SHINE Modelling Protocol

A user guide for generating inputs, running models, and processing and visualizing outputs.

Version 2.0, 19 April 2026

Ms Samantha Howe
Dr Tim Wilson
Professor Tony Blakely
and the SHINE team

Recommended citation:
Howe S, Wilson T, Blakely T. SHINE Modelling Protocol: a user guide for generating inputs, running models, and processing and visualizing outputs [version 2.0]. 2026 Apr 19. Population Interventions Unit, Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, University of Melbourne. doi: 10.26188/32528262

Contents

Executive Summary

WHAT IS SHINE?

The Scalable Health Intervention Evaluation (SHINE) program has built a sophisticated data and simulation modelling infrastructure to estimate the health and cost impacts of population interventions. SHINE's infrastructure allows users to model interventions that change disease rates (e.g. incidence, case fatality, remission or severity) or risk factors (e.g. tobacco, salt intake). SHINE leverages a vast array of existing epidemiological and economic data, applying these as inputs to simulation models to predict intervention impacts on future health, health inequality and economic outcomes.

PURPOSE OF THIS PROTOCOL

This protocol describes the standardised methodologies used by SHINE, in order to support clear and practical processes for current and future modelling. The key users are SHINE team members and affiliates, who can use this protocol to build off currently available models, or when building a model from scratch. This protocol may also be utilised by:

- Reviewers of SHINE models seeking a more in-depth understanding than that routinely published.
- Researchers who may take guidance from components of this protocol, for example when conducting a cost-effectiveness analysis.

BRIEF OVERVIEW OF SHINE MODELLING

SHINE models all major risk factors and diseases into the future under 'business-as-usual (BAU)', with the general assumption that recent historic trends will continue. The next step is to conceptualize and specify intervention scenarios that 'act on top of BAU' through one or more of: changing risk factors distributions that lead to changes in disease incidence rates (by far the most common approach); changing disease incidence rates directly; changing disease case fatality and remission rates directly (e.g. from new treatments); and changing disease severity (e.g. an intervention that reduces severity of asthma attacks, but not necessarily incidence rates of asthma). These intervention scenarios are then simulated using a proportional multi-state life table (PMSLT), resulting in future differences in disease-specific morbidity and mortality, aggregated to differences in future all-cause morbidity and mortality (as health-adjusted life years and total deaths) under an intervention in comparison to BAU. The model also produces economic outputs, including

changes in total disease expenditure and population income earnings for an intervention compared to BAU (or another intervention).

For interventions that act through changing risk factor distributions, SHINE deploys three types:

- **Theoretical minimum**: instantaneously shifting the prevalence or distribution of a risk factor to its theoretical minimum level of risk within a population. Unlike comparative risk assessment in the context of burden of disease studies, which estimates the difference in (or attributable) disease burden today had there been no exposure to a harmful risk factor in the past, the SHINE model is prospective, determining the stream of future gains in health (or avoidable future burden) while allowing for time lags and competing morbidity and mortality.
- **Achieving targets**: altering the prevalence or distribution of a risk factor to achieve (over time) targets set by governments or other agencies (e.g. that smoking prevalence in Australia will be 5% by 2030).
- **Actual interventions**: conceptualising and specifying actual interventions that alter risk factor distributions (e.g. a tobacco mass media campaign or new tobacco tax). This usually leads to a cost-effectiveness analysis including the upfront costs of the intervention, and cost-offsets of downstream changes in disease expenditure due to changing disease epidemiology and population demography resulting from the intervention.

As of early 2026, the SHINE infrastructure can generate BAU disease models for any country by sex and age, and a growing number of risk factor BAU forecasts. Modelling infrastructure and capacity is more advanced for Australia, with:

- Socioeconomic status (SES) (using SEIFA quintiles) able to be incorporated for disease and risk factor models – allowing quantification of health inequality impacts of public health interventions.
- Disease expenditure data incorporated (using NZ estimates from Blakely et al. [1] and methodology developed by Grimshaw et al. [2]), allowing cost offsets or future changes in health system expenditure able to be estimated due to interventions changing future disease prevalence and survival.
- Income loss by disease data incorporated (using NZ estimates from Blakely et al. [3]), allowing estimates of future gains in working age income due to interventions changing future disease prevalence and survival.

Extensive use is made of Global Burden of Disease study consortium (GBD) data to generate the epidemiological and demographic inputs to SHINE BAU models. The Australian Institute of Health and Welfare (AIHW) provides disease expenditure data [4] and epidemiological data to disaggregate models by SEIFA. To allow rapid and easily update-able analyses, SHINE has and will continue to invest in building software and data pipelines that routinise the collation and processing of these inputs where possible.

Other aspects of SHINE modelling are bespoke (and sometimes time consuming), such as pre-PMSLT modelling of complex risk factors (e.g., tobacco and vaping dynamics allowing for illicit markets of each, or COVID-19 or pandemic pathogen dynamics in a population), or the inclusion of additional population heterogeneity (e.g., by remoteness and Indigenous status in addition to sex, age, and socioeconomic strata). Concpetualizing and specifying 'actual interventions' is also often a bespoke and time consuming process. For example, estimating how reducing the number or density of tobacco retail outlets in a population will change smoking initiation and net cessation (and vaping and illicit market supply) can be very complex.

This protocol outlines SHINE's standards, guiding principles, and methodologies that are used across models to ensure good comparability.

PROTOCOL STRUCTURE

The protocol is structured as follows:

- Part I: Introduction to public health modelling
- Part II: SHINE Business-as-usual model
- Part III: SHINE Intervention model
- Part IV: Linking the SHINE BAU and Intervention models
- Part V: SHINE model outcomes and reporting
- Appendix A: Systematic error and quantitative bias analysis
- Appendix B: Model heterogeneity
- Appendix C: Impact attribution in the PMSLT
- Appendix D: Disease time lags

In sections II - V, SHINE 'default' approaches and methodology are presented, alongside more complex approach where applicable. These relate specifically to how the BAU scenario is specified, in terms of risk factor complexity and population heterogeneity. There are three main categories, as follows:

- Australia default model guidelines: sex by age by SES risk factor models using a standardised approach for risk factor parameterisation (including the forecasting and mediation approaches).
- Australia complex model guidelines: variations in risk factor specification (using a non-standard approach for forecasting/mediation) or population heterogeneity (e.g., stratification of model inputs by remoteness or Indigenous status in addition to sex, age, and SES)
- Other country default model guidelines: sex by age only models that can be applied to different country settings, with the standard risk factor parameterisation approach.

Acronyms

ABDS Australian Burden of Disease Study.

ABM agent-based model.

ABS Australian Bureau of Statistics.

ACMR all-cause mortality rate.

AIHW Australian Institute of Health and Welfare.

APC annual percentage change.

BAU business-as-usual.

BPRF Burden of Proof Risk Function.

CEA cost-effectiveness analysis.

CFR case fatality rate.

CPI consumer price index.

CRA comparative risk assessment.

DALY disability-adjusted life year.

DW disability weight.

EKE expert knowledge elicitation.

GBD Global Burden of Disease.

HALY health-adjusted life years.

HIIC Health Intervention Impact Calculator.

HTA health-technology assessment.

ICER incremental cost-effectiveness ratio.

IDI Integrated Data Infrastructure.

IHME Institute For Health Metrics and Evaluation.

MCDA multi-criteria decision analysis.

NHMRC National Health and Medical Research Council.

NMB net monetary benefit.

OECD Organisation for Economic Co-operation and Development.

PAF population-attributable fraction.

PIF potential impact fraction.

PMSLT proportional multi-state life table.

PPP purchasing power parity.

PY person-years.

pYLDs prevalent years lived with disability.

QALY quality-adjusted life year.

QBA quantitative bias analysis.

RCT randomised controlled trial.

RF risk factor.

RR relative risk.

SEIFA Socio-Economic Index For Areas.

SES socioeconomic status.

SEV summary exposure value.

SHINE Scalable Health Intervention Evaluation.

TMREL theoretical minimum risk exposure level.

WHO World Health Organization.

YLDs years lived with disability.

YLLs years of life lost.

Glossary

allocative efficiency Efforts to improve overall health system efficiency or cost-effectiveness of funded preventive and healthcare interventions across all or many types of interventions effective many diseases and/or risk factors.

attenuation Reduction in the effect of an intervention over time, or diminishing returns from an intervention, for the population it is reaching, without a change in the number of people being reached (i.e., attrition).

attributable burden Burden from a given risk factor in a specific timeframe (usually a single year), that would be alleviated if the risk factor had never existed above its theoretical minimum risk exposure level.

attrition Reduction in uptake or reach of an intervention over time (but with the same effect size for those who are still impacted - i.e., no attenuation in effect).

avoidable burden Future burden that could be alleviated if a risk factor were reduced to its TMREL (or some less harmful level than the current distribution/prevalence).

business-as-usual (BAU) The future projected demographic, epidemiological, health expenditure, and income profile of the population, estimated as that which would occur if historic trends continued. Also referred to as 'status quo'.

cohort lifetable Non-static lifetable that projects mortality rates for people born in a given year, to give that birth cohort's life expectancy and other lifetable-derived outputs. Life expectancy is therefore based on year of birth (as opposed to period lifetables, where life expectancy is fixed by age for all years of interest).

comparative risk assessment (CRA) Burden of disease analysis that attributes health burden (i.e. disability adjusted life years; DALYs) to different risk factors. This is done by comparing current disease burden with a counterfactual scenario in which past risk factor distributions shifted to the theoretical minimum exposure level (TMREL). This is therefore a measure of attributable burden.

confounding bias due to the presence of a confounder, which is associated with both the exposure and outcome of interest. The confounder does not lie on the causal pathway between A and B.

consistency Stipulates that an exposure/intervention is specific enough that variation (e.g., by applying the intervention to a different population, or at a different time) won't affect the outcome. It requires that interventions be clearly defined. Consistency is one of three pre-requisites to identify causal effects in empirical data (along with positivity and exchangeability).

cost-utility analysis (CUA) estimating and comparing the cost per unit of health gained from an intervention against a comparator (which could be a 'no intervention' or business-as-usual scenario, or another intervention), with summary metrics of health or 'effect' such as quality/disability/health-adjusted life years.

cost-effectiveness analysis (CEA) estimating and comparing the cost per unit of health gained from an intervention against a comparator (which could be a 'no intervention' or business-as-usual scenario, or another intervention), using natural units of health or 'effect', such as blood pressure, deaths, hospital days. Note that CEA is also often used as an umbrella term to describe each of cost-effectiveness, benefit, and utility analyses.

cost-benefit analysis (CBA) estimating and comparing the cost of an intervention against a comparator (which could be a 'no intervention' or business-as-usual scenario, or another intervention), including both actual costs, and health impacts or 'effect' component converted into monetary terms. For example, net monetary benefit, or a cost-benefit ratio.

disability-adjusted life years (DALY) Burden of disease metric, used to estimate health loss in a specific period (usually a single year). It combines years lived with disability (YLD) in the current year with future years of life lost (YLL) due to premature deaths occurring in the current year.

discounting the process of determining the present value of future financial costs or other outcomes (e.g. QALYs/HALYs).

exchangeability exchangeability is one of three pre-requisites to identify causal effects in empirical data (along with positivity and consistency).

friction cost approach In health economics, the friction cost approach assumes that reduced work capacity by a given individual (due to morbidity/premature death) will be picked up by others, e.g., by hiring someone previously unemployed, or by existing employees taking on the work of the lost individual. This differs to a human capital approach.

health-adjusted life year (HALY) General definition: umbrella term for DALYs and QALYs. SHINE definition: unique burden of disease metric, measured as the sum of future morbidity-adjusted (using disability weights) life years.

human capital approach (HCA) In health economics, the human capital approach takes the value of an individuals current earnings (by sex, age, and any other considered factors), and assumes that this rate of earning is lost when work capacity is reduced due to morbidity or premature death in the working age population. This differs to a friction cost approach.

Incremental cost-effectiveness ratio (ICER) Measure of the relative cost and health benefit of two scenarios, usually an intervention vs. BAU. The ICER is calculated as the main output of a CEA and CUA.

information bias Bias arising due to poor measurement of the exposure, outcome, or other variables (including confounders). Also called measurement or misclassification error. It can be non-differential (bias that is equally likely among cases and controls, or exposed and unexposed) or differential (bias that is more unequally likely among cases and controls, or exposed and unexposed).

magic wand Hypothetical intervention scenario that reduces a risk factor's distribution or prevalence to zero. It provides an estimate of the maximum future benefit (in terms of avoidable burden) a population could receive from removing harmful risk factor exposure.

Monte Carlo simulation Modelling approach that predicts outcomes based on varying inputs, e.g., those with an uncertainty distribution. The Monte Carlo model randomly samples each input from their distributions many times, providing output with a distribution reflecting the uncertainty of all inputs.

multi-criteria decision analysis (MCDA) an analysis of a decision problem in which there are multiple objectives or criteria that are appraised through a systematic process whereby stakeholders assign scores and weights to different criteria. In public health, MCDA methodology can be applied in health technology assessments or other priority setting/decision making tasks, to incorporate different criteria such as health benefit, costs, equity impacts, patient values.

net monetary benefit (NMB) Type of cost-benefit analysis (CBA) in which the total cost of scenario A (e.g., an intervention) is subtracted from the total cost of scenario B (e.g., the no intervention/business-as-usual/comparator scenario); the total cost of a scenario is the net of the actual cost and the health or 'effect' measure converted into monetary terms.

period lifetable Static lifetable that applies a constant mortality rate by sex and age in a given calendar year (i.e., period), and then estimates life expectancy (and other lifetable outputs) as though people lived out their lives with these rates. It is usually constructed using the lowest sex by age mortality rates from any country, considered to be the 'ideal' life expectancy. Period lifetables are what is generally used in burden of disease analyses.

positivity Stipulates that both the exposed and unexposed groups in a study must have a non-zero chance of having each confounder sub-group present. In contemporary counterfactual frameworks for causal inference in epidemiology, positivity is one of three prerequisites to identify causal effects in empirical data (along with consistency and exchangeability).

potential impact fraction (PIF) Difference in population impact under two differing risk factor exposure levels (BAU vs. intervention exposure). In SHINE modelling, the PIF results in (usually) a change in disease incidence under an intervention. (Note: often called a 'population impact fraction)..

quality-adjusted life year (QALY) Burden of disease metric, measured as the sum of future quality-adjusted (using utility weights) life years.

Quantitative Bias Analysis (QBA) Group of methods used to quantify, and adjust, for residual confounding or bias in observational analyses.

scenario In the context of SHINE simulation modelling, a situation projected into the future by a proportional multi-state lifetable model. Business-as-usual is scenario, as are interventions.

selection bias Bias that relates to how the sample population of a study was obtained, allocated, and followed up. It occurs when the association between the exposure and outcome differs for those who participate compared to those who are eligible to participate. The main effect is from the joint association, or dependency, between the exposure and outcome - i.e., when participation in a study differs by both the exposure and outcome.

summary exposure value (SEV) Standardised risk factor exposure metric, with range 0-100%. Allows comparison across risk factors and populations.

technical efficiency Maximising health outputs from a fixed set of inputs - e.g., optimising resources in a hospital; or using the minimum inputs for a desired output - e.g., the most cost-effective use of funds to reduce ischaemic heart disease burden.

willingess to pay In health economics, this is the amount that society is willing to spend to receive healthcare or avoid poor health outcomes.

Part I

INTRODUCTION TO PUBLIC HEALTH MODELLING

Chapter 1

Rationale for public health modelling

1.1 What is public health intervention modelling, and why do it?

Public health intervention modelling involves prospectively estimating the health and economic impacts of a proposed intervention, compared with a relevant comparator (often the 'business-as-usual (BAU)' scenario).

What is a public health intervention? There is no single standard definition of a public health intervention. Public health interventions are often understood as actions aimed at primordial and primary prevention, such as immunisation programs or measures targeting risk factors for poor health (e.g., tobacco consumption, physical inactivity). However, they may also include secondary prevention activities, such as early detection and screening interventions (e.g., colorectal cancer screening), as well as pharmacological interventions (e.g., cholesterol lowering treatment for those at high risk of a cardio-vascular disease event). More broadly, public health interventions are usually applied at the population level, although improving equity in access to preventive actions and healthcare services is a core function of public health. Many public health interventions also involve organised efforts outside of the healthcare system, for example government agencies enacting policy or regulation.

What are health and economic impacts? Public health modelling can consider a wide range of health, economic, and broader societal impacts. When comparing an intervention with a comparator scenario, a commonly used summary measure is the cost per unit of health gained. The unit of health considered may be a 'natural' measure, (e.g., blood pressure, tobacco use prevalence, or number of deaths) or a combined morbidity and mortality metric such as the quality-adjusted life year (QALY).

How does this relate to policy formation? Policy decisions are influences by many factors, including budget constraints, the political cycle, the influence of advocacy groups, and the available evidence for a policy's impact. With limited health budgets, governments must prioritise which healthcare and public health interventions to fund across many diseases and risk factors - decisions that relate to allocative efficiency. Public health intervention modelling is a critical input to prioritisation and decision making – but it is not *sufficient*, as other factors are also critical to decision making (e.g. societal preferences and values).

1.2 Burden of disease analyses only take us so far

Many decisions in healthcare are supported by estimates of disease burden. Burden of disease analyses quantify health loss for each disease or condition in a specified period (usually a year) using disability-adjusted life years (DALY). DALYs combine two components:

- Years of life lost (YLL) due to premature mortality, based on a counterfactual of optimal global life expectancy which is used to show how many years of life are lost from disease; and
- • Years lived with a disability (YLD), representing morbidity experienced in that same year.

These metrics allow health loss to be compared over time, across sub-population groups (typically by sex and age), and across diseases. Burden of disease analyses can also be extended to include comparative risk assessment (CRA), which attributes current DALYs to past adverse exposure to risk factors. CRA asks the counterfactual question: "How much lower would current health loss be (in DALYs) in a given year if the population (by country or even globally) had never been exposed in the past to harmful levels of a given risk factor?" For example, how many DALYs would not have occurred today if the world never had tobacco? Whilst valuable, CRA is limited in policy relevance because it estimates attributable burden, not future avoidable burden or health gain – that is, it looks backwards rather than forwards.[5]

The Global Burden of Disease (GBD) is the largest producer of burden of disease analyses, publishing updates regularly since 1991. The latest release (GBD 2023), provides trends spanning more than 30-years. The GBD methodology has been expanded on and adapted over time to improve its breadth and accuracy. Other countries, including Australia, produce their own burden of disease analyses. The Australian Burden of Disease Study (ABDS) draws on GBD methods, with some methodological adaptations.

1.2.1 Attributable vs. avoidable burden

Attributable burden, as estimated through CRA, is a 'rear view mirror' measure: it quantifies current disease burden caused by past exposure to harmful risk factors. In contrast, avoidable burden looks forward. It measures the future health loss that could be prevented under a hypothetical intervention scenario compared with no (or another) intervention. [5] This perspective is far more relevant for contemporary policy making and decision-making.

A useful example is tobacco.

- A CRA of tobacco in a given population estimates current tobacco-attributable DALYs due to current and past smoking behaviour. Looking at past smoking accounts for the fact that there is a time-lagged relationship between tobacco smoking (the exposure) and disease risk. This type of analysis can be used to compare the contribution of tobacco, in comparison to other risk factors, to disease burden, and show changes in tobacco-related burden over time. However, the analysis is inherently limited as the counterfactual scenario is smoking never having existed, which is something that cannot be changed.
- A prospective, avoidable-burden analysis instead estimates future health gains if tobacco use were reduced or eliminated today, compared with a defined baseline scenario (which may be a constant future smoking rate, or the rate expected based on past trends).

Avoidable-burden modelling also explicitly incorporates time lags in risk factor-disease relationships. For instance, smoking may take years to cause cancer, and risk in former smokers declines gradually toward

baseline over several years. Prospective modelling can reflect these dynamic transitions, offering a more realistic and policy-relevant assessment of intervention impacts

1.2.2 Public health simulation modelling: the third leg of the stool

Burden of disease analysis can be visualised in terms of a three-legged stool. Burden of disease studies, produced largely by the GBD, have focused on two of these 'legs':

1. Health loss, in DALYs
2. Attribution of this health loss to past risk factor exposure

Public health simulation modelling acts as a complimentary next by focusing on the future, as opposed to a single point in time (or multiple points in time through a cross-sectional analysis). In doing this, simulation models consider avoidable, as opposed to attributable, burden.

Simulation modelling is therefore an important source of information on the potential future benefits of interventions. It can be used to assess impacts on diseases, or general 'all-cause' morbidity and mortality. This can provide valuable information for both researchers and policy makers.

The COVID-19 pandemic provided a key example of the value of modelling for public health decision making. Many governments around the world utilised evidence from simulation modelling to guide policies in the rapidly changing pandemic environment, including around optimal lockdown durations, mask use, and vaccination roll-out.

Simulation modelling is also highly useful for evaluating the long-term effects of new chronic disease treatments, screening tools, and other prevention strategies, on both health and costs.

However, it is important to be aware of the limitations of modelling. The commonly used phrase "all models are wrong, but some are useful", coined by George Box, talks to the idea that while it is not possible to perfectly recreate reality with a model, they can still provide valuable information. But how can we assure that a model is providing value? Saltelli et al.[6] expand on this in their paper 'Five ways to ensure that models serve society: a manifesto'. The key messages in this paper revolve around transparency in both the methods and results of a model. All models contain assumptions. Being explicit about these assumptions, and testing the impact of these assumptions on modelled results, is essential for transparency and to guide future work. It is also important that a modelling analysis doesn't provide conclusions that are beyond what has actually been modelled, or shy away from addressing the uncertainty in modelled results. Further, providing clear results in plain language for end-users prevents misinterpretation.

Another important concept explored by Saltelli et al.[6] is complexity, or 'hubris'. It can be difficult to find the balance between parsimony and sufficient detail when modelling. Unfortunately, there is no single formula or approach to find the 'correct' level of complexity that addresses the key issues of the disease or system being modelling, while preventing an overly complex system that cannot be understood by end-users.

Chapter 2

Principles of public health modelling

2.1 How can you do public health intervention modelling?

Public health simulation models utilise existing data, including demographic and epidemiological information, along with estimates of the causal effects of interventions, to simulate the long-term health and cost impacts of those interventions. These models are particularly useful for examining the population-level impacts of interventions that cannot be examined in trial settings, or for extrapolating trial findings over a longer time horizons or to other populations.

There are many ways to do intervention modelling in public health. Model conceptualisation and development, including the type of model utilised, is generally driven by the research question at hand (though may also be influenced by the resources and knowledge of the modelling team). The below sections of this Chapter highlight the key considerations when conceptualising a model.

2.1.1 Model types

A variety of models can be used to evaluate the health (and broader) impacts of intervention scenarios in a population. Model types vary in complexity, and are chosen based on the research question at hand.

Multiple frameworks exist to guide model selection. Squires et al. [7] have produced a framework for developing the structure of public health economic models, with the aim of producing models that are useful for decision makers. Building on this work, Briggs et al. [8] have produced a guideline for selecting a simulation model structure for evaluating non-communicable disease public health interventions. This guide includes an update on Brennan's taxonomy of model structures [7, 9]). Under this taxonomy, simulation models can be categorised by key components:

- Interactions: These can be between groups or individuals in the model, or between model units and the environment
- Time: If or how temporality is handled by the model
- Population-aggregation: Cohort or population level vs. individual level
- Stochasticity: How the model deals with randomness and population heterogeneity (applies to individual-

level or microsimulation only)

- Population heterogeneity: How much stratification of cohorts (for macrosimulation) or parameterisation of individual characteristics (for microsimulation) is/should be accounted for (e.g. sex, age, ethnicity, SES, disease risk, etc).

There are pros and cons to each simulation model type. While an agent-based model may allow for more interactions and heterogeneity than a macrosimulation model, data demands and model uncertainty will likely increase with the greater number of inputs required to parameterise a complex microsimulation.[8] The type of model selected may therefore depend on the data available. Some examples of public health simulation models include:

Markov models are traditionally used to model transitions between 'healthy' and 'disease' states, they can also be used to model transitions between behaviour states, for example 'non-smoker' to 'smoker' to 'ex-smoker'. 'Markov' is a broad term that covers different types of model structure, from population-level macrosimulation, through to individual-level microsimulation. A key assumption of all Markov models in macrosimulation is that of the memoryless property: when a proportion of the population moves from state A to state B, no information from state A is transferred through to the new state – i.e., there is no 'memory' of the history of the cohort. However, there are ways to overcome this assumption, e.g., with 'tunnel states'.[10] In Markov microsimulation, the memoryless property is not an issue as it is possible to keep track of agents' history and use this data to alter future transition rates.

Agent-based simulation models (ABMs) are a type of microsimulation, used for infectious disease (e.g., Szanyi et al. [11]) and non-communicable disease modelling. An ABM is well suited to situations where interactions of agents with other agents or the environment are important, and when there is notable heterogeneity between agents to model. For example, tobacco control modelling can utilise ABMs to account for how behaviours relating to taking up/quitting smoking are impacted by ones surroundings (both the environment, such as how close one lives to a tobacco retailer, and the people in that environment, for example whether ones parents/friends currently smoke) – an example of this type of simulation is the Tobacco Town model.[12]

Lifetable models. There are multiple types of lifetable model, which can be categorised based on whether it is period or prospective, and use of states.

- Period vs. prospective

  Period lifetable: these are inherently cross-sectional, as they use a fixed set of mortality rates by age in a given period (e.g. a calendar year). However, they may still be used for comparisons over time, assuming that mortality rates remain fixed. Period life expectancy is most commonly produced by statistical agencies such as the Australian Bureau of Statistics (ABS).

  Cohort lifetables: these use age-specific mortality rates for a cohort over time. For example, the life expectancy of people born in 1940 would reflect the mortality rates that have actually occurred for that cohort. If being applied in simulation modelling, mortality rates would be forecast to reflect likely future changes in mortality rates and resultingly life expectancy.

  Note that both period and cohort lifetables can be extended to include morbidity, for example calculating health-adjusted life expectancy (HALE) using Sullivan's method.

- Use of states in intervention modelling

  Simple lifetable: used to show changes in all-cause morbidity and mortality in a population by sex and age, without being split further into different specific health states. May be period or cohort.

  Multi-state lifetable: used to estimate morbidity and mortality across multiple states, which can be defined for different diseases. This model structure can be extended to include people

with two or more diseases simultaneously, but can become unwieldy in terms of the many input parameters required and the structure (also known as state explosion). For example, a 5 disease model requires 32 disease states to fully represent all possible combinations (1 non-diseased or healthy state, five with 1 disease, 10 combinations of 2 diseases, 10 combinations of 3 diseases, 5 combinations of 4 diseases, and 1 for having all 5 diseases).

Proportional multi-state lifetable: utilised to estimate mortality and morbidity trends across multiple diseases, combining these into estimates of overall mortality and morbidity. A proportional multi-state life table (PMSLT) consists of multiple lifetables, with an individual table for each included disease, in additional to an all cause cohort lifetable. Unlike multi-state lifetables, diseases are assumed to be independent. This has the advantage of avoiding state explosion, at the possible cost of not incorporating any important joint effects of having two or more disease simultaneously that are not just the sum (or product depending on parameter) across separate diseases. The PMSLT is beneficial to use alongside simpler models, such as the Markov model, as it allows for multiple diseases to be measured, which a Markov model cannot handle (due to state explosion).[8]

2.2 Key model components

Regardless of the type of simulation model used, intervention modelling is generally guided by a research question of the following form:

Box 2.2.1. Simulation modelling research questions

What is the impact of [intervention] in comparison to [comparator scenario] on [outcome(s)] over [time horizon] for [population]?

Table 2.1: Key components of simulation models of interventions that answer research question 2.2.

Component	Description
Business-as-usual (BAU) or comparator scenario (a.k.a. 'status quo')	BAU is generally conceptualised as the future distribution or trends in epidemiological and demographic factors that are expected to occur if no action were taken beyond that which is currently in place. Depending on what is being modelled, this could include forecasting disease rates, mortality rates/life expectancy, risk factor distributions, and population characteristics (e.g., migration rates). It requires consideration of current and historic policies acting on the factors of interest and the likelihood of these continuing into the future. Sometimes, a BAU scenario be defined by static rates - this may be the case if historic trends indicate no change over time, or if there is limited data to parameterise a future trend. In the latter case, multiple BAU scenarios may be used with a range of plausible future trajectories.

Continued on next page

Component	Description
Intervention	This is the scenario being compared to BAU (or to other interventions). Interventions can be conceptualised at various levels of complexity, ranging from complete removal of a risk factor from the population to detailed estimates of effect sizes of specific real-world interventions. Multiple interventions may be modelled and compared to BAU as a 'package' or ranked separately by their effectiveness.
Population	This is the group to whom the modelled intervention is to be applied. It may, for example, include an entire country, or a specific sub-population (e.g. people who currently smoke). The inclusion of population heterogeneity (i.e., how risk factor and disease risk distribution vary by characteristics such as ethnicity or socio-economic status) can inform important equity considerations. Modelling an open cohort (e.g. including new births and migration) is arguably better for future projections of absolute changes in health burden, particularly when assessing long-acting policies or interventions targeting youth, but introduces additional complexity.
Outcome(s)	These are the metrics to be compared between the intervention and BAU scenarios. Outcomes can be conceptualised at several levels, from direct measures of health (e.g., the number of future deaths), measures of disease burden (e.g., health-adjusted life years gained), or costs (e.g., government expenditure on healthcare). Disease burden and economic outputs should be presented both discounted and undiscounted, with appropriate uncertainty intervals provided.
Time horizon	This is the time over which outcomes are measured. It is distinct from the intervention duration. There is no single universally accepted time horizon to use. It is important to model a long enough time horizon to capture the effects of the intervention, balanced with the increasing uncertainty of forecasts further into the future. Thorough modelling should present the results for several time horizons, with and without discounting, allowing the end-user to select the scenario that best matches their needs.

2.3 Health and economic modelling

In public health simulation modelling, monetary impacts are often presented in addition to health outcomes, either as a separate outcome, or combined with health outcomes in an economic evaluation.

Incorporating costs into modelling output, particularly from multiple perspectives, can support multi-criteria decision analyses.

2.3.1 Economic evaluations

Economic evaluations compare the relative cost of an intervention compared to a reference scenario, and the relative health impact. A key output of an economic evaluation is an incremental cost-effectiveness ratio (ICER). This metric gives the cost per additional unit of health being received, i.e., the incremental cost and benefit.[13]

$$ICER = \frac{cost_{intervention} - cost_{comparator}}{health_{intervention} - health_{comparator}} \tag{2.1}$$

where $health$ is some measure of health benefit related to an intervention, and $cost$ covers the cost of implementing an intervention and may also include downstream costs (depending on the perspective taken - discussed in Section 2.3.3).

There are multiple types of economic evaluation. These differ by how health benefits are measured:

- cost-utility analysis (CUA): health outcome is measured in terms of a combined morbidity and mortality metric, e.g., ICER = cost per DALY averted or QALY gained.
- cost-effectiveness analysis(CEA): health outcome is measured in 'natural' units, e.g., ICER = cost per death averted, or cost per 1mmHg reduction in systolic blood pressure.
- cost-benefit analysis(CBA): includes impacts beyond the health sector, and all impacts are monetized (e.g. HALYs monetized based on some willingness to pay). Rather than calculating an ICER ratio, a benefit-cost or return-on-investment ratio is calculated.[13]

Cost-effectiveness plane

One way to compare different interventions is through a cost-effectiveness plane (as shown in Figure 2.1). The cost-effectiveness plane has four quadrants, with the position of a new intervention in the plane determined relative to the comparator, which usually sits at the centre. The x-axis represents effectiveness (the health component), and the y-axis represents cost.

- Upper left quadrant (NW): *dominated* - the new intervention is more costly and less effective that the comparator
- Upper right quadrant (NE): the new intervention is more costly, but is also more effective, than the comparator
- Lower left quadrant (SW): the new intervention is cheaper, but is less effective, than the comparator
- Lower right quadrant (SE): *dominant* - the new intervention is cheaper and more effective than the comparator

Willingness to pay and cost-effectiveness thresholds

An ideal new intervention would be assessed in an economic evaluation as being dominant - i.e., be more effective and less costly than the current standard intervention. This would be a compelling case to implement

the intervention (noting that other factors that cost and effectiveness should be taken into account, e.g., acceptability). Conversely, a new intervention that is less effective and more costly than the current standard would not be implemented. However, a cost-effectiveness decision still needs to be made for those interventions that fall into the remaining two quadrants of the cost-effectiveness plane, particularly those that are more effective than the current standard intervention, but are more costly. This is where cost-effectiveness, or willingess to pay, thresholds are important.

In health economics, willingess to pay is a measure of how much society is willing to spend to receive a certain health benefit or avoid a poor health outcome. WTP is used in individual studies looking at specific treatments and outcomes. However, it can also be applied more generally as a cost-effectiveness threshold, for example considering how much individuals are willing to pay to receive an additional QALY. A cost-effectiveness threshold may be used to benchmark whether a given intervention being analysed is cost-effective or not, or how likely it is to be cost-effective.

Figure 2.1 shows a cost-effectiveness plane with a $50,000 per health-adjusted life year (HALY) gained WTP threshold (HALYs described in detail in Section 14.2.1 of this protocol). Four hypothetical interventions are shown on the plane. Interventions #3 and #4 sit in the lower right quadrant and are therefore dominant, i.e., they are both less costly and more effective than the BAU scenario. While interventions #1 and #2 are in the upper right quadrant and are therefore more costly than the comparator, both sit *below* the WTP threshold, and would therefore still likely be seen as cost-effective options (particularly #2, which has the greatest health gain of the four options).

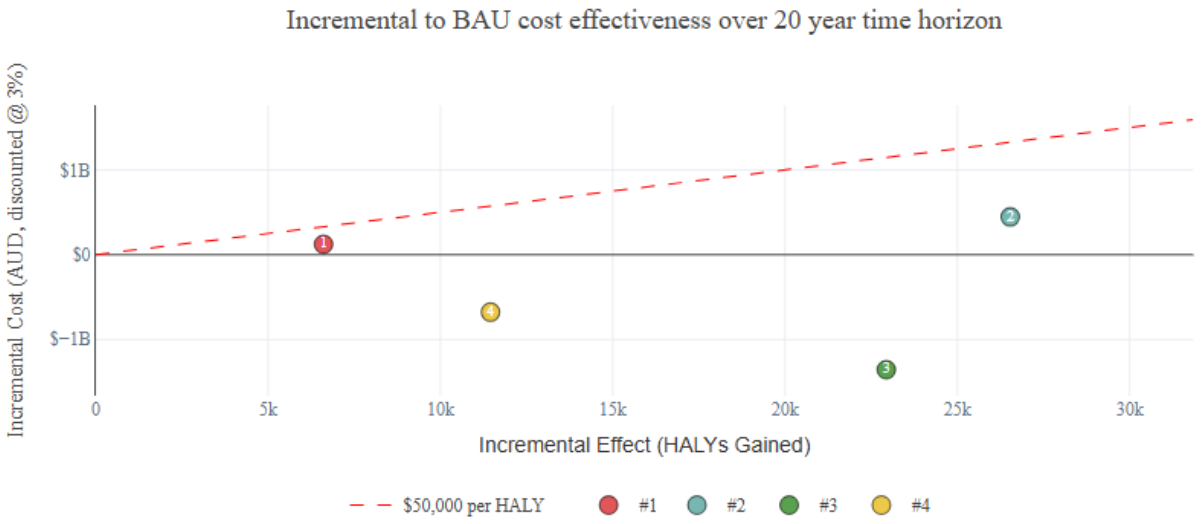

Figure 2.1: Example Cost-effectiveness plane, from the HIIC Tool. Produced using the SHINE HIIC Tool (see https://shine-hiic.com/dashboard)

2.3.2 Discounting

Discounting is the process of determining the present value of future costs (or other outcomes). It is particularly relevant when modelling interventions that take many years to generate their benefits.

Importantly, discounting is not the same as inflation. Discounting is based on the economic principle of

positive time preference, whereby society generally prefers to delay incurring costs into the future.[14] This principle applies to financial costs, as well as other consequences (or 'benefits'). For example, society would generally prefer to receive health benefits now as opposed to some later date the future.[14]

The present value of some future value (economic cost or other) incurred n years into the future, is calculated as

$$\text{Present Value} = \frac{\text{Value}_n}{(1+r)^n} \tag{2.2}$$

where r is the discount rate. E.g., If something costs $100 in 2033, at a 3% discount rate the present value in 2023 is calculated as Present Value $= \frac{100}{(1+0.03)^{10}} = 76.6$.

Different organisations have different standards or requirements for discounting of costs and consequences. While these recommendations generally pertain to health technology assessments that are required for new pharmaceuticals or medical technologies, the guidance can also be applied for health economic modelling. The PBAC in Australia recommends a discount rate of 5% in economic evaluations.[15] Comparatively, the National Institute for Health and Care Excellence in the UK recommends a 3.5% discount rate. There is also inconsistent advice as to whether both costs *and* health benefits in an economic evaluation be discounted, and if they should be discounted at the same rate.

Generally, for a given discount rate used in an economic analysis, sensitivity analyses with varying discount rates, or no discounting, would be undertaken. This is important present to allow greater comparability to other studies, as the discount rate used can cause large variation in modelled results.

2.3.3 Perspective

A variety of perspectives can be used when including economic costs in a public health simulation model. Table 2.2 summarises the differences between commonly used perspectives.

Table 2.2: Perspectives for costing/economic evaluation

Perspective	General definition
Health system	Direct costs within the healthcare system. This can include patient out-of-pocket (OOP) costs, and costs to the government or other funders (e.g health insurance company) for providing healthcare (including the costs of new public health interventions).
Patient	All costs incurred/received by patients. This can include direct medical costs such as that spent on medications and other treatments, as well as non-medical costs such as travel, and income gained/lost.
Government	This can include both government healthcare expenditure and other costs external to the healthcare system, for example taxes and pension payments. Note that in most economic evaluations, tax is not included as this is a 'transfer payment'.

Continued on next page

Perspective	General definition
Societal	Broadest perspective taken. Theoretically, a societal perspective includes all cost impacts that result either directly or indirectly from an intervention. In practice, this is difficult to do. It may be seen as an aggregation of the costs included in above perspectives.[14]. Key components often estimated when a societal perspective is taken are: • healthcare costs to government (including direct intervention costs) • OOP healthcare costs to patient • productivity costs • external/indirect impacts, e.g., on the environment or crime

2.4 Existing modelling guidelines

There is no optimal or single guideline for public health simulation modelling (hence the need for this protocol). Several relevant guidelines exist that provide partial assistance, summarised in Table 7.1.

Table 2.3: Guidelines relevant to public health simulation modelling

Guideline/report	Details
Reporting guidelines	
Guidelines for Accurate and Transparent Health Estimates Reporting (the GATHER statement, or simply 'GATHER') [16]	Reporting guidelines for syntheses of health information. Health estimates include disease epidemiology (incidence, prevalence), risk factor/health behaviour exposure (such as smoking prevalence or the distribution of blood pressure in a population).
Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement	Health economic evaluations reporting guidelines (note: not specifically modelling/public health).
Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) Statement	Reporting and validation guidelines for predictive modelling.
Overview, Design concepts and Details (ODD) protocol [17]	Reporting guidelines for agent-based and other microsimulation models.
Healthcare modelling guidance	

Continued on next page

Guideline/report	Details
International Society for Pharmacoeconomics and Outcomes Research (ISPOR)-Good Research Practice in Modelling Task Force reports, e.g. [18].	Health care modelling best practice Reports cover six areas: Model conceptualisation, State-transition modelling, Modelling using discrete event simulation, Dynamic transmission modelling, Model parameter estimation and uncertainty analysis, Model transparency and validation. Note that this is not specific to public health modelling.
ISPOR System, Interactions, Multi-level, Understanding, Loops, Agents, Time, Emergence (SIMULATE) tool [19]	Dynamic simulation modelling guidelines. May be useful if conducting complex modelling, e.g., agent-based modelling. This guideline is specific to dynamic simulation models of health care systems.
ISPOR Optimization Methods Emerging Good Practices Task Force reports, e.g. [20]	Guide to using optimisation methods. Useful for optimising a model input parameter.
BODE^3 protocols, e.g. [21]	Simulation modelling guidelines for disease burden research. Useful for guidelines on gathering model inputs and adjusting inputs for uncertainty. Note that BODE^3 is a precursor to SHINE.
Other specific guidance	
ACE-Prevention reports.	Classifying strength of evidence. Latest report: Vos et al. (2010) . Provides guidance on determining the strength of evidence. This is important to consider when parameterising intervention effect sizes.
Briggs et al. (2016)	Model selection guide. Specific to non-communicable disease modelling for public health.
Blakely et al. (2021)	Specific PMSLT guide. Important pre-reading for SHINE modelling. Note that this publication was produced prior to 'remission' being incorporated into the SHINE PMSLT.

Chapter 3

Uncertainty and heterogeneity in modelling

Public health modelling draws on many sources of primary and secondary research as input parameters. Uncertainty in these input parameters needs to be accounted for. However, there are also other sources of uncertainty in simulation models.

Four types of variability or uncertainty are defined in relation to modelling:

1. Stochastic uncertainty (micro-simulation only).
2. Model structure uncertainty
3. Heterogeneity (not so much uncertainty, but true variation in the population that we may want to explicitly model).
4. Input parameter uncertainty.

It is important that when forecasting or modelling the future, each source of uncertainty is taken into account explicitly.

3.1 Stochastic uncertainty

Stochastic uncertainty is the first-order randomness that results in different outcomes for identical individuals, even given the same odds.[23] It is inherent in simulations that track individual people or 'agents', e.g., agent-based or discrete-event simulation models, where two people with exactly the same characteristics and the same probability of something happening can still end up with different results, simply due to chance. For example, imagine a simulation where two 45-year-old men have identical profiles: same weight, diet, exercise habits, genetics. The model calculates that each has a 20% probability of having a heart attack in the next ten years. When the simulation is run, one man has a heart attack, and the other doesn't. This difference came from random chance (stochastic uncertainty) in the simulation. Importantly, stochastic uncertainty is *not* the uncertainty in the probability value itself. This is parameter uncertainty, discussed in Section 3.4.

3.1.1 Managing stochastic uncertainty

Stochastic uncertainty reflects that real-world outcomes vary and helps to quantify the range of possible outcomes. To address stochastic uncertainty and random noise in simulation modelling results, the model needs to be run many times to look at the average or distribution of outcomes. This is also known as Monte Carlo simulation (discussed in Section 13.3).

3.2 Model structure uncertainty

Model structure relates to the chosen simulation model and how it is constructed. For example, whether a linear or log-linear regression model is used to specify a risk factor trend. Or, how many/what states are included in a Markov model. Or, whether or not diabetes is also treated as a risk factor for ischaemic heart disease and stroke in addition to being its own outcome. It is often assumed that model structure uncertainty is large – perhaps often more than the cumulative effect of input parameter uncertainty.

3.2.1 Managing model structure uncertainty

Model structure decisions and assumptions need to be made for simplicity and feasibility. While a more complex model structure may be more realistic of the population being studied, it can also be increasingly complex to parameterise with necessary input data and difficult to analyse and interpret.

Structural assumptions should be stated clearly for end-users. Further, where possibly, structural assumptions should be tested in sensitivity analyses.

Model structure uncertainty can also be tested through comparison with completely different model structures. While generally not feasible for a single research team, such comparisons may be made to external published model outputs - though this may be limited by differences in input parameters used for each model.

3.3 Heterogeneity

Heterogeneity is the true variation that exists within a population. While this is not really uncertainty, it is often grouped with measures of uncertainty as it is can affect the accuracy of model output and conclusions drawn from modelling.

Heterogeneity can be categorised into two distinct types

Clinical heterogeneity, and similarly, risk stratification. These are differences in disease profiles or risk that may be modelled.

Socio-demographic heterogeneity. This is the differences in upstream factors that may affect treatment effects and measured health outcomes in a study or model. Sex, age, income, and ethnicity and some common examples.

3.3.1 Managing heterogeneity

Heterogeneity is always present in the population - the decision of whether to model heterogeneity across different dimensions depends on the research question and intervention.

There are also structural and data constraints that may limit the ability to include heterogeneity across more than by sex and age. For example, when using a Markov macro-simulation model, state explosion can occur with an increase number of states or strata. A greater level of heterogeneity can be achieved using a microsimulation model, but it may not be possible to obtain accurate data on the stratification of interest for a microsimulation.

3.4 Input parameter uncertainty

All research output is uncertain and has the risk of error. Epidemiologists and other quantitative researchers most often use incorporate this into research through confidence intervals, which measure random error around a point estimate. However, there are also systematic errors in most empirical studies, which are not captured in confidence intervals, but can be quantified with other measures (though these are less commonly applied in practice).

Results from quantitative research can have uncertainty from:

- Lack of **precision** = random error
  - Determined by study size (i.e., power of the study)
  - Generally quantified in the form of a confidence interval and p-value
- Lack of **validity** = systematic (i.e., non-random) error
  - Internal validity: impacted by three key sources of bias/error that result in poor accuracy: information bias, confounding, selection bias.
  - External validity, including:
    * generalisability to the population from which the study sample was obtained
    * transportability to populations other than that from which the study sample was obtained

Systematic error is discussed in more detail in Appendix A.

3.4.1 Managing input parameter uncertainty in modelling

Most input parameters derived from existing research will be at risk of the above uncertainty. Multiple approaches can be taken to address uncertainty in input parameters going into a public health simulation model, as discussed in Appendix A. The gold standard approach would be to conduct quantitative bias analyses for each parameter and use adjusted results in the simulation model. Inputs may also require further adjustment to account for external validity, though there are currently no standardised quantitative methods to address these factors. Such adjustments would likely both widen the confidence interval around point estimates, and shift the point estimate itself.

In reality, these analyses are often not feasible, particularly given the large number of inputs required for most simulation models. Most inputs are directly derived from the literature without adjustment, using the confidence interval from the input source as the uncertainty interval - each input can be drawn from its uncertainty distribution in each iteration of a Monte Carlo simulation, to show the impact of input parameter uncertainty on modelled outcomes. This, however, does not address the point estimate of the input parameter itself being an under/overestimate of the 'truth'.

Transparency around the source of input parameters, any adjustments made, and possible residual biases, is essential in any modelling analysis.

Chapter 4

SHINE

4.1 What does SHINE do?

Scalable Health Intervention Evaluation (SHINE) models the health and economic impacts of interventions that act through one or more of: changing risk factor distributions that lead to a change in disease rates (the major methodology of SHINE and focus of this protocol); changing disease incidence rates directly; changing disease case fatality or remission rates; and changing disease severity.

SHINE simulation modelling considers the future health and economic impact of interventions that target risk factors, or directly diseases. SHINE's multiple outputs lend themselves to a multi-criteria decision analysis (MCDA) approach[24, 25, 26, 27, 28, 29, 30] that is premised on the fact that individuals and society consider many costs and impacts (i.e. criteria) of an intervention, compared to many other interventions, before deciding which interventions to prioritise. Costs and impacts that SHINE routinely estimates for a given intervention include:

- Health adjusted life years (HALYs) gained – similar to a QALY, other than how morbidity conceptualised and specified
- Life expectancy and health adjusted life expectancy (HALE), either cohort or period, and the gap between the two (for assessing if interventions compress or expand morbidity)
- Deaths averted, either all ages or premature (usually <75 years)
- Changes in all-cause and cause-specific mortality rates
- Changes in all-cause and cause-specific morbidity rates
- Health inequality impacts e.g. differences between social groups in HALY gains per capita or per person year
- Changes in future health system expenditure – due to changing future disease prevalence, deaths, and severity
- Changes in future income earnings of the population due to changed disease rates
- Cost per HALY gained, usually from a health system perspective (i.e. net cost of intervention cost and changes in future health system expenditure, per HALY gained, for some time horizon [default 20 years in SHINE])

SHINE models are future-oriented, not cross-sectional, and explicitly model future changes in the population. SHINE can address complex issues for a number of risk factors and diseases by using existing databases, and routinely updating data through data infrastructure systems. SHINE exists primarily to improve population health through prevention (i.e. focusing on interventions that directly change disease incidence, or indirectly change incidence via changes to risk factor exposure). This protocol is therefore focused on modelling preventive interventions that act via risk factor pathways or directly impact disease incidence; however, it is important to note that SHINE infrastructure and methodology can also be utilised to model treatments (i.e., acting on diseases directly).

For interventions that act through changing risk factor distributions, SHINE deploys three levels of analysis:

- **Theoretical minimum**: instantaneously shifting the prevalence or distribution of a risk factor to its theoretical minimum level of risk within a population (in burden of disease terms, this is the theoretical minimum risk exposure level [5]). In SHINE, we also refer to these as **'magic wand'** interventions.
- **Targets**: altering the prevalence or distribution of a risk factor to achieve (over time) targets set by governments or other agencies (e.g. that smoking prevalence in Australia will be 5% by 2030 – noting that there are many ways to get to such a target over time, and through varying prevalence by different sociodemographic groups).
- **'Actual' interventions**: conceptualising and specifying interventions that alter risk factor distributions (e.g. a tobacco mass media campaign or new tobacco tax). This usually leads to a cost-effectiveness analysis (strictly CUA) including the upfront costs of the intervention, and cost-offsets of downstream changes in disease expenditure due to changing disease epidemiology and population demography resulting from the intervention.

SHINE scales in three ways:

- **Scaling up**: modelling packages of interventions targeting a particular risk factor (e.g. the combination of very low nicotine cigarettes, retail outlet reduction for tobacco and a tobacco free generation)
- **Scaling out**: comparisons of intervention impact across countries (e.g. running the same sodium reformulation intervention in multiple countries, in parallel)
- **Scaling down**: comparisons of intervention impact for specific sub-populations, estimating the impact of interventions on reducing health inequity (e.g. estimating the impact of cold housing by socioeconomic position to determine the ability and magnitude of such interventions to reduce social group inequality in health).

4.1.1 SHINE purpose

SHINE aims to both generate knowledge gain, and inform public health policy.

Knowledge gain

Specifically, SHINE aims to increase public health and economic understanding of intervention impacts. Much of epidemiology and health economics exists to determine whether some intervention or exposure has a positive or negative effect against some metric. But this research is rarely compared across interventions. For example, which has the largest impact on health gain and health system expenditure: a 10% reduction

in BMI, tobacco smoking, or air pollution? And how do these impacts vary over time into the future and by population sociodemographic factors? And a further example, which preventive interventions maximally compress morbidity (i.e. the gap between HALE and LE) and increase workforce income productivity? The latter example in particular is an essential contemporary public policy question, given ageing populations. The answer to these questions is either poorly or not at all understood by health experts and researchers. That is, SHINE (and other public health intervention modelling) has a major role to play in knowledge gain.

Inform policy

In parallel with knowledge gain is SHINE's intended role to inform public health policy. SHINE's approach to influencing policy is based on Kingdon's Multi-Stream Framework (MSF) of policy.[31] This framework posits that policy change happens when three things come together: agreement on the problem; agreement on the policy options (and preferably agreement on a preferred policy option); and political will and attention to enact a policy change.[31]

In this space, SHINE is largely a 'policy entrepreneur', assessing multiple policy response or interventions so that when a window of opportunity open for policy making there are pre-evaluated interventions (or the tools to rapidly inform which policy response option might be best or optimal). To be a policy entrepreneur, all three streams need to be brought together or "coupled".[32] SHINE mainly does this by: 1. Quantifying a problem (i.e., the burden due to a risk factor in a specific context), importantly in the context of other problems, and 2. Identifying a policy or strategies that could alleviate the problem, considering factors that are important to the population (e.g., a gain in health) and specifically to policymakers (e.g., healthcare expenditure). When 3. political opportunities align (largely beyond our control), then SHINE can contribute to political decision-making. SHINE can also shine a light on problems and seek to influence political players.

The SHINE work plan prioritises key risk factors for poor health, providing evidence to support decision making around specific risk factors and priority setting in the context of limited health funding (i.e. evidence to support allocative efficiency in decision making). To be valid at this allocative level requires a major emphasis on comparability, by which we mean the ability to confidently compare results from different evaluations because they have included the same (or similar) data inputs (e.g. disease rates all from the GBD), they have used the same methodology (e.g. unrelated disease expenditure costing; same time horizon), and they have used the same or similar model structure. Other teams, such as ACE-Prevention and BODE^3 have conducted analyses of multiple risk factors. These have been largely limited to Australia and Aotearoa New Zealand. SHINE builds on this work and aims to fills a gap in evidence by providing comparable analyses of the future impacts of interventions targeting many risk factors and diseases in an open cohort format. SHINE's contribution to comparability and overall value is furthered by the focus on scaling out to other countries and scaling down to provide sub-strata comparisons.

4.2 SHINE model structure

A generic SHINE model contains some key components, and follows a general structure, as shown in Figure 4.1.

The model consists of (usually, unless modelling interventions that act directly on a disease(s)) a risk factor specific model, and a PMSLT (the main structural component of SHINE models). The PMSLT is used to tally up the health and economic impacts on a population that result from changes to the distribution/prevalence of a risk factor, or direct changes to a disease, in a population. This occurs under an 'intervention' scenario,

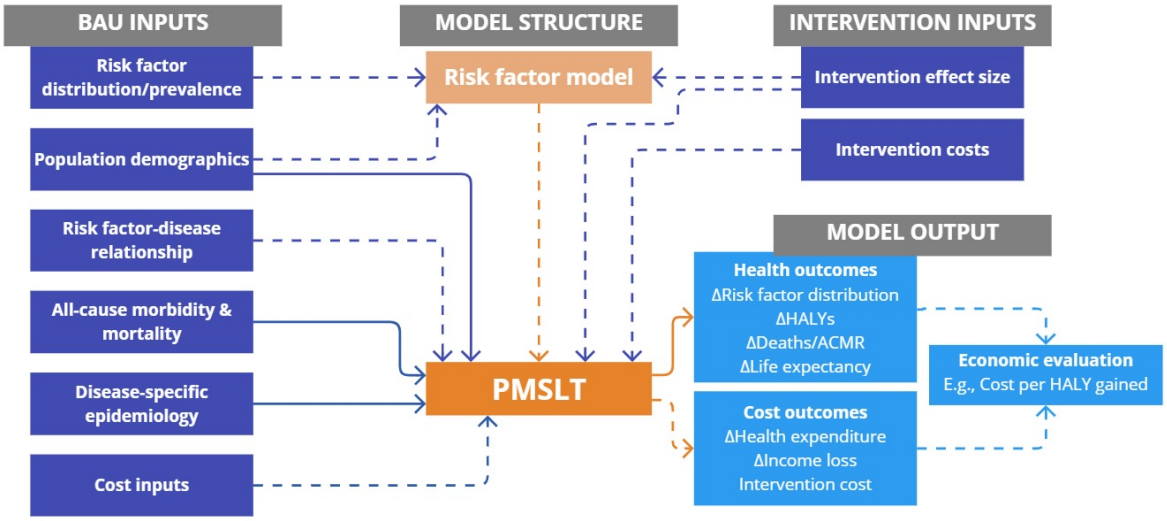

Figure 4.1: Conceptual generic SHINE model structure. Solid arrows reflect required relationships, dashed arrows reflect optional factors.

which is built on top of the BAU scenario(s) to test the impact of interventions on the population.

If modelling a risk factor, e.g., salt intake or smoking, an additional model may be used to estimate trends in the risk factor for a population over time (e.g., a simple regression model to estimate the distribution of salt intake over time), or possibly to model more complex behaviour changes and intervention impacts (e.g., using a Markov model to estimate smoking uptake and cessation trends).

4.3 SHINE modelling protocol

While this protocol is specific to the SHINE work program and modelling infrastructure, we also aim to fill a gap in existing protocols/guidelines pertaining to simulation modelling in health. This includes:

- Presenting a lifetable approach that appropriately and comprehensively accounts for competing risks when modelling health.
- Presenting a burden of disease metric, HALY, that combines components of QALY and DALY metrics to be most useful for scalable prospective disease modelling.
- Presenting an approach to forecasting coherent disease rates for simulation modelling.
- Providing guidance on the appropriate adjustment of intervention effect size estimates from existing literature, including adjustment of the central estimate and calculation of uncertainty.
- Providing guidance on non-traditional sources of evidence when modelling interventions that lack evidence to be drawn on for effect size estimation.

Part II

SHINE BUSINESS-AS-USUAL MODEL

Chapter 5

What is business-as-usual?

5.1 Introduction to BAU

The first key step when conceptualising a model is to define business-as-usual (BAU). In epidemiology and health economics, this is often referred to as the comparator scenario or status quo. In the context of clinical trials, the comparison group is generally those who are allocated to receive no treatment, or standard care. This can be more difficult to define for public health simulation modelling, where we are trying to simulate reality as opposed to a controlled clinical setting. This is particularly the case when specifying risk factors trends. For example, what is the future smoking prevalence under BAU? Is it that if we stopped all tobacco control measures and did nothing more? Or is it if we applied the same intensity of policies in the future as in the past (e.g., continuing with regular taxation increases and updates to health warning labels)?

The default position for SHINE is to define a *statistical BAU*, whereby trends seen in the recent past (e.g. last 10 to 30 years, depending on data availability) in risk factor, disease, and all-cause mortality rates are applied into the future. This approach involves both SHINE's own forecasting methods, in addition to sourcing existing projections for a given setting.

Importantly, the statistical BAU approach assumes that the force of change over time that has resulted from policies applied in the past and present will continue into the future.

When conceptualising BAU, it still needs to be considered what previous/existing interventions against the risk factor/disease of interest are in place in the given setting, and whether these are likely to continue at the same intensity into the future. This process can influence the interventions that can be modelled on top of BAU. For example, a tobacco taxation intervention to be modelled in the Australian setting should only be at a rate over and above the currently rate of excise tax increase that has occurred over the last 10-15 years in this setting. This is because the future BAU trend in smoking rates would be partially accounting for the effects of taxation, based on historic taxation policies.

Processes for BAU inputs are discussed in detail in Chapters 6 - 7. The relative impact of modelled interventions depends on how the BAU in parameterised.

5.2 Is there one true BAU?

Generally, SHINE uses a single BAU scenario as the comparator for different interventions. The SHINE BAU model holds multiple assumptions about future trends (or lack thereof) for a number of input parameters. This is addressed in all SHINE models by:

1. Parameterising all BAU inputs (except demographic factors, and occasionally risk factor trends e.g., for the SHINE Tobacco model) with uncertainty.
2. Projecting forward all time-varying trends for 20-30 years from the model base year (using an annual percentage change on base year values) then holding constant in subsequent years.

However, it is becoming increasingly clear that it is not always suitable to assume a single BAU, even with the above mechanisms in place. Therefore, at times SHINE utilises a third option:

3. Explicitly specify alternative BAU scenarios.

These scenarios would generally differ in the future risk factor trend and is used when a risk factor trend is unclear or minimally supported by data - e.g., an emerging risk factor such as e-cigarettes. This is particularly important if there is a wide range of plausible future trends – incremental differences in health impacts from an intervention compared to BAU can differ greatly if there is a plausible range of BAU scenarios. If this is the case, different scenarios should be presented clearly for end-users.

The following sections focus on the standard BAU parameterisation used by SHINE.

5.3 Components of a BAU scenario

The key components that formulate a standard BAU for any SHINE model are as follows:

- **Population demographics**: population forecasts are based on existing GBD population forecasts by sex and age. These forecast changes are the net of births, mortality and migration. We determine how much population size (by sex and age) changes into the future due to all-cause mortality rate forecasts (produced by SHINE), and the residual is the treated as the combined birth and net migration counts that we input in our open cohort simulation model. If using a closed cohort model approach, this birth and migration component can be 'turned off' so that only the base year population is counted.
- **All-cause epidemiological rates**: this includes all-cause mortality rates forecast into the future, and static all-cause morbidity rates. Mortality rates are based on GBD data from 1990-present, applying SHINE's own forecasting method. Morbidity rates are based on most recent GBD data; these are currently not forecast as all-cause morbidity rates by sex and age have tended to vary little over time in countries like Australia[1].

[1]Future versions of the SHINE PMSLT are likely to include forecast trends in all-cause morbidity rates, given there is some variation by sex and age – and especially between countries – in historic GBD data. This will make little difference to most SHINE analyses, except for interventions assessed by compression of morbidity (i.e. LE – HALE gaps)

- **Disease rates**: future disease (categorised as acute or chronic) trends are defined by incidence, case fatality, and remission rates acting on a base year prevalence. Static disability rates are also set for each disease. Like all-cause mortality, disease rates are based on historic GBD data (1990-present) - SHINE forecasts these factors then uses a calibration model to ensure that rates are coherent [33]. This is discussed further in Section 6.4.
- **Risk factor trends**: this involves determining the future distribution of a risk factor in the population or forecasting prevalence trends. For most risk factors, SHINE has developed a method that utilises GBD summary exposure value (SEV) forecasts as a key input. This approach, and alternative methods employed by SHINE, are described in Chapter 6.
- **Costs**: this includes heath system expenditure estimates by disease (a healthcare perspective), and income loss by disease (forming a partial societal perspective when added to the health system expenditure estimates). The SHINE default is to assume health expenditure and income rates are constant into the future – other than that due to inflation, which is handled by all monetary values being real values for the base-year of the model. SHINE has disease expenditure estimates for all OECD countries [2] and New Zealand-based estimates of income loss with disease that are scaled to Australia (based on [3]) – but not to other countries due to concerns about contextual variation. If robust health system disease expenditure and income decrements by disease are produced for other countries, they can be added to SHINE in future iterations.

5.3.1 Annual percentage changes

For epidemiological inputs that are forecast in a SHINE BAU model, we apply an annual percentage change (APC) on a base year value to produce trends over time, as opposed to defining yearly values for a given input. The APC acts for 20 years, following which there is a gradual reduction in the APC from 21 to 40 years down to no change in rates from year 41 onward (i.e. constant rates from year 41).

APCs are used to 1) reduce the number of inputs required for a model (for a given input for a given sex by age group only two values are required, being the initial rate and the APC, rather than requiring 20 values to represent each year), and 2) to produce smooth trends over time by applying a constant APC.

Initial rates and APCs are calculated within forecasting methodology, described in Section 6.4.2.

5.4 BAU infrastructure

SHINE uses an automated process to produce a number of BAU inputs for the PMSLT. The purpose of this infrastructure is two-fold: firstly, *formatting* the large number of inputs required for the PMSLT, and secondly, *processing* raw data in a standardised and efficient manner. We define processing as any 'pre-PMSLT' calculations, including regression and calibration.

The outputs of this automated process, which become inputs to the PMSLT, include:

1. Starting population: this is simply GBD country-specific population data by sex and age.
2. All-cause mortality rate forecasts: as a starting value and APC, produced from log-linear regression model forecasts of historic GBD data.

3. All-cause morbidity rates: static value based on most recent GBD data [2]
4. Acute disease values, comprised of:
    - Incidence and case fatality rates, as starting rates and APCs, produced from log-linear regression model forecasts of historic GBD data.
    - Static disability weight, derived from disease-specific YLD values from GBD.
5. Chronic disease values, comprised of:
    - Starting prevalence, produced from SHINE's 'chronic disease generator' (see Section 6.4.2).
    - Starting incidence, case fatality and remission rates, with APCs, produced from SHINE's 'chronic disease generator'
    - Static disability weight, derived from disease-specific YLD values from GBD.

Each of the above inputs is described in the following chapters.

5.5 Heterogeneity

5.5.1 Current modelling of heterogeneity with SHINE

Heterogeneity is a source of variation in epidemiological and economic variables, by sociodemographic, clinical, or other population characteristics (see Section 3.3).

The minimum level of heterogeneity in the SHINE BAU model is sex by age strata. For analyses in the Australian context, further heterogeneity by socioeconomic status (SES) is also possible with data available from the Australian Institute of Health and Welfare (AIHW) to disaggregate disease and all-cause morbidity and mortality rates across Socio-Economic Index For Areas (SEIFA) quintiles, within sex by age groups. If we want to understand the impact of interventions by SES, it is also important to consider possible differences in intervention effect sizes by SES.

It is also possible to incorporate other sources of heterogeneity, such as ethnicity, or geographic remoteness. The SHINE default when breaking sex by age strata into further heterogeneity is to disaggregate all demographic and epidemiological inputs, including risk factor distributions, disease incidence rates (usually remission and case fatality rates are assumed non-varying), and all-cause mortality and morbidity. However, we assume homogeneity in disease expenditure and income by heterogeneity strata (i.e. per annum disease costs assumed the same regardless of SES; i.e. average income and disease-specific income loss assumed constant by SES). Though this is not likely correct, particularly for income, there is a major risk to equity if income is modelled to vary by SES - interventions that prevent disease among higher income groups would be seen as more economically favourable as the gain in income would be higher than that for lower income groups.

When incorporating heterogeneity into model inputs, SHINE's default is to assume no time trends in rate ratios between additional strata. For example, if the incidence rate of ischaemic heart disease (IHD) has rate ratios of 1, 1.25, 1.5 , 1.75 and 2.0 from SES5 to SES1 quintiles – we hold these rate ratios constant into the future. But the 'envelope' trends in sex by age rates are retained, meaning IHD incidence rates will change over time in each strata.

[2] Note that in future versions of SHINE, regression smoothing and forecasting will be applied to produce all-cause morbidity rates, as with all-cause mortality

5.5.2 Future directions

Modelling heterogeneity is a demanding task, requiring that most model inputs vary by the additional strata. As a result, SHINE is currently limited to modelling heterogeneity by SES in Australia only. However, there are several future directions for incorporating additional heterogeneity into SHINE modelling:

1. **Heterogeneity by SES in other countries**. This will require collaboration with in-country experts to define the heterogeneity and source appropriate data.
2. **Other sources of heterogeneity and multi-heterogeneity models**. Addressing intervention impacts on other sources of health inequality than SES, such as that seen by ethnicity (e.g., Aboriginal and Torres Strait Islanders and non-Indigenous people in the Australian setting, or Māori and non-Māori in Aotearoa/New Zealand) is an important focus of SHINE modelling. The SHINE model has capacity to incorporate any source of heterogeneity for which there is data available across required input parameters - this will be done for specific projects where there is capacity and appropriate collaboration with relevant experts. While this is not the default, SHINE also has the capacity to model multiple sources of demographic heterogeneity simultaneously (while there are caveats to this). An example of such work is Howe et al. [34].
3. **Clinical heterogeneity/risk stratification**. This may be done in future for specific diseases where it is important to capture differences, for example genetic risk scores.

Chapter 6

Demographic & epidemiological inputs

6.1 Demographic components

SHINE models are, by default, run as open cohorts. They are initiated with a start year population (usually derived from GBD data), which can then be altered in each subsequent year with three components:

- Deaths
- Births
- Net migration (considering only international migration for default SHINE models)

SHINE uses two different options for demographic inputs:

Option 1 (default for sex by age models in any country):

- All-cause mortality rates forecast using historic GBD data (discussed below, Section 6.2.1)
- Future population counts are estimated using existing GBD population forecasts. With this approach, migration and births are not separated out individually - the population forecasts are adjusted for the already forecast mortality rates, leaving a combined births + migration estimate.

Option 2 (can be used for models with heterogeneity beyond sex and age):

- All-cause mortality rates forecast using historic GBD data (discussed below, Section 6.2.1)
- Yearly migration and birth counts are estimated with country-specific data sources, with data disaggregated by the stratum of interest.
  * For example, Australian analyses by SES involve first obtaining sex by age migration projections and birth counts from ABS population projections (based on Centre for Population forecasts). Migration counts are then disaggregated by sex, age, and SES using the most recent available historic data on net migration from the ABS (existing reports, or directly from Census data via the Australian Bureau of Statistics (ABS) TableBuilder platform), assuming no difference in the distribution by year. ABS census data is used to disaggregate birth counts by sex and SES.
  * For both components, values should be held constant after 20 years (to align with mortality rate and other epidemiological projections).

Note that the birth and migration components can be excluded to produce a closed cohort model. This may be done to estimate the impact of an intervention over the full remaining lifetime of a population, or to estimate changes in life expectancy expected under an intervention (discussed further in Chapter 15).

6.2 All-cause epidemiology

6.2.1 All-cause mortality

Country-specific all-cause mortality rate data, by sex and age, is sourced from the GBD (1990-present), and forecast using log-linear regression. These values can also be disaggregated by strata - e.g., by SEIFA strata for Australian analyses (see additional details in Appendix B).

6.2.2 All-cause morbidity

SHINE incorporates an all-cause morbidity rate, $morb$, which is the average morbidity experienced in a population on a scale from 0-1, across all causes. The morbidity rate is derived directly from the most recent GBD years lived with disability (YLDs) data, as

$$morb_s = YLD_s/N_s. \tag{6.1}$$

where s is a given sex by age stratum, and N is the population size of that stratum.

This morbidity rate is fixed over time under the BAU scenario [1]. Changes to population-level morbidity under an intervention scenario can be computed through changes in disease-level disability (see Section 6.4.5).

6.3 Risk factors

Future distribution or prevalence of a risk factor by sex and age (and possibly other variables) can be estimated for SHINE modelling in different ways:

Default: future risk factor distribution by sex and age derived from summary exposure value (SEV) forecasts from the GBD. These forecasts can then be disgagregated by SES (or other variable) sub-groups.

Other: for some specific risk factors or analyses, the risk factor distribution forecast with a regression model based on historic risk factor distribution data from a population health survey (e.g. tobacco smoking: forecasts are used to calibrate smoking uptake and quit rates in a Markov model); or alternatively a risk factor distribution may be held constant based on most recent population health survey data.

[1] Noting that in future iterations of SHINE, all-cause morbidity will be forecast with logistic regression.

6.3.1 Summary exposure values (SEV)

What are SEVs?

The Global Burden of Disease (GBD) forecasts risk factors using SEVs – these were first introduced in the 2015 GBD analysis [35], and aim to provide a standardised comparison of the relative burden of disease due to each risk factor. A SEV is a number between zero and one, where a SEV of 0 reflects a population all at or below the TMREL, and a value of 1 is where 100% of the population is at the "maximum" exposure level of the risk factor. Using (currently) 2019 GBD data, the SEV is calculated for each risk factor-outcome pair, rd, and sex by age strata, s, (Equation 6.2), then averaged to produce a single SEV for a given risk factor (Equation 6.3) as follows [36, 37]:

$$SEV_{rds} = \frac{\frac{PAF_{rds}}{1-PAF_{rds}}}{RR_{max} - 1} \tag{6.2}$$

$$SEV_{rs} = \frac{\sum_d SEV_{rds}}{N_{ds}} \tag{6.3}$$

where N_{ds} is the number of outcomes linked to the risk factor r.

The population attributable fraction, PAF_{rds}, is calculated for a risk factor-outcome pair, rd, then SEV_{rds} is calculated for the maximum relative risk level, RR_{max}. For categorical exposures, this is the relative risk at the highest exposure level observed globally, and for continuous variables it is the 99^{th} percentile of exposure.

This process allows for different risk factors with differing distributions to be compared within the same model.

SEVs & risk factors in SHINE modelling

For Scalable Health Intervention Evaluation (SHINE), a method has been developed that takes existing GBD SEV forecasts for a given risk factor and determines a parametric distribution of the risk factor for each sex by age by year by country group. This process is described in detail in Dhungel et al. [38].

For any risk factor with trends forecast with the GBD-derived SEV approach, the SHINE default is as follows:

1. Take the most recent year of published SEV values, currently including forecasts out to 2050. This is treated as the 'working truth'.
2. Use logistic regression to calculate an annual percentage change in the forecast SEV values.
3. Apply a linear reduction in the calculated APC each year to achieve an APC of 0% in 2050.
4. Convert the SHINE-adjusted SEV forecasts to parametric distributions as per Dhungel et al. [38].
5. (Optional - currently models in Australia only) Apply relative differences by SES, using SEIFA, to the risk factor distribution. Note that differences by SEIFA are assumed to be time invariant, unless there is sufficient high-quality data data on risk factor trends by SEIFA over time (by age and sex, or adjusted for these factors) to suggest otherwise.

This approach for forecasting differs to the approach taken for other epidemiological inputs in SHINE models (as described in previous sections of this chapter). Rather than applying APCs for 20-30 years then holding rates constant, we apply the 'detuning' step to the calculated APC (step 3 above), decreasing the APC from its actual value in equal annual increments until there is no APC by 2050. This rule is based on concerns about the plausibility of GBD SEV forecasts within country, by sex by age groups, appearing unstable and sometimes implausible or extreme in out-years; if our concerns regarding this are alleviated in future GBD SEV forecasts, we may change this 'rule' to be more similar to the approach taken for disease rates.

Once a risk factor distribution has been obtained via the SEV method, the forecasted distribution is converted into categories or 'bins' of exposure. Disease risk under an intervention scenario can then be calculated using the 'RR shift' method, which keeps the proportion of the population in each category of exposure the same, but alters the relative risk (RR) of each category and for each disease to match the category's new post-intervention mean.[39]

6.3.2 Complex risk factor models

Some risk factors include a multi-component structure. A key SHINE example is tobacco smoking. Forecasts of tobacco smoking prevalence are not input directly into the PMSLT. Instead, they are used to calibrate a Markov model, which is defined by smoking behaviour states and annual transitions between states. The Markov model output then goes into the PMSLT as an input.

6.3.3 Risk factor mediation

Mediation is the process by which the impact of an exposure on an outcome occurs either partially or completely via another risk factor that lies on the same pathway.

When the focus is 'just' on modelling interventions through a single risk factor, mediation does not need to be included in SHINE modelling, as GBD incidence rate ratios used to estimate the impact of a risk factor onto disease are the total effects of the risk factor (i.e., alreday capture mediated pathways).

However, when multiple risk factors are considered together, mediation factors need to be used. The inclusion and modelling of mediation pathways will vary depending on the nature of the intervention. Two example scenarios are provided below:

- **Modelling the effect of one intervention onto multiple risk factors**: RCTs of Tirzepatide (a GLP-1 agonist) have quantified not only the impact of Tirzepatide onto BMI, but also on to systolic blood pressure (SBP) and low-density lipoprotein (LDL), aamong other things. Modelling the resulting effect of BMI onto ischaemic heart disease (IHD) using BMI-IHD incidence rate ratios will capture some of the impact from changing SBP, but not the full effect of Tirzepatide $\rightarrow$ SBP $\rightarrow$ IHD. Conversely, if we included both BMI and SBP changes in the model, we would be double counting the part of the effect of SBP change that is due to BMI. In this case, the mediation factor for BMI onto SBP is known - this can be used to reduce the BMI $\rightarrow$ IHD incidence rate ratio.
- **Modelling a package of two or more interventions, onto multiple risk factors**: for example, modelling a dietary intervention that impacts body weight and an intervention targeting SBP as a package. This is more complicated to model. Ideally, and especially if the obesity intervention is large in its shift of BMI, we need to model the shift in SBP distribution due to just BMI shift, and then shift the SBP distri-

bution for the SBP intervention. This will only make a difference if the association of SBP with diseases increases more then linearly with SBP – meaning that not shifting the SBP from the BMI effect will lead to overestimating the incremental effect of the SBP treatment.

One additional consideration not included above is if BMI and SBP distributions are correlated – which they are due to the causal association, but actually the SBP among obese people may be even higher than non-obese people due to other (confounding) risk factors. SHINE has recently developed an approach to draw parameter distributions of one risk factor (e.g. SBP) by levels of another risk factor (e.g. BMI), when needed. This will be expanded upon in a later iteration of this protocol.

6.4 Diseases

6.4.1 Disease selection

SHINE's default is to include all diseases linked to a given risk factor unless structural issues exist in our primary data sources. An example of a structural issue is level 3 vs. level 4 disease classification – the GBD provides sufficient data for all level 3 diseases, but for some level 4 diseases, data available is not sufficient for SHINE to produce disease rate forecasts. These diseases are generally excluded (unless the level 3 outcome can be appropriately used instead).

Another example is a lack of risk factor-disease association data, which SHINE usually takes from the GBD (this data would generally only be missing for diseases that have minimal DALYs attributed to the risk factor) - these associations are discussed in Part IV of this protocol. At a minimum, diseases covering 95% total risk-factor attributable DALYs (sourced from IHME Results Tool [40]) should be included.

6.4.2 Chronic diseases

Chronic diseases in SHINE models each consist of a starting prevalence, with incidence, case fatality, and remission rates determining prevalence in each future year. This simple structure is shown in 6.1.

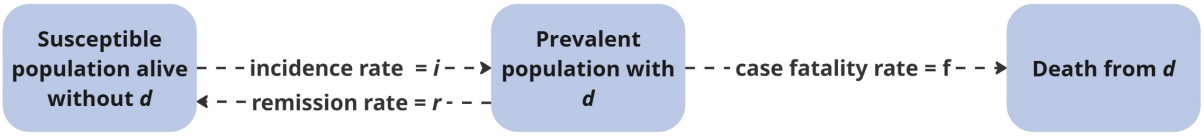

Figure 6.1: Simple chronic disease model.

Chronic disease inputs are parameterised using historic GBD data (1990-present). Two key issues arise when using this (or other) source data:

- Lack of input data on chronic disease rates, particularly remission. Disease prevalence data is generally the easiest to capture, e.g., through a national census or health survey, and is more reliable than available rate data.

- Lack of coherent rates. When forecast incidence, case fatality, and remission rates are combined to produce prevalence values, they do not generally 'match' directly forecast prevalence values - i.e., there is not coherence.

To get around these issue, and generate coherent forecasts of chronic disease rates, SHINE has developed a 'chronic disease generator' to smooth GBD chronic disease rate data, forecast and optimise cohort coherent rates.

The chronic disease generator infrastructure SHINE uses data science and optimisation techniques to generate coherent future disease rates by sex and age in all countries.[33] It has also been set up to incorporate calculation of acute disease rates, and all-cause mortality.

Chronic disease generator process:

1. Collect historic GBD data by country, sex, and age, from 1990-most recent year.
   - Disease Prevalence (P), Incidence rate (I), and Mortality rate (M) raw metrics taken directly from GBD.
   - Case fatality rate calculated as $f = M/P$
2. Solve remission rate, r, based on prevalence, incidence, and case fatality, and smooth all components.
3. Predict future trends with log-linear regression. The regression is used to generate initial estimates of annual percentage changes for each rate, with the base year rate being taken from GBD-derived parameters from Step 1.
   - Prevalence, p, and death counts, m, projected forward as calibration targets
   - Incidence, i, case fatality, f, and remission rates, r, (as base year rates with annual percentage changes) are the initialization parameters to be calibrated
4. Calibrate the disease rates using a stochastic optimization model. This involves iterative adjustments of the base year and annual percentage change values for each of incidence, case fatality and remission rates until they generate the calibration targets (prevalence and death count projections), forming a coherent dataset.

Once the required data is obtained (step 1), the process is fully automated in Python. For each disease rate, and for each sex by age group, a base year value and annual percentage change are produced. This model output acts as inputs for the proportional multi-state life table (PMSLT). The model can be run as a standalone disease rate generator (code available at: `https://github.com/population-interventions/CompartmentalDiseaseGenerator`). A more detailed description of this model is availale from Wilson et al. [33].

The forecasts can be further disaggregated, as described in Appendix B. This is currently only done for models in the Australian context, by SES.

Note that diseases are also characterised by a static disability rate, dw', based on most recent GBD YLD data (see Section 6.4.5).

6.4.3 Acute diseases

Acute diseases in SHINE models are parameterised much more simply than chronic diseases. Acute diseases are considered as instantaneous events, so are comprised of an incidence rate and case fatality rate. A disability weight is also assigned, as discussed in Section 6.4.5).

Acute disease inputs are generated from the Incidence, Deaths and YLDs (Years Lived with Disability) rate metrics reported by the GBD. Log-linear regression is applied to this data to produce base year values for each metric, as well as to calculate the annual percentage change (APC) from the slope of the regression. We denote these values by I_{BASE} and I_{APC} for incidence, M_{BASE} and M_{APC} for death rate, and YLD_{BASE} and YLD_{APC} for YLDs.

The PMSLT requires base rates and APCs for incidence, i_{BASE} and i_{APC}, and case-fatality rate, f_{BASE} and f_{APC}. These parameters are derived from the log-linear regression of the GBD data as follows.

$$\begin{aligned} i_{\text{BASE}} &= I_{\text{BASE}} \\ f_{\text{BASE}} &= M_{\text{BASE}}/I_{\text{BASE}} \\ i_{\text{APC}} &= YLD_{\text{APC}} \\ f_{\text{APC}} &= (M_{\text{APC}}+1)/(i_{\text{APC}}+1)-1 \end{aligned} \tag{6.4}$$

We set i_{APC} to YLD_{APC} because disability weights (see Section 6.4.5) are assumed to be time-invariant, but the GBD does not always completely adhere to this assumption. By using YLD_{APC} rather than I_{APC}, we assume that all changes in YLDs over time are due to changes in incidence, and ensure that the YLDs trend in our model matches that of the GBD. Note that the YLD regression output is not used beyond setting the incidence rate APC for acute diseases.

Some diseases have no incidence rate reported in the GBD. In this case we set the base incidence rate to 1, effectively reusing the case-fatality rate input as a mortality rate.

To derive f_{APC} we require that the projected mortality rate, $M_{\text{BASE}}(M_{\text{APC}}+1)^t$, be the product of the projected incidence and case-fatality rates, $f_{\text{BASE}}(f_{\text{APC}}+1)^t i_{\text{BASE}}(i_{\text{APC}}+1)^t$.

$$\begin{aligned} M_{\text{BASE}}(M_{\text{APC}}+1)^t &= f_{\text{BASE}} i_{\text{BASE}}(M_{\text{APC}}+1)^t \\ &= f_{\text{BASE}} i_{\text{BASE}} \left((i_{\text{APC}}+1)\frac{M_{\text{APC}}+1}{i_{\text{APC}}+1}\right)^t \\ &= f_{\text{BASE}} \left(\left[\frac{M_{\text{APC}}+1}{i_{\text{APC}}+1}-1\right]+1\right)^t i_{\text{BASE}}(i_{\text{APC}}+1)^t \end{aligned} \tag{6.5}$$

It follows that $f_{\text{APC}} = (M_{\text{APC}}+1)/(i_{\text{APC}}+1)-1$ (per equation 6.4).

6.4.4 Limitations to SHINE disease rate estimates

The SHINE chronic disease model generator, and related estimation of acute disease rates, has limitations. Two key limitatations are discussed below.

Issue 1: remission vs. death from other causes
Figure 6.2 shows a more complex and complete chronic disease model than that used by SHINE currently

to calibrate disease rates. This figure (unlike Figure 6.1), captures $mort(d')|d$, the mortality rate for all causes/diseases other than the chronic disease in question, among those with disease d; and $mort(d')|\bar{d}$, the mortality rate for all causes/diseases other than the chronic disease in question, among those *without* disease d. When $mort(d')|d = mort(d')|\bar{d}$, then disease mortality is said to be independent, e.g. the rate of dying of stroke or other diseases for people with ischaemic heart disease (IHD) is no different from the rate of dying of stoke or other diseases for people without prevalent IHD. The SHINE PMSLT assumes disease independence in its modelling. In reality, often $mort(d')|d > mort(d')|\bar{d}$. For example, people with IHD are more likely to die of stroke (and perhaps other diseases) that people without IHD.

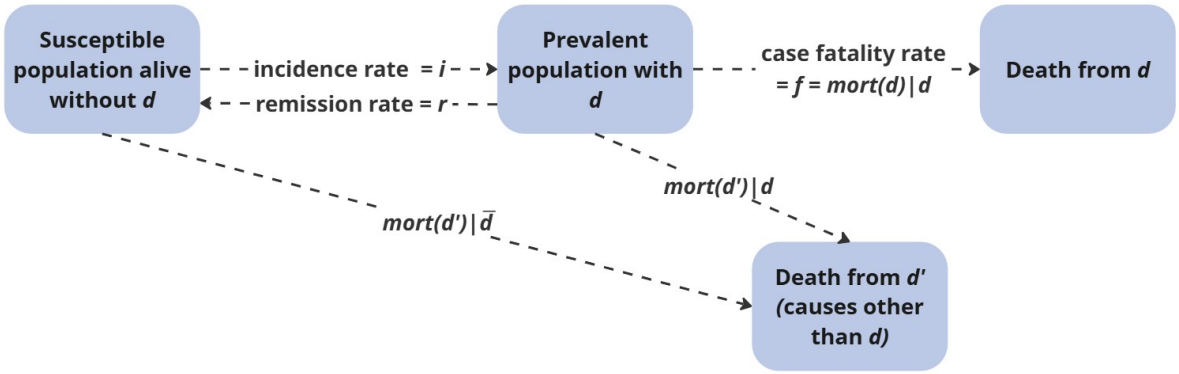

Figure 6.2: Complex chronic disease model.

The main issue with this assumption is in the solved remission rates. If the disease independence assumption is in fact true for a given disease, then the calculated remission rates are likely quite accurate. However, if there is a difference in $mort(d')|d$ and $mort(d')|\bar{d}$, then the remission rates solved for in the chronic disease generator actually reflect remission rates combined with excess mortality rates, e, where

$$\text{excess mortality rate}, e = mort(d')|d - mort(d')|\bar{d} \tag{6.6}$$

Note that the definition of excess mortality in equation 6.6 differs from that used in the GBD.

This is not a large issue currently with the SHINE model focusing on preventive interventions that target disease incidence. For the SHINE PMSLT to robustly model curative interventions that target remission, further work is required to determine for which diseases the assumption of independence holds, and adjust estimates to calculate 'true' remission for the diseases which the assumption does not hold. Future work may also evolve the chronic disease model to directly estimate excess mortality separately of remission, but this is currently beyond the scope of SHINE modelling.

Issue 2: lack of coherence with all-cause mortality rates
Related to the above is the fact that while SHINE disease forecasts are coherent at the specific disease level, there is not internal coherence between the disease projections and all-cause mortality. In other words, the mortality rate for all diseases do not specifically sum to the modelled all-cause mortality rate (ACMR). This approximation is deemed appropriate for the purposes of SHINE modelling, however, this is something to

be aware of. Note that the SHINE PMSLT does account for competing mortality (Chapter 11); this is 'just' not accounted for currently in disease forecast calibration.

6.4.5 Disease disability weights

SHINE models also incorporate morbidity for diseases. This is calculated as a disability rate, based on disease-specific disability weights for each sex by age group, derived from the most recent GBD study.

Disability weights, dw, are disease *sequelae*-specific values ranging from 0-1, where 0 is equivalent to death and 1 is full health.

The weights produced by the GBD (most recently updated by Salomon et al. [41]) were derived by combining two population surveys ([42] [43]). In each survey, paired comparison questions were used, involving providing lay descriptions of two health states to participants who then select which is the healthier state. Each of the health states define a disease at a specific severity level (though some diseases may only have one severity level). Person trade off (PTO) was then used to anchor the results from the paired comparison onto a 0-1 scale, where 0 is equivalent to perfect health, and 1 is death. Prior to being incorporated into YLDs for GBD analyses (and SHINE models), the disability weights are adjusted for a given disease by:

1. Calculating an average disability weights for a disease by averaging across each severity level/sequelae specified for the disease, weighted by the relative prevalence of each level of sequelae
2. Scaling down the disability weight to account for the fact that a proportion of the burden experienced by people with the disease is due to comorbidity (e.g., some of the burden experienced by people with diabetes is actually the result of commonly co-occurring cardiovascular diseases). This is done using a microsimulation model [44], and works by:
    (a) Randomly assigning simulants as having one or more diseases, based on the independent prevalence of each disease.
    (b) Calculating a disability weight for each simulant, multiplicatively across all diseases they have been assigned - this assumes that the disability associated with having two diseases is less than the sum of the disability weights for each disease.
    (c) Attribute a disability weight for each disease, based on the proportional contribution of the independent disability weight to the simulant disability weight. This results in disability weights that are adjusted (down) to account for the fact that disability comes individuals having >1 disease.

This process results in *disease* specific disability weights, denoted dw', which include both co-morbidity adjustment and may vary by sex and age based on different disease sequelae/severity distributions (noting that the prime symbol $'$ is used differentiate between the initial disability weights from Salomon et al. [41], dw, and the co-morbidity adjusted average disability weight for a given disease, by sex and age, dw').

Note that the GBD does not publish these disease-level, adjusted disability weights. They go directly into calculation of disease-level YLDs, for each disease, d, and sex by age by location (country) strata, s, as

$$YLD_{d,s} = dw'_{d,s} \times n_{d,s} \tag{6.7}$$

where n is the prevalent cases of the disease.

SHINE therefore incorporates disease-specific YLDs by first converting them back to comorbidity adjusted disability weights, as

$$dw'_{d,s} = YLD_{d,s}/n_{d,s} \tag{6.8}$$

The disability weight is assumed to be constant over time. This assumption is made given limited data on changes to YLD rates over time. This static morbidity rate approach inherently assumes that the distribution of disease sequelae/sub-levels are also static over time (given that these are incorporated into dw').

The modified disability weight is used as an input to the PMSLT (discussed in Chapter 11). Within the PMSLT, dw' is converted to disability at the population level, or disability rate dr, using disease prevalence (risk) for chronic diseases, or incidence for acute diseases.

Chapter 7

Cost inputs

7.1 Costs overview

SHINE takes a health system plus limited societal perspective for monetary costing. The two types of costing that included within this scope are:

1. **Cost offsets**. These are the differences in future health system expenditure between an intervention scenario and BAU (or another intervention if comparing multiple)
   - These differences can arise due to:
     - Changes in disease rates, which occur either as a direct result of an intervention acting on a disease rate, or indirectly via a risk factor
     - Changes in population size, for example as a result in people living longer under an intervention scenario.
   - There are two approaches to estimating cost offsets: related, and unrelated.
     - Related is, for example, where an intervention address CVD only includes changes in future health system costs for people with CVD.
     - Unrelated extends to include the changes in health system expenditure for people living longer with health system expenditure for longer and possibly other conditions they develop due to living longer [45]. SHINE uses an unrelated approach to cost offsets.
2. **Intervention costs**. This is the cost of both implementing an intervention, and maintaining that intervention for the duration it is in effect. These costs are discussed further in Part III of this protocol - the following chapter focused costs that exist under business-as-usual (and intervention scenarios).

SHINE's default approach is to calculate (health system) cost offsets, and intervention costs, when conducting cost-effectiveness analyses of interventions. If the perspective is expanded beyond the health system, other additional costs may be estimated, including:

- Income productivity. This is the most common additional cost included in a health system + limited societal perspective. Currently, SHINE has the capacity to include income gain as one measure of productivity for analyses in Australian and Aotearoa/New Zealand. This is discussed further in section 7.4
- Cost to carers

- Welfare or benefits costs (e.g., superannuation)

7.1.1 A note regarding CPI and PPP

When calculating cost from available data, it is likely that adjustment will be required to either:

- Convert dollar values to the base year of interest for a model
- Convert currency of source data to that of the country of interest for a model

These adjustments are made using consumer price index (CPI) and purchasing power parity (PPP) values, respectively, taken from the World Bank database. If both year and currency adjustment are needed (e.g. if you have data in 2021 AUD, and need 2025 USD for a simulation model in the US with a start year of 2025), the CPI adjustment (i.e., the within country inflation adjustment) is done first, followed by PPP adjustment to change currency. A multiplier combining these two values is calculated as follows.

To take raw cost data from start year, Y, and country, C, to a target year and country,

$$mult = (\text{CPI}_{\text{START C, TARGET Y}}/\text{CPI}_{\text{START C, START Y}}) \times (\text{PPP}_{\text{TARGET C, TARGET Y}}/\text{PPP}_{\text{START C, TARGET Y}}) \quad (7.1)$$

7.2 Existing costing protocols

Relevant protocols for conducting costing and economic evaluations are presented in Table 7.1. We draw on the knowledge from each of these guidelines to form the protocol for costing in SHINE modelling.

Table 7.1: Costing guidelines relevant to SHINE modelling

Guideline/report	Details	Relevance to SHINE
Drummond et al.[14]	Economic evaluation protocol [textbook].	Provides essential background on economic evaluation types, costing perspectives, and discounting.
WHO-CHOICE guideline.[46]	Economic evaluation guidelines.	Not specific to modelling. Provides guidance on considerations for intervention costing, which can be applied to SHINE modelling. Note that WHO-CHOICE is predominantly used for costing interventions targeting acute/communicable diseases, in low-income settings.

Continued on next page

Guideline/report	Details	Relevance to SHINE
BODE^3 protocol.[21]	Costing in economic modelling guidelines.	Directly relevant guidance to SHINE costing. Provides guidelines for calculating cost-offsets from a healthcare perspective, in addition to detailed guidance on intervention costing (including costs of making laws).
ACE-Prevention protocol.[22]	Cost-effectiveness in prevention protocol.	Provides guidance on presenting and ranking evidence, including cost-offsets or economic evaluation results alongside health outcomes.

7.3 Health system cost offsets

7.3.1 Conceptualisation

SHINE estimates disease-specific healthcare expenditure by sex, age, and disease phase, in Australia. Some diseases may incur greater costs in the final year prior to death, or conversely in the first year of diagnosis, relative to the 'in-between' or prevalent years of living with the disease. SHINE categorises three phases, of:

1. incident (first year from diagnosis)
2. prevalent (subsequent years until remission occurs, or up until the year prior to death occuring from the disease)
3. death (year of death due to the disease)

Importantly, disease-specific healthcare expenditure rates per person are assumed to be static into the future - changes in total health expenditure are driven by changes to population size/structure that occur under BAU, or changes in disease rates that occur under interventions. This implicitly assumes that there is no re-organisation of the healthcare system in the future that changes how expenditure is distributed to the treatment of different diseases or phases of diseases.

Changes to health expenditure under an intervention scenario can occur via two paths, either **direct** or **indirect**:

- Direct: a reduction in the incidence (or other factor, e.g., case fatality rate) of a measured 'Disease A' results in reduced healthcare expenditure on Disease A
- Indirect (or unrelated): a reduction in the incidence (or other factor) of Disease A results in fewer deaths from this cause, and increased life expectancy for those who would have died from Disease A under BAU. As a result, the population structure changes in future years relative to BAU, with an increase in size of the older population. Total healthcare expenditure is greater for older ages due to age-related diseases, so overall healthcare expenditure increases.

The changes in disease-specific expenditure under an intervention scenario compared to BAU are tallied up as a cost offset from the intervention.

The methodology described in the next section can currently be utilised to examine health system expenditure in Australia and other Organisation for Economic Co-operation and Development (OECD) countries (see SHINE publication Grimshaw et al. [2]), with the assumption that per person disease costs estimated for the Australian setting (with NZ data to disaggregate costs by disease phase) are applicable to all other OECD countries. The estimates are not applicable to other settings – a future aim of SHINE is to incorporate country-specific disease expenditure estimates for other settings.

7.3.2 Calculations

Input data
For a given disease, health expenditure is estimated by sex and age group. Health expenditure by disease for the Australian population is taken from the AIHW Health system spending on disease and injury in Australia, 2020-21 report [4]. This report covers expenditure on the following areas:

- Hospitals (private & public)
- General practitioner services
- Other health practitioner's services provided via the Medicare Benefits Scheme (including specialists, allied health)
- Dental services (note: SHINE excludes these costs currently as they are not available broken down by sex and age)
- Medical services (including medical imaging, pathology)
- Pharmaceutical Benefits Scheme medication

Importantly, this costing excludes the following areas:

- Unreferred health services that are not captured through Medicare (i.e., most allied health services, or unreferred specialist services paid for out-of-pocket only)
- Over-the-counter medication
- Healthcare expenditure items that are not assigned to individual diseases: e.g., public and community health, patient transport, administration costs, and research. The proportion of healthcare expenditure that is assigned to diseases in the AIHW report series is approximately 70% of total healthcare expenditure in Australia.

The data from AIHW is manipulated in the following process, before being input into the PMSLT. This approach is adapted from that developed by Grimshaw et al. [2].

1. **Estimate per person disease expenditure**

   GBD disease rate data is combined with the disease expenditure estimates for each sex by age group, to produce per person costs for each disease. Note that the AIHW and GBD use slightly different disease groupings for some outcomes. For example, the AIHW outcome 'back pain and problems' corresponds to two GBD outcomes, 'low back pain' and 'neck pain'. Therefore, prior to combining the AIHW & GBD datasets, costs for disease groups where an AIHW outcome corresponds to >1 GBD outcome are disaggregated based on the relative prevalence of the GBD outcomes.

2. **Disaggregate per person disease expenditure estimates by disease phase**

Disease costs are split into three phases: incident, prevalent, or death year. This is based on a regression analysis from Blakely et al. [1], which estimated estimated age and disease coefficients for the outcome of health system expenditure, in New Zealand. In this study, average expenditure for each disease by phase was compared to total average expenditure for a disease to obtain ratios for incidence, $r_{\text{incidence}}$, prevalence, $r_{\text{prevalence}}$, and death year, r_{death}, phases of a given disease.

These ratios are applied to the average per person cost from Step 1 to generate disease phase costs as follows:

$$\text{scalar}_{\text{d}} = \text{total cost}_{\text{d}} / ((r_{\text{incident,d}} \times \text{incident cases}_{\text{d}}) + (r_{\text{prevalent,d}} \times \text{prevalent cases}_{\text{d}}) + (r_{\text{death,d}} \times \text{death count}_{\text{d}}))$$

$$\text{incident cost}_{\text{d}} = \text{scalar}_{\text{d}} \times r_{\text{incidence, d}} \times \text{incident cases}_{\text{d}} \tag{7.2}$$

$$\text{prevalent cost}_{\text{d}} = \text{scalar}_{\text{d}} \times r_{\text{prevalence, d}} \times \text{prevalent cases}_{\text{d}} \tag{7.3}$$

$$\text{death cost}_{\text{d}} = \text{scalar}_{\text{d}} \times r_{\text{death, d}} \times \text{death count}_{\text{d}} \tag{7.4}$$

Note that not all diseases that SHINE models were included in Blakely et al. [1] - for remaining diseases, equal ratios were assumed for incident, prevalent, and death year disease costs. Note also that for acute diseases, only incident & death year cost values are used.

3. **Estimate per person average health expenditure**

   For the main lifetable in the PMSLT, average per person healthcare expenditure (across all causes) is required. Total health expenditure is estimated as the sum of expenditure across all diseases from [4]. Note that this is not the *actual* total healthcare expenditure in Australia, as the costs covered are 'just' those that the AIHW has assigned to specific diseases (for the purposes of SHINE modelling, this is sufficient, as we are focused on the change in *disease-specific* healthcare expenditure that occurs under intervention scenarios). Per capita estimates are then made using population counts.

4. **Disaggregate per person average healthcare expenditure estimates by life phase**

   Health expenditure at the total/all-cause level is disaggregated into two categories:

   - average health expenditure if in the last year of life
   - average health expenditure if *not* last year of life

   This is calculated as:

$$\text{health expenditure}_{\text{not last}} = \text{health expenditure}_{\text{total}}(qx \times r + (1 - qx)) \tag{7.5}$$

$$\text{health expenditure}_{\text{last}} = \text{health expenditure}_{\text{not last}} \times r \tag{7.6}$$

   where qx is age and sex-specific mortality risk [1], and r is the ratio of health expenditure in last year vs. not last year of life from Blakely et al. [1], for each sex by age group.

5. **(Optional) Estimate expenditure for other OECD countries**

[1] Mortality risk is derived from mortality rate, m, as $qx = 1 - e^{-m}$

To estimate both disease-specific and total health expenditure for other OECD countries, three sub-steps are required:

(a) **Estimate total health system expenditure by country**. Healthcare expenditure from the OECD database is available by country across the following areas: Curative and rehabilitative care, Long-term care (health), Governance and health system and financing administration, Other health care services unknown, Medical goods (non-specified by function), Ancillary services (non-specified by function), and Preventive care. To align total health expenditure (as closely as possible) with the expenditure areas estimated by AIHW, Governance and health system and financing administration, Long-term care (health), and Preventive care are subtracted from the total estimates by country.

(b) **Convert per person disease-specific costs to relevant currency**. PPP values from the World Bank database are used to converted the per capita estimates for each disease & disease phase to each country's currency.

(c) **Scale disease-specific estimates by total expenditure for each country**. The per capita disease cost estimates for each country are multiplied by incident, prevalent, or death year case numbers (using GBD disease rate data) and combined into an estimate of total expenditure. This is compared to the actual total expenditure to produce a scalar. The per capita costs are multiplied by this scalar to produce final estimates for each sex, age, disease, disease phase, and country.

Part IV of this protocol details how these costs are incorporated into the PMSLT, and the calculations under an intervention scenario.

7.4 Income productivity

7.4.1 Conceptualisation

For analyses in Australia and New Zealand, SHINE incorporates estimates of disease-specific income loss, as a way of estimating changes in population productivity. This is the only social cost currently incorporated into SHINE modelling.

The income loss estimates are drawn from an analysis of linked health and tax data in New Zealand [3] - using the same source data and methodology as the analysis used for health expenditure by disease [1]. The average income loss by disease in New Zealand is assumed to also approximate income loss by disease in Australia. SHINE does not apply this approach for other settings currently.

The estimates of income loss do not include sick leave entitlements funded by employers, as this is captured as received income.

Income may be lost due to poor health (morbidity), or from premature death in the working age population. Examples of how disease-related income lost may occur in SHINE models are provided in the below figure. In Figure 7.1a, the BAU scenario involves reduced work capacity due to a period of poor health from a disease. In the intervention scenario, the disease does not occur, resulting in an increase in income earned during that period. The example in Figure 7.1b shows a loss of work due to poor health while of working age, followed by premature death prior to age 65 under BAU. If, under an intervention, the disease did not occur,

income would be gained both during the period of morbidity, and the period of lost work due to premature mortality.

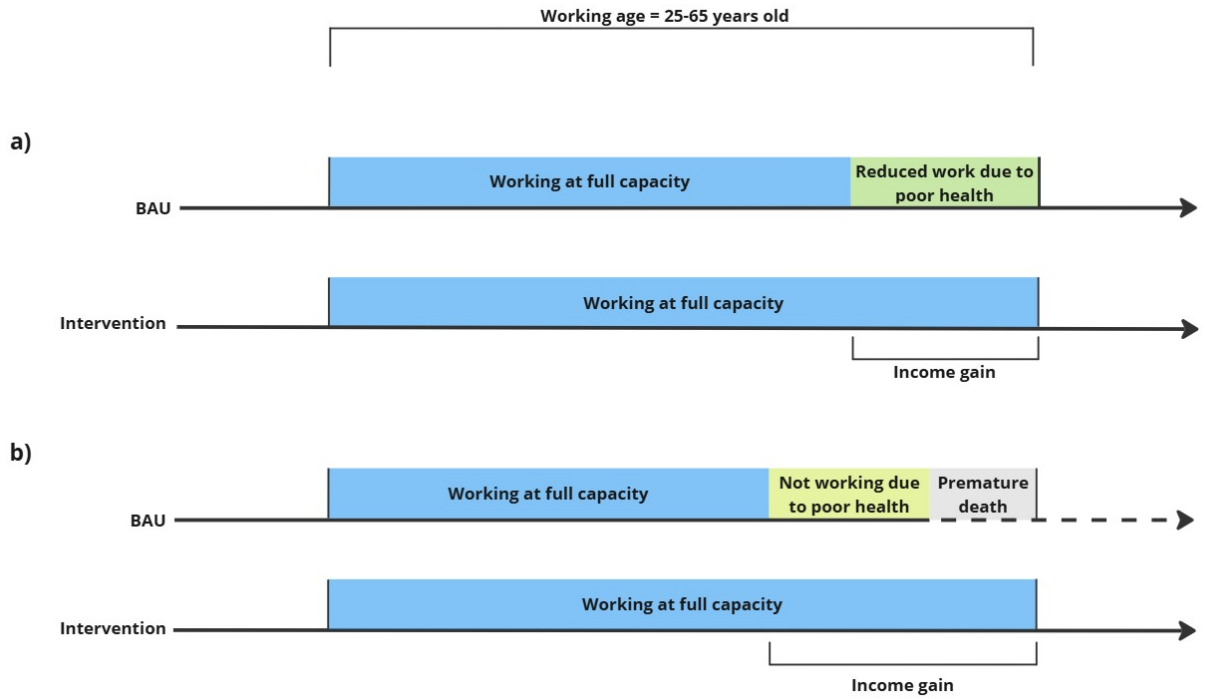

Figure 7.1: Conceptual diagram of income loss calculation in SHINE modelling.

Under BAU, the average income by sex and age is estimated for the population, utilising employee income data [47] spread across the whole population. Disease-related income loss is then tallied under an intervention scenario, and subtracted from the BAU values. This difference in income between intervention scenarios and BAU in SHINE is determined by what proportions of the working age population live longer with less disease (and hence an increase in aggregate income).

The working age is considered as 25-64-year-olds in the model, with no income beyond the age of 65 (per [3]). While this does not cover the full working age population, it is deemed sufficient for default analyses, as most chronic diseases do not result in significant morbidity or mortality for 16-25 year olds, and the 65+ year old population contributes to a minor proportion of income (<3% in Australia in 2023).[47] These assumptions can be relaxed and explored - an example of this is analyses conducted by SHINE that explore income changes if improved population health resulted in citizens working beyond the age of retirement (65 years). This is discussed in Section 7.4.3

The SHINE approach has some similarity to a human capital approach in that income decreases (usually) with disease, but deviates from a human capital approach that simply assumes any death before the age of 65 (or another 'retirement' threshold) loses full income earnings until age 65. Rather, the SHINE approach allows for deferred and competing mortality – and thus estimates income productivity gain proportionate to the increased person years (weighted by diseases present) lived up to age 65 in each future year. This approach is not the same as a friction approach, which sets an assumed period after leaving employment (through death or early retirement) before another person steps into the same job with no ongoing loss of

income productivity.

7.4.2 Calculations

Sex, age, and disease phase-specific income and income loss are calculated in the following process.

1. **Obtain average income data by sex and age**

   Average income is calculated by sex and age category. Importantly, the average for a given sex by age group averages across all labour force strata of employed, unemployed and not in the paid labour force. For analyses in Australia, income data is sourced from the most recent Characteristics of Employment survey, produced by the ABS. Sex by age average income per person is assumed to be static - any changes to total income levels in the future come from changes to population size/structure (and intervention effects).

2. **Disaggregate average income by life phase**

   As with total average health expenditure, income data is split into two categories or phases, of:

   - average income if in the last year of life
   - average income if *not* last year of life

   Total average income is split into the two categories based on all-cause mortality rates in Australia, and ratios of income in last year vs. not last year of life derived from Blakely et al. [3], as

   $$\text{income}_{\text{not last}} = \text{income}/(qx \times r + (1 - qx)) \tag{7.7}$$

   $$\text{income}_{\text{last}} = \text{income}_{\text{not last}} \times r \tag{7.8}$$

   where qx is age and sex-specific mortality risk [2], and r is the ratio of income in last year vs. not last year of life from Blakely et al. [3], for each sex by age group.

   If required, these values are scaled using a CPI ratio to produce dollar values for the base year of a given SHINE model.

3. **Estimate disease-specific income loss**

   Linked data from New Zealand by Blakely et al. [3] is used as the source of disease-specific income loss. This analysis uses the same datasets and approach as the health expenditure by disease estimates from [1]; the difference is that this regression model has income as the outcome. The disease coefficients are used to estimate the following age group and sex-specific estimates, in 2018 NZ dollars:

   - average income loss in first year of given disease
   - average income loss in prevalent years of disease
   - average income loss in last year if dying of that disease

   Note that for acute diseases, only first & last year income loss values are used.

   Rather than using PPP and CPI to scale the 2018 NZ dollar values to Australian dollars in the year of interest for SHINE modelling, age by sex specific ratios are calculated by dividing the two average income values calculated in Step 2 by the corresponding values from Blakely et al. [3]. These ratios are applied to the income loss estimates for each disease as scalars, resulting in estimates of disease-specific income loss in AUD in a relevant base year.

[2] Mortality risk is derived from mortality rate, m, as $qx = 1 - e^{-m}$

Part IV of this protocol details how these costs are incorporated into the PMSLT, and the calculations under an intervention scenario.

7.4.3 Extending productivity estimates

The income loss approach to estimating productivity changes, described above in section 7.4 can be extended on in a number of ways in future iterations of SHINE. This includes:

- Estimates of absenteeism and presenteeism. If valid estimates of average absenteeism and presenteeism can be made by disease or condition, and an intervention impacts absenteeism and presenteeism only through changing disease prevalence, these could be added to the PMSLT – broadening the capture of the 'productivity' concept.
- Sick leave payments. If valid estimates of average sick leave payments by employers by disease status can be sourced, these could be added.

One extension to the income loss estimates that can currently be incorporated involves considering future changes to the age of retirement (and superannuation eligibility) under an intervention. Changes to population health in the future may occur under interventions with significant impacts on a risk factor, particularly if the risk factor is linked to multiple diseases and is a major contributor to overall burden in a population. It is plausible that such an improvement to population health may not only improve workforce productivity in the current working age population, but also increase the average age of retirement. Changing the official age of retirement (or more specifically, age of eligibility for superannuation) in a population would change income received by a population, and consequently government revenue via increased income tax revenue. This concept was considered in a SHINE analysis of tobacco endgame interventions in A/NZ.[48] In this analysis, the age of retirement was increased commensurately with the reduction in population-level morbidity. Specifically, the morbidity rate for 65-year-olds under BAU was identified, and the age that this rate occurred under the intervention was found. The retirement age in the model was then shifted to the age that was equivalent to 65-year-old health status under BAU. The shift was less than 1 year (approximately 7 months), however, it still resulted in large increases in income for the population, and resultingly income tax revenue to the government. However, it did also decrease superannuation payments, which were calculated under the BAU and intervention models – as a larger proportion of the population was alive at post-retirement age under BAU, more superannuation payments were made by the government.

7.5 Other costs

Different sources of government revenue/expenditure, other than health expenditure, may be included in future iterations of SHINE costing. Examples are:

- Income tax: this is linked directly to the income loss estimates. Increases in population income due to improved health would lead to an increase in income tax received by the government, using average tax rates. If health and tax data for Australia (or any other country) can be linked, it would be possible to more exactly specify income tax changes by disease.
- Other sources of taxation: e.g., if looking at the impact of tobacco control interventions where tobacco purchasing behaviour is affected, changes in tobacco excise tax and goods and services tax may be incorporated. This data is sourced from government databases.

- Carer costs: if valid estimates of average carer costs by disease can be sourced, they too could be included for an additional social cost.

SHINE may present the above costs alongside other monetary impacts in an analysis of health and economic intervention effects. However, tax is not generally considered in an economic evaluation as this is a transfer payment.

Part III

SHINE INTERVENTION MODELS

Chapter 8

Conceptualisation of interventions

A note on interventions vs. scenarios

Scalable Health Intervention Evaluation (SHINE) models and compares multiple scenarios. Scenarios include business-as-usual (BAU), theoretical minimum exposure level (a.k.a. 'magic wand'), target exposure level, and actual interventions. The BAU scenario is always modelled, with one or more of the other scenarios modelled 'on top'.

Generally, SHINE models a risk factor under a BAU scenario in comparison to other scenarios in the order of theoretical minimum→ target(s)→ intervention(s). This means that target outcomes can be compared back to the theoretical minimum, indicating how effective a specific target would be against the gold standard or best possible outcome. Actual interventions can then be compared to both the target and theoretical minimum, to see which policy or policy package is 'best' in terms of population-level outcomes, and importantly equality impacts.

In this and the following chapters, we discuss the conceptualisation and parameterisation of *actual* interventions. By 'actual' interventions, we mean an intervention that is feasible to actually roll out in a population, and is likely to include graduated rollout and uptake, possibly attrition of participation or attenuation of effect, and likely to involve costing of the intervention – allowing a cost-effectiveness analysis. Primary prevention examples include: reformulation of sodium in packaged food; tobacco cessation programs; taxing sweetened sugary beverages; encouraging physical activity. Secondary and tertiary prevention is also possible to model.

8.1 Intervention overview

Interventions are input into the model as an 'effect size' that changes the prevalence of a risk factor, or the exposure-disease RR for a given exposure category. These effect sizes may apply to the whole population of interest, or may be different for specific subgroups, depending on the research question and conceptualisation of the intervention. Some interventions will also require their own intervention-specific model (e.g. how changes in sodium change systolic blood pressure, varying by sex, age and starting blood pressure, e.g. how a price change in high-salt foods will impact consumption of both high-salt and all other foods using own- and cross-price elasticities).

SHINE generally sources intervention effect sizes from existing primary literature, obtained via rapid review. Following the standard quality of evidence hierarchy, these reviews aim to obtain inputs from meta-analyses of randomised trials, single randomised trials, meta-analyses of observational studies (e.g., cohort and case control) or single observational studies. However, for some interventions, the traditional hierarchy of evidence may not be applicable. This is more common for preventive interventions than for pharmaceutical interventions targeting an existing disease/behaviour. Preventive interventions that act on a large scale cannot often be tested in a controlled setting, and therefore RCTs may not exist, or if they do, may not be appropriate.

When conducting a review, sources of bias need to be identified and if required adjusted for.

In cases where an effect size for the intervention of interest cannot be sourced directly, parallel or indirect data may be 'cross-walked' to the setting or population of interest (e.g., if no heterogeneity in lung cancer incidence by Indigenous status is available, but findings on heterogeneity by socioeconomic status (SES) exist – then the SES findings may be used as a proxy/to inform for heterogeneity by Indigenous status). Where this is not possible, for example in determining the effect size for an intervention that is yet to be evaluated, techniques to obtain reasonable consensus such as expert knowledge elicitation may be necessary. [49]

Intervention conceptualisation often occurs concurrently to BAU conceptualisation, as the impact point of the intervention can influence the specificity required in BAU.

8.2 Intervention components

Key considerations for a given intervention are detailed in Table 8.1.

Table 8.1: Key intervention considerations for SHINE models

Component	Detail
Target population	The intervention may be acting on a sub-set of the population (e.g., those who smoke). There may also be heterogeneity in the interventions effect (e.g., by age). These factors are important to consider when conducting a rapid review, and considering results of a meta-analysis (or single primary analysis) – are the results transportable to the population of interest?
Impact point in risk factor or disease pathway	If acting on a risk factor: consider whether this is impacting uptake of a behaviour (e.g., smoking uptake) or changing the distribution of a risk factor in the population. If acting directly on a disease: alter incidence, case fatality rate, or severity.
Effect size and uncertainty	[See protocol in Chapter 9].

Continued on next page

Component	Detail
Effect lag-time	Lag-times exist between: a) the intervention and change in it's direct target (e.g., risk factor), and b) the risk-factor and disease impact. Component a) is determined in the intervention specification. Component b) is incorporated into the BAU model (so will carry through to the intervention model).
Intervention duration	Intervention duration is how long the intervention is applied, which may be brief (e.g. advice in GP clinic) through to indefinite (e.g. tax on sweetened sugary beverages). Note that the intervention duration differs from the 'time horizon' over which the impacts of intervention are tallied up.
Attrition and attenuation of effect	Attrition: reduction in the effect of the intervention resulting from a change in participation. Attenuation: reduction in the effect of the intervention due to waning impact of the intervention itself.
Unintended consequences	E.g., for an intervention reducing the salt content of foods, will consumers 'compensate' for possible flavour change by incorporating their own salt to products, or altering intake of other harmful foods? E.g., for a tobacco control intervention mandating that only low-nicotine content cigarettes be available for (legal) purchase, may the impact on quit rates be 'dampened' by some individuals instead sourcing illicit market tobacco with normal nicotine content? These effects are not always possible to include in the model. If this is the case, it should be made clear as a limitation, and consideration for policy makers/future research to address.
Intervention costs	[See costing protocol in Chapter 10]

8.3 Modelling multiple interventions

Interventions may be modelled independently and compared against each other. However, SHINE may also model multiple interventions together as a package. An example of this can be found in Ait Ouakrim et al. [50].

Unless an effect size can be sourced (as described in next chapter, 9) for a package of policies, SHINE's general approach is to combine effect sizes for each intervention component multiplicatively, as

$$\text{Policy Package} = 1 - (1 - INT_A) \times (1 - INT_B) ... \times (1 - INT_N) \tag{8.1}$$

For example, if examining multiple interventions impacting the incidence of the same disease, a potential

impact fraction (PIF) is estimated for each intervention, then the combined impact on disease incidence for the combined interventions A, B, and C would be $1 - (1 - \text{PIF}_\text{A}) \times (1 - \text{PIF}_\text{B}) \times (1 - \text{PIF}_\text{C})$.

Before combining the interventions, one needs to consider interactive effects and where they occur in the causal chain. The 'simplest' scenario is where each intervention acts through separate risk factor pathways, and those risk factors are independently distributed in the population – but jointly impact multiple diseases. In this case, the above type of multiplicative formula using PIFs will work.

However, if the risk factors intervened on cluster within individuals (e.g. people with high systolic blood pressure (SBP) also are more likely to have high LDL cholesterol), and/or the interventions act on shared pathways (e.g. two interventions both reducing SPB; e.g. one intervention reducing BMI and one reducing SBP, which is also impacted by the change in BMI), then additional work to determine the likely combined pathway may be required.

Chapter 9

Sourcing intervention effect sizes

9.1 Intervention sourcing summary

The protocol employed by SHINE to source effect sizes is summarised in Figure 9.1 below. This protocol is used for parameterising interventions, as well as other non-standard inputs where required.

The following steps are defined within the protocol:

1. **Define the research question**
    - This should be done with standard PICOT format: Population, Intervention, Comparator, Outcome, Time. Refer to Table 8.1 for key considerations.
2. **Define the inclusion and exclusion criteria**
    - This is determined based on the research question.
3. **Conduct an initial rapid search, of the appropriate study type**.
    - The traditional hierarchy of evidence [51] defines RCTs as having the highest quality of evidence. However, this hierarchy is not always applicable to public health research and related modelling of preventive interventions. As shown in Figure 9.1, the selected evidence type may be RCTs, but observational studies or alternative forms of evidence may be more appropriate.
    - For more novel interventions that have not been implemented and evaluated, indirect/parallel evidence or consensus approaches are more likely to be required.
4. **Evidence identified: select the most appropriate evidence, and adjust if required**
    - Adjusting evidence to improve internal validity may require quantitative bias analysis, as described in Appendix A.
    - Additional adjustment may be required if appropriately transportable results are not available. This is detailed below.
5. **If traditional/direct literature is not available, look for alternative sources of evidence**. This includes indirect/parallel evidence, or expert consensus, both discussed below.

After the intervention effect size has been determined, it is applied in a SHINE model as follows:

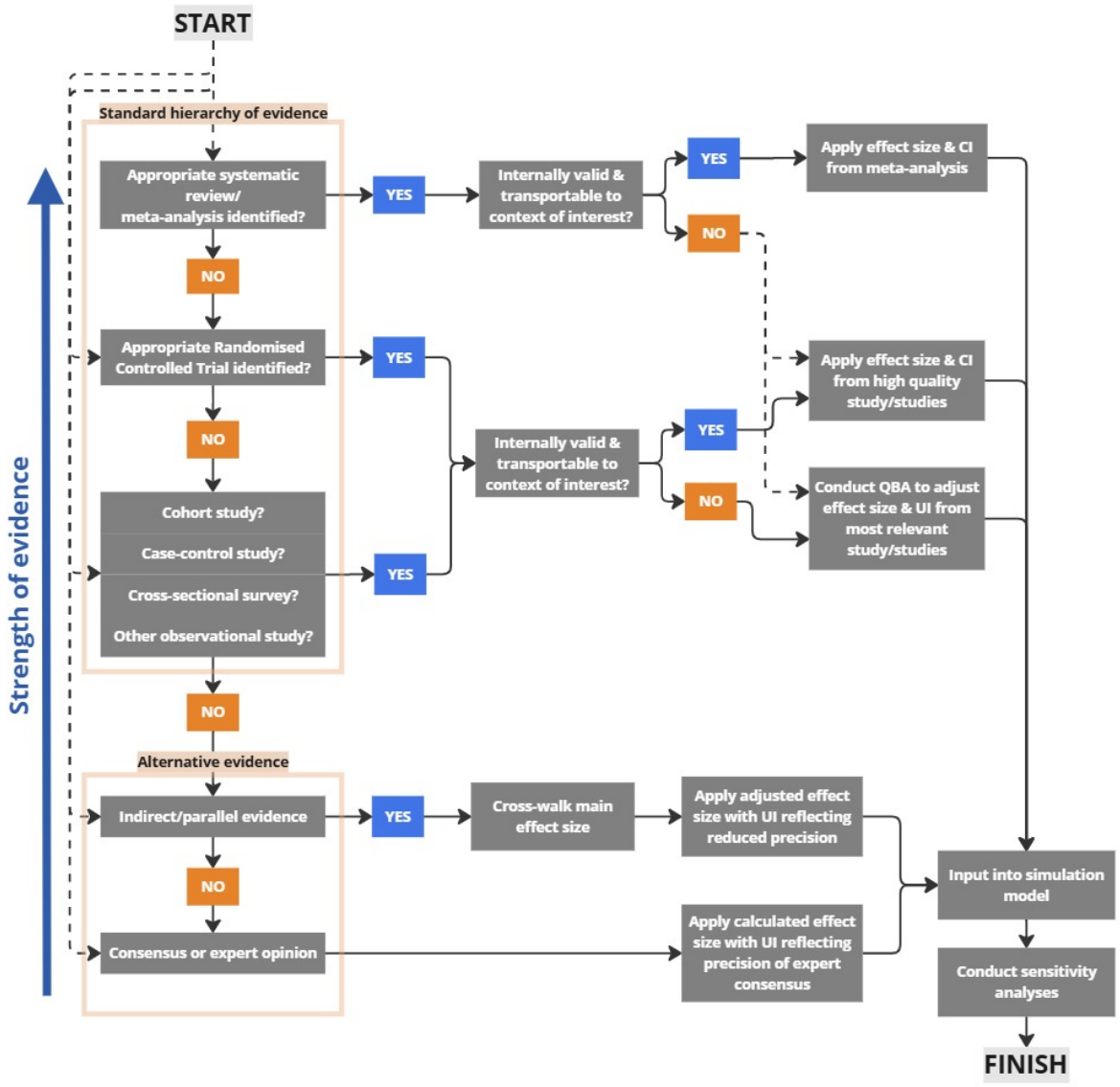

Figure 9.1: Intervention specification process

1. **Run Monte Carlo simulation model** (see Chapter 13) with the intervention parameterised as a main effect size with uncertainty as a probability distribution
2. **Run appropriate sensitivity analyses**. This includes:
    - Univariate sensitivity analysis: Tornado plot (see Chapter 15)
    - Scenario analyses: e.g., if a select number of studies were included from a meta-analysis to parameterise an intervention (due to biases present in remaining studies), run the simulation model with the full meta-analysis results (with all studies included in the effect size & uncertainty interval) and see the impact on modelled output

9.2 Assessing generalisability and transportability

Generalisability refers to the ability of research findings to be applied to the broader population from which the study sample was taken. **Transportability** refers to the ability of research findings to be applied to a partially/completely different setting (time, place, and person) to that from which the study sample was taken.[52]

Transportability is particular important to consider in prospective simulation modelling - the population being modelled is a future population, which is inherently different to a past/present population from which intervention effect sizes are drawn from. The difference between external validity when talking about generalisability vs. transportability are shown in Figure 9.2

A number of factors should be considered when assessing the generalisability and transportability of study findings:

- Internal validity (detailed in Section 3.4):
    - Selection bias
    - Information bias
    - Confounding
- External validity:
    - Consistency: exposure and intervention are specific enough that variation won't affect the outcome (also important for interval validity).
    - Effect measure modification (on the relative or absolute scale). This can be considered in an observational study or RCT by inclusion of sub-strata analyses, and may help transportability for modelling. E.g., if the impact of an intervention varies by age, starting SBP, or other characteristic, and the distribution of these characteristics vary in our population we are modelling, we need the effect sizes by those characteristics to calculate the weighted 'average' to apply to our population – or model our population by strata of these covariates each with their own effect size.
    - Stable unit treatment assumption: this requires that exposure/treatment for one individual is not impacted by exposure/treatment received by other individuals. This is often violated in infectious diseases interventions, but also in some more NCD interventions (e.g. contagion effects of interventions to change alcohol consumption).

Transportability is particular important to consider in relation to SHINE modelling - given that we are modelling the future, the population of interest is inherently (at least partially) separate to an existing target

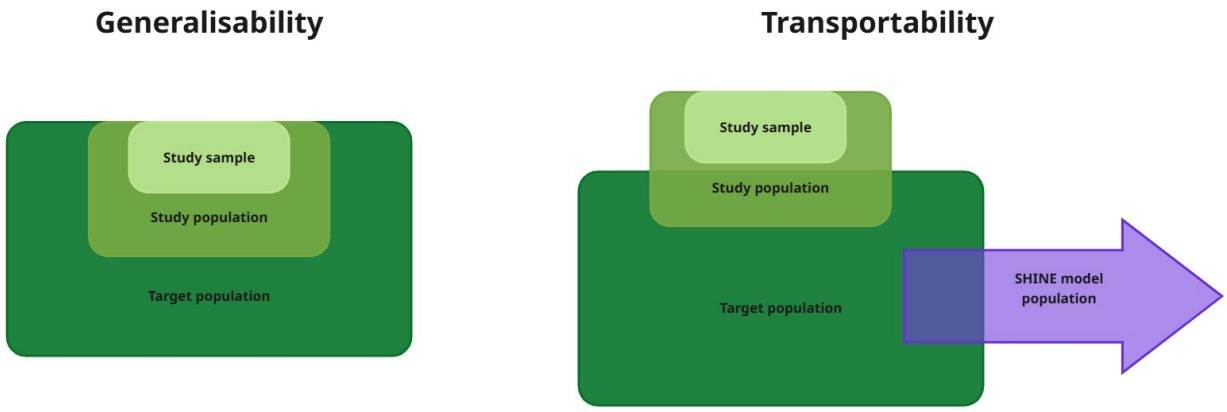

Figure 9.2: The difference between generalisability and transportability, in the context of SHINE simulation modelling

population (Figure 9.2). SHINE considers changes to the future population when parameterising the BAU model, e.g., through forecasts trends in disease rates and all-cause mortality (rather than assuming these are stable at the level currently seen in the population), but seldom considers how actual effect sizes (e.g. incidence rate ratios) might vary into the future.

SHINE's approach to addressing transportability in a study (meta-analysis or individual study) that estimates an intervention effect size is as follows:

1. Define the population: target population, study population, and study sample
2. List the assumptions around people, place and time, which are likely to impact the effect size
3. Identify a plausible shift in the effect size and uncertainty interval based on the identified assumptions.
    - Any shift in the main effect size requires strong theoretical or empirical evidence, or strong and transparent expert judgement. Depending on the situation, the adjusted effect size may be used as the main analysis, with the unadjusted effect size shown as a sensitivity analysis, or the other way round.
    - Widening of uncertainty intervals are discussed in Chapter 13.2.1.

9.3 Alternative sources of evidence

Why might alternative sources of evidence be needed? This is generally the case for novel interventions, for which no analyses exist from which to extract an effect size.

Methods to apply in these situations include:

1. Using indirect or parallel evidence from similar interventions
2. Obtaining input parameter from consultation with experts. A formal expert consultation process that may be used is expert knowledge elicitation.

9.3.1 Indirect and parallel evidence

The use of parallel and indirect evidence, when higher quality evidence is not available, is based on definitions by Vos et al. [22], as part of the Assessing Cost-Effectiveness in Prevention (ACE-Prevention) study, and previous related works by Haby et al. [53], and Swinburn et al. [54].

> **Parallel evidence** designates that which examines the intervention/exposure of interest on a related, but different, outcome. For example, using evidence from the effectiveness of a social marketing campaign on tobacco control to estimate the impact of a similar strategy on obesity prevention.[54]
>
> **Indirect evidence** is that which does not specifically look at the exposure-disease relationship of interest but indicates the positive or negative effect of the exposure/intervention.[54] Swinburn et al. use the example of continued advertisement of particular food products being an indirect source of evidence that such marketing campaigns increase the purchasing of these foods.[54]

9.3.2 Expert knowledge elicitation

Expert knowledge elicitation (EKE), also called 'structured expert elicitation' (SEE) , is a scientific method that allows quantification of unknown statistical parameters, using expert judgement.[49] Multiple formal protocols exist for eliciting information in an unbiased manner.[49] [55] The commonly used SHELF protocol [56] provides an online tool for combining elicited judgements (see `https://shelf.sites.sheffield.ac.uk/home`).

This is a time-consuming process and is therefore should only be used in SHINE modelling for model inputs that have a high impact on measured outputs.

Some key components of EKE to be aware of are:

- Collection of judgement:
  - Group judgement: a group of experts collectively discuss and provide judgements (similar to a Delphi process).
  - Individual judgements: individual responses collected anonymously (prevents 'group think' and some strong individuals influencing views of others, but conversely prevents discussion to clarify).
  - Combined approach: sometimes, EKE will consist of multiple rounds to combine both approaches.
- Aggregation of judgements:
  - Mathematical aggregation: used to pool individual responses. With this approach, an additional consideration is whether weighting is used, for example by using a 'seed question' to determine expert reliability in making judgements.
  - Behavioural aggregation: used to achieve consensus from group , on the distribution of the outcome (both width and shape).
- Uncertainty in judgements: regardless of the method chosen for collection or aggregation of judgements, uncertainty should be incorporated into judgements.
  - A distribution may be determined a priori, or responses may be fit to a distribution after being collected. the SHELF protocol has an online tool that can be used to fit a distribution for individual EKE responses (see `https://shelf.sites.sheffield.ac.uk/home`).

- It is also important to consider whether there are correlations, if multiple judgements are being collected - does an experts judgement about one quantity affect their subsequent judgement about other quantities? [49]

Chapter 10

Intervention costing

10.1 Introduction to intervention costs

10.1.1 Cost types

- Initial vs. ongoing costs
  - Initial costs: these are the set-up costs that occur within the first year of intervention.
  - Ongoing costs: this generally relates to yearly running costs, monitoring and evaluation
- Related vs. unrelated costs
  - Related costs: e.g., cost of pharmaceutical intervention; monitoring and evaluation of health promotion program; screening program set-up costs
  - Unrelated costs: for the purposes of SHINE modelling, unrelated costs are those captured in health system expenditure and income changes under an intervention scenario vs. BAU. These are the costs that are not directly resulting from the intervention, but instead are downstream consequences of changes to disease morbidity and mortality [45].
- Fixed vs. variable costs
  - Fixed cost: does not change by population affected, e.g., cost of legislation
  - Variable costs: unit cost that changes based on the number of people affected/targeted, e.g., personnel required to carry out screening program (will vary by the number of people eligible for screening)
- Patient vs. program costs [57]
  - Patient costs: costs associated with providing the service to individual patients.
  - Program costs: administrative costs of a program or treatment, not linked to individuals.

10.1.2 Costing approaches

- Top-down vs. bottom-up
  - Top-down: start with costs of a program, then if needed break this down into components

  - Bottom-up: aggregate costs of different components of a program
- Macro vs. micro-costing
  - Macro: costs for a patient/group of patients considered on a large scale, e.g., average cost per hospital stay, cost of a round of treatment. Macro-costing is generally used for a top-down approach.
  - Micro: individual units of each component of treatment relating to an individual estimated. E.g., considering the individual procedures/services in hospital, including provider time. Micro-costing is generally used for a bottom-up approach.

10.1.3 Costing components

Table 10.1 lists some of the key cost categories when costing different types of public health intervention.

Table 10.1: Intervention cost categories by intervention type

Intervention type	Cost categories [1]
Health promotion campaign/program	Initial costs: • Program development • Training Ongoing costs: • Program run costs • Evaluation & monitoring • Marketing/advertising
Regulation	Initial costs: • Cost of passing legislation Ongoing costs: • Enforcement costs • Evaluation & monitoring
Screening	Initial costs: • Program development • Training Ongoing costs: • Program run costs • Evaluation & monitoring • Marketing/advertising

Continued on next page

Intervention type	Cost categories [1]
Treatment	Initial costs: • Ongoing costs: • Health services • Medication costs (government & patient out-of-pocket cost) • Evaluation & monitoring

[1] This table draws largely on the BODE3 protocol by **?**].

Most interventions modelled by SHINE are preventive, targeting upstream risk factors. Therefore, health promotion & regulation are the more commonly costed intervention types.

10.2 Existing intervention costing protocols

There is not comprehensive costing protocol for preventive health interventions or chronic diseases. Most existing protocols are focused on infectious disease and are most commonly utilised in low and middle-income country settings, for example the WHO-CHOICE program protocol and related OneHealth tool, and the Global Health Cost Consortium protocol.

10.2.1 WHO-CHOICE

WHO-CHOICE uses a bottom-up approach with a health system perspective, estimating the unit cost of resources required for healthcare delivery then multiplying these unit costs by the number of resources used/required.[46] Costs are broken down into program and patient costs.

WHO-CHOICE provides a costing database for program costs. The database includes country-specific, unit-level estimates for inpatient and outpatient costs (per visit, or per hospital day, respectively), for 193 countries.[58] The tool is based on standardised methods, supporting comparison across countries.

While predominantly applied for communicable disease intervention costing, WHO-CHOICE has also been used for costing of non-communicable disease targeting interventions, using the same bottom-up approach. [59]

10.3 SHINE costing approach

10.3.1 Conceptual approach

Economic evaluations in SHINE modelling may be conducted in collaboration with health economics groups. However, SHINE sometimes conducts intervention costing and subsequent economic evaluations 'in-house'.

Note that intervention costing is not conducted for magic wand or target scenario modelling, as these are 'just' hypotheticals, for which health outcomes and cost offsets are considered. Economic evaluation, and specifically cost-utility analysis, comes into play for modelling actual interventions.

For cost offsets (see Chapter 7), SHINE takes a top-down macro-costing approach - aggregate health expenditure costs are estimated for each disease based on existing health system expenditure estimates from AIHW (noting that AIHW estimates these costs using a mix of top-down and bottom-up costing approaches [4]). For *intervention* costing specifically, it may be necessary to use a micro-costing approach, estimating individual component costs for a single patient, or a mix of top-down and bottom up costing. For example, if estimating the cost of a new healthcare service, a top-down approach may be used to allocate the total estimated cost of the service to individuals, and a bottom-up approach to estimate medication costs and other patient expenditure.

10.3.2 Method

1. List all components to be costed for an intervention. This includes initial costs, and ongoing costs. The cost categories to be included generally depends on the type of intervention - different costs will be incurred for a treatment (e.g., a new pharmaceutical) vs. a screening program vs. a health promotion campaign. The categories are further detailed in Table 10.1.
2. Review the literature - the strategy for the literature search largely depends on the intervention at hand, and in particular whether this is an intervention that has been implemented in other settings already.
    - Existing interventions: review the literature for existing economic evaluations; search government databases for information on the existing intervention.
    - Novel interventions: determine whether there are comparable existing interventions that can be used as a proxy.
3. Set uncertainty around included costs and consider sensitivity analyses.

For steps two and three above, the methods detailed in Chapter 9 can be applied.

Example sources are provided in Table 10.2, noting that these are mainly relevant to the Australian context.

Table 10.2: Intervention cost sources

Cost categories	Sources
Overall program costs	Overview of funding allocated to new health programs provided in Federal Budget (https://budget.gov.au/content/02-health.htm). Federal funding provided to state/territory governments to provide services/programs are listed on the Federal Financial Relations system (see: https://federalfinancialrelations.gov.au/)
Media campaigns	Federal Government advertising expenditure (annual reports produced by the Department of Finance, see: https://www.finance.gov.au/publications/reports; note that state/territory governments also produce similar reports.)
New legislation	Cost of passing a new bill estimated in the New Zealand context by Wilson et al. [60]
Medical treatment	Pharmaceutical Benefits Scheme (PBS) listed medication costs to government and patients available on PBS website (https://www.pbs.gov.au/pbs/home Medicare Benefits Scheme (MBS) provider service costs to government available on MBS website (https://www.mbsonline.gov.au/)

Part IV

SHINE SIMULATION OF HEALTH AND ECONOMIC DIFFERENCES BETWEEN INTERVENTION SCENARIOS AND BAU

Chapter 11

PMSLT

11.1 PMSLT overview

A proportional multi-state life table (PMSLT) is used to estimate mortality and morbidity across multiple diseases simultaneously, combining these into estimates of overall mortality and morbidity.[61] A PMSLT consists of individual lifetable for each included disease, and a population-level all-cause lifetable. Importantly, the PMSLT can be used to assess *changes* in overall (and disease-specific) mortality and morbidity as a result of interventions that act on one or multiple diseases. The PMSLT operates as a hybrid Markov model, as it allows for multiple disease lifetables to be evaluated before combining their effects additively, rather than tracking joint states, which a standard Markov model cannot handle without state explosion (i.e., an exponential increase in the number of model states required).[8]

The number and type of disease lifetables included depend on what is being modelled. Most risk factors have multiple linked diseases. The PMSLT draws from key methodology of burden of disease analyses and comparative risk assessments for the lifetable calculations. It has been constructed to support both health and economic inputs being incorporated, and is described in further detail in Blakely et al. [3].

In the main all-cause lifetable, sex by age (by SES, and possible other strata) mortality rates are used to estimate the number of people alive within a cohort in a given year and person-years lived. This is combined with morbidity rates (pYLDs), to calculate health-adjusted life years (HALYs). These calculations are made for each individual year being modelled. Unlike most burden of disease analyses, which use period lifetables, the PMSLT is a cohort lifetable. Mortality rates input in the lifetable are projected forward using the chronic disease generator (detailed in Chapter 6).

The PMSLT runs multiple scenarios in parallel. The business-as-usual (BAU) is a scenario, as are the interventions, and a model may have multiple BAUs. An intervention is modelled by modifying the lifetable relative to a BAU, which is commonly done by modfying disease rates. These diseases may be chronic or acute. The intervention may act on a disease(s) indirectly via a risk factor, which then acts on a BAU rate such as disease incidence via a population impact fraction. Alternatively, the intervention may act directly on a disease, for example a pharmaceutical intervention acting on a disease's case fatality rate. The model sums the differences in disease morbidity and mortality rates between the intervention and BAU scenarios across all affected diseases. These summed differences are then applied to the main lifetable, where they modify all-cause BAU rates. The difference between the all-cause BAU lifetable and the all-cause intervention lifetable are calculated in each year of the model and for each strata (sex by age or birth cohort by SES)

and presented as a number of outputs. The PMSLT structure is shown in Figure 11.1 below, focusing on an intervention acting on a risk factor.

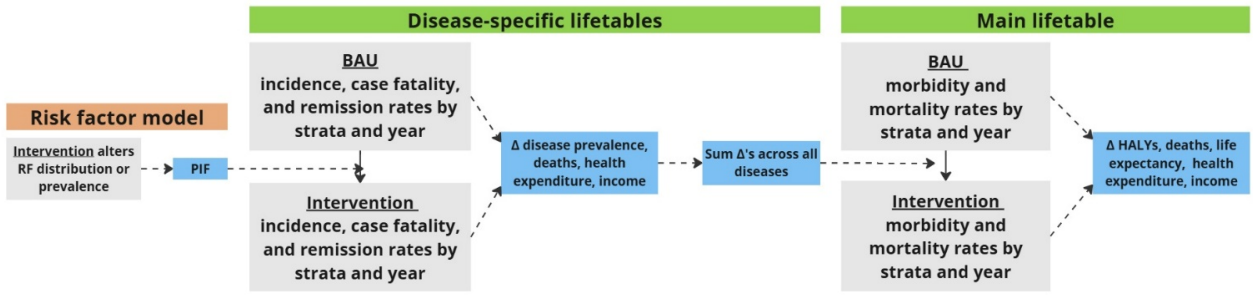

Figure 11.1: Conceptual generic PMSLT calculations. BAU = Business-As-Usual; PIF = Population Impact Fraction.

11.2 Health-adjusted life years

The health-adjusted life year (HALY)s is one of the key outputs of SHINE models. Before describing in detail the lifetable calculations that occur within the PMSLT to produce HALYs and other outcomes, we first give an overview here of what HALYs are.

11.2.1 Conceptual description of HALYs

HALYs measure population health over time, capturing both morbidity and mortality. This metric combines elements of both quality-adjusted life years (QALYs) and disability-adjusted life years (DALYs). Like QALYs, HALYs quantify years lived within a given time horizon, with each year weighted by health status. However, HALYs make use of the disease-specific disability weights produced (primarily) by the GBD, as are used for DALYs, rather than using utility weights used for QALYs. This is largely because of ease of availability of disability weights (or more specifically, YLDs produced using disability weights) for all diseases measured by the GBD. In comparison, utility is generally not calculated as a disease-specific metric. Instead, it is a holistic measure that considers multiple facets or dimensions of health for a given population, including physical, mental, and social wellbeing. Utility estimates are usually measured for specific populations and are context specific (i.e., depends on geographic location, cultural values, population characteristics), so can vary substantially across studies.

HALYs take a 'health gain' perspective where higher values indicate better health. This is the inverse of DALYs where higher values indicate greater burden. Consequently, morbidity is subtracted from 1 so that a year in perfect health equals 1 HALY. In SHINE modelling, HALYs are estimated for both intervention and BAU scenarios, with HALYs gained from the intervention (i.e., ΔHALY) being the metric of interest.

11.2.2 Calculating HALYs

HALYs are calculated as

$$HALYs = \sum_{s,t} PY \times (1 - morb) \tag{11.1}$$

where s is the stratum and t is the year. Under an intervention scenario, HALYs can change due to changes in either the person years, PY, and/or the morbidity rate, $morb$.

Morbidity calculation under the BAU scenario is discussed in detail in Chapter 6. Under an intervention scenario, all-cause morbidity is calculated as

$$morb_{\text{INT}} = morb_{\text{BAU}} + \sum_{d} \Delta dr_{s,d} \tag{11.2}$$

Disease-specific disability rates, dr, are calculated as

$$dr_{s,d(\text{chronic})} = dw'_{d,s} \times p_{d,s} \tag{11.3}$$

$$dr_{s,d(\text{acute})} = dw'_{d,s} \times i_{d,s} \tag{11.4}$$

Disease-specific disability rate changes under an intervention can occur as a result of change in prevalence of the disease (which occurs via changes to disease incidence, case fatality, or remission rates). Disease-specific disability weights, dw', are typically fixed, unless the intervention works via a change in severity (e.g., by applying a 20% reduction in the disability weight). Details on disability weights as used in SHINE models is provided in Section 6.4.5.

11.3 Main lifetable

11.3.1 Inputs

The main lifetable defines several metrics for each cohort:

- All-cause mortality rate (ACMR), split into:
  - Start year ACMR
  - Annual percentage change on ACMR
- Morbidity rate (fixed value over time under BAU)
- (Optional) Income, split into:
  - Average income in last year of life
  - Average income in all other years

- (Optional) Health system expenditure, split into:
  - Average expenditure rate in last year of life
  - Average expenditure rate in all other years

The model ACMR changes over time through an annual percentage change. All other rates above are set as static in each future each of the model. If the model is an open cohort, birth and migration rates are also included and may change over time (see Section 11.3.3.

The minimum population breakdown is sex by age - all factors listed above must be broken down across these strata. For a model that includes further heterogeneity, e.g., SES, the starting population, ACMR and disability rate inputs are also disaggregated (for SES analyses in Australia, this occurs through an automated system that takes in AIHW data - see Section 6.4.2). Note that income and health expenditure are only ever disaggregated by sex and age.

11.3.2 Population calculations

The main lifetable tracks the total population within each year. The model is initiated with a starting population (defined as the population at the end of the year prior to the base year, $\text{pop}_{\text{end t = -1}}$, broken down by sex and single year of age from 0-109 inclusive. For each year, starting with the base year, the population count changes through the ACMR (and births and migration for open cohort models).

The age of each cohort increases at the start of the year. The population at the end of the year is calculated by looking up the appropriate ACMR by age and sex (and other strata). The model also calculates the person-years (PY) by assuming that deaths are uniformly distributed within each year. The disability rate is then applied to compute HALYs.

$$\text{deaths} = \text{pop}_{\text{start}} \times (1 - e^{acmr}) \tag{11.5}$$

$$\text{pop}_{\text{end}} = \text{pop}_{\text{start}} - \text{deaths} \tag{11.6}$$

$$\text{PY} = (\text{pop}_{\text{start}} + \text{pop}_{\text{end}})/2 \tag{11.7}$$

$$\tag{11.8}$$

The main output of the PMSLT is the difference in HALYs under an intervention scenario in comparison to BAU, each cumulated over 10-, 20-, and 40-year time horizons (as well as 110-year time horizon for a closed cohort sensitivity analysis).

11.3.3 Births and migration

As discussed in Chapter 6, SHINE open cohort models infer birth and migration rates from population projection produced by the Global Burden of Disease (GBD). The BAU scenario computes the births and migrations required to match the projection (accounting for the already forecast ACMR) and then applies these same birth and migration counts to the other scenarios. The underlying assumption is that the intervention will not modify either of these rates.

Half the births and migration are added at the start of the year, after the age of the cohort increases. The other half are added at the end of the year, and in particular, after person-years are calculated. This split reflects the assumption that births and migration are spread uniformly over the year.

11.4 Chronic disease model

A chronic disease is modelled via a set of differential equations that regulate flow rates between the healthy, prevalent, and dead population. The flow rates correspond to incidence, case fatality, and remission rates for the disease (as depicted in Figure 6.1 in Chapter 6).

Let S, C and D denote the three states susceptible, prevalent (or living with the condition), and death from the disease. Let i, f and r be the incidence, fatality, and remission rates. Then a chronic disease is modelled as

$$\frac{dS}{dt} = -iS + rC \tag{11.9}$$

$$\frac{dC}{dt} = iS - (f + r)C \tag{11.10}$$

$$\frac{dD}{dt} = fC. \tag{11.11}$$

These differential equations are for fixed i, f and r, however these rates vary by age and intervention scenario in the PMSLT. To account for this, the system of differential equations are only used to evolve the state of a cohort from the start of a year to the end of that same year. This allows the cohort to experience different rates in subsequent years, as it increases in age.

11.4.1 Inputs

Chronic disease models are initialised with a base year prevalence, broken down by age and sex (and optionally additional strata). A model with births uses the initial prevalence at age zero to initialise new cohorts.

The chronic disease model requires the following parameters:

- Incidence rate
  - Start year incidence rate
  - Annual percentage change on incidence
- Remission rate[1]
  - Start year remission rate
  - Annual percentage change on remission
- Case fatality rate
  - Start year case fatality rate
  - Annual percentage change on case fatality
- Disability weight (static over time)
- (Optional) Income loss (static over time), split into:
  - Average income loss in the first year of having the disease (incident year)
  - Average income loss in the year in which death occurs from the disease

[1] Note that the way remission is calculated in SHINE that it likely also includes excess death rates of dependent diseases. See section 6.4.4

  - Average income loss in all other years living with the disease (prevalent years)
- (Optional) Health system expenditure (static over time), split into:
  - Average expenditure rate in the first year of having the disease (incident year)
  - Average expenditure rate in the year in which death occurs from the disease
  - Average expenditure rate in all other years living with the disease (prevalent years)

Chapters 6 and 7 described how these inputs are collated for the BAU scenario.

11.5 Acute disease model

The acute disease model is used for diseases that are too short for the model's one year time-step, for example lower respiratory tract infections. These diseases are treated as instantaneous events and do not have a prevalent population. This acute disease model is simple and flexible so can be used to model any short event which causes health or economic outcomes among a population.

The incidence rate of acute disease is technically irrelevant, i.e., the acute disease model is invariant under multiplying incidence by X and dividing all other rates by X. This allows us to include disease for which the GBD does not report incidence, such as aortic aneurysm, by setting the incidence rate to 1 and setting the case fatality rate to the fatality rate reported by the GBD. It would be equally valid to set the incidence rate to the fatality rate, and set the case fatality rate to 1.

11.5.1 Inputs

The acute disease model requires the following parameters:

- Incidence rate, split into:
  - Start year incidence rate
  - Annual percentage change on incidence
- Case fatality rate, split into:
  - Start year case fatality rate
  - Annual percentage change on case fatality
- Disability weight (static over time)
- (Optional) Income loss per event (static over time), split into:
  - Average income loss if the acute disease results in death
  - Average income loss if the acute disease does not result in death
- (Optional) Health system expenditure per event (static over time), split into:
  - Average expenditure rate if the acute disease results in death
  - Average expenditure rate if the acute disease does not result in death

Chapters 6 and 7 described how these inputs are collated for the BAU scenario.

11.6 Morbidity and mortality changes under interventions

Interventions act on one (or more) of the disease rates shown in Figure 11.1 - the incidence, severity (i.e., the disability weight of the disease), or fatality rate. The impact point of an intervention depends on its conceptualisation. Generally, an intervention will have resulting impacts on both disease morbidity and mortality (unless the intervention 'just' impacts disease severity, and there are no differences in case fatality by severity level modelled). These changes are described below.

11.6.1 Mortality

Main lifetable

The ACMR applied to a given cohort for a single year of an intervention, INT is

$$acmr_{\text{INT,s}} = acmr_{\text{BAU}} + \sum_{d \in \text{Diseases}} (m_{\text{BAU,d,s}} - m_{\text{INT,d,s}}) \tag{11.12}$$

where $acmr_{\text{s}}$ is the population-level ACMR of stratum s, and $m_{\text{d,s}}$ is the disease-specific mortality rate of disease d and stratum s.

Disease lifetables

The disease-specific mortality rate of a chronic disease is derived from the deaths caused by the disease each year. Recall that a chronic disease has three states, S, C, and D (Figure 6.1), and a set of differential equations that evolve these states from the start of the year to the end. Then

$$m_{\text{d,s(chronic)}} = \ln\left(1 - \frac{D_{\text{end}} - D_{\text{start}}}{S_{\text{start}} + C_{\text{start}}}\right). \tag{11.13}$$

This converts the proportion of the cohort that died during the year into an instantaneous mortality rate. The logarithmic transformation ensures the rate is consistent with continuous-time survival analysis.

The disease-specific mortality rate of an acute disease is

$$m_{\text{d,s(acute)}} = i \times f. \tag{11.14}$$

11.6.2 Morbidity

Main lifetable

The morbidity rate applied to a given cohort for a single year of an intervention is

$$morb_{\text{INT,s}} = morb_{\text{BAU,s}} + \sum_{d \in \text{Diseases}} (dr_{\text{BAU,d,s}} - dr_{\text{INT,d,s}}) \tag{11.15}$$

where $morb_s$ is the population-level morbidity rate of strata s, and $dr_{\text{d,s}}$ is the disease-specific disability rate of disease d.

Disease lifetables

The disability rate of chronic disease is defined by the modified disability weight (Section 11.2.2), multiplied by the average prevalence of the disease over the year to produce a population-level rate. Recall that a chronic disease has three states, S, C, and D, and a set of differential equations that evolve these states from the start of the year to the end. Then

$$\begin{aligned} dr_{\text{d,s(chronic)}} &= dw' \times p \\ &= dw' \times \left(\frac{C_{\text{start}}}{S_{\text{start}} + C_{\text{start}}} + \frac{C_{\text{end}}}{S_{\text{end}} + C_{\text{end}}} \right) /2. \end{aligned} \tag{11.16}$$

The disability rate of an acute disease is defined by the modified disability weight, dw' (Section 6.4.5), for that disease multiplied by disease incidence (Equation 11.3)

11.7 Breakdown of health gain by diseases

The PMSLT can be run with any number of diseases - e.g., for tobacco smoking, over 30 diseases are modelled. The default for SHINE is to output morbidity and mortality changes aggregated across all disease affected by an intervention - this occurs automatically within the lifetable calculations as shown in the previous section. However, we may want to disaggregate the effects of an intervention into the HALY gains *attributable* to each disease. To calculate this the PMSLT could be rerun for each individual disease, but this is resource intensive and disregards the diminishing returns inherent in intervening on many diseases simultaneously. SHINE solves these issues by calculating attributable health gains to individual disease within a single model run using a disease attribution method.

Disease attribution calculations exploit the linearity of mortality and morbidity shifts shown in Section 11.6 (i.e., there is not interaction between diseases). This linearity is specific to the BAU scenario, so the disease attribution method can only determine the difference between intervention scenarios and BAU. We do not use this approach to calculate the total health impact of a disease within a single scenario.

Because mortality and morbidity changes are additive across diseases (see Section 11.6), the contribution of each disease can be isolated by calculating what the intervention effect would have been if only that disease were affected. The calculations required for this system are explained in Appendix B.

11.8 Economic changes under interventions

SHINE disaggregated healthcare costs and disease-related income effects into different phases of disease, as discussed in Chapter 7. Briefly, costs are separated into whether an individual is in their last year of life or not. For chronic diseases, costs are also separated into year of diagnosis (i.e. incident year) vs. other non-death years (i.e. prevalent years).

The PMSLT calculations treat health expenditure and income loss identically in a mathematical sense, so we will use generic *cost* or v to refer to both in the following calculations.

11.8.1 Main lifetable cost calculations

The main lifetable cost calculation is as follows. For a given year, cost for a sex by age (by other strata) group is calculated as

$$cost = v_{\text{survival}} \times pop_{\text{end}} + v_{\text{death}} \times deaths \tag{11.17}$$

The survival cost rate, v_{survival}, is the cost incurred for living through the year, while the death cost rate, v_{death}, is the cost incurred by the population that dies during the year. The source of these values is detailed in Chapter 7.

11.8.2 Disease lifetable cost calculations

Incident & prevalent year costs

The change in costs due to changes in disease rates is conceptually similar to changes in disability rate (see Section 11.6.2).

Let v_{INT} be the cost rate under an intervention scenario, for both incident and prevalent cases (or 'non-death year' cases). For a given cohort within a single year,

$$v_{\text{survival,INT}} = v_{\text{survival,BAU}} + \sum_{d \in \text{Diseases}} \left(v_{\text{survival,BAU},d} - v_{\text{survival,INT},d}\right) \tag{11.18}$$

The cost rate of an acute disease is

$$v_{\text{survival},s,d(\text{acute})} = i \times v_{\text{incident}} \tag{11.19}$$

where $v_{s,d}$ is the disease-specific cost rate of disease d in scenario s (where s is either BAU or an intervention scenario). The BAU scenario uses the v_{survival} input files directly, while the intervention may modify disease incidence and therefore the cost.

The non-death year cost rates of chronic disease are broken down into first year of diagnoses (incident) and ongoing (prevalent) cases. Chronic diseases have three states, S, C, and D, which are evolved from the start to the end of the year, and an incidence rate i. Then

$$v_{\text{survival},s,d(\text{chronic})} = v_{\text{incident}} \times \left(1 - e^{-i}\right) \frac{S_{\text{start}}}{S_{\text{start}} + C_{\text{start}}} \tag{11.20}$$

$$+ v_{\text{prevalent}} \times \max\left\{0, \frac{C_{\text{end}}}{S_{\text{end}} + C_{\text{end}}} - \left(1 - e^{-i}\right) \frac{S_{\text{start}}}{S_{\text{start}} + C_{\text{start}}}\right\}. \tag{11.21}$$

The first term accounts for the incident cases by calculating the incidence risk and applying it to the susceptible cases at the start of the year. The second term accounts for the ongoing cases by tallying the cases that reached the end of the year, minus those that became prevalent this year. This second term is an approximation as some incident cases will die in the same year.

Death year costs

The cost attributable to deaths in the intervention, $v_{\text{death,INT}}$, cannot be calculated without knowing which diseases are responsible for the change in death rate. Breaking down the impact of the intervention by disease is covered in Appendix **??**, as is a derivation of $v_{\text{death,INT}}$. An outline and justification of this system is given below.

Consider two diseases, A and B. The health system cost in the year of death from A, $v_{death,A}$, is $\$100$, in comparison to a cost, $v_{death,B}$, of $\$500$ in the year of death from B. The average population-level health system cost in the year of death, $v_{\text{death,BAU}}$, is $\$200$. Now consider an intervention that doubles the remission rate of A and B, which just so happens to cause a reduction in total deaths from $d_{\text{BAU}} := 100$ to $d_{\text{INT}} := 80$. The cost savings under this intervention, for $v_{\text{death,INT}}$ specifically, depends on how many deaths from disease A vs. disease B were averted.

Consider the case in which all the averted deaths are of disease B, saving $v_{death,B}$ per averted death, so

$$\begin{aligned} v_{\text{death,INT}} &= (d_{\text{BAU}} \cdot v_{\text{death,BAU}} + (d_{\text{INT}} - d_{\text{BAU}}) \cdot v_{death,B})/d_{\text{INT}} \\ &= (100 \cdot \$200 - 20 \cdot \$500)/80 = \$125. \end{aligned}$$

Alternately, if all the averted deaths were due to A, then

$$\begin{aligned} v_{\text{death,INT}} &= (d_{\text{BAU}} \cdot v_{\text{death,BAU}} + (d_{\text{INT}} - d_{\text{BAU}}) \cdot v_{death,A})/d_{\text{INT}} \\ &= (100 \cdot \$200 - 20 \cdot \$100)/80 = \$225. \end{aligned}$$

Essentially, averting deaths from a disease that is more costly to the health system drives $v_{\text{death,INT}}$ down, while averting deaths from a disease that costs the health system less drives $v_{\text{death,INT}}$ up. So the PMSLT calculates death costs directly from the per-disease savings of averted deaths, and only calculates $v_{\text{death,INT}}$ at the end, purely as an output.

11.8.3 Actual intervention costs

The above calculations describe the downstream changes in healthcare and/or income that occur under an intervention scenario as a result of changes to disease rates.

When modelling *actual* interventions (as opposed to targets or 'magic wand' scenarios) additional costs may be incorporated for the intervention itself.

We could just add on intervention costs to PMSLT output. However, intervention costs are generally parameterised with uncertainty. Therefore, these additional costs are added on to v_{INT} in each iteration of the model.

11.8.4 Discounting

The concept of discounting was discussed in Section 2.3.2 of this protocol. SHINE uses the following rules for discounting in different situations:

1. Health outcome analysis only: 0% discount rate applied in main analysis; 3% discount rate in supplementary analysis.
2. Health outcome analysis and cost offsets: 0$ discount rate applied in main analysis; 3% discount rate to health outcomes and costs applied in supplementary analysis.
3. Economic evaluation: 3% discount rate to health outcomes and costs applied in main analysis; 0% and 6% discount rates in supplementary analysis.

11.9 Comparison of SHINE to CRA

In Chapter 1 of this protocol, we introduced the differences between classic comparative risk assessment (CRA) and prospective simulation modelling. A key difference is that the latter estimates *avoidable* morbidity and mortality changes under interventions that target risk factors, in comparison to BAU, whereas CRA analyses are cross-sectional and instead consider *attributable* burden from a specific risk factor, by comparing population health under a hypothetical scenario in which the risk factor never existed above the theoretical minimum risk exposure level (TMREL).

SHINE models are primarily used to estimated future avoidable morbidity and mortality under interventions, in comparison to BAU. While not done in standard analyses, it is possible to approximate attributable burden for a risk factor in a SHINE model.

11.9.1 Implementing a CRA with the PMSLT

A CRA estimates health differences under a hypothetical scenario in which a risk factor never existed. To perform a CRA with the SHINE PMSLT, we construct this hypothetical as an intervention, then compare it to the BAU. There are two problems to solve when doing this.

1. The PMSLT has time lags, both between risk factor exposure and disease incidence, and with ongoing prevalence of chronic disease.
2. The PMSLT models changes in population size due to reduced death rates, while CRAs ignore this effect.

The first problem is addressed by modelling the years leading up to the cross-sectional year of the CRA, t. Ideally the model would start in year $t - 110$, the year of birth of the oldest cohort, but in practice we just need to allow enough time for the chronic diseases to reach the new risk-free steady state. The time required depends on the rates of the specific disease and the lag times of the associated risk factor, but will usually be around 20 to 40 years.

The second problem is addressed by overwriting the population counts of the intervention with that of the BAU, at the start of each year. This will remove population from the intervention in each year, as removing a risk factor tends to result fewer deaths. Conceptually we are removing a portion of the population at the end of each year, effectively treating the lives saved as immortal and no longer having any trackable health state. Without this adjustment, the PMSLT may report more deaths under the intervention, in year t, due to deferred deaths from earlier years.

Chapter 12

Risk factor models

12.1 Potential Impact Fraction (PIF)

For any intervention that acts via a risk factor, such as BMI, a potential impact fraction (PIF) is calculated to result in change to disease rates.

The PIF is similar to the population-attributable fraction (PAF) calculated in CRA studies. The PIF translates to a change in a disease parameter (depending on what is specified by the relative risk (RR) – e.g., may be a change in disease incidence, or case fatality rate), resulting in differences in disease morbidity and mortality comparing a target/intervention scenario to BAU. The disease-specific changes are summed to then produce a change in all-cause morbidity and mortality in the main lifetable, output as mortality and health-adjusted life year (HALY) differences between the target/intervention scenario and BAU.

The PIF is calculated as

$$PIF_{sdt} = \frac{\sum_{j=1}^{n} P_{sjt} RR_{sdj} - \sum_{j=1}^{n} P'_{jt} RR_{sdj}}{\sum_{j=1}^{n} P_{sjt} RR_{sdj}} \tag{12.1}$$

where s is the specific strata (sex by age [by other included strata e.g., socioeconomic status]), d is the disease, t is the year (time-step), j is the health/behaviour state (out of a total n states), P is the proportion of the specific strata in each state j under BAU, compared to the target or intervention scenario (P'), and RR is a relative risk linking a specific state j to the disease d. The PIF translates to a change in a disease parameter (depending on what is specified by the RR– e.g., may be a change in disease incidence, or case fatality rate), resulting in differences in disease morbidity and mortality comparing a target/intervention scenario to BAU. The disease-specific changes are summed to then produce a change in all-cause morbidity and mortality in the main lifetable, output as mortality and HALY differences between the target/intervention scenario and BAU.

Each state j may represent states in a Markov model (with $RR = 1$ for each disease in the 'healthy' or 'unexposed' state, and $RR \neq 1$ for each disease in the state(s) representing exposure to a risk factor above the TMREL). Alternatively, j can represent bins/categories of a risk factors distribution with a proportion of the population in each category, such as systolic blood pressure.

12.2 Disease risk

Relative risk (RR) associations between risk factor exposure and disease incidence (e.g. between smoking and lung cancer) are sourced from the GBD (using the most up to date RRs available).

12.2.1 Disease selection

SHINE's default is to model all diseases that have been associated with exposure to a risk factor, per the GBD. If this is not possible, the aim should be to include diseases contributing to at least 95% of current burden attributed to a risk factor in the population of interest (prioritising the diseases contributing to the most individual burden).

For diseases contributing to a low level of risk factor-attributable burden, where disaggregated disease data by SES (or other sources of heterogeneity that are of interest) is not available, it may be appropriate to still include the disease, but with no heterogeneity beyond sex by age.

Any associated diseases omitted from a risk factor model, or without required heterogeneity, should be clearly stated in any report/publication.

12.2.2 Disease relative risk values

Whilst RRs may in reality change over time (e.g. they have steadily increased over time for the association of smoking with diseases, due to the never smoking population getting healthier and healthier), the default for SHINE is to posit no change in RRs into the future.

Where available, SHINE uses the Burden of Proof analyses from the GBD as the source of disease RR values (described below). These RRs generally differ by sex and age. SHINE does not attempt to disaggregate RRs by another other strata.

Note that the parameterisation of a risk factor exposure in SHINE models needs to align with the units in which the risk factor-disease dose response curves are calculated. For standard SHINE models, the risk factor distribution and disease RRs are both taken from the GBD (with the risk factor distribution estimated indirectly from SEV projections); therefore, the risk factor distribution is already set up in the same format as the disease RR values. However, this is not always the case. For smoking models, SHINE currently measures smoking prevalence by strata using national data, but the dose-response curve for smoking-related disease risk is in units of smoking intensity (either pack-years or cigarettes smoked per day, depending on the disease). Therefore, smoking intensity needs to be estimated for the smoking population so that the smoking population can be assigned an appropriate disease risk level.

Burden of proof and star rating

The SHINE default is to use disease incidence relative risks for each risk factor from the GBD, specifically using data resulting from the Burden of Proof Risk Function (BPRF) method.[62] The GBD conducts systematic reviews for each risk factor included in their analyses (88 risk factors as of the 2023 GBD [63]), to determine

risk factor-disease pairs and estimate effect sizes for each. The BPRF method has been incorporated into the meta-analytic process since the 2021 GBD iteration. It has two key purposes[62]:

1. Adjust for bias within individual studies included in the meta-analysis
2. Adjust for between study bias, allowing for heterogeneity between studies to be accounted for

The GBD produces two sets of relative risks RRs for each RF-disease pair: one that has been adjusted only for the within-study bias only, and the other that has been adjusted for both components listed above. The later produces RRs with wider uncertainty intervals, due to the inclusion of between study heterogeneity.

As part of the BPRF analysis, the GBD also produces a five-star rating system, based on the strength of association between each risk factor and disease pair. One and two star associations are equivalent to minimal/low evidence for an association. However, this rating does not necessarily indicate that an association does not exist. It may be the result of a lack of standard exposure definition between studies, resulting in measurement error and between-study differences, or due to a low number of studies being available.[64] As a sensitivity analysis for SHINE modelling, only diseases with a 3+ star rating are modelled, excluding those with limited evidence for an association.

The output of the BPRF process, including dose response curves and star ratings for each risk factor-disease pair, can be visualised using an interactive tool on the IHME website: `https://vizhub.healthdata.org/burden-of-proof/`.[65] SHINE directly downloads dose response curve data from the IHME for risk factor modelling. RR values for continuous risk factors are linearly interpolated to produce RRs for every appropriate unit increase in exposure to a risk factor.

12.3 Time lags

There are generally time delays between change in exposure (to risk factors) and resulting change in disease incidence, and this can vary for different disease groups. Many analyses exist that quantify increased risk (i.e. incidence rate ratios) of disease for increasing exposure to deleterious risk factors (e.g. systolic blood pressure, tobacco, diet), but these relative risks are for (usually poorly stated durations of) medium- to long-term exposure to this level of risk factor. There are few studies that answer the much harder question of how quickly disease rates change following change in risk factor level. One clear exception is tobacco, where Hoogenveen et al. (2008) [66] have estimated the rate at which excess rates of disease decrease over time when people quit smoking. The SHINE-Tobacco model makes explicit use of these decay functions (discussed further below). To our knowledge, no such decay functions exist for other risk factors.

Some evidence also exists for the time lag in the effects of changing BMI onto disease. Aminian et al.[67] harnessed the natural experiment of bariatric surgery, which dramatically changes BMI, to examine changes in cancer incidence rates. Whilst the authors did not find statistical evidence of non-proportional hazards, to the eye (and consistent with prior theoretical expectation) incidence rates for those who did and did not receive bariatric survey did appear to be similar in the first two years, then increasingly diverge (refer to Figure 2 in Aminian et al[67]). Note that bariatric surgery also has endocrine effects on the hepatic-gut system – so we cannot state that the impacts are purely through BMI change. Nevertheless, the evidence is compelling that an intervention (and BMI change) can see divergence of cancer incidence rates from two years.

It is important for these time lags to be incorporated into the PMSLT model - not including time lags will result in short-term health gains after risk factors changes being overestimated.

12.3.1 Conceptualisation of SHINE time lags

The BODE^3 PMSLT model (i.e., the precursor to the SHINE PMSLT) introduced the concept of allowing for time lags by using an average of (1 – PIF) estimates in previous years. SHINE builds on this approach, specifying time lags with two components:

1. Delay in decay 'start', Y_{lag}: this is how many years (or other unit time) before a change in risk factor starts seeing any change in disease incidence rates.
2. Duration of decay 'lookback', L: this is how many years into the past exposure is relevant, and therefore for how many years weighted averaging of past (1 – PIF) is required.

For example, if Y_{lag} = 3 years and L = 20 years, then exposure in a lagged time window of 3 to 23 years ago impacts disease incidence rates today.

We also include a 'half cycle correction' in our method. State transition models (e.g. Markov and lifetables) often include a half cycle correction to reflect that events over the course of some cycle length neither happen at the beginning or end of the cycle, and can be mathematically approximated as occurring mid-cycle. From a more public health perspective, if we say "the sodium reduction intervention started in 2023", what do we mean? It started abruptly and completely on the first day of 2023, the last day of 2023, or was steadily rolled out during 2023 to be at full effect by the year's end? We assume the latter, which can be mathematically approximated as an intervention taking abrupt and full effect mid-cycle. (More sophisticated intervention implementation, such as steady reductions in sodium over ten years, will need more explicit modelling in an intervention model preceding the PMSLT. Here, we consider only the half cycle correction for an intervention occurring in a given calendar year.)

SHINE categorises time lags by disease, rather than by risk factor. Three categories are defined:

1. Long time-lag diseases (e.g. cancers):
    - Delay in decay 'start': median 2 years (range: 1-3 years)
    - Duration of decay: median 20 years (range: 13-29 years)
2. Short time-lag diseases (e.g. cardiovascular diseases):
    - Delay in decay 'start': 0 years (no uncertainty)
    - Duration of decay: median 5 years (range: 3-8 years)
3. Nil or zero time-lag diseases:
    - Delay in decay 'start': 0 years (no uncertainty)
    - Duration of decay: 0 years (no uncertainty)

The assignation of diseases to long, short and nil time-lag is shown in Appendix D Table D.1.

There are many caveats and notes regarding this system:

- There is poor data and research on which to make the assumptions about delay and decay durations, but without doubt this system is 'closer to the truth' that simply assuming nil or zero time-lags for all diseases. This system will need improving in the future as more information becomes available.

- This system only differentiates by disease or condition; in the future, there may be enough theory, data and information to specify time-lags differently by risk factor-outcome pairs.
- Even within the long duration time-lags, more of the impact of risk factor change occurs soon after risk factor change than out to 20 years. Put another way, there is more weighting of (1 – PIF) more recently in the past. See below for further quantification of this weighting.
- The length of the delay for long duration time lags, and the length of the decay for both short and long duration time lags, are highly uncertain. We include uncertainty in these lengths, as specified above.

More details are provided in Appendix D.

12.3.2 SHINE Tobacco time lags

The SHINE Tobacco model uses a different approach to incorporating time lags, given the larger body of evidence on tobacco related risk. Two components to SHINE Tobacco time lags are considered:

1. Disease risk for those currently smoking: we do not incorporate time lags for those who currently smoking. This is because smoking-disease incidence risk ratios used in this model differ by 'pack-years', which incorporates duration of smoking. Hence, these RRs are assumed to reflect current disease risk.
2. Disease risk for those who have quit smoking: this draws on the decay function developed by Hoogenveen et al.[66], which defines disease-specific decay rates. Figure 12.1 is adapted from Hoogenveen et al. [66], and shows the reduction in risk of several diseases over a 20-year timeframe post-quitting smoking. SHINE applies this decay function for each tobacco-related disease for 20 years. Given that disease risk is still >RR=1 with tihs exponential decay, a linear reduction from year 21-30 post-quitting is then applied so that there is no excess disease risk for those who formerly smoked from 30-years post-smoking.[68]

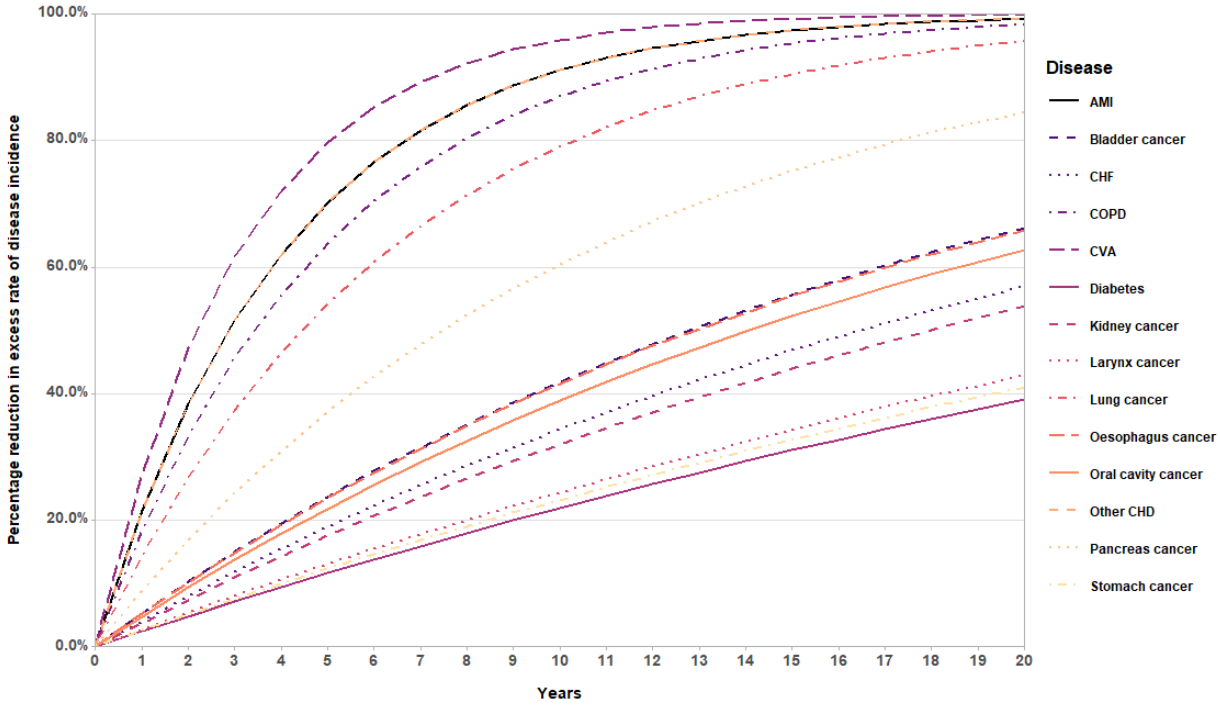

Figure 12.1: Cumulative percentage reduction in risk from that of people who smoke to that of people who never smoked among 50-year olds, after quitting tobacco smoking (Data source: Hoogenveen et al.[66])

Chapter 13

Running the model

13.1 Time horizon

There are multiple time horizons to be aware of when running a SHINE model:

- Time horizon for BAU parameter forecasts

  Most BAU parameters that vary by calendar year are forecast from the base year (t=0) for 20 years, with a gradual reduction in the annual percentage change from years 21-40, following which rates are static (Section 5.3.1). Note that there are some exceptions (e.g., risk factor forecasts - see section 6.3.1).

- Time horizon for the intervention

  Note that the intervention may not be active for the full duration that the model is run (below)

- Time horizon for tallying model outputs

  The model is run by default for 40-years, from the first year of the intervention. The main results will then focus on the first 20 years (see next Chapter). E.g., If the base year of a model is 2023, but the intervention start year is 2025, then the model is run until 2064, with the main results focusing on 2025-2044 inclusive. Note that in this example, the BAU model forecasts are held constant after 2042.

13.2 Input parameter uncertainty

When inputting parameters with uncertainty into the model, three characteristics need consideration:

- If a value less than (or sometimes greater than) zero is impossible, and one is positing large uncertainty (e.g., standard deviation >20% of the mean), it may be appropriate to log or otherwise transform the variable so that a value of less than zero is impossible to draw.
- Second, and related, the reference value may be one (e.g., relative risks), requiring the analyst to consider if uncertainty applies to the difference to one or can cross one. If the former, then a beta distribution may be appropriate (as it samples values from 0-1). If the latter, a log-normal distribution for a ratio

may be appropriate.

- Lastly, when drawing from the uncertainty distribution of a given parameter, it is necessary to decide whether to correlate this variable with the draws from other (see section 13.3.1).

SHINE default distributions by input type are shown in Table 13.1. Note that these are not a hard and fast rules.

Table 13.1: Key input distribution & uncertainty in SHINE models

Input	Default distribution type	Uncertainty source	Notes
Disease parameters: Incidence, case fatality rate, remission rate	Log-normal	Incidence and case fatality rate standard deviations (SD) derived from regression model on historic GBD data. Remission rate SD set 0.1.	No uncertainty specified around disability rate.
Disease prevalence	Beta	Prevalence distribution only applies in model base year. In future years, prevalence is based on drawn rates	
BAU cost parameters: Healthcare expenditure, income	Gamma	SD set at ±10% of mean value for all BAU costs	
Risk factor-disease relative risks	Depends on Burden of Proof Risk Function (BPRF) [usually log-normal]	SD derived from GBD BPRF	Star-rating system used for sensitivity analysis
Intervention effect size	Depends on intervention	Sourced from source study or adjusted/set (see Chapter 8, and Section 13.2.1 below).	
Intervention costs	Gamma	Sourced from source study or adjusted/set (see Chapter 10, and Section 13.2.1 below)	

13.2.1 Setting input parameter uncertainty

For BAU input parameters, SHINE mostly bases uncertainty intervals from the source data (e.g., GBD disease data confidence intervals). However, for some inputs, particularly intervention effect sizes, uncertainty around the effect size needs more consideration. There are multiple options for this, depending on the source of data, and its accuracy.

Options:

1. Use confidence interval around effect size, provided in meta-analysis/individual study, without any

further adjustment. This should only be done when no major sources of bias are identified

2. If necessary, widen or set the uncertainty from meta-analysis/individual study. SHINE uses the general rules below, adopted from the BODE^3 program:
    - Low uncertainty: +/- 5% of the expected or median value of the input parameter as the standard deviation (SD) for the input's probability distribution
    - Moderate uncertainty: +/- 10% SD
    - High uncertainty: +/- 20% SD
    - Very high-level uncertainty: +/- 30%-40% SD
    - Transported evidence: double the uncertainty defined in study.
3. If using a regression model to estimate an input parameter, apply bootstrapping to estimate uncertainty around each value.
4. If conducting an expert knowledge elicitation (Section 9.3.2), incorporate ranges into expert questions, in order to estimate group uncertainty from individual response.

13.3 Monte Carlo simulation

SHINE uses Monte Carlo simulation to incorporate input parameter uncertainty. The SHINE default is to draw 2000 combinations of input parameters for the PMSLT. Within each iteration, each parameter is drawn from its probability distribution (see Table 13.1) and the model then runs deterministically on each cohort in parallel. To reiterate, there is no within-population variation except for that deterministically captured by splitting the population into heterogeneous cohorts (such as by age, sex, and Socio-Economic Index For Areas (SEIFA)).

The 2000 draws of the Monte Carlo method approximate the probability distribution of the model outputs. SHINE typically reports the median and 95% uncertainty interval of these distributions, i.e., the50th, 2.5th, and 97.5th percentiles.

13.3.1 Correlations

SHINE PMSLT uses rank correlation to avoid incompatibilities between differently shaped input parameter distributions. For two inputs A and B that are 100% correlated, if A is drawn from the 38th percentile of its distribution, then B will also be drawn from the 38th percentile of its distribution. The model imposes linear correlation in percentile-space, regardless of the shape of A and B. Partial and negative correlation is also available via the use of a correlation matrix.

Positive correlations in SHINE models tend to reduce the uncertainty interval of outputs. This is not an intrinsic property of correlation, rather, it is due to a model structure that causes most inputs to pull in the same direction. The result is that variation in correlated draws pulls more strongly, while uncorrelated draws somewhat cancel out.

Part V

MODEL OUTPUTS AND REPORTING

Chapter 14

Standard model outputs

14.1 Outcomes summary

Outcomes able to be produced by the SHINE PMSLT model are listed below. Each outcome is presented as the difference between the intervention scenario(s) and BAU (or alternatively comparisons between different intervention scenarios).

- Health-adjusted life years
- Deaths
- All-cause mortality rates
- Risk factor distribution/prevalence
- Health expenditure
- Income
- Intervention cost
- Cost-effectiveness (combining HALYs or deaths with total costs)
- Life expectancy change (closed cohort time horizon only)

As a default, model outputs are shown as aggregate values over 20-40 years, with options of single year breakdowns, and extended time horizons where necessary. Model outputs are also by default presented as an aggregate across all diseases impacted by the intervention, however, any of the outputs discussed in the following sections can also be presented for individual diseases (using the disease breakdown system discussed in Section 11.7.

In the following sections, examples of the different outcomes listed above are presented graphically, drawing mainly on the SHINE Health Intervention Impact Calculator (HIIC) Tool. See Box 14.1 for a summary of the HIIC.

Box 14.1.1. SHINE Health Intervention Impact Calculator (HIIC)

The SHINE HIIC is a web tool that provides pre-run PMSLT output, allowing users to generate health, health expenditure and income changes from various changes to future disease rates, either directly or via changes to risk factors.
As of early 2026, the tool is available for Australia, with pre-run PMSLT output for 100+ diseases, and 4 risk factors (systolic blood pressure, BMI, LDL cholesterol, and fasting plasma glucose). Users can define interventions that target disease incidence, case fatality, or remission rates, or alternatively target one of the 4 risk factors, for any sex by age group.
Intervention costs can also be input by the user, with the tool supporting economic evaluations of interventions (combining disease health expenditure changes with intervention costs for a cost-effectiveness analysis from a health system perspective).
The tool can be found at: `https://shine-hiic.com/`

14.2 Health outcomes

The key outputs of the model is the difference (Δ) in HALYs and deaths under a BAU vs. intervention scenario. These outcomes can be produced for any SHINE model, regardless of the risk factor and/or diseases included. They are therefore suitable for answering questions of allocative efficiency, with comparisons across risk factors, sub-populations, and countries, all possible.

14.2.1 HALYs

For HALYs, the following outputs would be presented:

- Cumulative HALYs gained under the intervention vs. BAU for the following timeframes (from the first year the intervention is implemented):
  - 0-10 years
  - 10-20 years
  - 0-20 years (sum of above two timeframes)
  - 20-40 years
- Yearly HALYs gained under the intervention vs. BAU

Figure 14.1 shows yearly HALYs gained under an intervention vs. BAU.

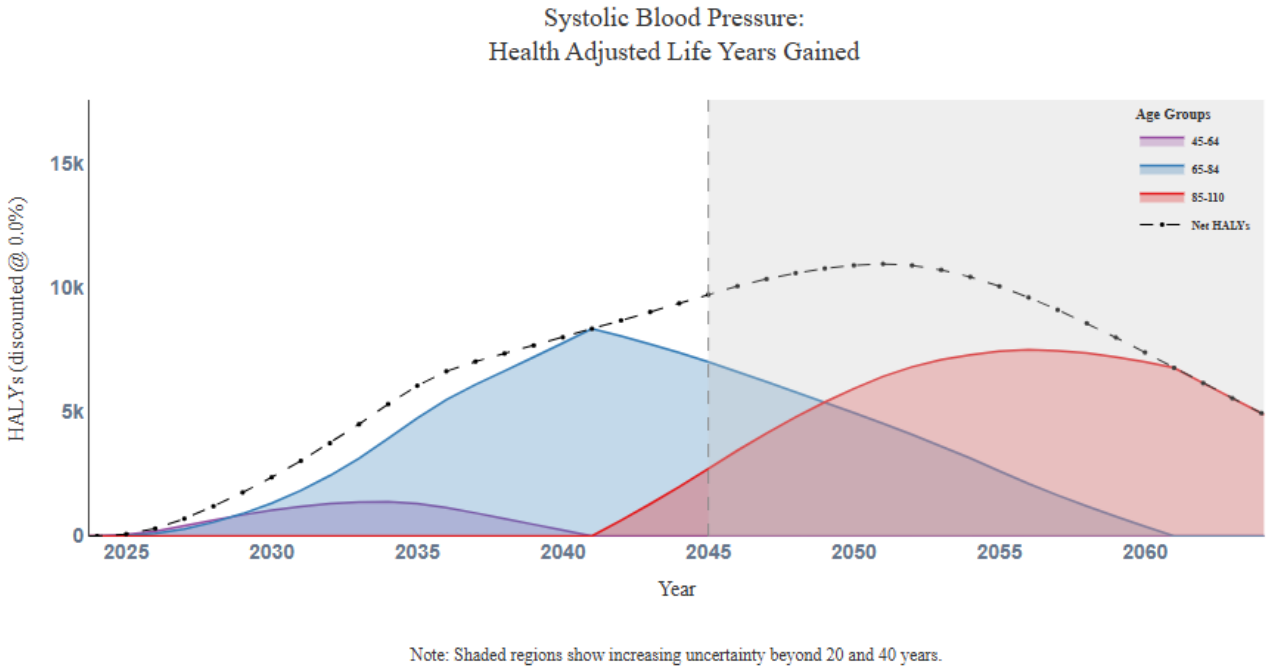

Figure 14.1: Yearly HALYs gained, resulting from a 10% shift in systolic blood pressure towards the theoretical minimum risk level, for the 45-64 year old cohort in 2025, implemented over 10 years. (Source: SHINE HIIC tool).

14.2.2 Deaths averted

The difference in deaths under an intervention scenario and BAU are simply calculated as the difference in population size between the two scenarios, occuring as a result of changes in all-cause or disease-specific mortality under an intervention scenario (see Section 11.6.1).

Deaths averted (or incurred, if an intervention has a negative impact on a population), are presented alongside HALYs, as:

- Cumulative deaths averted under the intervention vs. BAU for the following timeframes (from the first year the intervention is implemented):
  - 0-10 years
  - 10-20 years
  - 0-20 years (sum of above two timeframes)
  - 20-40 years
- Yearly deaths averted under the intervention vs. BAU

Importantly, deaths are those averted, not avoided. SHINE models include competing mortality - if an intervention prevents death at age 40 due to a given disease, that person will still die later due to the underlying all-cause mortality rate built into the main lifetable (that increases with age). If a given intervention reduces the risk of mortality from a given disease/group of diseases, in the short-medium term (depending on the time lag between the risk factor and disease) we would see this reflected in fewer deaths occurring in the

intervention scenario compared to BAU. However, if we continued to follow the population, we would eventually see more deaths occurring in the intervention scenario, as those people who did not die due to the disease of interest age and eventually die from other causes. An example of this is shown in Figure 14.2. The black dashed line in this figure shows the net deaths averted under a hypothetical systolic blood pressure reduction intervention compared to BAU - after approximately 30 years, the yearly deaths averted changes to negative, i.e., more deaths occuring each year in the intervention scenario. This occurs only for 85+ year olds, representing the older population who died at an earlier age in the BAU scenario, so did not see deaths for this age group.

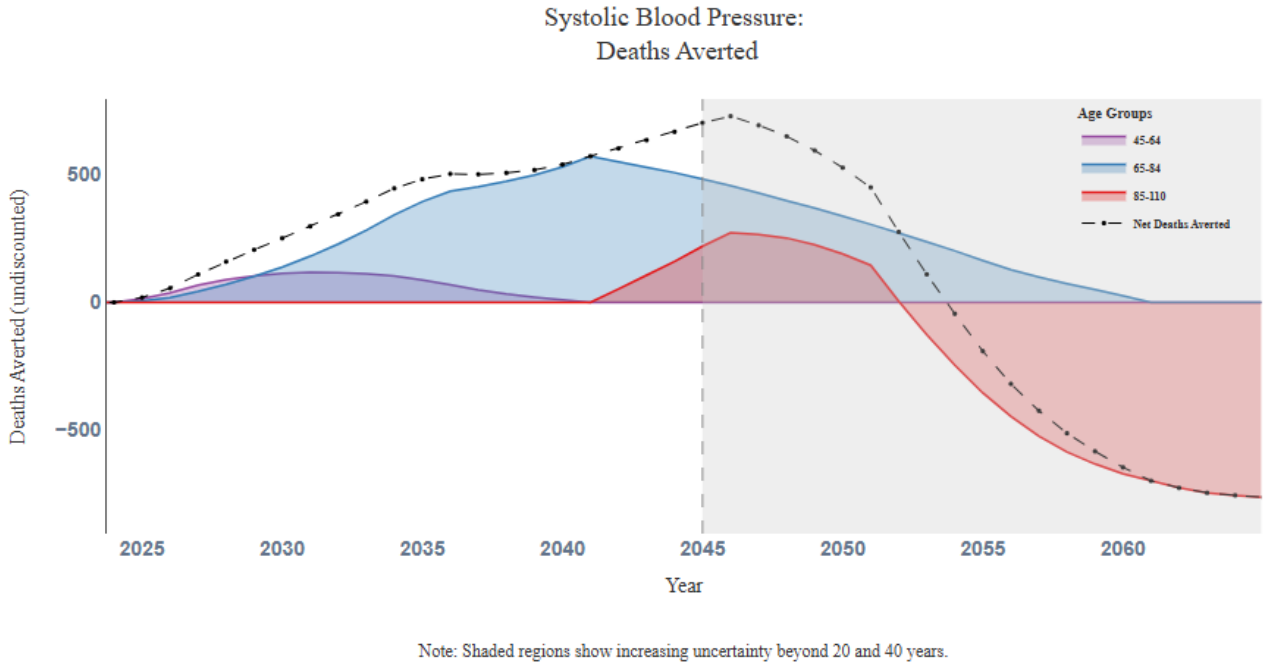

Figure 14.2: Yearly deaths averted, resulting from a 10% shift in systolic blood pressure towards the theoretical minimum risk level, for the 45-64 year old cohort in 2025, implemented over 10 years. (Source: SHINE HIIC tool).

14.3 Risk factor distribution change

For some interventions (specific interventions - the third level of SHINE 'intervention' that can be modelled), it may not be clear prior to running the model the effect it will have on a risk factors distribution, and hence risk factor distribution changes can themselves be a key SHINE model output. This is the case for complex models such as the SHINE-Tobacco model, where the intervention acts via a separate Markov model. In this model, the risk factor exposure of interest is smoking prevalence, but the Markov model is defined by smoking uptake and quit rates (which have been calibrated under the BAU scenario to reflect smoking prevalence forecasts). An intervention may be parameterised as a 2x increase in smoking quit rates for 30+ year olds - the Markov model (run alongside the PMSLT) will output the resulting change in smoking prevalence in each future year under this intervention in comparison to BAU.

14.4 Cost outcomes

Note that currently, SHINE only produces cost outcomes for analyses in Australia and New Zealand, given that much of the evidence draws on analyses in the New Zealand context, and direct data from Australia.

For both health expenditure and income productivity, the following output are presented:

- Cumulative cost change under the intervention vs. BAU for the following timeframes (starting from the first year the intervention is implemented):
  - 0-10 years
  - 10-20 years
  - 0-20 years (sum of above two timeframes)
  - 20-40 years
- Yearly cost changes under the intervention vs. BAU

When 'just' presented these cost offsets (i.e. without conducting an economic evaluation), SHINEs default is to presented undiscounted results as the main analysis, with discounted results (3% discount rate) presented as Supplementary results.

14.4.1 Health expenditure

Health expenditure estimates generally follow a similar trend to deaths averted. Under an effective intervention, a larger and older population occurs, and as a result increases in healthcare expenditure is seen. This is shown in Figure 14.3, drawing on the same hypothetical SBP intervention as shown above.

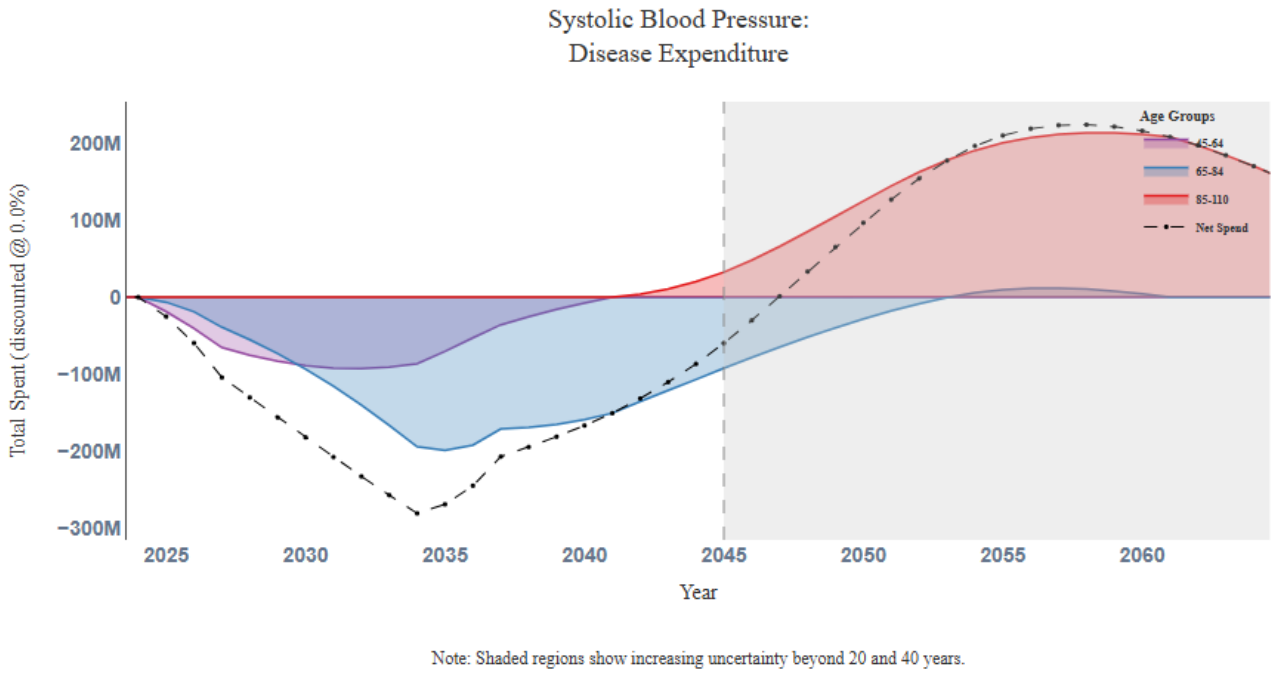

Figure 14.3: Yearly health expenditure savings, resulting from a 10% shift in systolic blood pressure towards the theoretical minimum risk level, for the 45-64 year old cohort in 2025, implemented over 10 years. (Source: SHINE HIIC tool).

14.4.2 Income productivity

Income changes under intervention scenarios compared to BAU only impact ages 25-64. Hence, in the example in Figure 14.4, income gain resulting from a systolic blood pressure reduction for 45-64 year olds in 2025 only affected this cohort while in that age bracket, rather than carrying through to older ages as seen for healthcare expenditure above.

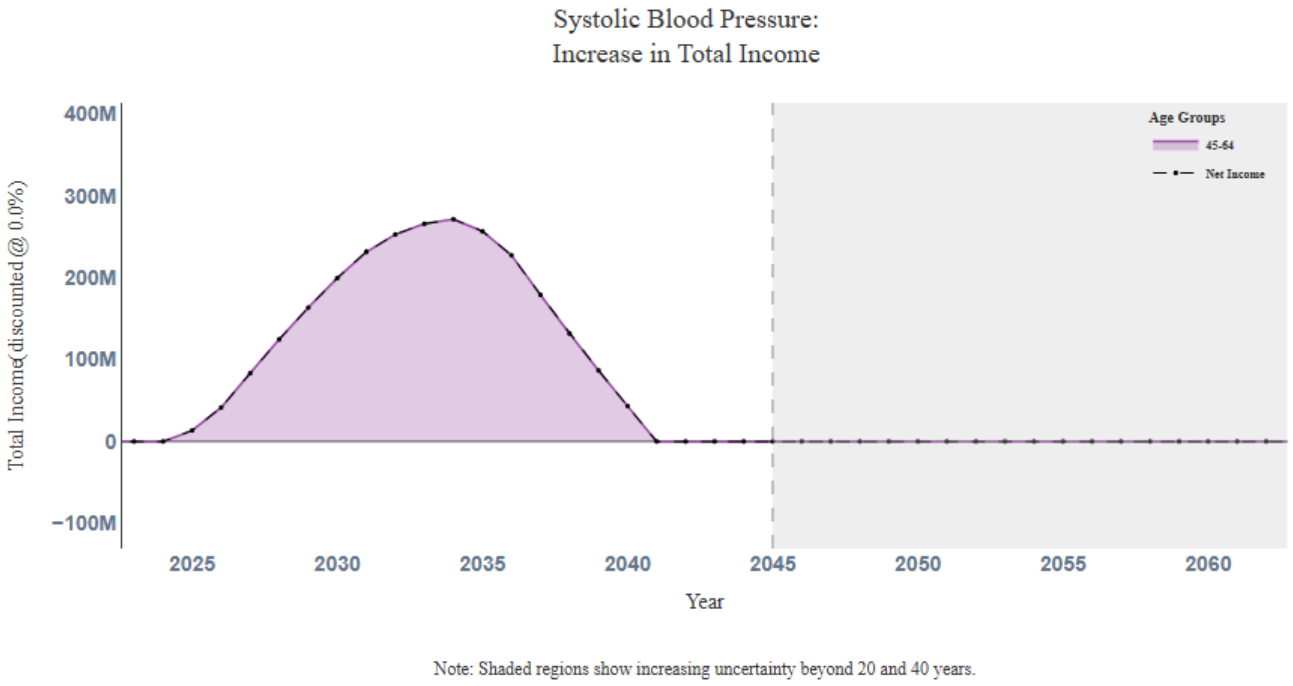

Figure 14.4: Yearly income gained, resulting from a 10% shift in systolic blood pressure towards the theoretical minimum risk level, for the 45-64 year old cohort in 2025, implemented over 10 years. (Source: SHINE HIIC tool).

14.5 Sensitivity analyses

A number of sensitivity analyses may be undertaken for intervention modelling, particularly for intervention effect sizes that have been transported from other settings (Chapter 9). These analyses are specific to the intervention at hand. However, there are also a number of standard sensitivity analyses that are run for all SHINE modelling projects. These are listed below:

- Risk factor-disease pair subset: based on the GBD Burden of Proof analyses, RF-disease pairs with 3+ star ratings are used as a sub-set in a sensitivity analysis.
- Extended time horizon: 20-40 year, and (optionally) lifetime (110 year) output. Lifetime model outputs are produced using a closed cohort (i.e., no births or net international migration), for the purpose of showing changes to life expectancy under an intervention. This is most relevant for interventions that affect youth and young people. This is discussed further in the next chapter (15).
- Multiple one-way sensitivity analyses of uncertain model inputs, presented as a Tornado plot (discussed below, Section 14.5.1).

14.5.1 Tornado plots

For all SHINE modeling analyses, multiple one-way-sensitivity analyses are run presented in a Tornado plot.

Tornado plots are used to show the impact of parameter uncertainty, from each uncertain parameter, to overall model output uncertainty.

The model is run multiple times, each with a single draw. For each model run, a different input parameter is varied to either its 2.5th or 97.5th percentile, with all other values held at their median value.

The tornado plot can help to determine the variables that contribute to the most uncertainty in modelled outcomes, therefore guiding where, if any, more evidence is needed to come to a decision about an intervention - i.e., a value of information analysis.[69]

An example Tornado plot is shown below (Figure 14.5), taken from a paper led by Ait Ouakrim et al. [50].

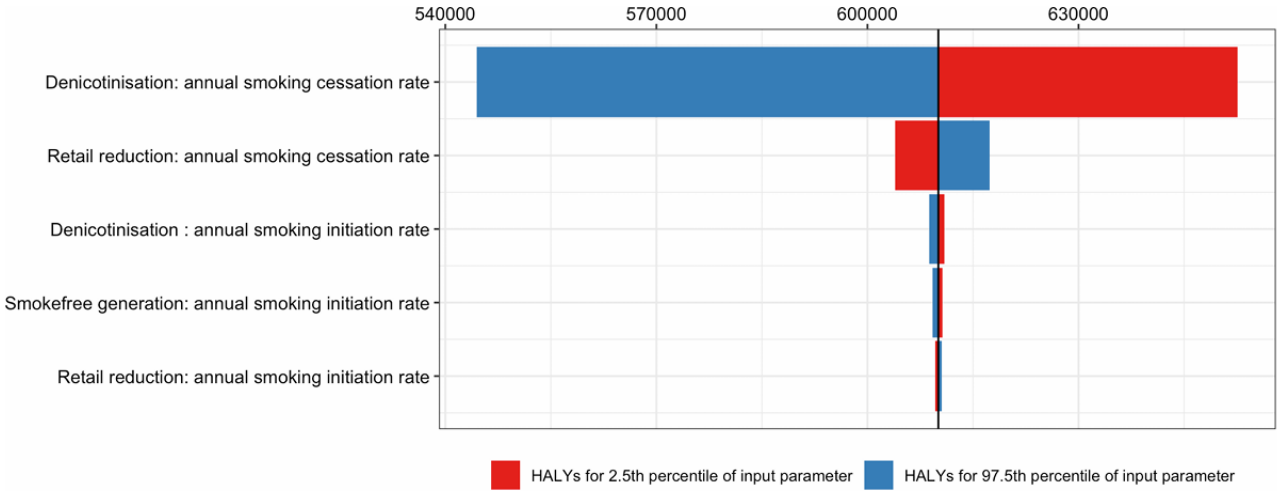

Figure 14.5: Tornado plot example, showing variation in HALYs gained under five intervention scenarios, when the effect size for each intervention is individually varied to the upper and lower 95% uncertainty interval value. (Source: Ait Ouakrim et al. [50]).

In this example, uncertainty around each parameterised intervention is shown. However, the Tornado plot can also be used to show uncertainty for BAU model parameters, including disease rates and risk factor-disease incidence RR values. For risk factor models where the exposure is linked to many diseases, it is not feasible to run sensitivity analyses varying the incidence, case fatality, and remission rates for each individual disease. In these cases, it is standard practice for SHINE to include in the Tornado plot for the three top diseases (those that have the greatest attributable burden from the risk factor of interest).

Disease rates can be varied simultaneously for all diseases, but this will produce uncertainty that is greater than the main analysis.

Chapter 15

Additional outputs

15.1 Life expectancy

Life expectancy (LE) at birth can be produced by running a closed cohort model for the full lifetime (110 years maximum) of a cohort of interest. This can then be used to show how life expectancy may change for a cohort under an intervention scenario. For example, if looking at an intervention starting in 2025 that targets 15-year-olds, life expectancy may be compared for the cohort born in the year 2000, under the intervention vs. BAU.

Life expectancy at birth, or remaining life expectancy at each subsequent age for a cohort, is calculated in a given year as the cumulative sum of future years lived over the remaining lifespan of the cohort, divided by the number alive in the current year. See Blakely et al.[61]) for an example of these lifetable calculations.

Importantly, this is cohort life expectancy; under BAU, sex and age-specific mortality rates, which determines the number of people alive in each year, change by calendar year based on projections of historic trends (using log-linear regression - Chapter 5). Comparatively, most burden of disease studies will produce period life expectancy estimates, assuming that current sex and age-specific mortality rates remain constant into the future.

15.1.1 HALEs and morbidity gap measurement

Health-adjusted life expectancy (HALE) estimates the duration of life lived at full health. It can be estimated using Sullivan's method (see ABDS report for more details [70]). This involves calculating HALYs, then using this as the denominator of life expectancy calculation (rather than total life-years). A further example is provided in Blakely et al.[61].

The difference between life expectancy and health-adjusted life expectancy represents the number of years lived at less than full health, or a morbidity gap. An analysis may assess whether an intervention 'compresses' morbidity [71] - i.e., whether the intervention increases the duration or proportion of life lived at full health to a greater extent than total life expectancy is improved.

LE and HALE may be calculated and compared at birth, in a closed cohort, or for the remaining lifespan of a particular cohort (again, in a closed cohort model).

15.2 Age-standardised results

The SHINE default is to age-standardise using the World Health Organisation (WHO)-produced standard population, to support between-country comparisons. [72]

15.2.1 Age-standardised HALY

While age-standardised HALYs are calculated within the PMSLT automatically, it is important to understand the process and what is actually being produced, given that unlike ACMR, HALYs are not a rate.

Age-standardised HALYs are calculated in the following process:

1. Taking age-specific HALYs from the intervention scenario and BAU, and person-years from BAU only. These values are aggregated across a desired time frame (e.g., 0-10 years, 10-20 years) and any strata for which age-standardised comparisons are not wanted (e.g., sex).
2. Dividing the HALYs under both scenarios by the person-years under BAU. Using the BAU person-years for both scenarios because if the intervention reduces mortality, the person-years will increase under the intervention, resulting in the HALYs being penalised when divided by a larger denominator.
3. Taking a weighted average over a standard population.
4. Multiply by 100,000 to present age-standardised HALYs per 100,000 person-years (under BAU).

Producing age-standardised HALYs is important when making comparisons across modelled sub-groups that may have different sizes and age-structures, e.g., across SEIFA strata. Both the relative and absolute difference in age-standardised HALYs should be presented.

15.2.2 All-cause mortality rates

While the main output of interest is usually deaths in absolute terms, comparisons may be made between sub-groups of different sizes and age distributions by using age-standardised all-cause mortality rates.

The all-cause mortality rate can be presented at multiple points in time, e.g., at 10, 20 and 40 years post-intervention start date, but they are not tallied over time like with age-standardized HALYs above, given that it is a rate.

Inequality gap

Inequality gaps can be examined through 'just' comparing the ACMRs themselves at multiple timepoints, or this can be extended to see how differences in the ACMR across sub-strata change over time. Consider a hypothetical intervention that reduces salt intake. Under BAU, in the base year, the ACMR in the most disadvantaged SEIFA strata is approximately 1.5 times that in the most advantaged strata. Under the intervention, ACMRs change for both stratum, but the reduction is more pronounced in the SES1 (most disadvantaged) group, for which there was higher average salt intake under BAU. As a result, 20-years after the intervention

first starts, the gap in all-cause mortality rates reduces from a relative difference of 1.5 to 1.2. In other words, a 20% reduction in the ACMR gap between SES1 and SES5.

An example of an ACMR gap analysis is shown below (Figure 15.1), taken from a paper led by Ait Ouakrim et al., which examined the impact of tobacco control interventions for the Māori and non-Māori populations in Aotearoa New Zealand.[50]

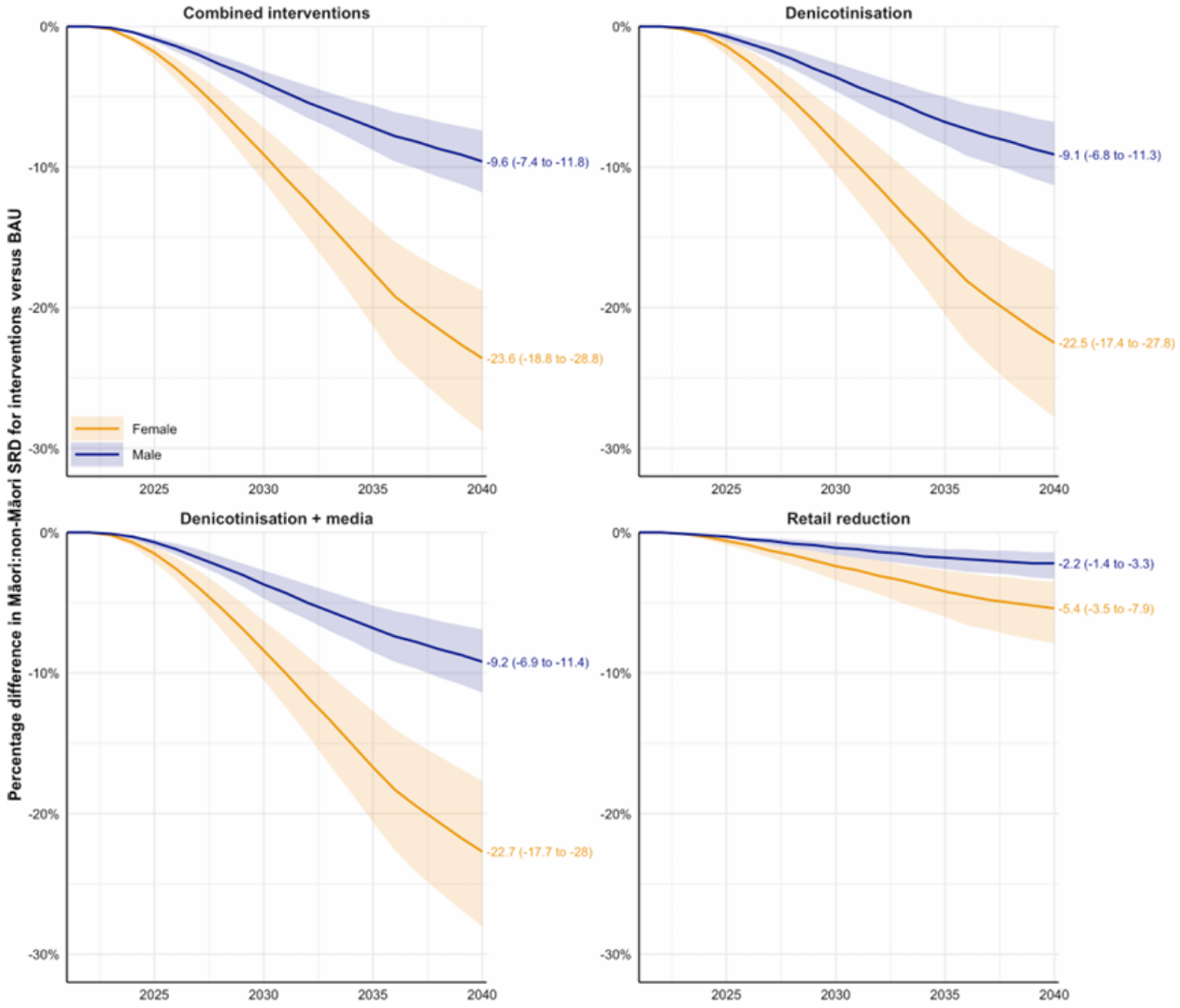

Figure 15.1: Change in all-cause mortality rate gap under intervention scenarios compared to BAU, by sex and Indigenous status (Māori vs. non-Indigenous), over time (Source: Ait Ouakrim et al. [50]).

In Figure 15.1, the 'Denicotinisation + media' intervention results in a 22.7% reduction in the gap in age-standardised all-cause mortality rates by 2040, relative to BAU. Importantly, this calculation takes into account changing mortality rates, and possible changes to the ACMR gap, occurring under BAU. The 22.7% difference is a reduction in the gap compared to the gap under BAU in 2040.

15.3 Economic evaluation

Economic evaluation outputs produced by the model can include:

- Cost per [risk factor distribution change]
- Cost per HALY gained
- Cost per death averted
- Net monetary benefit

As a default, economic evaluation costs are presented at a 3% discount rate. In Supplementary results, the results are presented at 0% and 6% discount rates. If the health outcome is HALYs gained, this is also discounted at the same rate as costs.

SHINE's default for economic evaluations is a health system perspective, including healthcare expenditure related to diseases, and specific intervention costs.

15.3.1 Cost-effectiveness plane

The results of a cost-utility or cost-effectiveness analysis can be presented on a cost-effectiveness plane. An example from the SHINE HIIC tool is shown below in Figure 15.2. This example compares 5 hypothetical interventions, all targeting the cohort of 45-64 year olds (both male and female) in 2025 for a 10 year duration, with varying intervention costs as follows:

- 10% reduction in body mass index (BMI), with a cost of $10 million AUD each year
- 10% reduction in systolic blood pressure (SBP), with a cost of $10 million AUD each year
- 10% reduction in LDL cholesterol, with a cost of $10 million AUD each year
- 10% reduction in ischaemic heart disease (IHD) case fatality rate, with a cost of $10 million AUD each year
- 10% reduction in ischaemic heart disease (IHD) case fatality rate, with a cost of $10 million AUD in the first year of implementation, and $100 million AUD in each subsequent year.

Health effects, and cost impacts (including offsets from changes to healthcare expenditure) are tallied over a 20 year period. Costs and HALYs are discounted at 3% per annum in this example.

The BMI targetting intervention is the most cost effective from a health system perspective - it is estimated to save $21,203 per HALY gained (as shown in the ICER in Figure 15.2). While LDL cholesterol is linked to both IHD and ischaemic stroke incidence, we can see from this graph that the health benefit of a 10% LDL cholesterol is less than a direct 10% reduction in IHD case fatality.

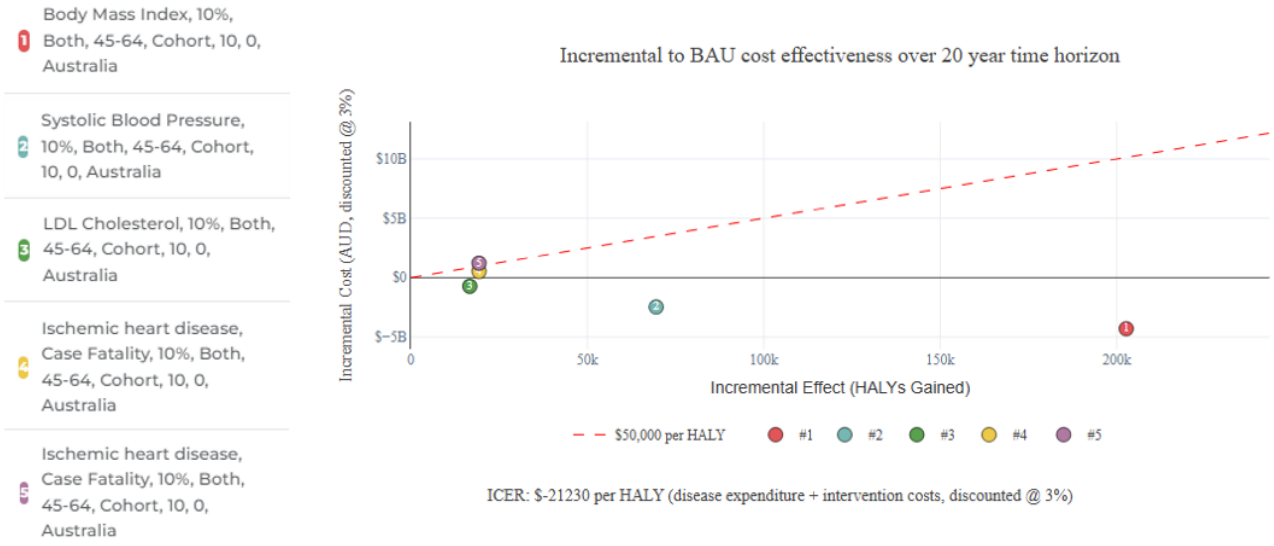

Figure 15.2: Comparison of cost-effectiveness across hypothetical cardiovascular disease/risk factor targeting interventions. (Source: SHINE HIIC tool).

Chapter 16

SHINE model validation and reporting

Validation and transparency are both important components to ensure the credibility of a simulation model.[18] Transparency has been discussed throughout this protocol. Here, we provide specific guidance for transparent reporting of SHINE modelling analyses, and introduce key concepts for SHINE model validation.

16.1 Validation

Validation considers how well a model is able to predict reality.

There are a number of types of validation, as discussed in the ISPOR modelling guidelines [18]:

- Face validity: using expert judgement to examine whether the model is functioning reasonably and producing reasonable output.
- Internal validity: comparing model outputs or components to the input data, i.e., ensuring that the model produces the information it was intended to.
- Cross validity: comparison of the outputs of a model to that from a different model
- External validity: comparing model outputs with real data, which was not used to as an input to the model.
- Predictive validity: comparing forecast results with the outcome that occurs in reality (at a later date)

16.1.1 SHINE approach to validation

SHINE uses multiple inputs, with models often including more than one component (e.g., a forecasting model or Markov model, alongside the PMSLT). Therefore, validation of multiple steps is generally required. A core part of SHINE's modelling process involves to conduct face validity checks at each stage of model development. This includes face validation of the model structure itself, the evidence/data used to inform the model, as well as the model output - under both BAU and intervention scenarios.

SHINE also conducts internal validity checks, to ensure that the model is producing outputs that are consistent with the input data. This ensures that the calculations within the model are occurring correctly.

External validation can be difficult for two reasons. First, there may be limited data available, so having an external source that was not used as the model input may be hard to source. Second, SHINE models the future - external validation would therefore either be of a model run in a historic time period to compare to real external data, or the comparison would be made to externally produced forecasts (i.e., be cross-validation).

Where possible, cross-validation should be considered when interpreting model findings (e.g., in the Discussion of a model paper). Cross-validation in this sense can be difficult, as inputs, time frames, and other decisions such as discount rate or the use of an open vs. closed cohort may differ between the different models to an extent where there cannot be valid comparisons made.

Where time and resources permit, collaboration with other research teams may be possible, in which key inputs and parameters are set, then different model structures are used to facilitate cross-validation. The 'Mount Hood Diabetes Challenge' does this for diabetes modelling [73, 74] - in 2025, SHINE was involved in an extension of the Mt Hood challenge for tobacco modelling (paper to come late 2026).

16.2 Reporting

Table 16.1 summarises key reporting recommendations to ensure transparency in SHINE modelling.

Table 16.1: SHINE model reporting guideline

Model/analysis component	Checklist item
Methods	
Model structure	Describe: • Model structure • Simplifications made in biological/other processes modelled
Model inputs (BAU)	List: • Input sources • Input uncertainty (distribution used, source of uncertainty/justification). This includes stating any parameters that do not have uncertainty, and why. • Heterogeneity. State which inputs include adjustment/disaggregation for heterogeneity. Justify any parameters that were not disaggregated across modelled strata.

Continued on next page

Model/analysis component	Checklist item
Model inputs (Intervention)	List: • Source • If intervention requires transporting (section 9.2): – List study sample, study pop., and target pop. – List threats to internal & external validity – Describe process used and adjustment made • Uncertainty (distribution, justification) • Specific population affected • Duration of intervention, and any attenuation of effect
Results	
Main results	For each outcome: • Report numeric outcomes, with large values (>10,000) rounded to four meaningful digits (e.g., 32,410 or 619,400). • State whether result is discounted/undiscounted • State timeframe
Sensitivity analysis	Include: • Discounted results • Tornado plot • Any sensitivity analyses for the intervention (only if modelling an 'actual' intervention)
Discussion	
Limitations	Limitations related to: • Model structure • BAU inputs • Intervention specification

Part VI

APPENDICES

Appendix A

Systematic error and quantitative bias analysis

A.1 Types of bias

There are three key types of bias in research, summarised below.

A.1.1 Selection bias

selection bias relates to how the sample population was obtained, allocated, and followed up. It occurs when the association between the exposure and outcome differs for those who participate compared to those who are eligible to participate. The main effect is from the joint association, or dependency, between the exposure and outcome - i.e., when participation in a study differs by both the exposure and outcome. There are different types of selection bias, for example

- Collider bias: type of selection bias caused by conditioning on a collider (or a descendent of a collider). Colliders are factors that are caused by both the exposure and outcome (see Figure A.1).
- M-bias: type of selection bias in which there is a collider that is caused by confounding co-variates (see Figure A.1

A.1.2 Information bias

information bias (also called measurement or misclassification error) is that arising due to poor measurement of the exposure, outcome, or other variables (including confounders).

It occurs when any of the below four parameters are less than 100%:

- Classification values:
  - Sensitivity (A/A+C): probability of a positive test among those with the outcome
  - Specificity (D/B+D): probability of a negative test among those without the outcome

	Outcome	No outcome
Test positive	A	B
Test negative	C	D

Table A.1: Classification of measurement error. Note that the 'outcome' could also be the measurement of an exposure.

- Predictive values:
  - Positive predictive value (PPV) (A/A+B): probability of a positive test result being a true positive
  - Negative predictive value (NPV) (D/C+D): probability of a negative test result being a true negative

There are two types of information bias:

- Non-differential: bias that is equally likely among cases and controls (or exposed and unexposed).
- Differential: bias that is more unequally likely among cases and controls (or exposed and unexposed).

A.1.3 Bias due to confounding

This is bias due to the presence of a confounder, which is associated with both the exposure (A) and outcome (B) of interest. The confounder does not lie on the causal pathway between A and B (see Figure A.1). If confounding is present in a study, then groups are not exchangeable, and therefore an identified association cannot be considered causal.

Residual bias: bias that is due to unknown, or unmeasured, confounders. While standard methods can be used to adjust for measured confounders, residual confounding needs to be addressed by contemporary approaches.

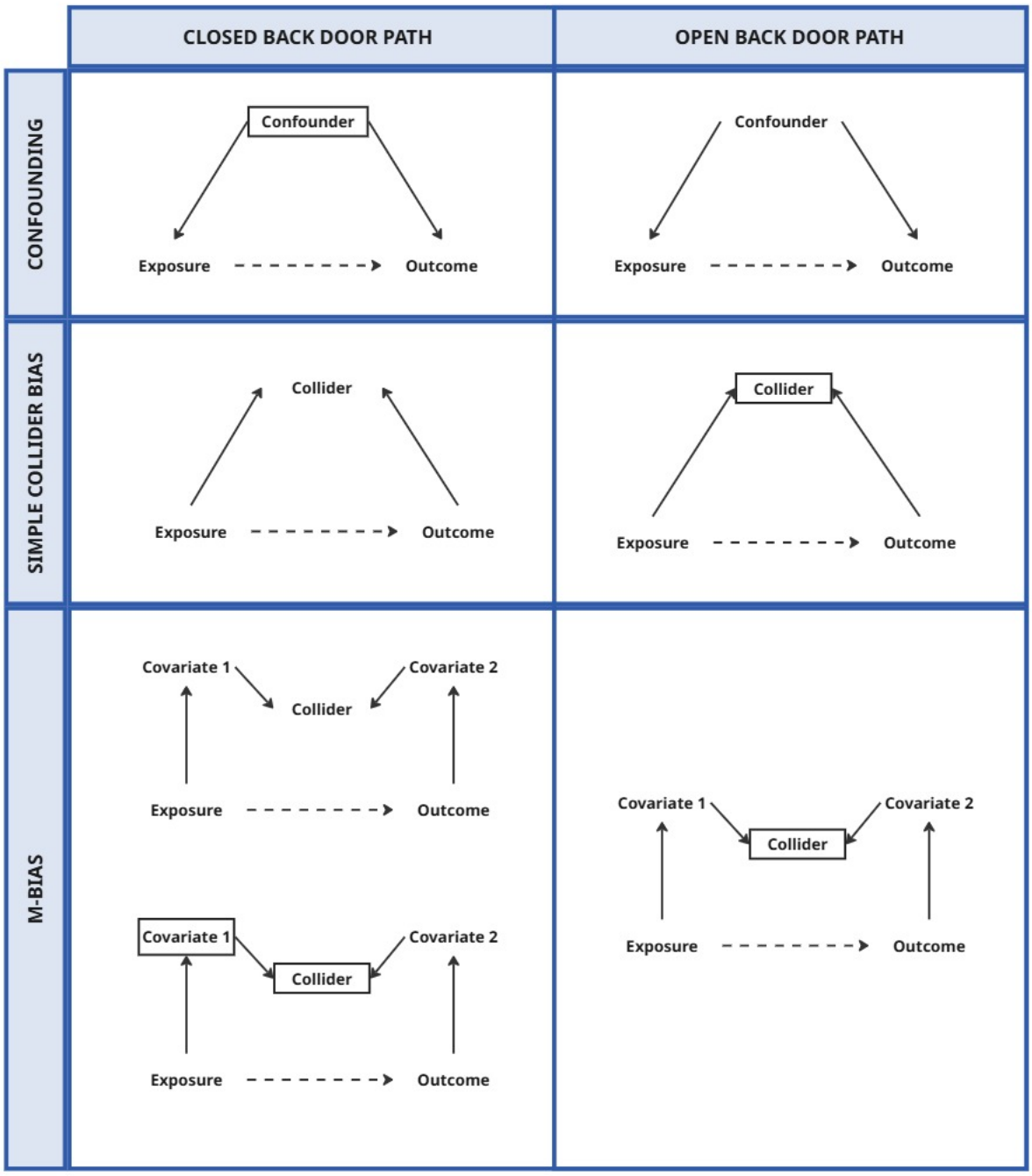

Figure A.1: Simple Directed Acyclic Graphs (DAG). Boxes represent variables that have been conditioned on.

A.2 Adjusting for bias

Table A.2 summarises available approaches to adjust for bias in analysis.

Table A.2: Bias adjustment approaches

Type of bias	Adjustment approaches
Selection bias	Quantitative bias analysis (QBA)
Information bias	QBA
Bias due to confounding	**Standard methods:** • Stratification • Standardisation • Restriction • Adjust in regression • Matching **Contemporary methods:** • G-methods – G-Computation – Inverse Probability (IP) weighting • Propensity scores • QBA

A.2.1 Quantitative Bias Analysis (QBA)

One such approach is **Quantitative bias analysis (QBA)**. QBA, often just referred to as sensitivity analysis, is a way to address the systematic error in quantitative research. It involves calculating both the direction and magnitude of systematic bias. QBA is detailed in the textbook from Lash et al. [75]. The authors have also provided instructional spreadsheets (access via website: `sites.google.com/site/biasanalysis/Home`) and code [76] that can be used to conduct these analyses. For this protocol, we provide a simple summary of the types of QBA and their application. Types of QBA:

- *Simple QBA*. Simple QBA involves adjustment of a single factor, e.g., adjusting for differential misclassification of the exposure, in an analysis. Bias parameters may be sourced from internal validation studies, or found from the literature (external validation). If not available through published literature, required parameters may also be estimated through expert opinion, or a range of best guess values may be used to determine the impact on the result of interest - this is multiple bias analysis.
- *Multiple bias analysis*. Multiple bias analysis involves adjusting for more than one factor. The order of adjustment is important to consider here: adjustment should occur in the reverse order in which the bias occurs, as measurement error $\rightarrow$ selection bias $\rightarrow$ confounding. However, this approaches becomes complicated if there are multiple sources of bias to be incorporated, particularly if there are multiple levels to a given input (e.g., non-binary exposure variable). This is where a probabilistic approach can be used.
- *Probabilistic bias analysis* This involves using a Monte Carlo simulation to draw from a distribution many times. This can be conducted for a single source of residual bias, or multiple (possibly correlated) in-

puts simultaneously. Importantly, the previous two approaches (simple or multiple bias analysis) do not involve calculating uncertainty around bias adjusted results. Probabilistic bias analysis is able to incorporate both the impact of bias on the central estimate, and uncertainty around that adjusted estimate. In probabilistic bias analysis, random error (i.e., the estimate of precision around the original result being analysed) can also be incorporated. The use of Monte Carlo simulation is also important if conducting an economic evaluation within a modelling study - rather than capturing a single ICER value, it can be calculated within each iteration of the simulation, and the probability of the intervention being cost-effectiveness estimated as the proportion of iterations in which the ICER fell below the set cost-effectiveness threshold.[77]

Appendix B

Heterogeneity

B.1 Overview

In the context of SHINE modelling, heterogeneity requires breaking down differences within a population such that the aggregate effect is unchanged. For example, the aggregate reported health-adjusted life year (HALY) in the business-as-usual (BAU) model should be the unaffected by whether all-cause morbidity rates are broken down by Socio-Economic Index For Areas (SEIFA). We implement heterogeneity to capture underlying differences that change how a population responds to an intervention, even at the aggregate level, and to report changes in outcomes by these broken-down levels.

B.1.1 Simple heterogeneity

For models that incorporate heterogeneity beyond sex by age, e.g., with SES incorporated, each individual stratum runs in parallel, and almost entirely independently. Each sub-strata has its own lifetable of population counts and rates, its own chronic disease prevalences, and its own disease rates.

The 'simple' heterogeneity system involves incorporating one additional stratification beyond sex and age, most commonly socio-economic status (SES). To break down disease rates by SES, rate ratios, shown in Table B.1 as a_s, are applied to aggregate rates, r. This process occurs within the PMSLT (see Chapter 11).

Table B.1: Heterogeneity inputs.[1]

Parameter	Purpose
Broken down population (p_s)	The population by the broken down strata, $s \in S$. This would be the population counts in the main lifetable, or the prevalence in a chronic disease. This is part of the input.
Aggregate population (p)	The total population of the cohort when aggregated. This is just $\sum p_s$ and as such is not a model input.

Continued on next page

Parameter	Purpose
Aggregate rate (r)	The rate experienced by the aggregate population. This is part of the input.
Rate ratios (a_s)	A number for each strata, $s \in S$, such that $r_i : r_j$ is the rate ratio of strata i and j. This is part of the input.
Broken down rates (r_s)	The rate experienced by each strata, $s \in S$. This is the desired output.

[1] S is a set of strata (e.g. $S = \{\text{SES1}, \text{SES2}, \text{SES3}, \text{SES4}, \text{SES5}\}$).

Table B.1 defines the components required to solve a set of inputs for heterogeneity. The PMSLT solves for the rates multiplier, m, such that

$$pr = m \sum_{i \in S} p_i a_i \tag{B.1}$$

so $r_s = ma_s$. This calculation is repeated for each age, sex and year, since the input parameters p, p_s, and r tend to vary over time. The interventions use the value of r_s calculated in the BAU, using BAU population counts, because r_s is effectively a model input. The rate ratios, a_s, are not necessarily expected to hold in non-BAU scenarios, and we do not need them do, as their only purpose is the derivation of the underlying BAU rates, r_s.

Heterogeneity on rates that cause changes in p (e.g. death or remission rates) use a modified form of (B.1) to solve. For example, in the case of all-cause mortality rates, we find m such that

$$p(1 - e^{-r}) = \sum_{i \in S} p_i (1 - e^{-ma_i}).$$

This can be solved by simple root finding techniques, as it is near-linear for realistic rates, and is solved by Andersen et al. [78].

B.1.2 Disease heterogeneity

For analyses in Australia, AIHW data on , SIHNE incorporates differences in each disease and all-cause factor by SEIFA, based on Australian Burden of Disease Study (ABDS) data. The heterogeneity system ensures that rates (incidence, remission, and case fatality) remain coherent with the GBD-derived prevalence projections, at the disaggregated and aggregate level. Note that the heterogeneity by SEIFA is assumed to be non-time varying.

B.1.3 Complex heterogeneity

Complex, or multi-dimensional, heterogeneity can also be incorporated into SHINE modelling, e.g., disaggregation of disease inputs by SES *and* geographic remoteness. This is an extension of the above automated process, requiring additional manipulation of GBD data for additional risk factors beyond SES.

This multi-dimensional heterogeneity model works with the assumption that there is no difference in relative risks for variable A (e.g., SES) across levels of variable B (e.g., remoteness), and that any observed crude variation is the result of different population sizes. This is a data intensive process, as information is required for each input variable (including disease rates, all-cause morbidity and mortality) across each strata of interest. Currently, this has only been applied for the SHINE Tobacco modelling in Australia [**?**].

Heterogeneity can be done along multiple dimensions (strata) simultaneously, however, we tend to only have rate ratios each individual strata. This causes a problem because populations can vary in size and distribution over multiple strata. For example, imagine that there are two strata; what people put on their toast and what they drink. We might be told that the rate ratio for jam is 1.5 that of butter, and the rate ratio of water is 1.5 that of milk. What is the rate ratio for people who have both jam and water, is it 2.25? It would conceivably be 2.25, if the population were split evenly between the four strata breakdowns, however the population may not be evenly split. If everyone in the population either has both jam and water, or both butter and milk, then the independent rate ratios of 1.5 were actually just measuring the same thing. In this case we would want people who eat butter and milk to have a rate ratio of 1.5 over jam and water.

Multi-dimensional heterogeneity solves this problem by using an optimisation process that tries to make the ratios across each strata match the targets, while trading the off against each other. It turns out we only need to find rate ratios across all strata breakdowns, and then use these ratios on the whole cohort in exactly the same way as single-dimensional heterogeneity (see Section B.1.1).

Let $\mathcal{S} = \{\mathbf{S}_1, \ldots, \mathbf{S}_n\}$ be the strata on which to apply heterogeneity. We will consider a single **Year** $\times$ **Age** $\times$ (**[Strata]** $- \mathcal{S}$) cohort, C, since heterogeneity happens independently across the aggregate strata.

The inputs for C are an aggregate target, T, the populations of the individual strata breakdowns, and target rate ratios $R_i = \{r_{i,1}, \ldots, r_{i,|\mathbf{S}_i|}\}$, for each $\mathbf{S}_i \in \mathcal{S}$. The aggregate target T is handled the same way as in single-dimensional heterogeneity , once we have a coherent set of rate ratios across all breakdowns of the strata in $\mathcal{S}$.

First define the set of strata breakdowns,

$$\mathcal{B} := \prod_{\mathbf{S}_i \in \mathcal{S}} \mathbf{S}_i,$$

where $\prod$ is the Cartesian product. We will need a rate ratio, t_B, for each full strata breakdown $B \in \mathcal{B}$. Let $v_{i,j}$ be the j^{th} variety of $\mathbf{S}_i \in \mathcal{S}$, $c(v_{i,j}) := \{B : B \ni v_{i,j}\}$ be the breakdowns that contain $v_{i,j}$, and V be the set of all $v_{i,j}$. Recall that for each variety $v_{i,j}$ of each strata, we have a target rate ratio $r_{i,j}$.

We can now define a "reasonable job" as follows. Let P_B be the underlying population of $B \in \mathcal{B}$, then the aggregate rate ratio for $v_{i,j} \in V$ is

$$a_{i,j} := \frac{\sum_{B \in c(v_{i,j})} P_B t_B}{\sum_{B \in c(v_{i,j})} P_B}.$$

This is just the population-weighted average of the rate ratios over the strata. We want to compare these to our targets, $r_{i,j}$, but to avoid biasing the comparison by our choice of reference group, we first weight our rate ratios before comparing them. So define the weighted rate ratios

$$a'_{i,j} := \frac{|\mathbf{S}_i| a_{i,j}}{\sum_{j \in [|\mathbf{S}_i|]} v_{i,j}},$$
$$r'_{i,j} := \frac{|\mathbf{S}_i| r_{i,j}}{\sum_{j \in [|\mathbf{S}_i|]} r_{i,j}}.$$

Then the objective to be minimised, the notion of "reasonable job", is

$$m := \sum_{\mathbf{S}_i \in \mathcal{S}} \frac{1}{|\mathbf{S}_i|} \sum_{v_{i,j} \in \mathbf{S}_i} \ln\left(\frac{a'_{i,j}}{r'_{i,j}}\right)^2.$$

We have the objective, now all we need are the constraints and the algorithm. We have chosen a single fudge factor, $f \in \mathbb{R}^+$, to optimise over, with aggregate rates defined by

$$t_B := \left(\prod_{v_{i,j} \in B} r_{i,j}\right)^f.$$

For example, if we have **Sex** and **SEIFA** as our heterogeneity strata ($\mathcal{S} = \{\mathbf{Sex}, \mathbf{SEIFA}\}$) and the following rate ratios.

	Female	Male		SES1	SES2	SES3	SES4	SES5
Sex	1	1.4	Seifa	1	1.1	1.2	1.3	1.4

Then t_B for each $B \in \mathcal{B}$ is defined by our fudge factor, f, as follows.

	SES1	SES2	SES3	SES4	SES5
Female	1^f	1.1^f	1.2^f	1.3^f	1.4^f
Male	1.4^f	1.54^f	1.68^f	1.82^f	1.96^f

Let P_B, the populations for our example, be the following.

	SES1	SES2	SES3	SES4	SES5
Female	100	90	160	80	140
Male	120	70	50	110	130

Then we have seven formulas for $a_{i,j}$.

$$\begin{aligned}
a_{\text{SES1}} &= (100 + 120 \times 1.4^f)/(100 + 120) \\
a_{\text{SES2}} &= (90 \times 1.1^f + 70 \times 1.54^f)/(90 + 70) \\
a_{\text{SES3}} &= (160 \times 1.2^f + 50 \times 1.68^f)/(160 + 50) \\
a_{\text{SES4}} &= (80 \times 1.3^f + 110 \times 1.82^f)/(80 + 110) \\
a_{\text{SES5}} &= (140 \times 1.4^f + 130 \times 1.96^f)/(140 + 130) \\
a_{\text{Female}} &= (100 + 90 \times 1.1^f + \ldots)/(100 + 90 + \ldots) \\
a_{\text{Male}} &= (120 \times 1.4^f + 70 \times 1.54^f + \ldots)/(120 + 70 + \ldots)
\end{aligned}$$

The weighted average, a'_{B} is taken within each group, so for example,

$$a'_{\text{SES1}} = \frac{5a_{\text{SES1}}}{\sum_{i\in\{1,\ldots,5\}} a_{\text{SES}i}}$$

Then f is optimised to minimise

$$(\ln(a'_{\text{SES1}}/r'_{\text{SES1}})^2 + \ln(a'_{\text{SES12}}/r'_{\text{SES2}})^2 + \ldots)/5 + (\ln(a'_{\text{Female}}/r'_{\text{Female}})^2 + \ln(a'_{\text{Male}}/r'_{\text{Male}})^2)/2$$

The result turns out to be $f \approx 1.0289$, so the final t_B is as follows.

	SES1	SES2	SES3	SES4	SES5
Female	1	1.103	1.206	1.310	1.414
Male	1.414	1.559	1.705	1.852	1.999

Once we t_B across all breakdowns, we can treat the breakdowns as belonging to a single strata, and do single-dimensional heterogeneity as per Section B.1.1. Note that $f = 1$ if the populations are all the same size, so the optimisation process only required to take correlations in population size across strata into account.

Appendix C

PMSLT impact attribution

C.1 Overview

The PMSLT is generally run with more than one disease, and often includes many. For example, SHINE Tobacco modelling includes 31 tobacco-related diseases.[68] The PMSLT outputs changes under an intervention scenario, in terms of morbidity and mortality, as the combined effect across all included diseases.

However, we may want to also disaggregate the effects of an intervention into the health-adjusted life year (HALY) gains attributable to each disease. To calculate this the PMSLT could be rerun for each individual disease, but this is computationally expensive and discounts the diminishing returns inherent in intervening on many diseases simultaneously. SHINE solves these issues by calculating attributable health gains to individual disease within a single model run.

Impact attribution calculations exploit the linearity of mortality and morbidity shifts shown in Section 11.6. Keep in mind that this linearity is only ever local to the space around the business-as-usual (BAU), so the impact attribution *only* ever breaks down differences between the intervention and BAU. At no point do we determine the total health impact of a disease within a single scenario.

The following sections describe impact attribution calculations in detail. See Box C.1 for some generic symbols used throughout.

Box C.1.1. Symbol glossary

- $\approx$ means approximately equal, and is often paired with an explanation or constraint for when the approximation holds.
- $:=$ means "defined to be". It is used when a new quantity is being defined in equation form. It is distinct from $=$ in the sense that equality is making a mathematical claim, while $:=$ is introducing notation.
- $\in$ means "is an element of (a set)". For example, $S \in \{\text{INT}, \text{BAU}\}$ means consider S to be either INT or BAU, whatever follows holds for both assignments of S.
- $'$, as in x', is just a way to name variables. The variable x' will be closely related to x, by convention.

C.2 Deaths

The first step of attribution is to determine how much each disease is responsible for changes in deaths. Fix a particular year-of-birth cohort C. For year y and scenarios $S \in \{\text{BAU}, \text{INT}\}$, let $p_{S,y}$ be the population at the end of year y and $m_{S,y}$ be the deaths during year y. Denote the difference in deaths between the scenarios, for year y, by

$$\Delta m_y := m_{\text{INT},y} - m_{\text{BAU},y}.$$

Let D be a set of diseases. The goal of attribution is to find $m_{d,y}$, for $d \in D$ and each year y, such that

$$\begin{aligned} \Delta m_y &= \sum_{d \in D} m_{d,y} \\ [\Delta m_y]_{d\text{-only model}} &\approx m_{d,y}. \end{aligned}$$

where $[\ldots]_{d\text{-only model}}$ is a hypothetical model that only includes disease d. In other words, the change in deaths attributed to disease d, $m_{d,y}$, should approximately match the total Δm_y for a model that only includes disease d. Note that this agreement can only ever be approximate, due to competing mortality in the full model.

Let $a_{S,y}$ be the all-cause mortality rate of C in scenario S. As per Section 11.3.2, the main lifetable evolves as follows.

$$\begin{aligned} m_{S,y} &= p_{S,y-1}\left(1 - e^{-a_{S,y}}\right) \\ p_{S,y} &= p_{S,y-1} - m_{S,y}. \end{aligned}$$

The value of $a_{\text{BAU},y}$ is sourced from input files, while

$$a_{\text{INT},y} := a_{\text{BAU},y} + \sum_{d \in D} \Delta a_{d,y}$$

where $\Delta a_{d,y}$ is the disease-specific mortality rate shift, as per Section 11.6.1. Then, expanding the above,

$$\begin{aligned} \Delta m_y &= p_{\text{INT},y-1}\left(1 - e^{-a_{\text{INT},y}}\right) - p_{\text{BAU},y-1}\left(1 - e^{-a_{\text{BAU},y}}\right) \\ &= p_{\text{INT},y-1}\left(1 - e^{-(a_{\text{BAU},y} + \sum_{d \in D} \Delta a_{d,y})}\right) - p_{\text{BAU},y-1}\left(1 - e^{-a_{\text{BAU},y}}\right) \end{aligned}$$

Now consider just the first year of the model, year 1. The model requires initial population data, which we notate here as the model state at the end of year 0. The initial population in the intervention and BAU are the same, so $p_{\text{INT},0} = p_{\text{BAU},0}$. Therefore,

$$\begin{aligned} \Delta m_1 &= p_{\text{INT},0}\left(1 - e^{-a_{\text{INT},1}} - \left(1 - e^{-a_{\text{BAU},1}}\right)\right) \\ &= p_{\text{INT},0}\left(1 - e^{-(a_{\text{BAU},1} + \sum_{d \in D} \Delta a_{d,1})} - \left(1 - e^{-a_{\text{BAU},1}}\right)\right) \\ &= p_{\text{INT},0}\left(e^{-a_{\text{BAU},1}} - e^{-(a_{\text{BAU},1} + \sum_{d \in D} \Delta a_{d,1})}\right) \\ &= p_{\text{INT},0} e^{-a_{\text{BAU},1}}\left(1 - e^{-\sum_{d \in D} \Delta a_{d,1}}\right) \\ &\approx p_{\text{INT},0} e^{-a_{\text{BAU},1}} \sum_{d \in D}\left(1 - e^{-\Delta a_{d,1}}\right) =: \Delta m_1' \end{aligned}$$

with the final step depending on $x \approx 1 - e^x$ for $x \approx 0$. This approximation is sound because the *difference* in population-level disease-specific mortality rate between scenarios tends to be quite small.

We now have an approximate decomposition of Δm_1 by the effect of each disease:

$$\Delta m_1' = \sum_{d \in D} p_{\text{INT},0} e^{-a_{\text{BAU},1}} \left(1 - e^{-\Delta a_{d,1}}\right)$$

To deal with the approximation, and to avoid compounding errors, we scale the sum by $\Delta m_1 / \Delta m_1'$. The result is the solution

$$m_{d,1} = p_{\text{INT},0} e^{-a_{\text{BAU},1}} \left(1 - e^{-\Delta a_{d,1}}\right) \frac{\Delta m_1}{\Delta m_1'}$$

which satisfies $\Delta m_1 = \sum_{d \in D} m_{d,1}$, as required.

C.2.1 Deaths in subsequent years

In general, the model runs for more than one year, which complicates matters since $p_{\text{INT},y} \neq p_{\text{BAU},y}$ for years $y \geq 1$. How do we get around this? Conceptually, the solution is to use death attribution from previous years to make diseases responsible for the future death of individuals who would have died earlier in the BAU. The result is zero long term net death attribution in a closed cohort, as expected. This works because changes in population size are entirely driven by changes in mortality,

$$p_{\text{INT},y} = p_{\text{BAU},y} + \sum_{i \leq y} \Delta m_i,$$

and we have a breakdown of Δm_i by disease. Denote the change in population caused by disease d, by the end of year y, by

$$\Delta p_{d,y} := \sum_{i \leq y} m_{d,i}$$

and note that

$$\sum_{d \in D} \Delta p_{d,y} = \sum_{d \in D} \sum_{i \leq y} m_{d,i} = \sum_{i \leq y} \Delta m_i = p_{\text{INT},y} - p_{\text{BAU},y}.$$

With the identity above, we can derive the general form of death attribution. Starting at the definition of Δm_y,

$$\begin{aligned}
\Delta m_y &= p_{\text{INT},y-1} \left(1 - e^{-a_{\text{INT},y}}\right) - p_{\text{BAU},y-1} \left(1 - e^{-a_{\text{BAU},y}}\right) \\
&= p_{\text{INT},y-1} \left(1 - e^{-a_{\text{INT},y}}\right) - \left(p_{\text{INT},y-1} - \sum_{d \in D} \Delta p_{d,y-1}\right) \left(1 - e^{-a_{\text{BAU},y}}\right) \\
&= p_{\text{INT},y-1} \left(1 - e^{-a_{\text{INT},y}}\right) - p_{\text{INT},y-1} \left(1 - e^{-a_{\text{BAU},y}}\right) + \sum_{d \in D} \Delta p_{d,y-1} \left(1 - e^{-a_{\text{BAU},y}}\right).
\end{aligned}$$

We know how to handle the first two terms since they match terms from the calculation of Δm_1, so

$$\begin{aligned}\Delta m_y &= p_{\mathsf{INT},y-1}e^{-a_{\mathsf{BAU},y}}\left(1-e^{-\sum_{d\in D}\Delta a_{d,y}}\right)+\sum_{d\in D}\Delta p_{d,y-1}\left(1-e^{-a_{\mathsf{BAU},y}}\right)\\ &\approx p_{\mathsf{INT},y-1}e^{-a_{\mathsf{BAU},y}}\sum_{d\in D}\left(1-e^{-\Delta a_{d,y}}\right)+\sum_{d\in D}\Delta p_{d,y-1}\left(1-e^{-a_{\mathsf{BAU},y}}\right)\\ &= \sum_{d\in D}\left(p_{\mathsf{INT},y-1}e^{-a_{\mathsf{BAU},y}}\left(1-e^{-\Delta a_{d,y}}\right)+\Delta p_{d,y-1}\left(1-e^{-a_{\mathsf{BAU},y}}\right)\right)=:\Delta m'_y.\end{aligned}$$

This gives us $\Delta m'_y$ expressed as the sum of contributions from each disease, and the two terms of the sum even have neat interpretations. The first term should be familiar from the previous section; it is the difference, due to the change in death rate, felt by the intervention population. The second term counts the change in background deaths felt due to changes in the population size.

Note that interventions will tend to reduce the direct deaths caused by disease, increasing the population, so $\Delta p_{d,y}$ will generally be positive. This makes the second term positive which implies that preventing deaths in past years causes those deaths to be attributed to the disease in future years. This is entirely as expected, and in fact, if you run a full closed cohort simulation, then $m_{d,y}$ hits zero in the final year, because there is no change to disease rates that prevents death in the long run.

In any case, we can now solve for $m_{d,y}$, using the same renormalisation trick as before.

$$m_{d,y} = \left(p_{\mathsf{INT},y-1}e^{-a_{\mathsf{BAU},y}}\left(1-e^{-\Delta a_{d,y}}\right)+\Delta p_{d,y-1}\left(1-e^{-a_{\mathsf{BAU},y}}\right)\right)\frac{\Delta m_y}{\Delta m'_y}$$

C.2.2 Direct deaths

Direct death attribution tracks the deaths directly averted by the disease, without following up on the eventual fate of those saved. This is a dangerous output as it over-reports the effect of an intervention. Nevertheless, it exists, and is required as a part of economic analysis, so here it is.

Let Δc_y be the difference in 'direct deaths' for year y, defined as follows

$$\Delta c_y := p_{\mathsf{INT},y-1}\left(1-e^{-a_{\mathsf{INT},y}}\right)-p_{\mathsf{INT},y-1}\left(1-e^{-a_{\mathsf{BAU},y}}\right),$$

and let us find $c_{d,y}$ such that $\Delta c_y = \sum_{d\in D} c_{d,y}$ and $\Delta c_1 = \Delta m_1$, because direct deaths should agree in the first year.

The difference between Δc_y and Δm_y is that the former assumes that the intervention is only coming into effect at year y, i.e. $p_{\mathsf{INT},y-1} = p_{\mathsf{BAU},y-1}$. Appropriately, Δc_y is exactly the single-year case, which we have already solved, so

$$\Delta c_y \approx p_{\mathsf{INT},y-1}e^{-a_{\mathsf{BAU},y}}\sum_{d\in D}\left(1-e^{-\Delta a_{d,y}}\right)=:\Delta c'_y$$

Then with renormalisation

$$c_{d,y} = p_{\mathsf{INT},y-1}e^{-a_{\mathsf{BAU},y}}\left(1-e^{-\Delta a_{d,y}}\right)\frac{\Delta c_y}{\Delta c'_y}.$$

C.3 Health-adjusted life years

Attributing HALYs depends on death attribution since deferred deaths generate HALYs. However, disease-specific disability rate is truly linear, so the remaining steps are simpler. For year y and scenario S, let $b_{S,y}$ be the disability rate in year y and $h_{S,y}$ be the HALYs accrued in year y, for a fixed year-of-birth cohort C. Denote the difference in HALYs accrued between an intervention and its BAU in year y by

$$\Delta h_y := h_{\text{INT},y} - h_{\text{BAU},y}.$$

As with deaths, we need to find $h_{d,y}$, for each disease $d \in D$ and year y, such that

$$\Delta h_y = \sum_{d \in D} h_{d,y}$$
$$[\Delta h_y]_{d\text{-only model}} \approx h_{d,y}$$

As per Section 11.3.2, HALYs are calculated as

$$h_{S,y} = \left(p_{S,y} + \frac{m_{S,y}}{2}\right)(1 - b_{S,y}).$$

As per Section 11.6.2, the disability rate in an intervention is

$$b_{\text{INT},y} := b_{\text{BAU},y} + \sum_{d \in D} \Delta b_{d,y}$$

where $\Delta b_{d,y}$ is the change in disability rate caused by disease d. To retain our sanity in what follows, define the *ability rate* $b'_{S,y} := 1 - b_{S,y}$, so

$$b'_{\text{INT},y} = 1 - b_{\text{BAU},y} - \sum_{d \in D} \Delta b_{d,y} = b'_{\text{BAU},y} - \sum_{d \in D} \Delta b_{d,y}.$$

This is all linear, so we can just crank the handle.

$$\begin{aligned}
\Delta h_y &= \left(p_{INT,y} + \frac{m_{INT,y}}{2}\right) b'_{INT,y} - \left(p_{BAU,y} + \frac{m_{BAU,y}}{2}\right) b'_{BAU,y} \\
&= \left(p_{INT,y} + \frac{m_{INT,y}}{2}\right) \left(b'_{\text{BAU},y} - \sum_{d \in D} \Delta b_{d,y}\right) \\
&\quad - \left(p_{INT,y} + \frac{m_{INT,y}}{2} - \sum_{d \in D} \left(\Delta p_{d,y} + \frac{m_{d,y}}{2}\right)\right) b'_{BAU,y} \\
&= -\sum_{d \in D} \Delta b_{d,y} \left(p_{INT,y} + \frac{m_{INT,y}}{2}\right) + \sum_{d \in D} \left(\Delta p_{d,y} + \frac{m_{d,y}}{2}\right) b'_{BAU,y} \\
&= \sum_{d \in D} \left(\left(\Delta p_{d,y} + \frac{m_{d,y}}{2}\right) b'_{BAU,y} - \Delta b_{d,y} \left(p_{INT,y} + \frac{m_{INT,y}}{2}\right)\right)
\end{aligned}$$

This gives us a decomposition of Δh_y into contributions from each disease d.

$$h_{d,y} = \left(\Delta p_{d,y} + \frac{m_{d,y}}{2}\right)(1 - b_{BAU,y}) - \Delta b_{d,y} \left(p_{INT,y} + \frac{m_{INT,y}}{2}\right).$$

The first term is the HALY gain due to changes in population size, while the second term is the HALY gain due to a reduction in the disability rate. Note that the second term contains $p_{INT,y}$, which makes up the

difference between the BAU and intervention disability rates in the first term. Using $(1 - b_{INT,y})$ in the first term would double count the HALYs gained by individuals who live to enjoy the lower disability rate of the intervention.

The equation otherwise tells us what we already know. Averted deaths generate HALYs, and $b_{d,y}$ tends to be negative, so HALYs will increase over time.

C.4 Cost attribution

As per Section 11.8.1, total cost for scenario S is in year y is

$$\text{cost}_{S,y} = v_{S,y}p_{S,y} + t_{S,y}m_{S,y}$$

where $v_{S,y}$ is the cost rate for surviving a year, and $t_{S,y}$ is the cost per death during the year. The survival cost rate is calculated similarly to the disability rate, so the attribution of this cost is similar to what we have seen before. The death cost rate is unique in that the PMSLT does not calculate it. Rather, attribution is used to determine the change in total death cost, and this change is then used to calculate a rate.

Cost attribution for survival and death are calculated separately, and then combined to give the total cost difference attributable to each disease.

C.4.1 Survival cost

Define the survival cost and the survival cost delta

$$\begin{aligned} s_{S,y} &:= v_{S,y}p_{S,y} \\ \Delta s_y &:= s_{INT,y} - s_{BAU,y}. \end{aligned}$$

As before, the goal of attribution is to find $s_{d,y}$ such that

$$\begin{aligned} \Delta s_y &= \sum_{d \in D} s_{d,y} \\ [\Delta s_y]_{d\text{-only model}} &\approx s_{d,y} \end{aligned}$$

where $s_{d,y}$ is the difference in survival cost attributable to disease d in year y. The change in the survival cost rate is

$$v_{\text{INT},y} = v_{\text{BAU},y} + \sum_{d \in D} s_{d,y}$$

where $s_{d,y}$ is the change in survival cost rate caused by disease d. The calculation for per-disease survival cost rate within a scenario can be found in Section 11.8.2. This is the same setup as for HALYs, except simpler

because we no longer have to worry about person years. Here are the calculations again for completeness.

$$\begin{aligned}\Delta s_y &= p_{INT,y} v_{INT,y} - p_{BAU,y} v_{BAU,y} \\ &= p_{INT,y} \left(v_{\text{BAU},y} + \sum_{d \in D} \Delta v_{d,y} \right) - \left(p_{INT,y} - \sum_{d \in D} \Delta p_{d,y} \right) v_{BAU,y} \\ &= \sum_{d \in D} \Delta v_{d,y} p_{INT,y} + \sum_{d \in D} \Delta p_{d,y} v_{BAU,y} \\ &= \sum_{d \in D} \left(\Delta p_{d,y} v_{BAU,y} + \Delta v_{d,y} p_{INT,y} \right)\end{aligned}$$

Therefore

$$s_{d,y} = \Delta p_{d,y} v_{BAU,y} + \Delta v_{d,y} p_{INT,y}.$$

C.4.2 Death cost

Death costs are calculated directly via attribution, then later divided by the deaths each year to generate an average rate. This is necessary because the average death cost is increased by interventions that reduce deaths attributable to diseases that are relatively cheap to die from. Indeed, if a mix of cheap and expensive diseases are in the model, and an intervention reduces the deaths attributable to one of them, then it is impossible to know whether the cost per death should increase or decrease without knowing the type of death averted.

Recall that $c_{d,y}$ is the direct deaths attributable to disease d in year y, as per Section C.2.2. Let $\ell_{d,y}$ be the death cost rate of disease d in year y. Note that this cost cannot vary by scenario. We define the difference in direct death cost, between the intervention and BAU, to be

$$\Delta t_y^{\text{direct}} := \sum_{d \in D} c_{d,y} \ell_{d,y}.$$

This is the cost difference due to direct deaths, but it ignores the costs incurred by deferred deaths. By definition there are $m_{\text{INT},y}$ deaths in the intervention, and there are

$$\Delta c_y := \sum_{d \in D} c_{d,y}$$

direct deaths, so we are yet to account for the expenditure of $(m_{\text{INT},y} - \Delta c_y)$ deaths. We can apply the population-level death cost rate, ℓ_{BAU}, since these are the deaths that were unaffected by the intervention, implying that the distribution of death type matches that of the deaths in the BAU. So the total death cost of the intervention is

$$t_{\text{INT},y} = \left(m_{\text{INT},y} - \Delta c_y \right) \ell_{\text{BAU},y} + \Delta t_y^{\text{direct}}.$$

Now to calculate the new death expenditure rate, $\ell_{\text{INT},y}$, we divide the total cost by the number of deaths

$$\ell_{\text{INT},y} = \frac{t_{\text{INT},y}}{m_{\text{INT},y}}.$$

The PMSLT routinely outputs $\ell_{\text{INT},y}$ and $t_{\text{INT},y}$, yet the former, the rate, is derived from the total.

Now that we have $\ell_{\text{INT},y}$, attributing the death cost is relatively straightforward. We want to find $t_{d,y}$, for $d \in D$ and each year y, such that

$$\Delta t_y = \sum_{d \in D} t_{d,y}$$
$$[\Delta t_y]_{d\text{-only model}} \approx t_{d,y}.$$

and, by definition,

$$\begin{aligned}
\Delta t_y &:= t_{\text{INT},y} - t_{\text{BAU},y} \\
&= (m_{\text{INT},y} - \Delta c_y)\, \ell_{\text{BAU},y} + \Delta t_y^{\text{direct}} - m_{\text{BAU},y} \ell_{\text{BAU},y} \\
&= (m_{\text{INT},y} - m_{\text{BAU},y} - \Delta c_y)\, \ell_{\text{BAU},y} + \Delta t_y^{\text{direct}} \\
&= (\Delta m_y - \Delta c_y)\, \ell_{\text{BAU},y} + \Delta t_y^{\text{direct}}.
\end{aligned}$$

Sections C.2.1 and C.2.2 let us expand Δm_y and Δc_y to

$$\begin{aligned}
\Delta t_y = \Bigg(& p_{\text{INT},y-1} e^{-a_{\text{BAU},y}} \left(1 - e^{-\sum_{d \in D} a_{d,y}}\right) + \sum_{d \in D} \Delta p_{d,y-1} \left(1 - e^{-a_{\text{BAU},y}}\right) \\
& - p_{\text{INT},y-1} e^{-a_{\text{BAU},y}} \left(1 - e^{-\sum_{d \in D} a_{d,y}}\right) \Bigg) \ell_{\text{BAU},y} + \Delta t_y^{\text{direct}}.
\end{aligned}$$

As you might expect, Δc_y cancels out with the direct component of Δm_y, giving us

$$\begin{aligned}
\Delta t_y &= \sum_{d \in D} \Delta p_{d,y-1} \left(1 - e^{-a_{\text{BAU},y}}\right) \ell_{\text{BAU},y} + \Delta t_y^{\text{direct}} \\
&= \sum_{d \in D} \left(\Delta p_{d,y-1} \left(1 - e^{-a_{\text{BAU},y}}\right) \ell_{\text{BAU},y} + c_{d,y} \ell_{d,y}\right)
\end{aligned}$$

which decomposes to

$$t_{d,y} = \Delta p_{d,y-1} \left(1 - e^{-a_{\text{BAU},y}}\right) \ell_{\text{BAU},y} + c_{d,y} \ell_{d,y}.$$

The attribution is much like those of previous sections, except with deaths instead of population. The first term is the cost attributed to changes in the population size, using the death rate and cost of the BAU. The second term takes the difference in deaths within the current year into account. note that the first term will often be positive, while the second term will often be negative.

Appendix D

SHINE time lag specification

D.1 Disease time lags

Diseases are assigned one of the above durations at GBD's level 2 (and occasionally level 3) designation, as shown in Table D.1 below.

Table D.1: Disease time lags

Cause - Level 2 Cause - Level 3	Category
HIV/AIDS and sexually transmitted infections	Nil
Respiratory infections and tuberculosis	Nil
Respiratory infections and tuberculosis	Nil
Neglected tropical diseases and malaria	Nil
Other infectious diseases	Nil
Maternal and neonatal disorders	Nil
Nutritional deficiencies	Nil
Neoplasms	Long
Cardiovascular diseases	Short
Chronic respiratory diseases	Long
Digestive diseases	Long
Neurological disorders	Long
Mental disorders	Nil
Substance use disorders	Nil

Continued on next page

Cause - Level 2 Cause - Level 3	Category
Diabetes and kidney diseases	-
Diabetes	Short
Kidney diseases	Long
Skin and subcutaneous diseases	Nil
Sense organ diseases	Short
Musculoskeletal disorders	Short
Other non-communicable diseases	-
Congenital birth defects	Not applicable
All other level 3 causes and conditions within 'other NCD'	Short
Transport injuries	Nil
Unintentional injuries	Nil
Self-harm and interpersonal violence	Nil

D.2 Short time lags

We first define the duration of decay, L. For a median value of $L = 5$, with an uncertainty distribution on the ln-normal scale with SD 0.2 (i.e. approximately 20%), and considering the probability mass from 0.5 beneath to 0.5 above each integer value, and only considering integer values with a probability mass > 0.01, we arrive at six vales of L that are eligible to include (3, 4, 5, 6, 7 and 8).

Second, we assign the probability of each of these six integer values for use in the PMSLT. Because we discarded integer values with a probability mass < 0.01, the total probability mass from ln(2.5) to ln(8.5) for a median of ln(5) and SD of 0.2 on the ln-scale is less than 1.0, namely 0.9974. We therefore divide the probability mass for each integer value by 0.9974 to arrive at the probabilities of each of the six integer values being drawn in an iteration of the PMSLT, Pr(L) shown in the second row of Table D.2.

Table D.2: Short time lags

Length of decay 'lookback' window, L	3	4	5	6	7	8
Probability of L being drawn in PMSLT, Pr(L)	0.0372	0.2630	0.3856	0.2230	0.0738	0.0174
Rate, r, used to generate weights for each year, Y, within length L, where r = -ln(1-L)/0.95	0.9986	0.7489	0.5991	0.4993	0.4280	0.3745
Year lag, Y	**Weighting of (1-PIF) by years since current year-cycle, W_y**					

0	0.4137	0.3288	0.2725	0.2325	0.2028	0.1797
-1	0.4035	0.3816	0.3516	0.3223	0.2959	0.2726
-2	0.1487	0.1804	0.1931	0.1956	0.1929	0.1875
-3	0.0341	0.0853	0.1061	0.1187	0.1257	0.1289
-4		0.0239	0.0583	0.0721	0.0819	0.0887
-5			0.0184	0.0437	0.0534	0.0610
-6				0.0149	0.0348	0.0419
-7					0.0126	0.0288
-8						0.0108

Third, consistent with Hoogenveen et al.[66] we assume that the transition from effect on disease incidence from one level of risk factor to the new level of risk factor occurs with an exponential rate decay. Because an exponential rate decay asymptotes towards zero over time, but never achieves zero, we focus on the rate necessary to see the risk fall by 95% of that between levels of the risk factor. We can then solve this rate, r, given set options for L, as

$$r = -ln(1 - L)/0.95. \tag{D.1}$$

Fourth, we use this rate to determine the weights for $(1 - PIF)$ for each year in the past. Note that the number of years is one more than length L, because of the mid-cycle correction meaning we mathematically start from the middle of the current year (year 0). For example, when $L = 5$ we generate weights for years $Y = 0$ to $Y = -5$ inclusive – which is six years, including half of the first (0) and last (-5) year. The initial estimate of the weight, W, for each year is given by:

$$W_Y^{initial} = e^{(min(0,Y+0.5)r)} - e^{(min(0,Y-0.5)r)} \tag{D.2}$$

For example, for $L = 5$, the initial weight estimated for $Y = -3$ years is:

$$W_Y^{initial} = e^{(min(0,-3+0.5)0.5991)} - e^{(min(0,-3-0.5)0.5991)} \tag{D.3}$$

$$= 0.1008 \tag{D.4}$$

The actual $W_Y^{initial}$ value in Table D.2 is slightly higher at 0.1061. This is because (as above) we used a rate that sees a 95% reduction in risk over L, so all initial estimates need scaling up by 1/0.95 so that the sum of $W_Y^{initial}$ in each column of Table D.2 equals 1.0.

Figure D.1 shows these weights graphically for the six possible values of L, for the short duration time lags used in SHINE. Note that:

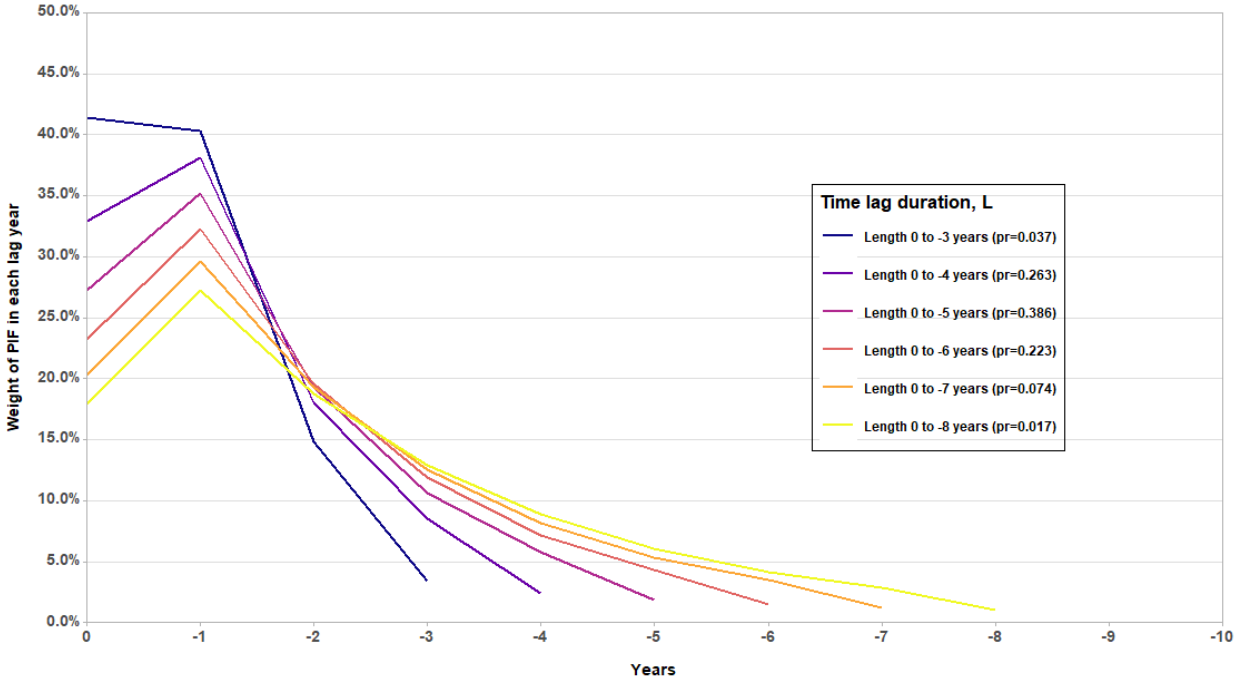

Figure D.1: Weighting of disease time lag duration options: short lag diseases.

- The usually lower weight values in year 0 for each patter of L, due to the use of a half-cycle correction.
- The probability of each of the six possible patterns being used is given in parentheses in the legend. Accordingly, the 'median' pattern for $L = 5$ has the highest probability of being drawn in the PMSLT (i.e. probability of 0.386). This mechanism is how we include input parameter uncertainty.

D.3 Long time lags

The delay in the start of decay means that disease incidence rates start to change 1 to 3 2 years after change in the given risk factor, with a probability of 0.7946 for 2 years, a probability of 0.0754 for 1 year, and a probability of 0.1300 for 3 years.

There are 17 possible integer vales of L, ranging from 13 to 29. Table D.3 shows these 17 patterns within the scenario of the drawbar being 2 years. The derivation of values in the table is the same as above for short duration time lags. Noteworthy additional points here include:

- The probability of each of the 17 patterns being drawn, conditional on already having drawn a value of 2 years for the drawbar, are shown in the second row of Table D.3.
- As expected, weights only commence in year -2

Table D.3: Long time lags

L	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29

$Pr(L)$	0.0159	0.0303	0.0491	0.0693	0.0872	0.0997	0.1050	0.1031	0.0954	0.0838	0.0704	0.0569	0.0445	0.0337	0.0249	0.0180	0.0128
r	0.2304	0.2140	0.1997	0.1872	0.1762	0.1664	0.1577	0.1498	0.1427	0.1362	0.1302	0.1248	0.1198	0.1152	0.1110	0.1070	0.1033
Y	**Weighting of (1-PIF) by years since current year-cycle, W_y**																
0																	
-1																	
-2	0.1146	0.1068	0.1000	0.0941	0.0888	0.0840	0.0798	0.0760	0.0725	0.0693	0.0664	0.0637	0.0612	0.0589	0.0568	0.0548	0.0530
-3	0.1931	0.1822	0.1725	0.1637	0.1557	0.1485	0.1419	0.1359	0.1303	0.1252	0.1204	0.1160	0.1120	0.1081	0.1046	0.1012	0.0981
-4	0.1533	0.1471	0.1412	0.1357	0.1306	0.1257	0.1212	0.1170	0.1130	0.1092	0.1057	0.1024	0.0993	0.0964	0.0936	0.0910	0.0885
-5	0.1218	0.1188	0.1157	0.1126	0.1095	0.1065	0.1035	0.1007	0.0980	0.0953	0.0928	0.0904	0.0881	0.0859	0.0838	0.0817	0.0798
-6	0.0967	0.0959	0.0947	0.0933	0.0918	0.0901	0.0884	0.0867	0.0849	0.0832	0.0815	0.0798	0.0782	0.0765	0.0750	0.0734	0.0720
-7	0.0768	0.0774	0.0776	0.0774	0.0770	0.0763	0.0755	0.0746	0.0736	0.0726	0.0715	0.0704	0.0693	0.0682	0.0671	0.0660	0.0649
-8	0.0610	0.0625	0.0635	0.0642	0.0645	0.0646	0.0645	0.0642	0.0639	0.0634	0.0628	0.0622	0.0615	0.0608	0.0601	0.0593	0.0585
-9	0.0484	0.0505	0.0520	0.0532	0.0541	0.0547	0.0551	0.0553	0.0554	0.0553	0.0551	0.0549	0.0546	0.0542	0.0537	0.0533	0.0528
-10	0.0385	0.0407	0.0426	0.0441	0.0454	0.0463	0.0471	0.0476	0.0480	0.0483	0.0484	0.0484	0.0484	0.0483	0.0481	0.0479	0.0476
-11	0.0306	0.0329	0.0349	0.0366	0.0380	0.0392	0.0402	0.0410	0.0416	0.0421	0.0425	0.0428	0.0429	0.0430	0.0430	0.0430	0.0429
-12	0.0243	0.0266	0.0286	0.0303	0.0319	0.0332	0.0343	0.0353	0.0361	0.0368	0.0373	0.0377	0.0381	0.0383	0.0385	0.0387	0.0387
-13	0.0193	0.0214	0.0234	0.0252	0.0267	0.0281	0.0293	0.0304	0.0313	0.0321	0.0327	0.0333	0.0338	0.0342	0.0345	0.0347	0.0349
-14	0.0153	0.0173	0.0192	0.0209	0.0224	0.0238	0.0250	0.0262	0.0271	0.0280	0.0287	0.0294	0.0300	0.0304	0.0309	0.0312	0.0315
-15	0.0064	0.0140	0.0157	0.0173	0.0188	0.0202	0.0214	0.0225	0.0235	0.0244	0.0252	0.0259	0.0266	0.0271	0.0276	0.0280	0.0284
-16		0.0059	0.0129	0.0144	0.0158	0.0171	0.0183	0.0194	0.0204	0.0213	0.0222	0.0229	0.0236	0.0242	0.0247	0.0252	0.0256
-17			0.0055	0.0119	0.0132	0.0144	0.0156	0.0167	0.0177	0.0186	0.0194	0.0202	0.0209	0.0216	0.0221	0.0226	0.0231
-18				0.0052	0.0111	0.0122	0.0133	0.0144	0.0153	0.0162	0.0171	0.0178	0.0186	0.0192	0.0198	0.0203	0.0208
-19					0.0048	0.0104	0.0114	0.0124	0.0133	0.0142	0.0150	0.0158	0.0165	0.0171	0.0177	0.0183	0.0188
-20						0.0046	0.0097	0.0106	0.0115	0.0124	0.0132	0.0139	0.0146	0.0153	0.0159	0.0164	0.0169
-21							0.0043	0.0092	0.0100	0.0108	0.0116	0.0123	0.0130	0.0136	0.0142	0.0148	0.0153
-22								0.0041	0.0087	0.0094	0.0101	0.0108	0.0115	0.0121	0.0127	0.0133	0.0138
-23									0.0039	0.0082	0.0089	0.0096	0.0102	0.0108	0.0114	0.0119	0.0124
-24										0.0037	0.0078	0.0084	0.0090	0.0096	0.0102	0.0107	0.0112
-25											0.0035	0.0074	0.0080	0.0086	0.0091	0.0096	0.0101
-26												0.0034	0.0071	0.0076	0.0082	0.0086	0.0091
-27													0.0032	0.0068	0.0073	0.0078	0.0082
-28														0.0031	0.0065	0.0070	0.0074
-29															0.0030	0.0063	0.0067
-30																0.0029	0.0060
-31																	0.0028

Rate, r, is the rate to generate weights for each year, calculated as $r = -ln(1 - L)/0.95$.
$Pr(L)$ is the probability of a given length, L, being drawn in the PMSLT.
Y is the years to full impact.

Bibliography

[1] Tony Blakely, Giorgi Kvizhinadze, June Atkinson, Joseph Dieleman, and Philip Clarke. Health system costs for individual and comorbid noncommunicable diseases: An analysis of publicly funded health events from new zealand. *PLoS Medicine*, 16(1):e1002716, 2019. doi: 10.1371/journal.pmed.1002716. URL https://doi.org/10.1371/journal.pmed.1002716. eCollection 2019 Jan.

[2] Samantha Grimshaw, Emily Bourke, and Tony Blakely. Estimates of non-communicable disease expenditure by disease phase, sex, and age group for all oecd countries. *Population Health Metrics*, 23(1):53, 2025. ISSN 1478-7954. doi: 10.1186/s12963-025-00418-5. URL https://doi.org/10.1186/s12963-025-00418-5.

[3] Tony Blakely, Finn Sigglekow, Muhammad Irfan, Anja Mizdrak, Joseph Dieleman, Laxman Bablani, Philip Clarke, and Nick Wilson. Disease-related income and economic productivity loss in new zealand: A longitudinal analysis of linked individual-level data. *PLoS Medicine*, 18(11):e1003848, November 2021. ISSN 1549-1676. doi: 10.1371/journal.pmed.1003848.

[4] Australian Institute of Health and Welfare. Health system spending on disease and injury in australia, 2020-21, 2023.

[5] Christopher JL Murray, Majid Ezzati, Alan D Lopez, Anthony Rodgers, and Stephen Vander Hoorn. Comparative quantification of health risks conceptual framework and methodological issues. *Population health metrics*, 1(1):1, April 2003. ISSN 1478-7954. doi: 10.1186/1478-7954-1-1. URL https://doi.org/10.1186/1478-7954-1-1. epublish.

[6] A. Saltelli, G. Bammer, I. Bruno, E. Charters, M. Di Fiore, E. Didier, W. Nelson Espeland, J. Kay, S. Lo Piano, D. Mayo, R. Pielke, T. Portaluri, T. M. Porter, A. Puy, I. Rafols, J. R. Ravetz, E. Reinert, D. Sarewitz, P. B. Stark, A. Stirling, J. van der Sluijs, and P. Vineis. Five ways to ensure that models serve society: a manifesto. *Nature*, 582(7813):482–484, 2020. ISSN 1476-4687. doi: 10.1038/d41586-020-01812-9. URL https://www.ncbi.nlm.nih.gov/pubmed/32581374.

[7] Hazel Squires, James Chilcott, Ronald Akehurst, Jennifer Burr, and Michael P. Kelly. A framework for developing the structure of public health economic models. *Value in Health*, 19(5):588–601, 2016. ISSN 1098-3015. doi: https://doi.org/10.1016/j.jval.2016.02.011. URL https://www.sciencedirect.com/science/article/pii/S1098301516000589.

[8] Adam D. M. Briggs, Jane Wolstenholme, Tony Blakely, and Peter Scarborough. Choosing an epidemiological model structure for the economic evaluation of non-communicable disease public health interventions. *Population Health Metrics*, 14(1):17, 2016. doi: 10.1186/s12963-016-0085-1. URL https://doi.org/10.1186/s12963-016-0085-1.

[9] Alan Brennan, Stephen E. Chick, and Ruth Davies. A taxonomy of model structures for economic evaluation of health technologies. *Health Economics*, 15(12):1295–1310, 2006. doi: https://doi.org/10.1002/hec.1148. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.1148.

[10] N. Hawkins, M. Sculpher, and D. Epstein. Cost-effectiveness analysis of treatments for chronic disease: using r to incorporate time dependency of treatment response. *Med Decis Making*, 25(5):511–9, 2005. ISSN 0272-989X (Print) 0272-989x. doi: 10.1177/0272989x05280562.

[11] Joshua Szanyi, Tim Wilson, Samantha Howe, Jessie Zeng, Hassan Andrabi, Shania Rossiter, and Tony Blakely. Epidemiologic and economic modelling of optimal COVID-19 policy: public health and social measures, masks and vaccines in victoria, australia. 32. ISSN 2666-6065. doi: 10.1016/j.lanwpc.2022.100675. URL https://www.thelancet.com/journals/lanwpc/article/PIIS2666-6065(22)00290-5/fulltext. Publisher: Elsevier.

[12] Ross A Hammond, Todd B Combs, Austen Mack-Crane, Matt Kasman, Amy Sorg, Doneisha Snider, and Douglas A Luke. Development of a computational modeling laboratory for examining tobacco control policies: Tobacco town. *Health & place*, 61:102256, January 2020. ISSN 1353-8292. doi: 10.1016/j.healthplace.2019.102256. URL https://europepmc.org/articles/PMC11410381.

[13] Mike Paulden. Calculating and interpreting icers and net benefit. *PharmacoEconomics*, 38(8):785–807, 2020. ISSN 1179-2027. doi: 10.1007/s40273-020-00914-6. URL https://doi.org/10.1007/s40273-020-00914-6.

[14] MF Drummond, MJ Sculpher, K Claxton, GL Stoddart, and GW Torrance. Methods for the economic evaluation of health care programmes, 2015.

[15] Disability Department of Health and Ageing. Guidelines for preparing a submission to the pharmaceutical benefits advisory committee, version 5.0. Report, Australian Government, 2016. URL https://pbac.pbs.gov.au/.

[16] Gretchen A Stevens, Leontine Alkema, Robert E Black, J Ties Boerma, Gary S Collins, Majid Ezzati, John T Grove, Daniel R Hogan, Margaret C Hogan, Richard Horton, Joy E Lawn, Ana Marušić, Colin D Mathers, Christopher J L Murray, Igor Rudan, Joshua A Salomon, Paul J Simpson, Theo Vos, and Vivian Welch. Guidelines for accurate and transparent health estimates reporting: the gather statement. *Lancet*, 388(10062):e19–e23, Dec 2016. ISSN 1474-547X. doi: 10.1016/S0140-6736(16)30388-9. URL https://doi.org/10.1016/S0140-6736(16)30388-9. Epub 2016 Jun 28.

[17] Volker Grimm, Steven F. Railsback, and Christian E. et al. Vincenot. The odd protocol for describing agent-based and other simulation models: A second update to improve clarity, replication, and structural realism. *Journal of Artificial Societies and Social Simulation*, 23(2):7, 2020. ISSN 1460-7425. doi: 10.18564/jasss.4259. URL http://jasss.soc.surrey.ac.uk/23/2/7.html.

[18] D. M. Eddy, W. Hollingworth, J. J. Caro, J. Tsevat, K. M. McDonald, and J. B. Wong. Model transparency and validation: a report of the ispor-smdm modeling good research practices task force-7. *Med Decis Making*, 32(5):733–43, 2012. ISSN 0272-989x. doi: 10.1177/0272989x12454579.

[19] Deborah A. Marshall, Lina Burgos-Liz, Maarten J. IJzerman, Nathaniel D. Osgood, William V. Padula, Mitchell K. Higashi, Peter K. Wong, Kalyan S. Pasupathy, and William Crown. Applying dynamic simulation modeling methods in health care delivery research—the simulate checklist: Report of the ispor simulation modeling emerging good practices task force. *Value in Health*, 18(1):5–16, 2015. ISSN 1098-3015. doi: https://doi.org/10.1016/j.jval.2014.12.001. URL https://www.sciencedirect.com/science/article/pii/S1098301514047640.

[20] William Crown, Nasuh Buyukkaramikli, Mustafa Y. Sir, Parveen Thokala, Alex Morton, and Deborah A. Marshall et al. Application of constrained optimization methods in health services research: Report 2 of the ispor optimization methods emerging good practices task force. *Value in Health*, 21(9):1019–1028, 2018. ISSN 1098-3015. doi: https://doi.org/10.1016/j.jval.2018.05.003.

[21] R Foster, T Blakely, N Wilson, and D O'Dea. Protocol for direct costing of health sector interventions for economic modelling (including event pathways) – burden of disease epidemiology, equity and cost-effectiveness programme (bode3) technical report number 12, 2012.

[22] Theo Vos, Rob Carter, Jan Barendregt, Cathrine Mihalopoulos, Lennert Veerman, Anne Magnus, Linda Cobiac, Melanie Bertram, and Angela Wallace. Assessing cost-effectiveness in prevention: Ace–prevention final report. Report, University of Queensland and Deakin University, Brisbane, Australia, September 2010. Prepared for the ACE–Prevention team.

[23] Andrew H. Briggs, Milton C. Weinstein, Elisabeth A.L. Fenwick, Jonathan Karnon, Mark J. Sculpher, and A. David Paltiel. Model parameter estimation and uncertainty: A report of the ispor-smdm modeling good research practices task force-6. *Value in Health*, 15(6):835–842, 2012. ISSN 1098-3015. doi: https://doi.org/10.1016/j.jval.2012.04.014. URL https://www.sciencedirect.com/science/article/pii/S1098301512016592.

[24] P. Thokala and A. Duenas. Multiple criteria decision analysis for health technology assessment. *Value Health*, 15(8):1172–81, 2012. ISSN 1098-3015. doi: 10.1016/j.jval.2012.06.015.

[25] P. Thokala, N. Devlin, K. Marsh, R. Baltussen, M. Boysen, Z. Kalo, T. Longrenn, F. Mussen, S. Peacock, J. Watkins, and M. Ijzerman. Multiple criteria decision analysis for health care decision making–an introduction: Report 1 of the ispor mcda emerging good practices task force. *Value Health*, 19(1):1–13, 2016. ISSN 1098-3015. doi: 10.1016/j.jval.2015.12.003.

[26] Trudy Sullivan. *Using MCDA (Multi-Criteria Decision Analysis) to prioritise publicly-funded health care*. Thesis, 2012.

[27] Paul Hansen and Nancy Devlin. Multi-criteria decision analysis (mcda) in healthcare decision-making, 2019-04-26 2019. URL https://oxfordre.com/economics/view/10.1093/acrefore/9780190625979.001.0001/acrefore-9780190625979-e-98.

[28] P. Gongora-Salazar, S. Rocks, P. Fahr, O. Rivero-Arias, and A. Tsiachristas. The use of multicriteria decision analysis to support decision making in healthcare: An updated systematic literature review. *Value Health*, 26(5):780–790, 2023. ISSN 1098-3015. doi: 10.1016/j.jval.2022.11.007.

[29] Hamed Taherdoost and Mitra Madanchian. Multi-criteria decision making (mcdm) methods and concepts, 2023.

[30] K. Marsh, I. Jzerman M, P. Thokala, R. Baltussen, M. Boysen, Z. Kaló, T. Lönngren, F. Mussen, S. Peacock, J. Watkins, and N. Devlin. Multiple criteria decision analysis for health care decision making–emerging good practices: Report 2 of the ispor mcda emerging good practices task force. *Value Health*, 19(2):125–37, 2016. ISSN 1098-3015. doi: 10.1016/j.jval.2015.12.016.

[31] John W Kingdon. *Agendas, Alternatives, and Public Policies*. Longman, 2003. ISBN 0673523896. URL https://nla.gov.au/nla.cat-vn912147.

[32] Rob DeLeo, Reimut Zohlnhöfer, and Nikolaos Zahariadis. Multiple streams and policy ambiguity. *Elements in Public Policy*, 02 2024. doi: 10.1017/9781009397926.

[33] Tim Wilson, Samantha Howe, Stephanie Khuu, Abraham Flaxman, and Tony Blakely. Specifying prospective compartmental models of chronic disease. *medRxiv*, 2026. doi: 10.64898/2026.01.14.26344075. URL https://www.medrxiv.org/content/early/2026/01/16/2026.01.14.26344075.

[34] Samantha Howe, Tim Wilson, Coral Gartner, Tony Blakely, and Driss Ait Ouakrim. A novel tobacco forecasting model by multiple sociodemographic strata in australia. *International Journal of Epidemiology [in press]*, 2025.

[35] S. E. Vollset, H. S. Ababneh, Y. H. Abate, C. Abbafati, R. Abbasgholizadeh, M. Abbasian, et al. Burden of disease scenarios for 204 countries and territories, 2022–2050: a forecasting analysis for the global burden of disease study 2021. *The Lancet*, 403(10440):2204–2256, 2024. doi: 10.1016/S0140-6736(23)02277-5.

[36] T. Vos, S. S. Lim, C. Abbafati, K. M. Abbas, M. Abbasi, M. Abbasifard, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. *The Lancet*, 396(10258):1204–1222, 2020. doi: 10.1016/S0140-6736(20)30925-9.

[37] Mohammad H Forouzanfar et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the global burden of disease study 2015. *The Lancet*, 388(10053):1659–1724, 2016.

[38] Bibha Dhungel, Timothy Wilson, Stephanie Khuu, and Antony Blakely. Forecasting risk factor distributions by sex and age in any country, 1990 to 2050: A parametric estimation framework using GBD summary exposure values and other inputs. 2 2026. doi: 10.26188/31310767.v3. URL https://figshare.unimelb.edu.au/articles/preprint/_p_dir_ltr_b_Forecasting_risk_factor_distributions_by_sex_and_age_in_any_country_1990_to_2050_A_b_p_p_dir_ltr_b_parametric_estimation_framework_using_GBD_summary_exposure_values_and_other_inputs_b_p_/31310767.

[39] Jan J Barendregt and J Lennert Veerman. Categorical versus continuous risk factors and the calculation of potential impact fractions. *Journal of Epidemiology & Community Health*, 64(3):209–212, 2010. ISSN 0143-005X. doi: 10.1136/jech.2009.090274. URL https://jech.bmj.com/content/64/3/209.

[40] Global Burden of Disease Collaborator Network. Global burden of disease study 2021 (gbd 2021) results, 2022.

[41] J. A. Salomon, J. A. Haagsma, A. Davis, C. M. de Noordhout, S. Polinder, A. H. Havelaar, A. Cassini, B. Devleesschauwer, M. Kretzschmar, N. Speybroeck, C. J. Murray, and T. Vos. Disability weights for the global burden of disease 2013 study. *Lancet Glob Health*, 3(11):e712–23, 2015. ISSN 2214-109X. doi: 10.1016/S2214-109X(15)00069-8. URL https://www.ncbi.nlm.nih.gov/pubmed/26475018.

[42] J. A. Haagsma, C. Maertens de Noordhout, S. Polinder, T. Vos, A. H. Havelaar, A. Cassini, B. Devleesschauwer, M. E. Kretzschmar, N. Speybroeck, and J. A. Salomon. Assessing disability weights based on the responses of 30,660 people from four european countries. *Population Health Matrics*, (1478-7954 (Print)), 2015.

[43] J. A. Salomon, Vos T., and Daniel R. et al. Hogan. Common values in assessing health outcomes from disease and injury: disability weights measurement study for the global burden of disease study 2010. *The Lancet*, (1474-547X (Electronic)), 2012.

[44] Alize J. Ferrari, Damian F. Santomauro, and Amirali et al. Aali. Global incidence, prevalence, years lived with disability (ylds), disability-adjusted life-years (dalys), and healthy life expectancy (hale) for 371 diseases and injuries in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the global burden of disease study 2021. *The Lancet*, 403(10440):2133–2161, 2024. ISSN 0140-6736. doi: 10.1016/S0140-6736(24)00757-8. URL https://doi.org/10.1016/S0140-6736(24)00757-8. doi: 10.1016/S0140-6736(24)00757-8.

[45] Pieter van Baal, Alec Morton, David Meltzer, and Werner Brouwer. Future unrelated medical costs need to be considered in cost effectiveness analysis. *The European Journal of Health Economics*, 20(1):1–5, 2019. ISSN 1618-7601. doi: 10.1007/s10198-018-0976-0. URL https://doi.org/10.1007/s10198-018-0976-0.

[46] M. Y. Bertram, J. A. Lauer, K. Stenberg, and T. T. T. Edejer. Methods for the economic evaluation of health care interventions for priority setting in the health system: An update from who choice. *Int J Health Policy Manag*, 10(11):673–677, 2021. ISSN 2322-5939. doi: 10.34172/ijhpm.2020.244. URL https://www.ncbi.nlm.nih.gov/pubmed/33619929.

[47] (ABS) Australian Bureau of Statistics. Working arrangements, 2023, 2024. URL https://www.abs.gov.au/statistics/labour/earnings-and-working-conditions/working-arrangements/latest-release.

[48] Driss Ait Ouakrim, Tim Wilson, Samantha Howe, Philip Clarke, Coral E. Gartner, Nick Wilson, and Tony Blakely. Economic effects for citizens and the government of a country-level tobacco endgame strategy: a modelling study. pages tc–2023–058131, . ISSN 1468-3318. doi: 10.1136/tc-2023-058131.

[49] Anthony O'Hagan. Expert knowledge elicitation: Subjective but scientific. *The American Statistician*, 73:69–81, 03 2019. doi: 10.1080/00031305.2018.1518265.

[50] D Ait Ouakrim, T Wilson, A Waa, R Maddox, H Andrabi, S Mishra, J Summers, C Gartner, R Lovett, R Edwards, N Wilson, and T Blakely. Tobacco endgame intervention impacts on health gains and māori:non-māori health inequity: a simulation study of the aotearoa/new zealand tobacco action plan. . ISSN 1468-3318. doi: 10.1136/tc-2022-057655. URL https://pubmed.ncbi.nlm.nih.gov/36627213/. Publisher: Tob Control.

[51] (WHO) World Health Organization. Evidence, policy, impact: Who guide for evidence-informed decision-making. Report, WHO, 2021. URL https://apps.who.int/iris/handle/10665/350994.

[52] Irina Degtiar and Sherri Rose. A review of generalizability and transportability. *Annual Review of Statistics and Its Application*, 10(Volume 10, 2023):501–524, 2023. ISSN 2326-831X. doi: https://doi.org/10.1146/annurev-statistics-042522-103837. URL https://www.annualreviews.org/content/journals/10.1146/annurev-statistics-042522-103837.

[53] M. M. Haby, T. Vos, R. Carter, M. Moodie, A. Markwick, A. Magnus, K. S. Tay-Teo, and B. Swinburn. A new approach to assessing the health benefit from obesity interventions in children and adolescents: the assessing cost-effectiveness in obesity project. *Int J Obes (Lond)*, 30(10):1463–75, 2006. ISSN 0307-0565 (Print) 0307-0565. doi: 10.1038/sj.ijo.0803469.

[54] B. Swinburn, T. Gill, and S. Kumanyika. Obesity prevention: a proposed framework for translating evidence into action. *Obes Rev*, 6(1):23–33, 2005. ISSN 1467-7881 (Print) 1467-7881. doi: 10.1111/j.1467-789X.2005.00184.x.

[55] M. Soares, A. Colson, L. Bojke, S. Ghabri, O. U. Garay, J. K. Felli, K. Lee, E. Molsen-David, O. Morales-Napoles, V. A. Shaffer, and M. J. IJzerman. Recommendations on the use of structured expert elicitation protocols for healthcare decision making: A good practices report of an ispor task force. *Value Health*, 27(11):1469–1478, 2024. ISSN 1524-4733. doi: 10.1016/j.jval.2024.07.027. URL https://www.ncbi.nlm.nih.gov/pubmed/39505473.

[56] John Paul Gosling. *SHELF: The Sheffield Elicitation Framework*, pages 61–93. Springer International Publishing, Cham, 2018. ISBN 978-3-319-65052-4. doi: 10.1007/978-3-319-65052-4_4. URL https://doi.org/10.1007/978-3-319-65052-4_4.

[57] Benjamin Johns, Rob Baltussen, and Raymond Hutubessy. Programme costs in the economic evaluation of health interventions. *Cost Effectiveness and Resource Allocation*, 1(1):1, 2003. ISSN 1478-7547. doi: 10.1186/1478-7547-1-1. URL https://doi.org/10.1186/1478-7547-1-1.

[58] M. Y. Bertram, K. Stenberg, C. Brindley, J. Li, J. Serje, R. Watts, and T. T. Edejer. Disease control programme support costs: an update of who-choice methodology, price databases and quantity assumptions. *Cost Eff Resour Alloc*, 15:21, 2017. ISSN 1478-7547 (Print) 1478-7547. doi: 10.1186/s12962-017-0083-6.

[59] M. Y. Bertram, D. Chisholm, R. Watts, T. Waqanivalu, V. Prasad, and C. Varghese. Cost-effectiveness of population level and individual level interventions to combat non-communicable disease in eastern sub-saharan africa and south east asia: A who-choice analysis. *Int J Health Policy Manag*, 10(11):724–733, 2021. ISSN 2322-5939. doi: 10.34172/ijhpm.2021.37.

[60] N. Wilson, N. Nghiem, R. Foster, L. Cobiac, and T. Blakely. Estimating the cost of new public health legislation. *Bull World Health Organ*, 90(7):532–9, 2012. ISSN 1564-0604. doi: 10.2471/BLT.11.097584. URL https://www.ncbi.nlm.nih.gov/pubmed/22807599.

[61] T. Blakely, R. Moss, J. Collins, A. Mizdrak, A. Singh, N. Carvalho, N. Wilson, N. Geard, and A. Flaxman. Proportional multistate lifetable modelling of preventive interventions: concepts, code and worked examples. *Int J Epidemiol*, 49(5):1624–1636, 2020. ISSN 1464-3685. doi: 10.1093/ije/dyaa132. URL https://www.ncbi.nlm.nih.gov/pubmed/33038892.

[62] Peng Zheng, Ashkan Afshin, and Stan et al. Biryukov. The burden of proof studies: assessing the evidence of risk. *Nature Medicine*, 28(10):2038–2044, 2022. ISSN 1546-170X. doi: 10.1038/s41591-022-01973-2. URL https://doi.org/10.1038/s41591-022-01973-2.

[63] Simon I. Hay, Kanyin Liane Ong, and Damian F et al. Santomauro. Burden of 375 diseases and injuries, risk-attributable burden of 88 risk factors, and healthy life expectancy in 204 countries and territories, including 660 subnational locations, 1990–2023: a systematic analysis for the global burden of disease study 2023. *The Lancet*, 406(10513):1873–1922, 2025. ISSN 0140-6736. doi: 10.1016/S0140-6736(25)01637-X. URL https://doi.org/10.1016/S0140-6736(25)01637-X.

[64] Luisa S Flor, Jason A Anderson, and Noah et al. Ahmad. Health effects associated with exposure to secondhand smoke: a burden of proof study. *Nature Medicine*, 30(1):149–167, January 2024. ISSN 1546-170X. doi: 10.1038/s41591-023-02743-4. URL https://doi.org/10.1038/s41591-023-02743-4.

[65] Institute for Health Metrics and Evaluation. Burden of proof, 2022. URL https://www.healthdata.org/data-tools-practices/interactive-visuals/burden-proof.

[66] R. T. Hoogenveen, P. H. van Baal, H. C. Boshuizen, and T. L. Feenstra. Dynamic effects of smoking cessation on disease incidence, mortality and quality of life: The role of time since cessation. *Cost Eff Resour Alloc*, 6:1, 2008. ISSN 1478-7547. doi: 10.1186/1478-7547-6-1. URL https://www.ncbi.nlm.nih.gov/pubmed/18190684.

[67] Ali Aminian, Rickesha Wilson, Abbas Al-Kurd, Chao Tu, Alex Milinovich, Matthew Kroh, Raul J Rosenthal, Stacy A Brethauer, Philip R Schauer, Michael W Kattan, Justin C Brown, Nathan A Berger, Jame Abraham, and Steven E Nissen. Association of bariatric surgery with cancer risk and mortality in adults with obesity. *JAMA*, 327(24):2423–2433, Jun 2022. ISSN 1538-3598. doi: 10.1001/jama.2022.9009. URL https://doi.org/10.1001/jama.2022.9009. PMID: 35657620.

[68] S Howe, Tim Wilson, Coral Gartner, Sandra Eades, Tony Blakely, and Driss Ait Ouakrim. Health and equity impact of achieving 5 *SSRN*, 2026. doi: 10.2139/ssrn.6100921.

[69] B. Groot Koerkamp, T. Stijnen, M. C. Weinstein, and M. G. Hunink. The combined analysis of uncertainty and patient heterogeneity in medical decision models. *Med Decis Making*, 31(4):650–61, 2011. ISSN 0272-989x. doi: 10.1177/0272989x10381282.

[70] Australian Institute of Health and Welfare. Australian burden of disease study: Methods and supplementary material 2018. Report, AIHW, 2021. URL https://www.aihw.gov.au/reports/burden-of-disease/abds-methods-supplementary-material-2018.

[71] J. F. Fries. The compression of morbidity. 1983. *Milbank Q*, 83(4):801–23, 2005. ISSN 0887-378X (Print) 0887-378x. doi: 10.1111/j.1468-0009.2005.00401.x.

[72] OB Ahmad, Boschi-Pinto C, AD Lopez, R Lozano, and M Inoue. Age standardization of rates: A new who standard. Report, World Health Organisation, 2001.

[73] J. B. Brown, A. J. Palmer, P. Bisgaard, W. Chan, K. Pedula, and A. Russell. The mt. hood challenge: cross-testing two diabetes simulation models. *Diabetes Res Clin Pract*, 50 Suppl 3:S57–64, 2000. ISSN 0168-8227 (Print) 0168-8227. doi: 10.1016/s0168-8227(00)00217-5.

[74] James Altunkaya, Xinyu Li, Amanda Adler, Talitha Feenstra, Adam Fridhammar, Mi Jun Keng, Mark Lamotte, Phil McEwan, Andreas Nilsson, Andrew J. Palmer, Jianchao Quan, Harry Smolen, An Tran-Duy, William Valentine, Michael Willis, José Leal, and Philip Clarke. Examining the impact of structural uncertainty across 10 type 2 diabetes models: Results from the 2022 mount hood challenge. *Value in Health*, 27(10):1338–1347, 2024. ISSN 1098-3015. doi: 10.1016/j.jval.2024.06.010. URL https://doi.org/10.1016/j.jval.2024.06.010.

[75] Timothy Lash, Matthew Fox, and Aliza Fink. *Applying Quantitative Bias Analysis to Epidemiologic Data*. Statistics for Biology and Health. Springer, 01 2009. ISBN 9780387879604. doi: 10.1007/978-0-387-87959-8_8.

[76] M. P. Fox, R. F. MacLehose, and T. L. Lash. Sas and r code for probabilistic quantitative bias analysis for misclassified binary variables and binary unmeasured confounders. *Int J Epidemiol*, 52(5):1624–1633, 2023. ISSN 1464-3685. doi: 10.1093/ije/dyad053. URL https://www.ncbi.nlm.nih.gov/pubmed/37141446.

[77] Mark J. C. Nuijten. Incorporation of statistical uncertainty in health economic modelling studies using second-order monte carlo simulations. *PharmacoEconomics*, 22(12):759–769, 2004. ISSN 1179-2027. doi: 10.2165/00019053-200422120-00001. URL https://doi.org/10.2165/00019053-200422120-00001.

[78] P. Andersen, A. Mizdrak, N. Wilson, A. Davies, L. Bablani, and T. Blakely. Disaggregating proportional multistate lifetables by population heterogeneity to estimate intervention impacts on inequalities. *Population Health Metrics*, 20(1), 2022. doi: 10.1186/s12963-022-00282-7. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-85122926681&doi=10.1186%2fs12963-022-00282-7&partnerID=40&md5=49d0c55151c6f0e1e176419fc7c918bc. Export Date: 28 January 2025; Cited By: 2.