genetic predisposition to develop diabetes resulting from obesity, and are therefore at greater risk than non-Indigenous Australians.

However, others maintain that there are ongoing concerns regarding the evidence for this hypothesis (Paradies, Montoya & Fulleton 2007) and uncertainty still surrounds the issue of increased Indigenous risk due to genetic predisposition. Studies exploring race and ethnicity are frequently confounded by a variety of social processes, including education, employment, and socioeconomic status. Despite statistical adjustment for these factors, study results potentially confuse residual confounding for genetic effect (Cooper & Freeman 1999). Furthermore, while racial variation can be observed in the genetic risk of some rare conditions such as thalassaemia, sickle cell anaemia and cystic fibrosis (Karter 2003, Cutting 1997), self-identified Indigenous status may serve as a poor marker of genetic background and ignore the genetic diversity found within the Aboriginal and Torres Strait Islander population.

Significantly, while some of the studies listed above may suggest differences in certain relative risk estimates by ethnicity, none of the published studies relate to Aboriginal and Torres Strait Islander people. In view of the above, it seems reasonable not to use separate relative risk estimates for the risk-outcome pairs used in GBD 2010 for the Indigenous estimates.

Are the GBD 2010 risk-outcome pairs and exposure distributions appropriate for the Indigenous population?

As the risk-outcome pairs included in GBD 2010 were chosen based on a strict set of inclusion criteria which included large-scale epidemiological studies and evidence to support the generalisability of effect sizes to populations other than those included in the available studies, it would seem appropriate for these to be included in the Indigenous component of the study. However a consideration is whether any additional risk-outcome pairs should be added.

As with the national component, good quality Indigenous Australian data (e.g. from the 2012-13 AATSIHS and epidemiological studies) may support the inclusion of particular risk-outcome pairs in the Indigenous component that were excluded from GBD 2010. These may differ slightly to those included in the national component, depending on the risk factors agreed to be included in the Indigenous component.

By definition, the theoretical minimum risk exposure distributions from GBD 2010 should be applicable to all population groups including Indigenous Australians.

Is adjustment for Indigenous under-identification required?

The main data source that is expected to be used for calculating risk factor exposure estimates for the Indigenous population is the 2012-13 AATSIHS. This data source does not require adjustment for under-identification as the survey is weighted to the total Indigenous estimated resident population. Whether adjustment for under-identification is applicable (and possible) for other data sources used in the risk factor analysis could be assessed by the AIHW.

How to deal with small number issues?

There may be some risk factors we wish to include in the Indigenous analysis, for which there are relatively small numbers when data are disaggregated by age and sex. Small numbers will affect the robustness of the estimates produced.

In such cases, modelling strategies could be used to derive estimates for each age group and sex (e.g. by applying the all-ages Indigenous exposure rate or Indigenous/non-Indigenous rate ratio across each age-group). Some methods employed to calculate estimates for the Maori population in the recent New Zealand BoD study may be adopted for the Indigenous estimates.

Intended approach

The following outlines our intended approach for the risk factor estimates for the Australian Burden of Disease Study, subject to further consultation with the Expert Advisory Group and the Indigenous Reference Group:

- Use Australian data for risk factor exposure estimates for the Australian and Indigenous populations.
- Assess on a case-by-case basis whether to use multiple data sources or the single best data source for each risk factor, according to a protocol to be developed with the assistance of the EAG. For some risk factors, best estimates may be obtained by combining data from two or more sources, while for others, good quality estimates could be obtained from a single data source such as the AHS/AATSIHS.
- Adopt GBD specific methods for estimating exposure for selected risk factors (e.g. smoking based on smoking impact ratio; ozone exposure; asbestos and airborne pollutants).
- Adopt GBD risk-outcome pairs, effect sizes and TMREDs for those risk factors to be used in the Australian, Indigenous and sub-national analysis.
- Consider methods for calculating uncertainty intervals based on measurable sources of error (e.g. using relative standard error estimates generated from survey data and meta-analysis).
- Add certain risk-outcome pairs to the national and Indigenous analyses where good Australian or Indigenous Australian data are available (e.g. unsafe sex and HIV).
- Calculate the attributable burden for the combined set of risk factors included in the Australian and Indigenous studies (method still to be determined).

Additional considerations specific to the Indigenous estimates:

- For risk factor exposure data sourced from the 2012-13 AATSIHS, adjustment for under-identification is not required. For other data sources used, assess whether it is appropriate and possible to adjust for under-identification.

Appendix 6A

Table 6A.1: Risk-outcome pairs and theoretical minimum exposure definitions used in GBD 2010

Risk factor	Outcomes	Exposure definition	Theoretical minimum exposure
Unimproved water	Intestinal infectious diseases	Proportion of households using unimproved water sources	All households use improved water
Unimproved sanitation	Intestinal infectious diseases	Proportion of households using unimproved sanitation	All households use improved sanitation
Ambient particulate matter pollution	Lower respiratory infections, trachea, bronchus and lung cancers, IHD, ischaemic stroke, COPD,	Ambient concentration of particles with an aerodynamic diameter smaller than 2.5 µm, measured in µg/m ²	5.8-8.8 µg/m ²
Household air pollution from solid fuels	Lower respiratory infections, trachea, bronchus and lung cancers, IHD, ischaemic stroke, COPD, cataracts	Proportion of households using solid fuels for cooking	All households use clean fuels for cooking
Ambient ozone pollution	COPD	Ambient concentration of ozone in air	33.3-41.9 parts per billion
Residential radon	Trachea, bronchus and lung cancers	Residential radon, measured in Bq/m ³	10 Bq/m ²
Lead exposure	Cardiovascular and circulatory diseases, mental and behavioural disorders, chronic kidney disease	Blood lead (measured in µg /dL) and bone lead (measured in µg/g)	Bone lead level expected for age-specific cumulative exposure to blood lead of 0.09652 µmol/L
Non-exclusive breast feeding	Diarrhoea, lower respiratory infections and other infectious diseases; nutritional deficiencies; neglected tropical diseases and malaria	Proportion of children younger than 6 months with predominant, partial, or no breastfeeding	All children exclusively breastfeed for 6 months
Discontinued breast feeding	Diarrhoea, lower respiratory infections and other infectious diseases; nutritional deficiencies; neglected tropical diseases and malaria	Proportion of children aged 6-23 months with discontinued breastfeeding	Continued breastfeeding until 2 years
Childhood underweight	Diarrhoea, lower respiratory infections and other infectious diseases; nutritional deficiencies; neglected tropical diseases and malaria	Proportion of children less than -3SD, -3 to -2 SDs, and -2 to -1SDs off the WHO standard weight-for-age curve	Proportion of WHO reference population in each SD range
Iron deficiency	Iron-deficiency anaemia	Haemoglobin, measured in g/L	Country-specific
Vitamin A deficiency	Intestinal infectious diseases; measles; vitamin A deficiency	Proportion of children with serum retinol concentration less than 70 µmol/L	No childhood vitamin A deficiency
Zinc deficiency	Intestinal infectious diseases; lower respiratory infections	Proportion of children with inadequate zinc intake based on estimated mean daily amount of absorbable zinc per head in the food supply compared with mean physiological requirements	No inadequate zinc intake

(Continued)

Table 6A.1 (Cont.): Risk-outcome pairs and theoretical minimum exposure definitions used in GBD 2010

Risk factor	Outcomes	Exposure definition	Theoretical minimum exposure
Tobacco smoking	Cardiovascular and circulatory disease; cancers; chronic respiratory diseases; diarrhoea, lower respiratory infections and other infectious diseases; tuberculosis	Smoking impact ratio for cancers and chronic respiratory disease; 10-year lagged tobacco smoking prevalence for all other causes	No tobacco smoking
Second-hand smoke	Cardiovascular and circulatory disease; cancers; diarrhoea, lower respiratory infections and other infectious diseases	Proportion of children and non-smoking adults reporting exposure to second-hand smoke	No second-hand smoke exposure
Alcohol use	Cardiovascular and circulatory disease*; mental and behavioural disorders; cancers; cirrhosis of the liver; neurological disorders; diarrhoea, lower respiratory infections and other infectious diseases; tuberculosis; unintentional injuries; intentional injuries; transport injuries	Average consumption of pure alcohol and proportion of the population reporting binge consumption of 0.06kg or more of pure alcohol on a single occasion	No alcohol consumption
Drug use	Mental and behavioural disorders; cancers; cirrhosis of the liver; intentional injuries; HIV/AIDS and tuberculosis; ; other communicable, maternal, neonatal and nutritional disorders;	Proportion of the population reporting use of cannabis, opioids, and amphetamines; proportion of the population reporting use of injecting drugs	No use of cannabis, opioids, or amphetamines; no use of injecting drugs
High fasting plasma glucose	IHD, ischaemic stroke, diabetes, CKD, tuberculosis	Fasting plasma glucose, measured in mmol/L	Mean 4.9-5.3mmol/L
High total cholesterol	IHD, ischaemic stroke	Total cholesterol, measured in mmol/L	Mean 3.8-4.0mmol/L
High blood pressure	Cardiovascular and circulatory diseases; CKD	Systolic blood pressure, measured in mm Hg	Mean 110-115 mm Hg
High body-mass index	cardiovascular and circulatory disease; diabetes, urogenital, blood and endocrine diseases; cancers; musculoskeletal disorders	Body-mass index, measured in kg/m ²	Mean 21-23 kg/m ²
Low bone mineral density	Falls	Standardised bone mineral density measured at the femoral neck	90 th percentile of NHANES-III cohort by age
Diet low in fruit	Cardiovascular and circulatory disease, cancers	Dietary intake of fruit	Mean 300g/day
Diet low in vegetables	Cardiovascular and circulatory disease, cancers	Dietary intake of vegetables	Mean 400g/day
Diet low in whole grains	IHD, ischaemic stroke, diabetes	Diet intake of whole grains	Mean 125g/day
Diet low in nuts and seeds	IHD	Dietary intake of nuts and seeds	Mean 114g/week

(Continued)

Table 6A.1 (Cont.): Risk-outcome pairs and theoretical minimum exposure definitions used in GBD 2010

Risk factor	Outcomes	Exposure definition	Theoretical minimum exposure
Diet low in milk	Colorectal cancer,	Dietary intake of milk	Mean 450g/day
Diet high in red meat	Colorectal cancer, diabetes	Dietary intake of red meat	Mean 100g/week
Diet high in processed meat	Colorectal cancer, diabetes, IHD	Dietary intake of meat preserved by curing, smoking, salting or the addition of chemical preservatives	No dietary intake of processed meat
Diet high in sugar-sweetened beverages	Cardiovascular and circulatory disease, cancers, musculoskeletal disorders, diabetes, CKD	Dietary intake of beverages with ≥ 50 kcal per 226.8g serve, excluding 100% fruit and vegetable juices	No dietary intake of sugar-sweetened beverages
Diet low in fibre	Colorectal cancer, IHD	Dietary intake of fibre from all sources	Mean of 30g/day
Diet low in calcium	Colorectal cancer, prostate cancer	Dietary intake of calcium from all sources	Mean of 1200mg/day
Diet low in seafood omega-3 fatty acids	Death by IHD	Dietary intake of eicosapentaenoic acid and docosahexanoic acid, measured in mg/day	250mg/day
Diet low in polyunsaturated fatty acids	IHD	Dietary intake of omega-6 fatty acids from all sources	Substitution of present saturated fatty acid intake to a mean intake of polyunsaturated fatty acids of 12% of energy
Diet high in trans fatty acids	IHD	Dietary intake of transfats from all sources	Mean of 0.5% of energy
Diet high in sodium	Cardiovascular and circulatory diseases, CKD, stomach cancer	24 hour urinary sodium, measured in mg/day	Mean of 1000mg/day
Physical inactivity and low physical activity	Brest cancer, colorectal cancers, diabetes, IHD, ischaemic stroke	Proportion of the population in specified categories of physical activity	All individuals are highly active
Occupational risk factors	Chronic respiratory disease, cancers, lower back and neck pain, sense organ diseases, unintentional injuries, transport injuries	Proportion of population ever exposed, based on the distribution of the population in nine relevant industries	No occupation exposure to relevant carcinogens
Childhood sexual abuse	Alcohol use disorders, unipolar depressive disorders, intentional self-harm	Proportion of population who have ever experienced unwanted non-contact, contact abuse, or intercourse with an older person when aged 15 years or younger	No childhood sexual abuse
Intimate partner violence	Abortion, unipolar depressive disorders, intentional self-harm, interpersonal violence	Proportion of the population who have ever experienced acts of physical or sexual violence by a present or former partner since age 15 years	No intimate partner violence

* Low level alcohol exposure is assumed to have a protective effect on some cardiovascular and circulatory conditions

Source: Lim et al. 2012.

Table 6A.2: Risk-outcome pairs and theoretical minimum exposure definitions used in 2003 Australian Burden of Disease study

Risk factor	Outcomes	Definition	Theoretical minimum exposure
High blood pressure	IHD, stroke, hypertensive heart disease	Level of usual systolic blood pressure	115mm Hg
High blood cholesterol	IHD, stroke	Level of usual total blood cholesterol	3.8mmol/L
High body-mass index	IHD, stroke, hypertensive heart disease, diabetes, osteoarthritis, endometrial, kidney, colon and breast cancers	Body mass index	21kg/m ²
Low fruit and vegetable consumption	IHD, stroke, colorectal, gastric, lung and oesophageal cancers	Fruit and vegetable intake per day	600g intake per day for adults
Osteoporosis	Fractures from falls, striking or crushing accidents, other unintentional injuries	Bone mineral density of the femoral neck	Males 1.107g/cm ² Females 1.018g/cm ² With osteoporosis defined as ≥ 2.5 SD below these means
Physical inactivity	IHD, stroke, breast and colon cancers, diabetes	Four categories: inactive, insufficient, recommended, and highly active	All in 'highly active' group
Past smoker	COPD, cancer of the mouth, oesophagus, lung, pancreas, larynx, bladder, kidney, stomach and uterus.	Past smoker	No smoking
Current daily smoker	IHD, stroke, peripheral vascular disease, Parkinson disease, pneumonia, fire injuries, macular degeneration	Current daily smoker	No smoking
Passive smoker	IHD, stroke	Passive smoker	No smoking
Maternal smoker, smoking whilst pregnant	Asthma, pneumonia (children), sudden infant death syndrome, otitis media, low birth weight	Maternal smoking, smoking whilst pregnant	No smoking
Alcohol	Cancer of the mouth and oropharynx, oesophagus, liver, larynx and breast, inflammatory heart disease, hypertensive heart disease, IHD stroke, alcohol dependence and harmful use, gallbladder and bile duct disease, pancreatitis, road traffic accidents, falls, fires/burns/scalds, drowning, machinery accidents, suffocation and foreign bodies, suicide and self-inflicted injuries, homicide and violence, occupational injuries	Average number of standard drinks per day	Low level drinking
Illicit drug use	HIV/AIDS, hepatitis B, hepatitis C, inflammatory heart disease, suicide and self-inflicted injuries, road traffic accidents		Abstinence

(Continued)

Table 6A.2 (Cont): Risk-outcome pairs and theoretical minimum exposure definitions used in 2003 Australian Burden of Disease study

Risk factor	Outcomes	Definition	Theoretical minimum exposure
Daily cannabis use	Schizophrenia		No cannabis use, or use less often than daily
Unsafe sex	Sexually transmitted diseases, abortion, cervical cancer, HIV/AIDS, hepatitis B, hepatitis C	Unprotected sex	Abstinence or only protected sex
Child sexual abuse	Anxiety & depression, alcohol, heroin or polydrug, benzodiazepine, cannabis, and other drug dependence & harmful use, suicide and self-inflicted injury	Non-contact only, contact only, intercourse	No child sexual abuse
Intimate partner violence	Anxiety & depression, alcohol, heroin or polydrug, benzodiazepine, cannabis, and other drug dependence & harmful use, suicide and self-inflicted injury, tobacco smoking, cervical cancer, syphilis, chlamydia, gonorrhoea, other sexually transmitted diseases, anorexia nervosa, bulimia nervosa, other eating disorders, falls, unintentional injuries, homicide & violence	Physical or sexual violence by current or previous partner	No history of sexual or physical violence by an intimate partner
Occupational exposures and hazards	All accidents, intentional and unintentional injuries, cancers, heart disease, neurological disorders, chronic respiratory disorders, renal disease, osteoarthritis, slipped disc, occupational overuse syndrome	Exposure in the workplace to disease-causing agents such as carbon monoxide, dyes, inorganic and organic dusts, pesticides, metals, metal fumes, petrochemicals, plastics, solvents, isocyanate and nitroglycerine or nitroglycerol	No exposure
Urban air pollution	Short term: cardiovascular respiratory and other deaths Long term: lung cancer, IHD, stroke, inflammatory heart disease, hypertensive heart disease and COPD	Exposure to particulate matter and/or oxygen (i.e. total population of cities of interest)	No exposure

Source: Beggs et al. 2007.

7 Data sources

Why are data sources important to BoD estimates?

The choice of data sources used in burden of disease studies has a direct impact on the YLL, YLD, risk factor and DALY estimates produced. The number, type, range and quality of data sources used will all have an impact on the accuracy of the resulting estimates.

This chapter discusses the main data sources used in GBD 2010, including those used for Australian estimates, based on current available information. An assessment of their suitability for the Australian BoD study is provided, noting whether updated data or alternative data sources are available.

Information about the data sources used in the 2003 Australian and Indigenous BoD studies is also provided. Differences in the approach taken between these studies and GBD 2010 in the identification and use of all available data sources (GBD 2010) versus the use of the single best data source (2003 Australian BoD studies) is discussed.

A number of likely data improvements to Australian estimates produced in GBD 2010 are discussed including updated vital registration data, updated health survey data, and data linkage capabilities. A preliminary list of data sources for conditions and risk factors for which the Australian Health Survey is likely to be inadequate to produce prevalence estimates is included.

As GBD 2010 did not include estimates for Indigenous Australians, a discussion of relevant data sources for updated Indigenous estimates is provided. Issues that need to be considered in identifying data sources for Indigenous estimates are also discussed in this chapter. This includes data for specific conditions or risk factors not being available for the Indigenous population; small sample size and small number issues; and adjustments for under-identification.

GBD 2010 data sources

GBD 2010 had a heavy emphasis on systematically identifying all published and unpublished data sources for mortality, risk factor exposure and disease prevalence estimates. All identified data sources were used in the estimation of DALYs, rather than choosing only those that were considered to have the best data quality. It is the AIHW's understanding that models were used to take into account the quality of the estimates produced which utilised information such as the number of observations, the stability of the trends and the number of data points in the time series to predict variation due to measurement bias and variation in the 'true' prevalence rates. Other quality measures (i.e. scope, coverage) were included as covariates in DISMOD-MR as a way to account for differences in data quality between different data sources.

The strengths of this approach (to use all available data sources) is that estimates produced for a particular condition are based on all available information and therefore could be considered to be more accurate than estimates produced from a single data source. This approach also means that subjective judgements do not need to be made to decide which data sources are of better quality than others and therefore avoids selection bias. Weaknesses of this approach are that more data sources need to be sourced and are required to be inputted for each condition (more time consuming and labour intensive); a meta-analysis

tool (such as DISMOD-MR or METAXL) is required in order to handle multiple data sources and take into account data quality; and the inclusion of poor quality data sources may affect the accuracy of the results if quality is not adequately taken into account through the modelling process.

What types of data sources were used?

GBD 2010 used several different types of data sources for disease prevalence estimation. IHME indicated at the GBD Australian Launch (May 2013) that some types of data sources were considered to be better sources of information for particular diseases than others, and a strategy was developed for what was considered the best type of data source for capturing each disease and sequelae. For example, published literature was used for epilepsy; population surveys (as well as published literature) were used for mental health; and hospitalisation data were used for appendicitis.

The types of data sources used in GBD 2010 to produce YLL (mortality data), YLD (based on disease prevalence data) and risk factor exposure estimates are listed below.

Mortality

GBD 2010 identified and assembled a multi-year and multi-source array of input data on causes of death for each country based on a combination of the following data sources:

- civil/vital registrations with medical certification of deaths
- verbal autopsy
- cancer registers
- police reports
- surveys/censuses
- maternal mortality
- deaths in health facilities
- burial and mortuary
- sibling history.

The resulting global mortality data are believed to be incomplete and patchy necessitating modelling to fill these data gaps. For example, not all countries have comprehensive or even partial estimates of deaths and causes of death for the reference year (or for any year); and the population coverage (e.g. age, geography) and granularity of age and cause of death varies by country.

Disease prevalence

GBD 2010 sourced data on incidence, prevalence, readmission, duration of disease and excess mortality to generate prevalence estimates for each disease and sequela. The types of data sources used included:

- systematic reviews of published literature
- household surveys (e.g. interview questions, direct measurements, serological measurements and anthropometry from re-analysis of multiple household surveys) – for chronic diseases, functional health status

- population-based disease registers – for cancers, chronic kidney disease, multiple sclerosis, Parkinson’s disease and congenital abnormalities
- hospital discharges – for conditions requiring acute care e.g. stroke, myocardial infarction, appendicitis, pancreatitis, serious injuries
- outpatient visits – for conditions such as skin diseases and mental and behavioural disorders
- cohort or follow-up (longitudinal) – for causes such as impairment due to injury, and to provide information about remission rates, duration and mortality risks for many chronic diseases
- surveys of selected populations – school children for intellectual disability, nursing home residents for dementia, and mental health clinic attendees for schizophrenia
- antenatal clinic surveillance – for conditions such as HIV/AIDS and syphilis
- notifiable diseases databases – for conditions such as measles, pertussis, tuberculosis, leprosy, dengue, cholera, yellow fever
- indirect prevalence studies (combining data from treatment centres, police records, court records and survey data) – to estimate total number of drug users
- censuses – for denominators.

In addition to the above, GBD 2010 assembled available data on the distribution of severity for conditions specified as having more than one severity level. Importantly, these data drew upon population surveys conducted in the United States and Australia only.

As with mortality, the resulting global prevalence data are believed to be somewhat incomplete and patchy necessitating modelling to fill data gaps. For example, for some causes/sequelae, limited epidemiological data were available for some countries and some years.

Risk factors

For most risk factors included in GBD 2010 a systematic search was done to identify published and unpublished data sources. The search strategy covered survey databases such as the WHO Global Database on Child Growth and Malnutrition, general citation databases such as Google Scholar and PubMed, manual searching of reference lists of articles and conference abstracts, and contacting experts in the relevant fields. The search identified risk factor exposure estimates in:

- household surveys
- health examination and nutrition surveys
- community-based epidemiological studies.

The data were used to estimate risk factor exposure distributions.

What data sources were used specifically for Australian estimates?

Currently, with the exception of YLL estimates, little information is publically available on the specific data sources used to produce Australian estimates in GBD 2010. The visualisation tools on the IHME website have recently been updated to show the data sources used for YLL estimates, however the same detail is not included for YLDs or risk factors.

The Global Health Data Exchange (GHDx) (see Box 7.1) on the IHME website includes a list and description of Australian 'data set records'. However this is not the complete set of data sources used for Australia. The AIHW understands that IHME will be adding data sources to the GHDx and that a more complete list will be available on IHME's website in the future.

It should be noted, that where data for Australia were not available for a particular condition for GBD 2010, Australian estimates were produced from first borrowing data from the region (Australasia which includes New Zealand) if available, followed by data from the super-region (which includes similar high income countries).

Box 7.1: Global Health Data Exchange

The Global Health Data Exchange (GHDx) is a data catalogue created and supported by IHME. It was created as a dedicated place for anyone interested in global health and demography to quickly find and share information about data, along with actual datasets.

A dataset may be the output of an ongoing data collection system, where the dataset is an extract of the system: a selection of data from the system where the selection is based on some useful criteria; most common is the data on a particular topic for a particular year. The data in the dataset may stem from primary data collection (e.g. survey) or secondary data via aggregation or synthesis.

In the GHDx, a record is a catalogue entry for a particular dataset. A record consists of general metadata about the dataset, a citation and other source information, and information about where to obtain the dataset.

Table 7.1 below lists the Australian 'data set records' found on the Global Health Data Exchange (GHDx). The AIHW's assessment of what these data sources were used for and what years of data are listed on the GHDx for each source are also included.

A number of state and territory cancer incidence and mortality reports appear on the list, as well as vital registrations data (for mortality) and two global surveys (World Health Survey and World Values Survey). The only national data sources included on the GHDx are the Census of Population and Housing, the National Health Survey, the Longitudinal Study of Australian Children, and the Australian Studies of Health and Relationships. The Austcare Annual Report and the Australian Household Expenditure Survey are also listed on the GHDx, which are likely to be included for their use in health expenditure work rather than burden of disease estimates. IHME also utilised data from the World Health Organisation (WHO) for some Australian estimates, for example, mortality data held by WHO for YLL estimates.

Table 7.1: Australian data set records included on the Global Health Data Exchange for GBD 2010

Dataset	Content	Year/s of data used
Austcare Annual Report	Contains information on financial assistance for health	2004-2007
New South Wales Cancer Incidence and Mortality (ACIM) Books	Cancer incidence; mortality	1968-2005
New South Wales Cancer Incidence and Mortality Report	Cancer incidence ; mortality	2004-2008
New South Wales Cancer Incidence, Mortality, and Prevalence Report	Cancer incidence, mortality, and prevalence; cancer survival	2005
New South Wales Survival from Cancer Report	Cancer survival	1980-2003
Tasmania Cancer Incidence and Mortality Reports	Cancer incidence and mortality	1995-2006
Australia Household Expenditure Survey	Expenditure	1984; 1988-98
Australia Longitudinal Study of Children	Selected risk factors	2003-2010
National Health Survey	Disease prevalence; risk factors	1995-96
Population and Housing Census	Population denominators	1986, 2001, 2006
Australian Studies of Health and Relationships, Women's Survey	STI diagnosis and selected risk factors (e.g. sexual behaviours, drug use, intimate partner violence)	2001-2002
Vital Registration - Deaths (WHO database)	Number of deaths by cause	1950-2006
World Health Survey	Some disease prevalence	2003
World Values Survey	Questions in survey relate to social values and not prevalence of risk factors or disease	2005 (Australia)

A key data gap identified for Australian YLD estimates is that Australian hospitalisation data were not included in GBD 2010.

It should be noted that IHME are planning on updating country-level data for release of 2013 estimates, and release of 2014 estimates in 2015. At a minimum, it is expected that updated mortality data from WHO and published data from the 2011-12 AHS will be used to update Australian estimates.

What about data for Indigenous estimates?

GBD 2010 did not include data specifically on the Indigenous population.

Methodological differences from the 2003 Australian and Indigenous studies

In contrast to GBD 2010 which used all available data sources, the 2003 Australian and Indigenous BoD studies identified the single best data source for each disease and risk factor. The strength of this approach is that the estimates are based on high quality data. A weakness is that this requires judgments to be made about data quality.

Similar to GBD, particular types of data sources were considered to be better sources of information for particular diseases than others.

The full list of data sources used in the 2003 Australian BoD study and the 2003 Indigenous BoD study can be found at Appendix 7A and 7B of this chapter.

These include disease registers, surveillance and notification systems (e.g. notifiable diseases and cancer registers), health service utilisation data (e.g. hospitalisations), and population health surveys (e.g. National Health Survey). In addition to these data sources, epidemiological studies and indirect estimation were used to derive incidence and prevalence estimates for some diseases; and risk factor models and attributable fractions from published studies were used to derive estimates for some risk factors.

YLL estimates in both the Australian and Indigenous 2003 BoD studies used data from the AIHW National Mortality Database. The main data sources used for disease incidence and prevalence estimates in both studies were the National Notifiable Disease Surveillance System, hospitalisation data, and the 2004-05 National Health Survey/ National Aboriginal and Torres Strait Islander Health Survey (NHS/NATSIHS). The NHS/NATSIHS were the main data sources used for risk factors in both the Australian and Indigenous BoD studies. AusDiab was another key source used for the Australian study.

In the 2003 Indigenous BoD study, data on directly observed Indigenous health events in health surveys, administrative collections and epidemiological studies were used where available. However, where this information was not available, ratios of the difference between Indigenous and total population rates were used as proxy measures of disease occurrence, such as hospital and mortality data.

Applicability of the GBD 2010 data sources to the Australian context

Which data sources used in GBD are relevant to the Australian BoD Study?

The Australian data set records included on the GHDx were examined and assessed to determine whether updated data are available, whether they include an Indigenous identifier, whether sub-national estimates could be produced, whether an alternative data source exists that will provide more relevant and/or up-to-date information, and their overall relevance to an updated Australian BoD study. This information is presented in Table 7.2 below.

Of the data sources listed, three are considered relevant to the Australian BoD Study, with each having more recent data available since GBD 2010. These are the Longitudinal Study of Australian Children (for selected risk factors), the Census (for population denominators), and the National Health Survey (the 2011-12 survey being known as the Australian Health Survey, for disease prevalence and risk factors). Mortality and cancer incidence would be sourced from the national data in the AIHW National Mortality Database and the AIHW Australian Cancer Database, respectively, rather than state and territory cancer incidence and mortality reports or vital registration data held by the WHO.

Table 7.2: Assessment of Australian data set records included on the Global Health Data Exchange

Dataset	Updated data available?	Indigenous identifier	Sub-national	Relevance to Australian BoD/ alternative data sources
Austcare Annual Report	Yes	No	S/T	Content not relevant
New South Wales Cancer Incidence and Mortality (ACIM) Books	Yes - 2009	No	NSW	Alternative national data sources: Australian Cancer Database; National Mortality Database (NMD)
New South Wales Cancer Incidence and Mortality Report	No	No	NSW	Same as above
New South Wales Cancer Incidence, Mortality, and Prevalence Report	No	No	NSW	Same as above
New South Wales Survival from Cancer Report	Yes – 2002-2006	No	NSW	Same as above
Tasmania Cancer Incidence and Mortality Reports	Yes - 2009	No	Tas	Same as above
Australia Household Expenditure Survey	Yes	No	S/T	Relevant for health expenditure estimates only
Australia Longitudinal Study of Children	Yes- 2013	No	S/T	Yes – useful for selected risk factors
National Health Survey	Yes - 2011-13	Yes	S/T; Remoteness ; SEIFA	Yes – 2011-13 AHS will be used for selected prevalence and risk factors
Population and Housing Census	Yes - 2011	Yes	S/T; Sub-state; remoteness; SEIFA	Yes – for population denominators
Studies of Health and Relationships, Women's Survey	No	No	No	Alternative data source: Longitudinal Survey on Women's Health 2011
Vital Registration - Deaths (WHO database)	Yes - 2011	Yes	S/T	Alternative data source: AIHW NMD
World Health Survey	No	No	No	Not relevant – old data and does not provide any additional information to AHS
World Values Survey	No	No	No	Content not relevant

Which data sources used in GBD are appropriate for sub-national estimates?

GBD 2010 did not produce sub-national estimates for Australia. However it is likely that some data for Australian estimates were input at the state/territory level (given that some data sources listed on the GHDx were state and territory reports). The three data sources in the GHDx identified as relevant to the current Australian BoD Study each include a number of geographical variables that enable sub-national estimates to be produced. In order to produce comprehensive and reliable sub-national estimates of disease burden in Australia, more detailed data at various levels of geography (e.g. state/territory, remoteness, sub-state and local area of usual residence) would be needed.

Which data sources used in GBD are appropriate for Indigenous estimates?

As mentioned earlier in this chapter, GBD 2010 did not include data specifically on the Indigenous population. Despite three data sources listed on the GHDx including an Indigenous identifier, this information was not included in IHME's database to generate BoD estimates for Australia. Therefore Indigenous BoD estimates will not be able to use data from the GBD but will need to be obtained from relevant data sources independently. Information on relevant data sources for Indigenous BoD estimates is included later in this chapter.

Is the use of multiple data sources the best approach for national and/or Indigenous estimates?

IHME's approach to using multiple data sources is likely to be designed, at least partly, to fill data gaps and compensate for deficiencies in country-level data. The argument for applying this approach to estimate YLLs for Australia is much less persuasive due to the availability of high quality vital registration data. However, this approach may be appropriate or even necessary for some YLD estimates, particularly for Indigenous BoD estimates, as data deficiencies may be overcome using multiple data sources.

An assessment of the quality of each data source proposed for the Australian study would need to be undertaken. An objective set of criteria against which the quality of data sources can be measured in order to determine whether they should be included or excluded from analyses would be useful.

Likely data improvements to Australian estimates

In the absence of detailed information on the data sources used for Australian estimates in GBD 2010, likely improvements to Australian estimates from new or updated data sources can still be identified. These are discussed below. The combination of vital registrations data (mortality), disease registers, surveys, service use data and information from various epidemiological studies should provide data of reasonable quality for many of the conditions and risk factors.

Updated Census data

Final Estimated Resident Population (ERP) estimates from the 2011 Census were released by the ABS in June 2013. These will provide the denominator for some rates used as inputs to the calculations (such as when data from administrative data collections are used as the numerator) as well as the denominator for YLL, YLD and DALY rates.

Updated vital registrations data

GBD 2010 used death registrations data up until 2006 for Australian estimates which was supplied by WHO. The AIHW currently has access to final Cause of Death (CoD) up to 2009, a revised version for 2010 and a preliminary version of 2011 at the unit record level. See Chapter 4 for more detail on this.

Updated health survey data

GBD 2010 used data from previous National Health Surveys. The 2011-12 Australian Health Survey (AHS) will provide more recent prevalence estimates for some diseases and more recent risk factor exposure statistics for many of the listed risk factors included in GBD 2010. It will include health measurement data (including physical and biomedical measures) on key risk factors such as blood pressure, cholesterol, and height and weight.

Results from the 2011-12 AHS on the self-reported prevalence of key conditions and risk factors, physical measurements, physical activity, nutrition and biomedical measurements have already been released. ABS TableBuilder for the National Health Survey component of the 2011-12 AHS was made available in September 2013 which includes information on long-term health conditions, general diet, physical measurements, smoking and alcohol use. The physical activity TableBuilder dataset was released in December 2013. Further updates to these products will include more information, and CURFs are expected to become available around the middle of 2014.

Updated, new or additional data sources for conditions and risk factors that cannot be obtained from the AHS

While the 2011-13 AHS will be a very important data source of disease prevalence and risk factor exposure data, other data sources will also be needed. There are some conditions and risk factors that will not be able to be obtained from the AHS, either because they are not collected (e.g. occupational and environmental risks), or because the AHS (or surveys in general) is likely to be inadequate to produce accurate prevalence rates (either for particular age groups or for the population as a whole). Some examples of such risk factors and conditions are listed in tables 7.3 and 7.4 below with possible alternative data sources for these. This is not a comprehensive list at this stage. More detailed examination of data sources would need to be undertaken as part of the project work.

Table 7.3: Examples of risk factors for which the AHS may not provide adequate estimates, and potential alternative data sources

Risk factor	Potential data source/s
Environmental risks (e.g. unimproved water and sanitation, air pollution, other environmental risks)	State and territory Environmental Protection Agencies National Hazard Exposure Worker Surveillance Survey; Personal Safety Survey.
Occupational risk factors	National Hazard Exposure Worker Surveillance Survey 2008 National Pollutant Inventory
Child and maternal under-nutrition	Longitudinal Survey of Australian Children
Child sexual abuse	AIHW National Child Protection Data Collection
Intimate partner violence	Specialist Homelessness Services Data collection Australian Longitudinal Study on Women's Health 2011
Tobacco, alcohol and illicit drug use	National Drug Strategy Household Survey Alcohol & Other Drug Treatment Services NMDS
Breastfeeding	Australian National Infant Feeding Survey 2010
Smoking during pregnancy	National Perinatal Data Collection

Table 7.4: Examples of conditions for which the AHS may not provide adequate estimates, and potential alternative data sources

Condition	Potential data source/s
Mental health conditions	National Survey of Mental Health and Wellbeing National Residential Mental Health Data Collection National Community Mental Health Data Collection Residential Aged Care Data Collection
Cancer	Australian Cancer Database
STIs, other communicable diseases	National Notifiable Disease Surveillance System
HIV/AIDS	National Centre in HIV Epidemiology and Clinical Research
Diabetes	National Diabetes Register, National Diabetes Services Scheme Medicare/PBS
Diabetes mellitus sequelae, kidney disease	ANZDATA, linked ANZDATA and mortality data
Rheumatic heart disease, Acute Rheumatic Fever	Rheumatic heart disease registers
Congenital abnormalities	Australian Congenital Anomalies Monitoring System
Dental caries, periodontal disease	Child Dental Health Survey National Adult Oral Health Survey 2004-06
Stroke, myocardial infarction	National Hospital Morbidity Database, and/or linked data
Appendicitis, pancreatitis	National Hospital Morbidity Database
Injuries	National Hospital Morbidity Database
Trachoma	Australian Trachoma Surveillance report 2010
Alzheimer's	Survey of Disability and Carers

Data sources for risk factors

In addition to the 2011-13 AHS, other large-scale surveys such as the Longitudinal Study of Australian Children (LSAC), the National Infant Feeding Survey and the Australian Longitudinal Study on Women's Health may also provide good information on several maternal and child health risk factors. GBD 2010 used LSAC data from Waves 1 to 4 of the survey (Wave 4 was released in August 2011). Wave 5 is expected to be released in late 2013 which will provide further data on the two cohorts of children included in the Study.

The National Drug Strategy Household Survey (NDSHS) may be useful to provide additional information on tobacco, alcohol and illicit drug use, particularly for those aged under 15 years. Results from the 2011 NDSHS are currently available; and results from the 2013 NDSHS are expected in mid-2014.

Data on environmental exposures may be obtained from sources such as the state and territory Environmental Protection Agencies, the National Hazard Exposure Worker Surveillance Survey, and the Personal Safety Survey.

Data sources for causes

The National Hospital Morbidity Database and National Notifiable Disease Surveillance System (NNDSS) will be key data sources for some disease prevalence estimates. 2011-12 hospital data and 2012 NNDSS data are currently available for analyses.

The Australian Cancer Database (ACD) will be a key data source for cancer incidence/prevalence and survival statistics. 2010 ACD data are currently available for analysis. 2011 data are expected to be available in 2014.

The National Diabetes Register (NDR) was used in the 2003 Australian Burden of Disease study and will be a key data source for the incidence/prevalence of diabetes. 2011 NDR data are expected to be available for use in the first half of 2014.

It is unclear what data sources were used in GBD 2010 for Australian estimates for child sexual abuse and intimate partner violence. The AIHW National Child Protection Data Collection may provide more recent data on child sexual abuse, and the AIHW Specialist Homeless Services Collection may provide more recent data on intimate partner violence. As these data sets are held by the AIHW, unit record data could be used for analysis, after receiving necessary permissions.

Medicare data may provide useful data on prevalence of some chronic diseases, through MBS items related to chronic disease management (e.g. diabetes). However, data limitations means further validation is required. The limitations include: unavailability of specific diagnostic information (except in where it can be inferred from the item) and incomplete coverage of the population with the disease. Validation work may require data linkage to derive robust prevalence measures.

Administrative data sources held by the AIHW can be used with other data to provide useful information, including severity, for some causes. For example, the National Community Mental Health Care Database, the National Residential Mental Health Care Database and the Residential Aged Care data collection may provide additional information on the prevalence of some mental health conditions to supplement survey data.

The 2012 Survey of Disability and Carers (SDAC) covers both non-institutionalised and institutional populations, thereby providing prevalence estimates for some conditions largely affecting the elderly (e.g. Alzheimer's).

Whilst the AIHW holds many of the data collections mentioned above, it must be noted that in some cases, data suppliers (e.g. state/territory health departments) will need to give permission for the AIHW to release these data for the purposes of burden of disease work.

There may be cases where no Australian data are available for a particular condition to produce a prevalence estimate. In such cases, an option is to use the prevalence estimate for Australia used in GBD 2010, where appropriate.

Updated and more comprehensive epidemiological studies

A review of recent epidemiological studies specific to Australia that have been published since GBD 2010 for relevant causes is needed. As it will not be feasible within the timeframes of the project to undertake a wide and comprehensive review of the literature for all GBD causes, advice could be sought from the Expert Advisory Group and other relevant Reference Groups to identify relevant studies that should be examined. A protocol outlining criteria to assist with determining what studies should (and should not) be used in the national and Indigenous analyses (e.g. based on their sample size, scope, peer-reviewed, published would be useful.

Use of linked data

It appears that data integration (data linkage) methodologies were not used in GBD 2010, however there is potential to improve national and Indigenous disease estimates through data integration.

Linking data from different sources can enhance information available on the prevalence and comorbidities of specific diseases and provide more robust data. Data integration can provide information about health outcomes, to improve coverage within a data collection and to estimate Indigenous under-identification in particular data sets.

A small amount of linked data was used in the 2003 Australian BoD study. This included data from the Western Australian Data Linkage System and data from the Victorian Admitted Linked Episodes Database to estimate incidence and prevalence for specific conditions.

Australia's capability in this area has improved significantly since then. The AIHW has recently been accredited as a Commonwealth Integrating Authority, meaning that the AIHW has met stringent criteria covering project governance, capability, data management, and the protection of privacy and confidentiality. As an accredited integrating authority the AIHW is authorised to undertake data integration projects involving Commonwealth data, including those assessed as high risk projects. This may enable the integration of multiple data sources to produce enhanced estimates for the Australian Burden of Disease Study, where timelines permit this.

Data sources that may be useful to link for the BoD studies include hospitalisations, mortality, perinatal, cancer, MBS or PBS and DVA data. The AIHW has linked data from the National Death Index (mortality) with the cancer and diabetes registries, and has linked Medicare, hospital, mortality and perinatal data with other data collections for specific projects. There is a current project underway to assess whether prevalence measures for cardiovascular disease, diabetes and chronic kidney disease can be derived using linked hospital and deaths data for NSW and WA.

Some states and territories may also hold linked data files that could be utilised for the Australian BoD Study. For example, the Department of Health, Western Australia has a linked hospital and mortality data file which could potentially be requested for use in the Australian study to improve prevalence estimates for specific conditions such as stroke and chronic obstructive pulmonary disease (COPD).

The feasibility of using linked data for the purposes of the Australian BoD Study could be assessed. Due to the tight timeframes for this project and the length of time it takes to obtain ethics approval and data custodian approval for data linkage, it is not likely to be feasible to undertake new data linkage projects for use in the current Australian study, however already linked data held by the AIHW or the states and territories could be considered for inclusion, after obtaining ethics approval for such work.

Use of existing measures of uncertainty for some data sources

GBD 2010 generated its own measures of uncertainty around the data based on statistical models. For the Australian BoD Study, availability of measures of uncertainty (i.e. relative standard errors, confidence intervals) in the proposed data sources could be sought to contribute to uncertainty estimation. Most ABS surveys produce relative standard errors for estimates. However, the production of uncertainty intervals for all estimates is complex, and would need to be assessed for validity and feasibility within the timeframe allowed for the project. See Chapter 4 (Mortality) and Chapter 5 (Morbidity) for further information on uncertainty intervals and outlined approach to their use and calculation.

Sub-national estimates

Many of the proposed national data sources to produce updated BoD estimates for Australia include variables that enable disaggregation for analysis by geographic region and socioeconomic status. For each proposed data source for the Australian study, the availability of geographic variables to inform how the data support risk factor and disease prevalence estimates would need to be determined.

Population estimates using the Australian Statistical Geography Standard (ASGS) from the 2011 Census with backcasting at various levels of geography were released in June 2013 (state/territory) and August 2013 (Remoteness, postcode, SA1-4). These estimates are needed for the calculation of rates at the sub-national level. The specific details of these releases are documented in Appendix 7C of this chapter.

Indigenous estimates

The capacity to produce estimates of burden of disease for the Aboriginal and Torres Strait Islander population in the Australian study will be an important improvement on the Global Burden of Disease Study.

As noted above for national estimates, the extent of data sources available in Australia should provide reasonable estimates for the Indigenous population for at least some causes and risk factors.

Many of the proposed data sources include an Indigenous identifier and, of these, some are considered to have reasonable Indigenous data quality (noting that this varies by state and territory) and can therefore be used to produce estimates for the Indigenous population. For example, the AIHW National Mortality Database would be used to update the fatal burden estimates (YLL) by cause of death. The National Hospital Morbidity Database, the National Notifiable Disease Surveillance System and the Australian Cancer Database will be key data sources for some disease prevalence estimates.

The 2012-13 Australian Aboriginal and Torres Strait Islander Health Survey (AATSIHS) will provide recent prevalence estimates and risk factor exposure statistics for most diseases and risk factors included in GBD 2010. It will include, for the first time, national health measurement data (biomedical measures) on key risk factors such as blood pressure, cholesterol and body mass. The ABS released the first results from the AHS relating to Aboriginal and Torres Strait Islander peoples in November 2013. Results on nutrition and physical activity and results from the biomedical component are expected to be released in late 2014.

It is expected that Indigenous results from the AHS will be released in ABS Table Builder as well as in CURF (confidentialised unit record) form. It is not clear whether Table Builder will provide enough detail for the analyses required, and recent advice from the ABS is that a CURF for the AATSIHS is not expected to be released until March 2015. Access to the Master Unit Record file (MURF) held at the ABS data laboratory could be pursued as an option for timely access to detailed Indigenous data from the AHS.

The Longitudinal Survey of Indigenous Children (LSIC) provides comparable data to the LSAC for Indigenous children so is a potential data source for a number of risk factors. The survey covers breastfeeding, childhood underweight, dietary factors and exposure to second-hand smoke. Results from waves 1 to 4 of the LSIC are currently available for which

the AIHW has access to the unit record file. Data from Wave 5 is expected to be released in 2014.

Indigenous population estimates from the 2011 Census are required for denominators for rates used as inputs (when data from administrative data collections are used as the numerator) and for YLL, YLD and DALY rates. The final 2011 Estimated Resident Population figures for Indigenous Australians were released by the ABS in August 2013. Backcast and forward population projections from the 2011 Census are expected to be released in April 2014.

Table 7.5 presents a preliminary list of data sources that are of relevance to Indigenous BoD estimates. It does not include data sources which include an Indigenous identifier but the sample size is too small to generate robust estimates for the Indigenous population (e.g. SDAC, NDSHS). This list should be reviewed by the project team to determine which conditions and/or risk factors can be obtained from these data sources.

Table 7.5: Preliminary list of data sources of relevance to Indigenous BoD estimates

Source
Disease registers, surveillance, notifications
National Mortality Database
National Notifiable Diseases Surveillance System
HIV/AIDS National Registry, National Centre in HIV Epidemiology and Clinical Research
Australian Cancer Database
Rheumatic Heart disease Registries
Australian and New Zealand Register of Dialysis and Transplant Patients
Australian Trachoma Surveillance report 2010
Breastscreen Australia
National Bowel Cancer Screening Program data
National Perinatal Data Collection
National Child Protection Data Collection
Specialist Homelessness Services Collection
Deadly Ears (6 years population based screening in discrete communities and ENT data)
Australian Institute of Criminology Deaths in custody data
Population health surveys
Australian Health Survey 2011-13 (Indigenous component)
National Aboriginal and Torres Strait Islander Social Survey (2008)
Census of Population and Housing
Longitudinal Study of Australia's Children
Child Dental Health Survey
National Adult Oral Health Survey 2004-06
Indigenous Eye Health Survey 2008
Study of Aboriginal and Torres Strait Islander Child Oral Health in Remote Communities 2000-03
Western Australian Aboriginal Child Health Survey 2001 and 2002

(Continued)

Table 7.5 (Cont): Preliminary list of data sources of relevance to Indigenous BoD estimates

Source
Health Service utilisation
National Hospital Morbidity Database
Medicare database
BEACH survey
Alcohol and Other Drug-treatment Services NMDS
National Opioid Pharmacotherapy Statistics Annual Data
National Residential Mental Health Data Collection
National Community Mental Health Data Collection
Healthy for Life/National Key performance indicators
NT Child Health Check Initiative data collections (Child health checks, follow up dental, audiology and ENT services)

Issues to consider in identifying data sources for the Indigenous population

Despite the range of data sources relevant to Indigenous BoD estimates, there are a number of issues to consider in identifying and deciding on data sources to be used for Indigenous estimates. These are discussed below.

National data source does not include an Indigenous identifier

Not all data sources used for the Australian BoD study include an Indigenous identifier. Data sources that can provide risk factor exposure estimates (e.g. the National Infant Feeding Survey, the Australian Longitudinal Study on Women's Health; and the National Hazard Exposure Worker Surveillance Survey) do not include an Indigenous identifier. In such cases, alternative data sources could be sought.

No data source identified for Indigenous population

In cases where data for the Indigenous population cannot be obtained from either a national data source, or an Indigenous-specific data source, a literature review could be undertaken to identify whether this information can be obtained from recent epidemiological studies.

Where no data are available for the Indigenous population for certain conditions, modelling strategies could be considered to derive prevalence rates. Judgments about whether it is necessary or valid to carry over certain parameters from the national estimates to the Indigenous estimates will also need to be formed. For example, where prevalence data are not available for the Indigenous population for a particular condition, an option is to apply the Indigenous/non-Indigenous mortality rate ratios for the relevant condition to the prevalence rates for the total population (this approach was undertaken in the 2003 Indigenous BoD study).

Small number issues (e.g. Indigenous sample size too small to produce robust estimates)

In the case of surveys, not all data sources have a large enough Indigenous sample to provide robust estimates. Even if a dataset has an Indigenous status data item and Indigenous identification is reasonably complete, the numbers of Indigenous people included may be too small to support reliable estimation of the variables needed for BoD estimation. For example, the National Drug Strategy Household Survey and the SDAC have Indigenous sample sizes that are too small to produce robust estimates for the Indigenous population. However, the Indigenous component of the 2011-13 AHS has been specifically

designed to be able to provide reasonable Indigenous estimates and thus should be able to be used for most of the risk factors and diseases included in these surveys.

In the case of administrative data collections, particularly death registrations, the number of Indigenous cases reported for some conditions will be small, particularly when disaggregated by age and sex.

For mortality estimates and YLD estimates for causes affected by small numbers, it may be appropriate to combine multiple years of data. If small number issues still remain, then the modelling strategies may be considered to generate more robust estimates for the Indigenous population. For example, if numbers are large enough for all ages combined but not split by age-group then an option is to apply the total Indigenous/non-Indigenous rate ratio for a condition across all age groups. If numbers are large enough for a condition group but too small for a specific condition then an option is to apply the Indigenous/non-Indigenous rate ratio for the condition group to the lower-level condition. Both of these approaches were used in the New Zealand BoD study for the Maori population when numbers were small.

Need for Indigenous-specific data sources

Data sources specific to Indigenous Australians may provide valuable information for some diseases and risk factors, particularly those for which Indigenous data are not available from national data sources.

For example, the 2008 Indigenous Eye Health Survey may provide useful data on prevalence of specific eye conditions such as trachoma. The Deadly Ears Program, the Study of Aboriginal and Torres Strait Islander Child Oral Health in Remote Communities, and the Northern Territory Child Health Check Initiative data collections (dental, audiology and ENT services) may provide useful data on the prevalence of certain oral health and ear conditions among the Indigenous population.

Need for data sources on diseases/risk factors specific to the Indigenous population

There are some diseases and conditions that are more prevalent in Indigenous communities than in other communities which may require the identification of relevant data sources. Such conditions include glaucoma, trachoma, scabies, acute rheumatic fever and otitis media. The 2012-13 AATSIHS will provide useful data on some of these conditions, while other data sources may be needed to produce estimates for other conditions. For example, the 2010 Australian Trachoma report may provide data on the prevalence of trachoma in Indigenous communities. The Rheumatic Heart Disease registers in NT, WA and Qld may provide useful information on the incidence of ARF among the Indigenous population living in the areas covered by the registers.

There are also some risk factors which are of particular concern among the Indigenous population, such as petrol sniffing. This was not included in GBD 2010. Whether adequate exposure data and relative risk estimates can be obtained for such risk factors would need to be investigated.

For further information on conditions and risk factors relevant to the Indigenous population see Chapter 3.

Need for epidemiological studies on the pattern of Indigenous health and the evolution of disease among Indigenous people

The epidemiological relationships between some diseases and risk factors (e.g. co-morbidities, risk-outcome pairs, risk effect sizes) may differ for the Indigenous population. It

would be useful for the project team to canvass relevant epidemiological literature. This may inform the evidence base for relative risks and differences between the Indigenous and non-Indigenous populations.

Need for data sources on Indigenous under-identification

Surveys such as the AATSIHS are not affected by Indigenous under-identification as the survey is weighted to the total Indigenous estimated resident population.

In the case of administrative data, not all data sources are considered to have adequate levels of Indigenous identification in all states and territories. For example, mortality data for the Indigenous population are considered (by the ABS) adequate for reporting in NSW, Qld, WA, SA and NT only. Adjustments for under-identification can be made in some cases where information on the level of Indigenous identification is available from the relevant data source. In theory, we would like to know the pattern of identification across all relevant levels of disaggregation (e.g. cause, age, sex) for each contributing data set. But in practice, it is likely that estimates of under-identification will only be available for a small number of data sets and only at higher levels of aggregation (such as state or remoteness region).

The 2003 Indigenous BoD adjusted hospital data using the results of a study undertaken by the AIHW in 2005, and adjusted mortality data using an indirect method using Census population data.

Since then, the AIHW and ABS has made advances in the assessment of under-identification in various data sets, namely hospitalisation (AIHW) and mortality data (AIHW and ABS), for which recent studies have produced robust adjustment factors that can be applied to the data. These include:

- **ABS Census Data Enhancement Mortality Quality Study** - death records are probabilistically linked with Census records to derive factors for adjusting registered Indigenous deaths. This study was first completed using data from the 2006 Census and was repeated with data from the 2011 Census for which results were released in late 2013. From this, two sets of adjustment factors were published: the first at the national level by 3 broad age groups, and the second for selected states and territories. These adjustment factors could be used to adjust Indigenous mortality rates and YLL estimates for the current Indigenous BoD study.
- **AIHW Enhanced Mortality database** - registered deaths in the AIHW National Mortality Database (NMD) are linked with Indigenous death records from alternative data sources (residential aged care data, hospital data and neonatal death data from the National Perinatal Data Collection) to produce adjustment factors and life expectancy estimates. This study has been completed for the period 2001-2006 and is currently being repeated for the period 2007-2011. From this, adjustment factors may be available at various levels of disaggregation (age, sex, state/territory), which could be used to adjust Indigenous mortality rates and YLL estimates for the Indigenous component of the study.
- **AIHW Indigenous under-identification in hospitals data quality studies** - results of face-to-face interviews with patients in public hospitals in each state and territory are compared to the information recorded in the patients' administrative records to produce correction factors for Indigenous under-identification in hospital data. Such a study was first completed in 2010, and the most recent was published in May 2013 which used data for the period 2010-11. This latter study provided correction factors at the national, state/territory, remoteness and remoteness within jurisdiction levels. This data source

could be used to adjust Indigenous hospitalisation rates and YLD estimates for the Indigenous estimates.

In addition to the studies described above, there are a number of sub-national studies which may be useful for adjusting for Indigenous identification in particular data collections (AIHW 2013, forthcoming).

The ABS in collaboration with the Department of Health WA and the Telethon Institute for Child Health Research, has also recently completed an Indigenous identification data linkage project 'Getting our story right' which has linked a number of health data collections with education data collections to improve Indigenous status information across these collections. This has resulted in the calculation of an Indigenous status flag at the individual record level based on the Indigenous status information across multiple data sets. This could potentially be used to inform adjustments for under-identification in relevant data collections for the Indigenous analyses.

The best available information about patterns of Indigenous identification and about preferred methods of adjustment for under-identification based on recent studies undertaken, including information obtained from data linkage work, should be used.

Need for data on non-Indigenous Australians

The AIHW is still considering how best to produce estimates of the 'health gap' for the updated Indigenous study. Ideally data for non-Indigenous Australians is required to generate estimates for all diseases so that these can be compared to estimates for the Indigenous population.

This is likely to be feasible for YLL estimates which rely on a single data source which includes an Indigenous identifier; however there may be some conditions where a comparable disease prevalence estimate for the non-Indigenous population is unable to be generated from the data available.

Intended approach

The following outlines our intended approach to data sources for the Australian Burden of Disease Study, subject to further consultation with the Expert Advisory Group and the Indigenous Reference Group:

- Use the AIHW National Mortality Database as the data source for mortality estimates and YLL calculations.
- Assess on a case-by-case basis whether to use multiple data sources or the single best data source for each condition for YLD calculations and risk factor estimates, according to a protocol to be developed with the assistance of the EAG. This would include reviewing recent epidemiological studies specific to Australia that have been published since GBD 2010 for relevant causes, and considering their inclusion or exclusion based on the protocol developed.
- Use Australian data sources for compiling morbidity and risk factor exposure data. If data for a condition are not available for Australian estimates, then obtain the Australian prevalence rate from GBD 2010, where appropriate.
- Assess the feasibility of using linked data for the purposes of the national and Indigenous estimates.

Additional considerations specific to the Indigenous component of the Study

- Adjust for Indigenous under-identification in mortality data (using available adjustment factors, for example, from either the ABS Census Data Enhancement Data Linkage Study or the AIHW Enhanced Mortality Database linkage project). Adjust for Indigenous under-identification in hospital data using adjustment factors available from the 2011 AIHW hospital data quality study. Assess whether it is feasible to adjust for under-identification in other administrative data sources.
- Use national and Indigenous-specific data sources for compiling morbidity and risk factor exposure data. If data for a condition are not available for the Indigenous population, or numbers are small, then consider adopting modelling strategies to derive prevalence estimates for the Indigenous population.
- Review epidemiological studies on the relationships between some diseases and risk factors (risk-outcome pairs, risk effect sizes) to determine if these may differ for the Indigenous population.
- Consider how best to produce estimates of the 'health gap' between Indigenous and non-Indigenous Australians and whether this is feasible within the timeframes and resources allocated for the project.
- Where possible, collect and input data for both Indigenous and non-Indigenous Australians for 'health gap' measurement. However, it should be noted that non-Indigenous estimates are not planned to be reported or published as a separate output for the study.

Appendix 7A: Data sources used in Australian 2003 BoD study

Table 7A.1: Data sources for disease prevalence in 2003 AUS BOD

Data source	Custodian	Year	Diseases
Disease registers, surveillance, notifications			
National Notifiable Diseases Surveillance System	DoHA	2003	Communicable diseases, STIs
HIV/AIDS National Registry	NCHECR	2003	HIV/AIDS
National Cancer Statistics Clearinghouse	AIHW	2001	Malignant neoplasms
State and territory cancer registries	S/T cancer registries	1997	Breast cancer
Breastscreen Australia	AIHW	2001-02	Breast cancer
National Perinatal Data Collection	AIHW	2003	Low birthweight
Victorian Perinatal Data Collection		2001-02	Low birthweight
Queensland Perinatal Data Collection		2002	Low birthweight
National Diabetes Register	AIHW	2001	Diabetes mellitus
Australian and New Zealand Register of Dialysis and Transplant Patients	ANZDATA	2002	Diabetes mellitus sequelae
Victorian Cystic Fibrosis Screening Program		1989-1998	Cystic Fibrosis
ABS causes of death data set	ABS	2003	Motor neurone disease Anencephaly
WA Intellectual Disability Exploring answers database		1983-96	Intellectual disability
Congenital malformations Australia	AIHW	2001-02	Intellectual disability, congenital heart disease; digestive system malformation; other urogenital tract malformations; abdominal wall defect
Victorian Perinatal Data Collection, Births Defects Register		2001-02	Spina Bifida, digestive system malformation; renal agenesis
WA Births Defects Registry		2003	Other utogenous tract malformations
Health service utilisation data			
National Hospital Morbidity Database	AIHW	2002-03	Various
BEACH	AIHW	2000-01	Upper and lower respiratory tract infections; otitis media, peptic ulcer disease, skin ulcers
Alcohol and Other Drug Treatment Services NMDS	AIHW	2002-03	Stimulant dependence
WA Data Linkage system		1999-2003	Heart failure, ischaemic heart disease, stroke
Vic Linked Admitted episodes data		1996-2002	Heart failure, vascular insufficiency of intestine
Population health surveys			
National Gastro-enteritis survey		2001-02	Diarrhoea
1980 National Trachoma and Eye Health Program		1976-78	Trachoma sequelae; otitis media
National Health Survey	ABS	2001	Various
Australian Diabetes, Obesity and Lifestyle Study (AusDiab)		19990-2000	Deficiency anaemia Diabetes mellitus

(continued)

Table 7A.1 (continued): Data sources for disease prevalence in 2003 AUS BOD

Data source	Custodian	Year	Diseases
Risk Factor Prevalence Study		1989	Deficiency anaemia
2002 National non-melanoma skin cancer survey		2002	Non-melanoma skin cancer
National Mental Health and Wellbeing Study	ABS	1997	Alcohol dependence; benzodiazephrine dependence, cannabis dependence; psychotic disorders; anxiety and depression; bipolar disorder; personality disorders; ADHD
Australian Child to Adult Development Study		1990-96	Intellectual disability
Australian Longitudinal Survey on Women's Health	FaHCSIA	1996-2002	Urinary incontinence Menstrual problems
Survey of Disability, Ageing and Carers	ABS	1998	
Child Dental Health Survey	AIHW DSRU	2000	Dental caries
National Oral Health Survey	AIHW DSRU	1987-88	Dental caries; periodontal disease
South Australian Dental Longitudinal Study		1991-96	Dental caries
Adelaide Dental Study of Nursing Homes		1998	Dental caries
		1999	Edentulism
Longitudinal Survey of Dentist's Practice Activity		2002	Edentulism
National Dental Telephone Interview Survey		2002	Pulpal infection

Table 7A.2: Data sources for 14 selected risk factors in 2003 AUS BOD

Data source	Custodian	Year reported in 2003 study	Risk factors in 2003 study
National Health Survey	ABS	2001	Low fruit and vegetable consumption; Physical inactivity; Smoking; Alcohol
AusDiab study		2002	High blood pressure; High cholesterol; High Body mass index; Unsafe sex; Illicit drug use
Dubbo and Geelong Osteoporosis Study	Deakin University	2005	Osteoporosis
National Drug Strategy Household Survey	AIHW	2004	Illicit drug use
Women's Safety Survey		1996	Intimate partner violence
National Worker's Compensation Statistics Database		2003	Occupational exposures and hazards
National Coroners Information System		2003	Occupational exposures and hazards

Appendix 7B: Data sources used in 2003 Indigenous BoD study

Table 7B.1: Data sources used for disease incidence/prevalence in 2003 Indigenous Australian BoD study

Primary data source	Custodian	Year	Disease and injury categories
Disease registers, surveillance, notifications			
National Notifiable Diseases Surveillance System	DoHA	2001-03	Communicable diseases, STIs
National HIV Database and National AIDS Registry	NCHECR	2003	HIV/AIDS
Australia and New Zealand Dialysis and Transplant Registry (ANZDATA)	ANZDATA	2001-03	Nephritis & nephrosis
Australian mortality data	AIHW	2003	Malignant neoplasms Nephritis & nephrosis
National Perinatal Data Collection	AIHW	2003	Low birth weight
Health service utilisation data			
Australian hospital data	AIHW		Various
Bettering the Evaluation And Care of Health		2000-01	Otitis media
Western Australian Data Linkage System		1990-03	Stroke
Population health surveys			
National Trachoma Eye Health Program		1976-78	Trachoma Otitis media
National Health Survey	ABS	2004-05	Asthma; Alcohol dependence; Otitis media; Anxiety and depression
Western Australian Aboriginal Child Health Survey		2001-02	Anxiety and depression
NSW Health Survey		1997-98	
2004-06 National Oral Health Survey	AIHW DSRU	2004-06	Periodontal disease; Edentulism; Dental caries
South Australian Child and Adult Dental Health Surveys		1999-01	Dental caries

Table 7B.2: Data sources for 14 selected risk factors in 2003 Indigenous Australian BoD study

Data source	Custodian	Risk factor
DRUID (Cunningham et al. 2006); Wang & Hoy (Wang & Hoy 2003)		High blood pressure High cholesterol
2004-05 NATSIHS	ABS	High body Mass Insufficient fruit and vegetable consumption Physical inactivity Smoking Alcohol Illicit drug use
National Health Survey 1995	ABS	Passive smoking
Australian Mothers and Babies 2003	AIHW	Maternal smoking during pregnancy
1994 National Drug Strategy Household Survey	AIHW	Cannabis Use
National Child Protection data 2002-03	AIHW	Child Sexual abuse
Women's Safety Survey 1996		Intimate Partner Violence
International Violence Against Women Survey 2004		Intimate Partner Violence
Supported Accommodation Assistance Program 2005	AIHW	Intimate Partner Violence

Appendix 7C: Population estimates

Table 7C.1: Timing of release of population data by geography

Data available	Populations
20 June 2013	Final ERP 1991 to 2011 (national and state/territory)
30 August 2013	Final sub-state ERP 1991–2011 (SA1-SA4, RA, POA, SEIFA) Final ERP 2011 – Aboriginal and Torres Strait Islander Peoples
March 2014	2012 Revised ERP totals (SA1-SA4, RA, POA) 2013 Preliminary ERP totals (SA1-SA4, RA, POA)
April 2014	Backcast and projected estimates for the Aboriginal and Torres Strait Islander population 2001-2016 (age, sex, state/territory)
August 2014	2012 Revised ERP age/sex (SA1-SA4, RA, POA) 2013 Preliminary ERP age/sex (SA1-SA4, RA, POA)

SA1-SA4: Statistical Areas 1 to 4

RA: Remoteness

POA: Postcode

8 Logistics and implementation

The GBD 2010 study developed new methods to deal with specific issues relevant to the global study. These methods have been discussed in detail in the preceding chapters. The developments have been accompanied by the need to develop tools and infrastructure to meet the challenges imposed by the project's scope and objectives.

This chapter looks at aspects of logistics and infrastructure from GBD 2010 as they may apply to the current Australian project.

Overview of GBD 2010 logistics

GBD 2010 builds on the previous work that started in 1990, the updates/revisions to GBD between 1999 and 2004, and other national studies. The major extensions arising from GBD 2010 include country level estimates, comparable estimates across three points in time, and the inclusion of measures of uncertainty. GBD 2010 provides key information about health trends both globally and regionally and can be used to monitor health trends and inform policy.

What GBD 2010 achieved

GBD 2010 developed methods to estimate the burden of disease in 187 countries. Results were reported by sex and 20 age groups. These age groups included 3 distinct groups for those less than 1 year of age, 1–4 years and then 5-year age groups up to 80+. Estimates were provided for 3 points in time: 1990, 2005 and 2010.

To do this the GBD 2010 team created a database with 5 million mortality data points – the number of deaths for 187 countries x 235 causes x 40 age/sex groups x 3 years – and 6.5 million morbidity data points – prevalence estimates for 187 countries x 291 causes x 40 age/sex groups x 3 years.

The newly developed methods and extensive modelling has allowed data gaps to be filled and uncertainty around the burden of disease estimates to be measured for the first time. The use of three time points has enabled GBD 2010 to identify changes in health globally since 1990.

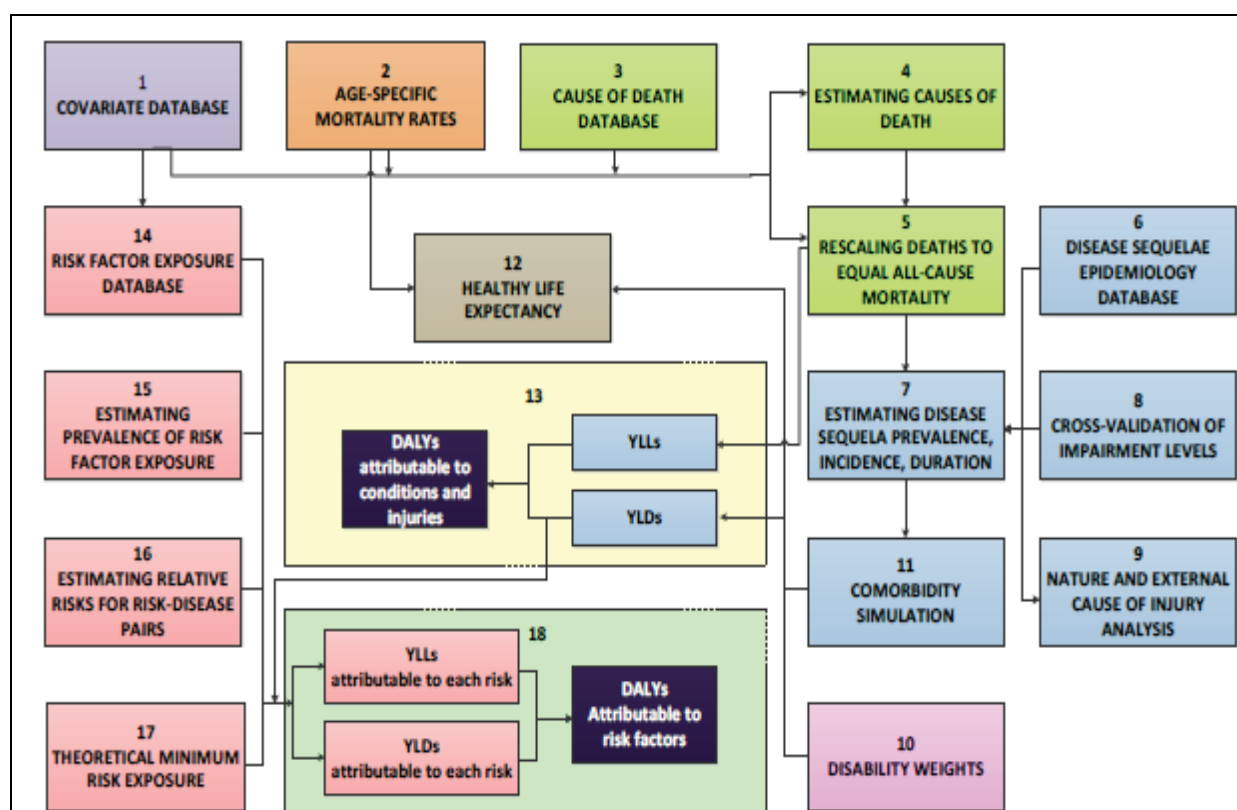
Extensive new visualisation tools were developed for GBD 2010, and are available on their website. They provide users with the opportunity to delve into the detail of the results, depending on their own requirements.

How they did it

GBD 2010 was a collaboration of 500 researchers from 300 organisations in 50 countries over a 5-year period, coordinated through the Institute for Health Metrics and Evaluation (IHME) at the University of Washington. The researchers were part of expert groups working on a particular component of the project.

As a result of the collaboration, the estimation methods were completely revised and improved, most notably estimations for the prevalence of an increased number of diseases and their sequelae are inferred from all available data.

Figure 8.1 shows the 18 components of the GBD 2010 study and their inter-relationships.



Source: Murray CJL, Ezzati M, Flaxman AD, et al. Supplement to GBD 2010: Design, definitions, and metrics. Lancet 2012; 380: 2063-66

Figure 8.1: Eighteen components of the GBD 2010 and the inter-relationships

The GBD2010 study aimed to include all available and relevant data sources, rather than relying on the best single data source for each component of the analysis. This was facilitated by sophisticated modelling procedures that were developed to fill data gaps. These models use epidemiological data from 'regions' and 'super regions' combined with an extensive covariate database to produce estimates of prevalence for countries where data are sparse.

The first step for modelling was to construct an extensive database of covariates that could be used in various models. The covariate database currently includes 343 covariates across 84 topic areas for the period 1980 to 2010, with some back to 1950. The covariates are used to inform the estimation process in many models in GBD 2010, and their role is largest for the models with the least amount of data available. Examples of common covariate topic areas used include education, measures of income, population density, rainfall, health system access, antenatal care, alcohol consumption and diabetes prevalence. A full list of the covariates is available in Web Table 4 of the supplement to: Murray CJL, Ezzati M, Flaxman AD, et al. GBD 2010: design, definitions, and metrics. Lancet 2012; 380: 2063-66.

Mortality (Years Life Lost)

Mortality information was gleaned from vital registration, birth histories, sibling histories and censuses and surveys to determine all-cause mortality. This created the envelope into which cause-specific mortality needed to fit. In addition, cause of death information came from a number of other sources including verbal autopsy, mortuary records, hospitals, and police reports. Due to the varied quality of the data, an ensemble statistical modelling

approach (operationalised in the analytical tool CODEm) was used to determine the number of deaths from each cause. A correction technique called CoDCorrect was then applied to fit the cause specific mortality into the all-cause mortality framework.

Morbidity (Years Lived with Disability)

The change to using all available data, instead of the best available, for all countries and all 1,160 disease sequelae was implemented through a new meta-regression analysis tool called DISMOD-MR.

DISMOD-MR (as used for GBD 2010) is a Bayesian meta-regression tool using a negative binomial statistical model. It used compartmental models of disease that describe how a population moves from healthy to diseased to death (or back to health) to create consistent estimates from many data sources. The process incorporates assumptions (expert 'priors' and study level covariates) and borrows strength from regional disease patterns and country level covariates. DISMOD-MR is being developed further for the next iteration of GBD.

Logistically, the vast amount of data (3 terabytes) put through DISMOD-MR required a large amount of computing power and a bank of 100 computers were established to provide the processing capability.

Uncertainty intervals

Each disease for each country was modelled using a combination of epidemiological and covariate data. The level of uncertainty around the estimate was quantified using simulation methods – repeated running of the models. GBD 2010 used 1,000 iterations of each model to determine the distribution of estimates from the model and thus the uncertainty of each estimate. Even with the large amount of computing power at their disposal, DISMOD-MR still took many hours to run.

Future GBD

Underlying the GBD approach is the concept of constant improvement. To that end IHME plan to update GBD methods as required and recast the whole time series with each update of the estimates. This approach reflects the need to use the best available methods to monitor and evaluate changes in patterns of health globally.

Differences from the 2003 Australian and Indigenous studies

As outlined in Chapter 1, in 2007 the AIHW published a report on the burden of disease and injury in Australia for the 2003 reference year. This report updated and expanded on the first Australian burden of disease study which was based on 1996 data and was released in 1999. The 2003 study included state, remoteness and socioeconomic group estimates based on small-area analysis, and disease projections.

The 2003 study adopted some methods from the 1996 study and some from earlier GBD studies. Some examples are:

- the cause list was largely derived from the 1996 Australian study with some additions and alterations due to the switch from ICD9 to ICD10

- the reference life table was the standard life table developed through the GBD process, not the Australian cohort life table as was used in the 1996 Australian study
- the best available data for each cause was used
- the disease models were built in Excel, and DISMOD-II was used to develop the full epidemiological picture for each cause. Note that DISMOD-II uses similar disease models as DISMOD-MR which describe how a population moves from healthy to diseased to death (or back to health) to create consistent estimates from many epidemiological measures.

The report of the 2003 study listed 6 key authors and 60 expert advisors, and was a collaboration between the AIHW and University of Queensland. This project was led by some of the key members of the GBD 2010 coordination team who are now at IHME.

In parallel to the national study, the first study of burden of disease for Indigenous Australians was undertaken based on 2003 data (with a report published in 2007). The Indigenous study produced burden of disease estimates for more than 170 diseases in the Indigenous population using national and Indigenous-specific data sources. The methods used were comparable with the national study so that valid comparisons were available between Indigenous Australians and all Australians.

Applicability of GBD methods to the Australian context

Much of what has been developed or built for GBD 2010 is to address the needs of the project in terms of scale and data availability. That is, the methods need to be able to provide estimates for many countries, causes and risk factors, even when there are large gaps in the data. As Australia does not have the same issues of massive scale and limited data availability and quality, the methodological developments implemented by IHME aimed at addressing these issues are not of themselves necessary for the current Australian project. In Australia, we do not have the same issues of data gaps, low data quality or data volume faced by the GBD team. Although Indigenous estimates and potentially some sub-national estimates may also be affected by data limitations, these are not likely to be on the same scale as for global estimates.

To undertake burden of disease analysis, all 18 components shown in Figure 8.1 need to be undertaken. However for the current Australian project, there are key aspects of the GBD 2010 methodology that can be estimated more simply than was necessary for GBD. As a result, if the approaches outlined in detail in the previous chapters are implemented, the Australian study would adopt a new approach—a combination of GBD methods and assumptions, and simpler country specific processes that include the ability to undertake Indigenous and sub-national analysis.

Whether or not DISMOD-MR is able to be used in the Australian study is also a consideration, given that IHME have advised they are continually updating this software tool to make refinements for GBD updates.

The GBD infrastructure has not produced population subgroup estimates to date. Therefore the Australian study will need to develop the required methods to produce Indigenous estimates, and make these methods integral to the Australian analysis which will also be of interest to other countries. A similar hybrid approach was taken by the New Zealand

Ministry of Health to produce Māori and non-Māori estimates for New Zealand. These are detailed in Appendix A.

The GBD study has released extensive visualisation products on their website, to provide users with access to detailed results. This approach, along with the form of other outputs from the Australian study, would need to be assessed in detail in consultation with key stakeholders and the Expert Advisory Group. The outputs will need to meet the needs of users (within available resources), including for policy-relevant information and more detailed analytical information in an appropriate form.

Design of the new Australian study

Relationship to GBD

A key role of the GBD study is to provide global estimates, and then break those down to region and country level to support international comparisons and benchmarking. Thus its priority is to fill gaps in the global data, and to maintain international comparability. The Australian study has a parallel role, though starting at the country level and then breaking the estimates of disease burden down within the country. Hence the priority for the Australian study is for the best quality country-level data, while maintaining as much international comparability as possible. While these roles and priorities can sometimes be at odds, it would be very beneficial if the Australian data used in future GBDs and Australian BoD studies were the same, in order to maximise comparability in the estimates.

It is desirable for Australian data to be added to the GBD database, including data that were not identified in the most recent GBD 2010 work and data that have recently become available, in order to ensure the best quality Australian estimates in future iterations of GBD. There is also a strong need to understand and validate the processing of Australian data in order to meet the need for transparency. To satisfy that need, it appears highly desirable that the Australian study undertake many of the analysis steps required to generate burden of disease estimates. Australian estimates generated through this process would then be compared with the Australian slice of the GBD estimates. If differences are found, then it may be desirable to undertake a systematic process to fully understand why they occur.

As noted above, there are challenges around the potential use of DISMOD-MR in the Australian study due to the software continuing to change. Further, it is not currently available for use in the Australian study. A viable alternative could be to follow the same approach as New Zealand – to use DISMOD II, the previous version (developed by WHO and freely available on their website). While this does not enable multiple datasets to be used, it may be possible to undertake meta-analysis prior to DISMOD II for the conditions where a multiple dataset approach will be useful.

It is expected that there would be a number of analytical projects the AIHW will do to further develop BoD methods for the Australian context (e.g. the mortality redistribution codes and analysis of SF-12 data for severity distributions). This work may contribute to the regional and global effort, and the results will be published for use beyond Australia if they are considered useful and relevant to other countries and globally.

Summary of overlap in methods between proposed approach and GBD

There are areas where it seems appropriate for the Australian study to accept the GBD 2010 approach without adjustment; these include the aspirational life table, no age weighting, not discounting for time and a prevalence approach to YLD calculation.

The areas where the Australian study may follow the GBD 2010 approach but in a less complex way:

- Cause of death analysis, where Australia has very good mortality data and there is limited need for modelling.
- Morbidity analysis, where less modelling will be required for some conditions where Australia has comprehensive data. For example, the estimation of burden for injuries will be less complex than in GBD as Australian has information on both the type of injury and cause of injury (external cause) from hospitalisation data.
- Use of multiple data sources, which the Australian study could adopt for some conditions; however, for other conditions, the single best data source is likely to result in the best estimate and therefore a meta-regression will not be required for these.

Proposed refinement to GBD methodology to suit the Australian context

There are 6 key areas where the AIHW is proposing to adapt the GBD 2010 methodology to better suit the Australian context:

1. Cause and risk factor lists

It is proposed that the GBD cause list be modified to better suit the Australian context. This may include some causes being aggregated (e.g. some communicable diseases not common in Australia), and some disaggregated (e.g. potentially splitting diabetes into Type 1 and Type 2). The changes to the cause list can occur either at the start of processing or just for reporting purposes.

There may also be a different set of risk factors included in the Australian study (some GBD risk factors are not relevant to Australia, e.g. indoor air pollution), though this will not result in differences in DALYs or disease rankings, but rather in the ranking of risk factors.

2. ICD mapping and 'garbage code' redistribution

The GBD 2010 approach is to redistribute 'garbage codes' using various methods that may change over time, and the same method appears to be used for all countries. For the Australian project, the term 'garbage codes' will not be used. Redistribution of such codes may be able to be further refined using local mortality and other data.

3. Estimating disease prevalence

GBD 2010 uses cross-walking (a data harmonisation technique) and DISMOD-MR to process the raw epidemiological data to model the prevalence of conditions. AIHW is proposing to use DISMOD-II with meta-analysis (when there are multiple data sources) to process the raw epidemiological data to provide a stable processing platform.

4. Disability weights and severity distribution

Further exploration of the disability weights may be required to increase the understanding of how they relate to Australian data and to support enhanced interpretation of the YLD results.

The GBD approach uses a number of methods to model severity distributions globally including from published literature and from surveys. AIHW propose to use Australian data sources, potentially hospitalisations, health surveys and disease registers, to develop models of severity for Australia with emphasis on potential differences by region and Indigeneity.

5. Comorbidity adjustment

Some analyses of Australian data would be useful to test the assumptions used in GBD 2010 in relation to their method for comorbidity adjustment which only accounts for independent comorbidity.

Management of differences between GBD and Australian study estimates

It is clear from the operational differences described above that there will likely be some differences in Australian estimates produced by GBD and by the Australian study.

The simpler modelling practices and other areas where the Australian study may diverge from GBD methods will be sources of difference between the estimates from the Australian study and the Australian estimates from the GBD process. Differences between GBD and Australian estimates may also appear where data differ, such as when one study has extra data sources or more up-to-date or detailed data.

9 Intended approach

The information in the previous chapters highlights a number of issues to be resolved and gaps in available information to undertake the Australian Burden of Disease Study, including national and Indigenous estimates, and sub-national estimates where valid. Each chapter also outlined suggested approaches to be further explored with the assistance of the Expert Advisory Group, Indigenous Reference Group and other relevant expert groups established for the project.

This chapter draws together all of the suggested approaches for the individual components of the burden of disease estimation process into an overall planned approach. Note that a separate decision tree has been developed for the national and the Indigenous components of the study which can be found in the Summary of this report.

The approach outlined in this chapter is in accordance with the broad principles and requirements underpinning the project:

1. Provide **national** estimates of fatal, non-fatal and total burden, as well as the attribution to specific risk factors, that are **up-to-date, of high-quality and meet Australia's needs**.
2. Provide **Indigenous** estimates, noting the extra challenges due to data limitations and differences in disease profiles. Because of the focus on the gap between Indigenous and non-Indigenous health, the aim will be to use comparable methods to those used for the national estimates as much as possible, though some differences are still likely due to differences in data sources, data quality and issues affecting Indigenous data.
3. Provide **sub-national** estimates (such as state/territory, regional, socioeconomic groups) where valid.
4. Maintain comparability with **GBD methods** as much as possible, with full clarity around any differences.
5. Provide **transparency** in the data sources, assumptions and methods used, with the ability to replicate the results.
6. Complete the work in an **efficient and flexible** manner, build national capacity, and set up the relevant infrastructure to enable efficient and timely **ongoing updates**.
7. Ensure **collaboration** with the various stakeholders including other burden of disease experts both nationally and internationally in order to contribute to global burden of disease work.

The details of the approach fall broadly into two categories: methodological and operational.

Methodological approach

Most of the discussion in the previous sections centred on methodological issues – that is, decisions around various methodological choices that will impact the final estimates. From these discussions, the following is intended for both the national and Indigenous components of the Australian BoD Study.

General

1. In order to maintain comparability as much as possible with GBD methods and ability for international benchmarking, use GBD 2010 methods where possible and/or

necessary. These can then be extended for national, Indigenous and sub-national analyses. Any innovations in methods implemented for GBD updates should be assessed for the Australian context as details become available, and implemented in the Australian study where appropriate.

2. Re-calculate the national estimates in the Australian study and compare them with the GBD estimates. This will provide clarity about differences between the estimates (see next point), and enable the production of sub-national and Indigenous estimates that are as consistent as possible with Australian estimates produced by GBD.
3. Conduct analyses using the GBD 2010 cause list as the starting point, adjusted where necessary for the Australian and Indigenous contexts.
4. Consider developing separate national and Indigenous reporting cause lists within the hierarchical structure of the analysis cause list, dependant on an assessment of need.
5. Report age groups for YLLs and YLDs based on what the data are able to support. For the national study, at a minimum, age-groups should reflect those used in GBD 2010 to enable comparison, and extended to older age-groups where possible. For the Indigenous study, age-groups should reflect what the data are able to produce robust estimates for (for example, the highest age-group is expected to be lower than for the national study due to small number issues in the older age groups).
6. No discounting for time or age-weighting (as per GBD 2010) for the standard analysis; however, system capacity may be built in to estimate YLL and YLDs using discounting and age-weighting to provide further functionality and usefulness.
7. Undertake work to determine the most appropriate methods for calculating uncertainty intervals for YLLs, YLDs and risk factor estimates based on measurable sources of error. For example, error derived from pooling multiple years of data or meta-analysis, relative standard errors produced from survey estimates, and confidence intervals produced around under-identification adjustment factors applied in Indigenous analysis could be incorporated into the calculation of uncertainty intervals.
8. Conduct sensitivity analyses to determine the relative importance of choices and assumptions where required. For example, this may include sensitivity analyses of the GBD 2010 disability weights, adjustments for Indigenous under-identification and redistribution of certain causes of deaths.

Additional considerations for the Indigenous analyses

9. Where possible, collect and input data for both Indigenous and non-Indigenous Australians where data are available for 'health gap' measurement. Non-Indigenous estimates will not be reported or published as a separate output for the study.

Mortality component

10. Adopt the standard reference life table defined in GBD 2010 for calculation of Australian and Indigenous YLLs.
11. Use Australian cause of death data held in the AIHW National Mortality Database to compile numbers of deaths for Australian, sub-national and Indigenous YLL estimates.
12. Investigate differences in death counts reported by year of registration versus year of occurrence, and between preliminary, revised and final versions of the mortality data over a number of years.

13. To avoid problems with small numbers and provide greater stability in YLL estimates, combine multiple years of data to derive counts for Indigenous deaths, deaths due to rare conditions, and for sub-national estimates. Simple modelling strategies may be used where numbers are still too small for stability.
14. Follow GBD methods for identifying and redistributing certain cause of death codes where appropriate. Analyses should be undertaken on Australian data sources to inform modifications where the GBD rules may not be suitable in the Australian or Indigenous contexts.

Additional considerations for the Indigenous study

15. Apply adjustment factors for Indigenous under-identification in mortality data at the lowest possible level. These could be derived from either the ABS's Indigenous Mortality Data Quality study or AIHW's Enhanced Mortality Database study.

Morbidity component

16. Adopt the GBD prevalence-based methodology for estimation of YLDs.
17. Assess on a case-by-case basis whether to use multiple data sources or the single best data source for each condition for YLD calculations, according to a protocol to be developed with the assistance of the EAG.
18. Use national and Indigenous-specific data sources for compiling morbidity data. If data for a condition are not available for Australian estimates, then consider using the Australian prevalence rate from GBD. If data for a condition are not available for the Indigenous population, or numbers are small, then consider adopting modelling strategies to derive prevalence estimates for the Indigenous population.
19. Use the disability weights developed for the GBD study, subject to targeted analysis of Australian data sources to support interpretation in the Australian and Indigenous contexts.
20. Adopt GBD method to adjust for comorbidity, subject to validation against Australian data sources to assess whether this method is suitable for the Australian and Indigenous contexts.

Additional considerations for the Indigenous study

21. Adjust for Indigenous under-identification in hospital data using adjustment factors available from the 2011 AIHW hospital data quality study. Assess whether it is feasible to adjust for under-identification in other administrative data sources.

Risk factor component

22. Adopt the GBD 2010 risk factor list as a starting point, adjusted where necessary for the Australian and Indigenous contexts. This will include consideration of the addition of some risk factors that are of particular interest in Australia and the exclusion of other risk factors which are not considered relevant to the Australian and/or Indigenous contexts.
23. Use national and Indigenous-specific data sources for compiling risk factor exposure data.
24. Consider how to incorporate social determinants and access to health services. Undertake work to investigate whether there is enough evidence to include these as risk

factors in the Australian and/or Indigenous studies. If not, then consider the feasibility of disaggregating results by social determinants rather than including as risk factors.

25. Assess on a case-by-case basis whether to use multiple data sources or the single best data source for each risk factor, according to a protocol to be developed with the assistance of the EAG.
26. Adopt GBD risk-outcome pairs and theoretical minimum risk exposure distributions, with extra risk-outcome pairs added where good Australian and/or Indigenous data exist, to calculate attributable burden.
27. Adopt GBD effect sizes unless there is compelling evidence that these would differ for the Australian or Indigenous populations.
28. Calculate attributable burden for the combined set of risk factors included in the Australian and Indigenous BoD studies (method to be determined).

Additional considerations for the Indigenous study

29. Adjustment for under-identification is not required for risk factor exposure data sourced from the 2011-12 AATSIHS. For other data sources used, assess whether it is appropriate and possible to adjust for under-identification.

Operational considerations

A number of the methodological decisions above rely on operational issues, such as the development of various protocols and input from the Expert Advisory Group. As such, the following operational factors should be considered for both the national and Indigenous components of the Australian Burden of Disease study:

30. Establish partnerships and communication channels between those undertaking related work in Australia and in other countries, including state and territory health authorities, universities and related organisations.
31. Consider options for outputs from the project which meet the needs of users (including policy makers) such as analytical reports, data visualisation tools and method papers. These will need to be considered in light of the resources and timeframes allocated for the project.
32. Make adjustments to the GBD 2010 cause list to form the Australian modified cause list (to be used in both the national and Indigenous components) using a defined protocol in consultation with the Expert Advisory Group and other relevant reference groups established for the project.
33. Make adjustments to the GBD 2010 risk factor list for the national and Indigenous analyses using a defined protocol in consultation with the Expert Advisory Group and other relevant reference groups established for the project.
34. Consider the feasibility of backcasting estimates using the same methods used for the updated national estimates. Also consider the feasibility of providing projections.

Appendix A New Zealand Burden of Disease study

The New Zealand Burden of Diseases, Injuries and Risk Factors Study, 2006–2016 (NZBD) was released by the New Zealand Ministry of Health on 8 August 2013. It includes estimates for 2006, and projections to 2011 and 2016, nationally and for both major ethnic groups (Māori and non-Māori) in New Zealand. It is available online at <
<http://www.health.govt.nz/nz-health-statistics/health-statistics-and-data-sets/new-zealand-burden-diseases-injuries-and-risk-factors-study-2006-2016#nzbdstudy>>

The NZBD currently includes four major components (listed below) with Māori -specific and topic-specific reports to be released at later dates:

- Health losses in New Zealand report which summarises the key findings, objectives and methods of NZBD and discusses implications for policy and monitoring
- Injury-related health loss report which provides more detailed methods and results for intentional and unintentional injuries
- Ways and Means report on methodology which provides more detail about the data sources and statistical and epidemiological methods used in NZBD
- Statistical annexe which is a series of excel pivot tables providing the full results from the study, including point estimates and standard errors, by age, sex and ethnicity.

The NZBD is independent of the Global Burden of Disease Study 2010. The two studies use different data sources and methods; however NZBD sourced many important standards from GBD 2010 including the reference life table for calculating YLLs, disability weights for calculating YLDs, and theoretical minimum risk exposure distributions and relative risks for risk factor analysis.

Methodological overview

The NZBD study provides estimates for 2006 as the reference year, with projections to 2011 and 2016 based on continuing demographic and epidemiological trends.

There is no sub-national analysis; however, estimates of DALYs, YLLs and YLDs by ethnicity (Māori and non-Māori) were calculated using direct standardisation against WHO world population. A summary of differences by ethnicity are included in the main report, expressed as rate differences and standardised rate ratios.

Data sources

The NZBD used a variety of data sources predominantly from New Zealand, of which an estimated 75% was considered “highly robust”. Of particular note was the use of the New Zealand Health Tracker to link different administrative databases using the National Health Index which is a person-level identifier assigned to every person who uses health and disability support services in New Zealand. Where possible and appropriate, standards were obtained from the GBD 2010 study.

Table X.1 summarises the data sources used in the NZBD.

Table A.1: Data sources for NZBD

Data	Source
Population estimates	New Zealand Health Tracker* population (derived from primary health organisation enrolment data); Statistics New Zealand: 2006 Census, series 5 projections for 2011 and 2016
Deaths	New Zealand Mortality Collection: average for 2005–2007, projections for 2011 and 2016
YLL weights	Reference life table from GBD 2010
YLD: incidence / prevalence	National Minimum Dataset (NMDS) – publicly funded hospital discharges New Zealand Health Tracker* Population-based surveys: 2004 Mental Health Survey (Te Rau Hinengaro), 2006/07 New Zealand Health Survey, and 2008/09 New Zealand Adult Nutrition Survey Disease registers: New Zealand Cancer Registry, New Zealand Birth Defects Registry, Public Health Surveillance Databases, the Australia and New Zealand Dialysis and Transplant Registry (ANZDATA) New Zealand and international research studies
YLD: duration / severity	New Zealand Health Tracker* New Zealand and international research studies
Disability weights	GBD 2010, NZBD expert panel
Risk factor exposure	Population-based surveys: 2004 Mental Health Survey (Te Rau Hinengaro), 2006/07 New Zealand Health Survey, 2008/09 New Zealand Adult Nutrition Survey, 2006 Census (smoking), alcohol consumption surveys, research studies
Theoretical minimum risk exposure distribution (TMRED)	GBD 2010, other burden of disease studies
Relative risks	GBD 2010, other burden of disease studies, meta-analyses of randomised controlled trials or prospective cohort studies

* New Zealand Health Tracker = anonymously linked Ministry of Health administrative data sets.

Assumptions

The NZBD was calculated assuming no age-weighting and no discounting, consistent with GBD 2010.

Condition list and health states

NZBD reports on 217 diseases and injuries grouped into 16 broad condition groups that more or less coincide with the ICD-10 chapters. There are also summary estimates for alternative groupings or classifications of some of these condition groups (for example, combining mental and neurological disorder groups, injury and musculoskeletal groups, or reclassification of cancers, injuries and infections according to organ systems).

Classification of most conditions from ICD-10 codes was straightforward; however, where there was debate about where to classify a particular condition the following criteria were applied:

- The condition group that makes the most sense from a policy and planning perspective
- The location of the condition in the ICD-10 classification and the GBD 2010 classification

Conditions were decomposed into a total of 470 individual health states. As there is no universally agreed classification of health states into health conditions, this was undertaken on the basis of expert clinical and public health advice. In order to utilise the GBD 2010 health state valuations (disability weights), the health state definitions were developed to

enable mapping of the GBD health state valuations to the New Zealand health state definitions.

Fatal burden (YLLs)

Deaths data were sourced from the New Zealand Mortality collection by year of registration (excluding non-residents). Mortality counts for 2006 were averaged from deaths registered from 2005 to 2007 to reduce the impact of annual fluctuations in the number of deaths in each category, with rates calculated using the corresponding New Zealand population data.

Deaths coded to unacceptable causes were classified into one of 5 types of garbage code: implausible causes, immediate causes, intermediate causes, ill-defined causes and causes that are risk factors rather than diseases. Specific redistribution algorithms were then applied to each type. These algorithms included proportionate redistribution within age-by-sex strata and statistical models, with or without data linkage. In total, 10.4% of deaths were redistributed.

The reference life table for calculation of YLLs was sourced from GBD 2010.

Non-fatal burden (YLDs)

Years lived with a disability was predominantly based on prevalence estimates, consistent with GBD 2010; however, acute conditions (that is, conditions with a duration less than 1 year) were based on incidence.

Prevalence of health states were generally calculated from the condition prevalence and estimates of the severity distribution. Incidence-based YLDs were also calculated for cost-effective analysis purposes.

DISMOD II (second generation disease modelling software) was used to estimate missing data and to force consistency between epidemiological estimates derived from different sources.

Despite some concerns about the legitimacy of the disability weights generated by GBD 2010, the NZBD study adopted the GBD 2010 disability weights with as little modification as possible to aid comparability and international benchmarking. Where NZBD health states could not be mapped to GBD sequelae, or for a small number of health states where the disability weight was considered problematic or implausible, panel-based expert judgement was used to interpolate disability weights for those health states.

Disability weights were adjusted for comorbidity using the multiplicative model as in GBD 2010, which assumes comorbid conditions are independently distributed. This assumption was tested using New Zealand and Canadian health survey data, and resulting estimates were found to be a reasonable approximation.

Risk factors

The NZBD study aimed to include all proximal risk factors that are potentially modifiable, rather than distal social and environment risks to health such as poverty, inequality and climate change. In order to be included, risk factors needed to:

- Be an area of policy interest for the New Zealand Ministry of Health
- Be potentially modifiable
- Have strong causal link and reliable estimates of the excess mortality and morbidity.
- Be measurable, with reliable estimates for the New Zealand population

- Have reliable estimates of theoretical minimum risk exposure distribution (TMRED).

NZBD identified 31 individual risk factors including dietary, metabolic or physiological, behavioural and substance use, as well as adverse health care events, diseases that are risk factors for other diseases and injury risk factors. Several major groups of risk factors were excluded: environmental/occupational risks; risk factors relating to mental health and social structural determinants of health; and micronutrient deficiencies.

As some risk factors mediate the effect of other risk factors, a hierarchical classification was not used. Instead, risk factors were classified into 10 largely independent risk factor clusters where

- Related risk factors are in the same cluster
- Risk factors included in the same cluster must not mediate the effect of any other risk factor in the same cluster.

Risk factor exposure distributions were predominantly obtained from New Zealand population health surveys, while TMRED and relative risks were obtained (where possible) from GBD 2010.

Uncertainty estimates

The NZBD undertook two types of uncertainty analyses to assess uncertainty for YLL, YLD and DALY estimates:

- Qualitative analysis by public health physicians of each condition to assess robustness based on:
 - YLL:YLD ratio
 - Source of data for prevalence (or incidence and duration) and severity distribution
 - Source of disability weights and consistency with disability weights for related health states.

This process enabled “red flagging” of a number of conditions whose burden estimates were not considered robust and highlighted the need for improved data collections.

- Quantitative analysis of uncertainty using standard statistical simulation methods to generate standard error (SE) estimates. These simulation methods were applied to known distributions of mortality data, prevalence data, and disability weights. This quantitative analysis does not fully capture all sources of uncertainty inherent in the NZBD. While uncertainty around data was fully captured, model uncertainty (arising from such things as classification of conditions, standards and modelling assumption) is not fully captured. Therefore, the uncertainty intervals presented should be considered underestimates of true uncertainty.

The NZBD study chose not to present uncertainty estimates in the main reports in order to assist clarity and readability. Instead, a table outlining the broad margins of uncertainty (expressed as +/- %) for major components was included in the introduction, along with an assessment of “robustness”. Full enumeration of uncertainty is included in the Statistical Annex.

Comparison of estimates by ethnicity

The NZBD includes estimates for both Māori and non-Māori sub-populations to compare the burden experienced by each group. Absolute and relative comparisons between the two

groups were made using differences and ratios of rates standardised to the WHO world population.

The facility for this comparison was incorporated into the study design. In particular:

- Mortality data included ethnicity variable.
- Garbage code redistribution was scaled to maintain age-sex-ethnicity mortality envelopes.
- Same standard life table (GBD 2010) was used for both ethnic groups, so that the weight given to a death depended only on the age, and not on any other personal characteristics.
- Data sources for YLDs were extracted, where possible, by ethnic group. Where this was not possible, regression or other smoothing techniques were used to derive the sub-population estimates. For prevalence estimates based on surveys, a survey weight was available by ethnicity grouping. These data were then entered by age-sex-ethnic group into DISMOD-2.
- The same disability weights were assumed for both Māori and non-Māori sub-populations; however, the severity distribution did vary by demographic characteristics, including ethnicity
- Potential risk factors were included if likely to have a significant impact on ethnic inequalities, particularly for diseases-as-risk factors which were considered if the disease and linked condition accounted for substantial health loss in at least one age-sex-ethnicity stratum.
- Similar synthetic life tables (used for healthy life expectancy (HALE) estimates) were built by DISMOD-2 for both Māori and non-Māori sub-populations.

The NZBD has noted that extending this methodology to other ethnic groups within New Zealand would be challenging due to data limitations.

Comparing NZBD and GBD 2010

The NZBD has devoted a chapter of the main report to comparing the NZBD and GBD 2010 studies, with a particular focus on validation of the NZBD for international benchmarking.

As previously noted, the NZBD is independent of GBD 2010 and uses different data sources and methods; however, NZBD sourced many important standards from GBD 2010 (see Box A.1). As a result, comparability between NZBD and GBD 2010 is only partial due to some differences in classifications, modelling methods, assumptions and data sources.

Box A.1: NZBD 2006–2016 compared with GBD 2010

- NZBD and GBD have similar, but not identical, lists of conditions and sequelae, and the case definitions sometimes differ.
- NZBD and GBD use different data sources for most epidemiological inputs: incidence, prevalence, severity distribution, case fatality, remission, risk factor exposure.
- GBD did not use New Zealand-specific data other than mortality and some risk factor exposure data, but modelled burden in New Zealand largely on Australian data.
- Health state valuations (disability weights) are largely similar across the two studies, but they sometimes differ, either because no equivalent health state exists in the GBD list, or because the GBD disability weight was considered implausible for the New Zealand context by the NZBD team and its expert advisors.
- Theoretical minimum risk factor exposure distributions and relative risks are largely similar across the two studies, although the NZBD included fewer risk factors overall and some different risk factors (e.g. adverse health care events, unsafe sex, diseases-as-risks).
- The periods of observation do not coincide exactly (2010 for GBD and 2006 for NZBD), although backcasts (GBD) and forecasts (NZBD) are available.

Source: NZBD Health Loss in New Zealand p48

Comparison of NZBD and GBD 2010 results

Direct comparison of estimates for individual years is not possible as GBD provide estimates for 2005 and 2010, while NZBD provides estimates for 2006 and 2011. However, comparison of the overall number of deaths, YLLs, YLDs and DALYs for 2005 from GBD 2010 and 2006 from NZBD showed close alignment with NZ estimates generally within the uncertainty bounds of the GBD estimates.

Comparison of all-cause crude and age-specific DALY, YLL and YLD rates for New Zealand from both studies and the GBD 2010 Australasia region also generally show good agreement, as do rank orders and proportion of DALYs of major specific conditions, especially for conditions that have clear case definitions and better data sources. The main area of disagreement between the two studies is for back disorders, with some lack of agreement also around alcohol use disorders and osteoarthritis. Reasons for this lack of agreement are postulated within the NZBD report.

An analysis of estimates for comparable risk factors also showed relatively high agreement. The main differences were around estimates for vegetable and fruit consumption, and to a lesser extent illicit drugs (which was likely based on Australian data) and low bone mineral density.

Despite these differences, the New Zealand Ministry of Health assert there is sufficient agreement between NZBD and GBD 2010 estimates for both the Australasia region and New Zealand for GBD 2010 to be used for trends over time, and to benchmark New Zealand's performance against peer group countries.

Strengths and weaknesses of the New Zealand approach

The approach adopted by New Zealand can be considered a ‘hybrid’ approach, whereby standards and methodologies from GBD 2010 were incorporated into a predominantly independent study.

One of the main advantages of this approach was the ability to design a study to suit New Zealand Ministry of Health policy and planning requirements using New Zealand data, while maintaining a sufficiently high level of comparability with GBD 2010 to enable international benchmarking and trends analysis. It also enabled analyses for Māori and non-Māori sub-populations which is not available within the GBD methodology.

While using GBD 2010 definitions and methods (where appropriate) enabled greater comparison between the two studies, it also impacted NZBD timelines as they waited for various components of GBD to be released.

The publication of two different sets of estimates is not considered a major issue. The New Zealand Ministry of Health have made clear comments about the expected sources of differences, and have provided clear guidance on how both studies should be used to inform policy planning and monitoring in New Zealand.

Applicability of New Zealand methods to the Australian context

There are a number of lessons to be drawn from the New Zealand approach in developing an approach for the Australian BoD studies:

- a ‘hybrid’ approach combining GBD 2010 methods and standards and country-specific data and definitions is possible and can yield similar estimates
- alignment with the GBD methods is important to maintain international comparability and benchmarking and, because the estimates turned out to be similar, could also be used to examine trends over time
- discussions of the variations of country-level estimates and rankings between the two approaches can be informative in “red flagging” those conditions where burden estimates do not agree, and can highlight the need for improved data collection
- like New Zealand, Australia has a high quality collection of data suitable for burden of disease analysis, removing reliance on regional estimates
- minor modification of cause and risk factor lists to better suit local health policy and planning can be accommodated within the GBD methodology
- disaggregation of national estimates by ethnicity by New Zealand provides some insight into undertaking the Indigenous component of the Australian study.

Consequently, a hybrid methodology should be considered a reasonable and valid approach to undertaking national and sub-national studies, and provides a useful template for calculating Australian burden of disease estimates.

Glossary

Age weighting – A method used to adjust the relative value of years lived at different ages, for example, to value a year lived by a young adult more highly than a year lived at older ages. Age weighting means that some age groups will have greater influence on the results than others.

Attributable burden – The burden attributed to a particular risk factor. It is the reduction in burden that would have occurred if exposure to the risk factor had been avoided. Compare **avoidable burden**.

Avoidable burden – The reduction in future burden that would occur if current and/or future exposure to a particular risk factor were avoided. Compare **attributable burden**.

Burden of disease and injury – Term referring to the quantified impact of a disease or injury on an individual or population, using the disability-adjusted life year (DALY) measure.

Comorbidity – When a person has two or more health problems at the same time.

Condition (health condition) – A broad term that can be applied to any health problem, including symptoms, diseases, and certain risk factors, such as high blood cholesterol and obesity. Often used synonymously with disorder or problem.

Cross-walking – a technique used in GBD 2010 to harmonise data to same reference definitions for modelling.

Data linkage (also referred to as data integration) – The bringing together (linking) of information from two or more different data sources that are believed to relate to the same entity, for example, the same individual or the same institution. The term is used synonymously with ‘record linkage’ and ‘data integration’.

Disability – In burden of disease analysis, any departure from an ideal health state.

Disability adjusted life year (DALY) – A year of healthy life lost, either through premature death or equivalently through living with disability due to illness or injury. It is the basic unit used in burden of disease and injury analysis.

Disability weight – A factor that reflects the severity of health loss from a particular **health state** on a scale from 0 (perfect health) to 1 (equivalent to death).

Discounting – A method used to adjust the relative value of years lived (or lost) in the future. It is based on the assumption that a year lived in the future is of less ‘value’ than a year lived now. ‘Discounting for future benefits’ is standard practice in economic analysis.

Effect size – a statistical measure of the strength of the relationship between two variables, which is relatively independent of sample size. For example, the **relative risk** or the odds ratio.

Garbage codes – for the purposes of burden of disease analysis, causes of death that are inappropriately coded, either because they should not be considered underlying causes of death, because they are likely to be intermediate causes in the chain of events leading to death, or because there is insufficient detail to ascertain an appropriate cause.

Health-adjusted life expectancy (HALE) – An estimate of the number of healthy years (free from disability or disease) that a person born in a particular year can expect to live based on current trends in deaths and disease patterns. Sometimes called ‘healthy life expectancy’.

Health states – groups of sequelae reflecting key differences in symptoms and functioning

Incidence – The number of new cases (of an illness or event, and so on) occurring during a given period.

Life expectancy (LE) – How long a person can expect to live on average given prevailing mortality rates.

Life table – a table that shows, for each age, the probability that a person of that age will die before their next birthday.

Population attributable fraction (PAF) – Relating to the health effects of a risk factor (or group of risk factors), the PAF is the percentage reduction in disease, illness, disability or death in a population that would occur if exposure to the risk factor was reduced to zero (or some other ‘ideal’ exposure distribution).

Prevalence – The number or proportion (of cases, instances, and so forth) in a population at a given time.

Relative risk – the risk of an event relative to exposure, calculated as the ratio of the probability of the event occurring in the exposed group to the probability of it occurring in the unexposed group.

Risk factor – Any factor which represents a greater risk of a health disorder or other unwanted condition or event. Some risk factors are regarded as causes of disease, others are not necessarily so. Along with their opposites, protective factors, risk factors are known as determinants.

Sequelae – consequences of diseases and injuries

Social determinants of health – the economic and social conditions (such as income, level of education and employment status) that influence health status

Uncertainty interval (UI) – A statistical term describing a range (interval) of values that are likely to include the ‘correct’ estimate. Limited or poor quality data create substantial uncertainty.

Vital registration – a system of recording important events in human life, such as births and deaths, including causes of death. Usually maintained by a government authority.

Years of life lost (YLL) – years of life lost due to premature mortality, calculated as the difference between the age at death and the reference life expectancy at that age.

Years lived with disability (YLD) – a measure of non-fatal disease burden, weighted for the severity of the condition.

References

- ABS (Australian Bureau of Statistics) 2013. Life tables for Aboriginal and Torres Strait Islander Australians, 2010–2012. Cat. no. 3302.0.55.003. Canberra: ABS
- AIHW 2011. The health and welfare of Australia's Aboriginal and Torres Strait Islander people, an overview 2011. Cat. no. IHW 42. Canberra: AIHW.
- AIHW 2013. Indigenous identification in hospital separations data–Quality report. Cat. no. IHW 90. Canberra: AIHW.
- Begg S, Vos T, Barker B, Stevenson C, Stanley L & Lopez AD 2007. The burden of disease and injury in Australia, 2003. PHE 82. Canberra: AIHW.
- Cooper RS & Freeman VL. 1999. Limitations in the use of race in the study of disease causation. *Journal of the National Medical Association*. 91(7):379-383.
- COAG (Council of Australian Governments) 2008. Intergovernmental Agreement on Federal Financial Relations: Canberra.
- Cutting GR. 1997. Genetic epidemiology and genotype/phenotype correlations. National Institute of Health. Program and Abstracts. NIH Consensus Development Conference on Genetic Testing for Cystic Fibrosis.
- Ezzati M, Vander Hoorn S, Rodgers AL, Lopez AD, Mathers CD & Murray CJL 2004. Potential health gains from reducing multiple risk factors. In: Ezzati M, Lopez AD, Rodgers A & Murray CJL (eds.) *Comparative quantification of health risks: global and regional burden of disease attributable to selected major risk factors*. Vol. 2. Geneva: World Health Organization.
- Hunt D, Blakely T, Woodward A & Wilson N 2005. The smoking-mortality association varies over time and by ethnicity in New Zealand. *International Journal of Epidemiology* 34:1020–8.
- IHME (Institute for Health Metrics and Evaluation) 2013a. The global burden of disease: generating evidence, guiding policy. Seattle, WA: IHME
- IHME 2013b. Global burden of disease (GBD) visualizations: GBD compare. Seattle, WA: University of Washington. Viewed 7 June 2013, <<http://viz.healthmetricsandevaluation.org/gbd-compare/>>.
- Karter A. 2003. Race, genetics, and disease - in search of a middle ground. *International Journal of Epidemiology*. 32:26-28.
- Lancet 2012. Global burden of disease study 2010. Viewed 21 June 2013. <<http://www.thelancet.com/themed/global-burden-of-disease>>
- Lee J, Heng D, Chia KS, Chew SK, Tan BY & K H 2001. Risk factors and incident coronary heart disease in Chinese, Malay and Asian Indian males: the Singapore Cardiovascular Cohort Study. *International Journal of Epidemiology* 30:983–8.
- Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H et al. 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 380:2224-60.

- Lopez A, Mathers C, Ezzati M, et al., editors 2006. Global burden of disease and risk factors. Washington (DC): World Bank.
- Mathers C, Vos T & Stevenson C 1999. The burden of disease and injury in Australia (full report). Canberra, AIHW.
- Murray C, Ezzati M, Flaxman A, et al. 2012. GBD 2010: design, definitions, and metrics. *Lancet* 380:2063–2066
- Murray C and Lopez A 1996. The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020. Harvard: Harvard School of Public Health.
- New Zealand Ministry of Health (NZ MoH) 2013a. Health loss in New Zealand: A report from the New Zealand Burden of Diseases, Injuries and Risk Factors Study, 2006-2016. Wellington: Ministry of Health.
- New Zealand Ministry of Health (NZ MoH) 2013b. Ways and Means: A report on the methodology from the New Zealand Burden of Diseases, Injuries and Risk Factors Study, 2006-2016. Wellington: Ministry of Health.
- Peto R, Lopez AD, Boreham J, Thun M & Heath C 1992. Mortality from tobacco in developed countries: indirect estimation from national vital statistics. *Lancet* 339:1268–78.
- Salomon J. 2010. New disability weights for the global burden of disease. *Bulletin of the World Health Organization* 88: 879.
- Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, Mokdad A et al. 2012. Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010. *Lancet* 380:2129–43.
- Schroeder SA 2012. Incidence, prevalence, and hybrid approaches to calculating disability-adjusted life years. *Population Health Metrics* 10:19.
- Shai I, Jiang R, Manson JE, Stampfer MJ, Willett WC, Colditz GA et al. 2006. Ethnicity, obesity, and risk of Type 2 diabetes in women. *Diabetes Care* 29:1585–90.
- Vasilevska M, Ross SA, Gesink D & Fisman DN 2012. Relative risk of cervical cancer in indigenous women in Australia, Canada, New Zealand, and the United States: A systematic review and meta-analysis. *Journal of Public Health Policy* 33:148–64.
- Vos T, Barker B, Stanley L & Lopez AD 2007. The burden of disease and injury in Aboriginal and Torres Strait Islander Peoples 2003. Brisbane: School of Population Health, The University of Queensland.
- WHO 2008. The global burden of disease: 2004 update. Geneva, WHO.
- WHO 2009. Global health risks: mortality and burden of disease attributable to selected major risks. Geneva, WHO.
- WHO 2013. About the Global Burden of Disease (GBD) study. Viewed 8 May 2013. http://www.who.int/healthinfo/global_burden_disease/about/en/index.html

List of tables

Table 1B.1: Australian Burden of Disease Study Expert Advisory Group (as at 25 February 2014)	12
Table 1C.1: Australian Burden of Disease Study Indigenous Reference Group (as at 15 May 2014)	13
Table 3.1: GBD 2010 cause-sequelae list: examples of hierarchical structure	23
Table 3.2: GBD 2010 risk factor list: examples of hierarchical structure	31
Table 3A.1: GBD 2010 hierarchical risk factor list.....	37
Table 5.1: Summary of the differences between incident and prevalent YLD.....	59
Table 5.2: Comparison of top five causes of YLD, by incidence and prevalence, Australia, 2003	63
Table 5.3: Disability weights for selected specific conditions, 2003 Australian and Indigenous studies and GBD 2010	64
Table 6A.1: Risk-outcome pairs and theoretical minimum exposure definitions used in GBD 2010	79
Table 6A.2: Risk-outcome pairs and theoretical minimum exposure definitions used in 2003 Australian Burden of Disease study	82
Table 7C.1: Timing of release of population data by geography	106
Table A.1: Data sources for NZBD	119

List of figures

Figure 1A.1:	Map of GBD 2010 regions and super-regions	11
Figure 2.1:	Per cent change in the value of a year due to discounting: no discounting and 3% discounting	16
Figure 2.2:	Inputs to a burden of disease study, prior to GBD 2010.....	17
Figure 2.3:	Inputs to a burden of disease study, GBD 2010.....	17
Figure 2.4:	A simple disease model	18
Figure 2.5:	A more complex disease model	18
Figure 3.1:	Number of causes in the various revisions of the GBD cause list and 2003 Australian burden of disease study cause list, disaggregated by different levels in the cause hierarchy	24
Figure 4.2:	Deaths by sex and age group, Australia 2011	45
Figure 4.3:	Indigenous deaths by sex and age group, Australia 2011	49
Figure 6.1:	The basic process for calculating the burden of disease attributable to risk factors.....	69
Figure 8.1:	Eighteen components of the GBD 2010 and the inter-relationships	108

The AIHW is undertaking a project to update burden of disease estimates for both Australia and Aboriginal and Torres Strait Islander people. This working paper, the first in a series describing the methods and decision-making processes of the project, describes the methods used in the 2010 Global Burden of Disease Study and assesses their applicability to the current Australian and Indigenous contexts.