Appendix A Technical notes

These numbered additional explanatory and technical notes refer to superscript references in the text of the main body of the report. References in these notes are included in the main reference list above.

Chapter 1

1. The second type of information is addressed through a range of statistical activity to monitor health expenditure, workforce numbers and infrastructure (AIHW 1998a, 1999b, 1999d); the third is still largely unexplored territory, although there are increasing numbers of cost-effectiveness studies for particular health interventions (Salkeld et al. 1994, Lave & Joshi 1996).

2. Different countries may have different values and wish to include different importance weights in the calculation of the burden of disease. It is nevertheless desirable to conduct cross-national comparisons and this requires the adoption of common criteria. Those incorporated in the WHO analysis would be widely regarded as reasonable and representative of a wide range of values.

3. When descriptive DALYs (describing current burden) are used in an evaluation setting (e.g. if disease incidence is decreased through an intervention or survival prospects improve), there is the issue that the YLL have been estimated against a standard ideal life expectancy rather than the actual health-adjusted life expectancy of the population group concerned. An example would be improved survival after breast cancer screening or mastectomy, where a QALY adjustment is required to the YLL recovered- life after mastectomy probably will not be valued at equivalent to full health. A simple pro-ration of the DALY might be the starting point for macro evaluation work (i.e. 10% reduction in incidence generates 10% reduction in the DALYs), but there will probably be a need to develop more sophisticated models for specific policy analyses. There may also be an equity argument to use the same ideal life expectancy for everyone—to avoid the situation where an intervention is less cost-effective for disadvantaged groups with lower current life expectancies.

4. Examples of health state profiles intended for use with health state valuations include the EuroQol with three levels on each of five dimensions (Dolan et al. 1996, Dolan 1997), the Health Utilities Index with 5 or 6 levels on eight dimensions (Torrance et al 1995, Furlong et al. 1998) and the AQOL with 4 levels on fifteen dimensions (Hawthorne & Richardson 1995).

5. The original 1992 version of DALYs asked public health practitioners to use a rating scale method to map disease sequelae into six disability classes, defined using word definitions related to activities of daily living and instrumental activities of daily living (Murray 1994). The final version of the Global Burden of Disease 1990 study (Murray & Lopez 1996a) used disability weights for disease sequelae derived directly using a deliberative approach with multiple person trade-off methods. Participants were instructed to evaluate the average individual with the (disease or injury) condition described, taking into account the average social response or milieu. The resulting preferences are probably influenced by perceptions of the average handicap (participation restriction) stemming from each condition.

6. The current revision process for the ICIDH has emphasised that participation restriction (formerly referred to as handicap), results from the interaction of impairments, functional limitations (formerly referred to as disability), individual and cultural beliefs and expectations, and the physical and social

environment (WHO 1999b). Murray (1996, page 33) argued that the DALY should attempt to capture the impact of disability rather than handicap on equity grounds. The disadvantage resulting from disability may be smaller in already disadvantaged population groups, since they have less advantage to lose, and so allocating resources to avert handicap rather than disability could exacerbate inequalities.

Mathers (1997c), Nord (1997) and Wolfson (1998) have argued that summary population health measures should relate to dimensions of health, such as impairments and activity limitations, that are intrinsic to the person or 'within the skin', rather than dimensions of health or broader wellbeing that are determined by the interaction between the individual and the social and environmental contexts. Here, 'within the skin' includes mental health and function as well as physical health and refers to functioning at the level of the body and individual (in the terms used by the draft ICIDH revision).

7. Aspects of the standard gamble, time trade-off and person trade-off methods approximate situations that frequently arise in health services. The standard gamble is similar to the choice faced by a patient with a serious condition for which the treatment could result in death, but if successful would leave the patient much better off. The time trade-off is similar to a patient having a chronic condition where the treatment is likely to improve but shorten life. The person trade-off is similar to the situation faced by a health planner allocating scarce resources between treatments for different conditions.

All these methods capture something more than pure health state preference or utility (Nord 1992, Nord et al. 1993). Ratings scales approaches tend to give preferences for mild health states that are too low (for example, the Quality of Wellbeing Scale values 50 dental pulp extractions as equivalent to saving a year of life). The standard gamble approach is affected by aversion to risk: some people are less willing to gamble with life than others. Time trade-off is influenced by the length of time being traded, as most people value years of life further into the future less than years closer to the present. Current person trade-off approaches are influenced by equity considerations (willingness to trade health in one group of people against that in another).

The majority of economists have argued that preferences should be obtained using a trade-off instrument which requires respondents to consider the 'cost' of good health in terms of what they are trading it off for. In particular, if we are to accept that the final metric gives us a trade-off between life and quality of life, then the trade-off should involve life. This narrows the options for the standard gamble, PTO and TTO (Richardson & Nord 1997, Brazier et al. 1999).

8. Nord (1994) and Murray and Lopez (1996a) have argued that for evaluation of health programs at the societal level and for assessment of burden of disease or health benefits at the population level, the person trade-off (PTO) is to be preferred to the standard gamble or time trade-off. This is because the PTO method measures preferences in terms closest to the uses to which the weights are to be put. These authors have argued that the PTO more directly attempts to measure social preferences for health states, rather than the average of individual preferences for health states. The two are not necessary identical. For example, a majority of individuals may have little individual preference for being fertile because they are past the reproductive stage or do not plan to have children. But they may place a greater social value on fertility because they value fertility for those who are of reproductive age and desire to have children.

9. The deliberative approach ensures that people understand the task they are being asked to perform, by asking the group to discuss and defend differences in the weights chosen by members. It does not require members of the group to reach consensus on the weights, but to ensure they have thought through the reasons for their choices and understood the questions posed to them (Murray and Lopez 1996a). In contrast, most studies by health economists have used an individual

questionnaire format that does not require explicit conceptualisation or group deliberation. A number of focus group studies, including some carried out by AIHW in 1991, have shown that many people do not understand the trade-off exercises correctly.

10. If the purpose is to obtain comparable values across a wide range of conditions for use in health policy applications, there are practical and theoretical problems in using groups of health professionals or people with particular health problems or disabilities. Each individual in a deliberative group is required to elicit preferences for a number of health states to ensure consistency and comparability of preferences across a range of health states. Individuals from either of these two groups do not have a comprehensive understanding of health states outside their own experience and so are not better placed than a general population to quantify social health state preferences:

- Health professionals may have a better understanding of health states in their area of expertise, but are no better placed than anyone else to evaluate disability states outside their professional fields.
- People with a particular health problem or in a particular disability state may be the best persons to understand that state but are no better placed than others to evaluate other disability states. Additionally, there is evidence that people with experience of a health problem tend to rate it less severely than do people who have not experienced the problem. This may reflect adaptation or more accurate knowledge.

The ethical and equity issues relating to the use of disability weights derived by people who have adapted to long-term health problems or disabilities has been discussed in detail by Murray (1996: 29–32). Additionally, some health economists have argued that we should generally use the 'insurance principle' according to which we make policy on the basis of before-the-facts assessments. Otherwise policy may be determined by people speaking too narrowly from their vested interest in a particular health problem. Given the opposite dangers of discrimination and ignorance of the states being assessed, however, it will be important to develop techniques to better describe health states for weighting exercises. This will provide a greater role for those who have directly experienced illness, impairment and disability by allowing their experiences to inform the weighting process. To date, the majority of writers have argued for the inclusion of a personal perspective (Brazier et al. 1999, Richardson et al. 1999).

11. This may reflect insufficient sample sizes to detect these differences or the general lack of comparable data on health state preferences. However, it is possible that there is reasonable cross-cultural agreement on what constitutes a severe or less severe health state, and on the contributions of different domains of health to the overall preference for the health state, if the health is defined in terms of 'within-the-skin' domains.

12. The use of health state preferences and summary measures for policy making, priority setting or resource allocation, e.g. in allocation based on marginal cost-effectiveness criteria, does not require us to maximise health outcomes. This is one option, but there are other options which society may prefer:

- We might give priority to the worst-off (Nord 1996).
- We could attach greater priority to large benefits than to the sum of many small ones, with life-saving counting the most of all. Thus an intervention which gave 40 DALYs to one individual might be preferred to an intervention which gave 1 DALY to 40 individuals.
- We could attach greater importance to giving everyone some benefits as opposed to larger benefits for a few. Richardson and Nord (1997) present some empirical evidence that Australians prefer more equally distributed benefits to less equally distributed benefits.

- Or we might attach less importance to life extension past a normal lifespan, thus attaching greater moral weight to achieving a 'fair innings' (Williams 1999).

It is useful to apply Rawls' principle of a veil of ignorance (Rawls 1971) in considering these options. An individual behind a veil of ignorance does not know who he or she is in a population and must choose one of the above approaches, or a combination, keeping in mind that he or she could be any member of the population, and experience any health problem. Wolfson (1998) has argued that summary measures assist us in making explicit these trade-offs between efficiency (maximising health outcomes) and equity (providing health benefits to all groups and reducing inequalities in health outcomes). They allow us not only to measure the burden of a health problem and the potential for health gain, but also to generate measures of the distributional impacts of health-related interventions. Equity concerns could then be addressed explicitly in any priority setting or resource allocation process, along with the potential to reduce the overall burden.

13. For ease of calculation, the DALY formulae use a continuous discounting function of the form e^{-rt} where r is the discount rate and t is time. The rate (3% in this study) is not precisely the same as the annual discount rate used in the discrete form of the discount function $(1+r)^{-t}$. With a continuous discount rate of 3%, the corrresponding annual discount rate is 2.96%.

14. A number of arguments have been advanced to support discounting in economic analyses (see Goodin 1982, Murray & Lopez 1996a). These include:

- pure time preference (impatience, moral urgency 'the currently sick deserve help' and moral myopia 'I want my cake now');
- uncertainty and risk ('I might be dead next year so I discount its value'—the world average death rate is about 1% per annum);
- diminishing marginal utility coupled with historical rising levels of consumption ('I will be better off next year and so will value marginal benefits less'); and
- opportunity cost of capital (without discounting society could always buy more benefit in the future by investing the money rather than spending it now).

15. The excessive sacrifice argument is that if there is a greater payoff through future investment than present (say, because technology is improving), then with zero discounting we would postpone all current spending resulting in an excessive sacrifice by the current generation for future generations.

16. Arguments against discounting future health gains (or losses) include:

- life does not lose value (to society) if it is in the future rather than the present (Goodin 1982);
- life cannot be valued in monetary terms so the usual opportunity cost arguments do not apply (Anand & Hanson 1997);
- if we are concerned about excessive sacrifice, we should build this in to our thinking as an equity principle directly, rather than discount (Parfit 1984); and
- the social discount rate may very well not be constant for every year into the future (Murray & Acharya 1997).

17. There are good arguments to use a 'social discount rate' rather than an opportunity cost of capital rate or an average of individual discount rates (which empirical studies show can vary from 0% to 10% or more). Individuals may have different concerns for public issues (including the future of their

children and descendents) than for private issues. It can also be argued that the time preferences of individuals are not relevant to the time preferences for a stable society.

18. This is a low positive rate that is probably at the lower limit of acceptability for those economists who are persuaded by the opportunity cost argument and at the upper end of acceptability for those wanting to avoid the excessive sacrifice problem (Murray & Acharya 1997).

19. The GBD incorporated age-weighting into the DALY using an integrable mathematical function that rises rapidly from zero at birth to a peak in the early twenties after which it steadily declines. This function has three parameters specifying its maximum amplitude, peak age, and the proportion of the age weight that is applied (so that the value for a year at birth can be set anywhere from zero (full age weighting) to one (uniform age weights). The amplitude was chosen so that total global DALYs were the same with and without age weights.

Chapter 2

20. The use of a standard life table to calculate the years of life lost due to a death at a given age achieves three objectives:

- deaths at the same age in any population subgroup contribute equally to the burden of disease;
- deaths at all ages contribute to the burden of disease (unlike the usual methods for calculating potential years of life lost to age 75); and
- deaths at a given age in different years result in the same years of life lost, so that changes in the burden over time are not confounded by changes in expected years of life lost.

Table A.1: Comparison of disability weights for GBD indicator conditions with Dutch weights

Global Burden of Disease Study		Dutch study	
Indicator condition	**Weight**	**Comparable condition**	**Weight**
Angina pectoris[(a)]	0.18	Angina	0.22
Late complications after STD infection	0.11	Infertility	0.19
Rheumatoid arthritis	0.21	Mild rheumatoid arthritis	0.21
Mild mental retardation	0.36	Mild mental handicap	0.21
Deafness	0.33	Severe hearing loss	0.37
Blindness	0.62	Severe vision loss	0.43
Down syndrome without cardiac malformation	0.41	Down syndrome without comorbid conditions	0.51
Paraplegia	0.67	Paraplegia	0.57
Unipolar major depression	0.62	Severe depression	0.76
Quadriplegia	0.90	Quadriplegia	0.86
Dementia	0.76	Moderate or severe dementia[(b)]	0.73
Active psychosis	0.72	Schizophrenia, several psychotic episodes	0.71
		Alcoholic psychosis	0.83

(a) Average of weights for mild stable angina and severe stable angina, assuming relative prevalences as modelled for Australia in this study.

(b) Average weight derived assuming relative prevalences of moderate and severe dementia as described in Appendix B.

Sources: Stouthard et al. 1997, Murray and Lopez 1996a.

21. For younger ages, it is necessary to project mortality rates beyond 2051. Gompertz curves were fitted to the observed and projected life expectancies at birth for males and females from 1966 to 2051 using the method of Rowland (1994) in order to project period life expectancies up to 2095. The asymptotic life expectancies at birth for Australian males and females are 84.7 and 87.4 years for males and females respectively. The asymptotic male/female difference is 2.7 years, very close to the 2.5 year difference used for the GBD standard life tables.

22. Twelve of the 22 indicator conditions used in the development of the GBD weights had comparable counterparts in the Dutch study. Table A.1 lists these conditions and the weights derived by each of the two studies.

23. Multiplicative multi-attribute functions provide much better fit to observed preference data than additive models (Furlong et al. 1998). A multiplicative model of the following form was fitted to the Dutch weights for 153 disease sequelae or stages:

$$\log(w) = d_{12} + d_{13} + d_{22} + d_{23} + d_{32} + d_{33} + d_{42} + d_{43} + d_{52} + d_{53} + d_{62} + d_{63} + s + p$$

where

d_{ij} = 1 if EQ-5D+ state is j on dimension i, 0 otherwise.

s = 1 if EQ-5D+ is 111111 but there is a disease present

p = 1 if the prognosis for the disease is uncertain (0 otherwise).

Annualised weights associated with a short duration disease in an annual profile were excluded. A small number of outliers were also eliminated from analysis. Nearly all of these were states described by a distribution of EQ-5D+ states for which the overall weight was not consistent with the mix of states.

The fitted regression model resulted in a single attribute weight slightly greater than 1 (on a scale where 1= good health) for the second level (some problems) in the third dimension (usual activities—work, family leisure). A final regression model was fitted in which this attribute weight was constrained to be equal to 1.

24. HUI3 levels have been mapped to EQ-5D+ levels through examining and matching as closely as possible the attribute-level definitions. There is no self-care dimension in the HUI3; the dexterity dimension in HUI3 has been mapped (approximately) to the self-care dimension. The HUI3 contains dimensions for vision and hearing loss whereas the EQ-5D+ does not. However, Dutch weights are available for 3 levels of hearing loss and 3 levels of vision loss and these have been used to include a comparison of the vision and hearing loss dimensions in Figure 2.4. The attribute levels are matched as shown in Table A.2.

25. The apparent close correspondence for vision and hearing loss weights is misleading. The vision and hearing dimensions of HUI3 have single attribute weights very consistent with the Dutch weights for mild, moderate and severe vision and hearing loss. However, HUI3 weights for mild hearing loss and vision loss are for conditions that are fully corrected by aids (spectacles, hearing aid). The Dutch weights are for the net impairment after correction.

Table A.2: Mapping of HUI3 levels to EQ-5D+ levels for Figure 2.5

Dimension	EQ-5D+ states	Comment
Mobility	No problems walking around	
	Some problems walking about	Average of ambulation states 3 and 4 (requires walking aids)
	Confined to bed	Average of ambulation states 5 and 6 (unable to walk alone even with aids + cannot walk at all)
Self-care	No problems washing or dressing	
	Some problems wash/dress	HUI3 dexterity level 2–4 (problems with fingers or hands)
	Unable to wash or dress	HUI3 dexterity level 5–6 (need help or unable to do most tasks)
Usual activities	No problems (work, family, leisure)	No comparable scale in HUI3
	Some problems	
	Unable to perform	
Pain/discomfort	No pain or discomfort	
	Moderate pain or discomfort	Average of pain states 2 and 3 (mild to moderate and moderate pain preventing activity)
	Extreme pain or discomfort	Average of pain states 4 and 5 (moderate to severe pain preventing activity and severe pain preventing activity)
Anxiety/ depression	Not anxious or depressed	Happy and interested in life
	Moderately anxious or depressed	Average of somewhat unhappy and very unhappy
	Extremely anxious or depressed	So unhappy that life is not worthwhile
Cognition	No problems cognitive function	
	Some cognitive problems	Somewhat forgetful, some problems with thinking and solving day to day problems
	Extreme problems	Average of states 5 and 6 (very forgetful, great difficulty or unable to solve day to day problems)
The following HUI3 dimensions are not in EQ-5D+ but Dutch weights for these states have been measured		
Vision	No problems with vision	
	Mild vision loss	Some difficulty reading newspaper, no difficulty recognising faces at 4m
	Moderate vision loss	Great difficulty reading newspaper, some difficulty recognising faces at 4 m
	Severe vision loss	Unable to read newspaper or recognise faces at 4m
Hearing	No problems with hearing	
	Mild hearing loss	Some difficulty in group conversation
	Moderate hearing loss	Great difficulty in group conversation, some difficulty one on one (average of HUI3 states 3,4)
	Severe hearing loss	Great difficulty one on one and unable to participate in group discussions (average of HUI3 states 5,6).

26. DISMOD© is a software program developed by the Burden of Disease Unit at the Centre for Health and Population Studies, Harvard, to assist disease experts to arrive at internally consistent estimates of incidence, duration and case fatality rates for the Global Burden of Disease Study. The program is based on a multi-state life table and uses various input parameters to derive consistent epidemiological estimates of disease incidence, duration and case fatality. Some of the input parameters are general (such as the age composition of the male or female population and the general mortality risk at each age) and others specific to the disease under consideration (such as instantaneous incidence and remission rates and cause-specific mortality risk). Outputs from the program include estimates of prevalence, average duration (before remission or death) and cause-

specific mortality by age. Because data on the prevalence of most conditions is easier to obtain than incidence rates, DISMOD is often used iteratively to find a set of incidence rates by age that match the observed prevalences, given estimates of remission rates and cause-specific mortality risk derived from population data or epidemiological studies.

27. For 2 conditions with weights w_1 and w_2, the weight for the comorbid state with both conditions is assumed to be

$$w_{12} = 1-(1-w_1)\times(1-w_2)$$

This is equivalent to assuming that the weights in QALY form (0=dead, 1=good health) are multiplicative. The combined weight is apportioned between the two conditions as follows:

a. Rank the conditions so that w_1 is the larger weight (more severe condition). The weight for this condition is taken to be w_1.

b. The comorbid weight attributed to the second condition is then the balance of the comorbid weight:

$$w_2^{adj} = w_{12} - w_1 = w_2 \times (1-w_1)$$

Example 1: if a person has ischaemic heart disease (weight 0.2) and diabetes (weight 0.07), then the adjusted weight for both conditions is 0.256 and the adjusted weight for diabetes 0.056.

Example 2: if a person has dementia (weight 0.44) and mild vision loss (weight 0.02), then the adjusted weight for both conditions is 0.45 and the adjusted weight for the vision loss is 0.01.

c. For 3 comorbid conditions, follow a similar procedure and sequentially attribute the additional weight to the second and third conditions (ranked in descending order of severity.

Example 3: if a person has dementia (weight 0.44), ischaemic heart disease (weight 0.2) and mild vision loss (weight 0.02), then the adjusted weight for all 3 conditions is 0.577 and the adjusted weights for the ischaemic heart disease and vision loss are 0.128 and 0.009 respectively.

28. Conditions for which comorbidity adjustments have been made at older ages are shown in Table A.3 below.

29. The IRSD is compiled initially at the Collector's District (CD) level, a census collection unit broadly equivalent in urban areas to a small group of suburban blocks, comprising approximately 250 dwellings (CDs in rural regions usually contain fewer dwellings). Lower IRSD scores are indicative of greater socioeconomic disadvantage. This study uses IRSD scores for Statistical Local Areas (SLAs), which in most cases correspond to council boundaries defined by Local Government Areas. IRSD scores for each SLA are constructed by taking the weighted average, using population counts from the 1986 and 1996 census, across all CDs comprising the SLA. In aggregate, SLAs cover the whole of Australia without gaps or overlaps.

30. The Gini coefficient is based on the Lorenz curve, and is widely used to measure income inequality in populations (Creedy 1996). The Lorenz curve can be used to examine the inequality in distribution of health outcome measures. In Figure 2.6, for example, the *x* and *y* ordinates could represent the cumulative proportion of people across small areas ranked in terms of decreasing mortality burden per capita and the cumulative total mortality burden respectively. If no inequality exists, the Lorenz curve corresponds to the diagonal line of equality. As the extent of inequality increases, so does the area between the line of equality and the Lorenz Curve. The Gini coefficient is defined as the area enclosed by the line of equality and the Lorenz Curve expressed as a proportion of the area below the diagonal and is bounded to range from zero (complete equality) to one (complete inequality).

Table A.3: Comorbidity adjustments for diseases with low disability weights and high prevalence at older ages

Category	Code	Prevalence at ages 65+	Disability weight	Comorbidity adjustment to weight	
				Age 65–74	Age 75+
Edentulism	S3	40.6%	0.004	0.946	0.872
Iron deficiency/mild anaemia	E2	2.7%	0.005	0.947	0.872
Osteoarthritis grade 2 (asympt.)	Q2	7.9%	0.010	0.953	0.873
Moderate anaemia	E2	0.6%	0.011	0.952	0.873
Vision loss—mild	K8c	7.1%	0.020	0.952	0.869
Hearing loss—mild 25–34 dB	K8d	24.7%	0.020	0.951	0.907
Urinary incontinence	O3	8.1%	0.025	0.955	0.915
Hearing loss—mild 35–44 dB	K8d	13.5%	0.028	0.937	0.898
Skin problems	P1, P2	1.6%	0.056	0.938	0.900
Non-melanoma skin cancer	F11	0.1%	0.058	0.938	0.900
Diabetes mellitus—cases	Ha	12.5%	0.070	0.951	0.918
Asthma	M2	5.6%	0.076	0.959	0.927
Hearing loss—moderate	K8d	13.4%	0.080	0.946	0.886
Angina	L2	5.1%	0.080	0.951	0.898
Osteoarthritis grade 2 (sympt.)	Q2	1.6%	0.140	0.942	0.889
Osteoarthritis grade 3 (asympt.)	Q2	6.1%	0.140	0.943	0.891
Melanoma	F10	0.5%	0.145	0.943	0.891
Hearing loss—severe	K8d	2.7%	0.153	0.976	0.864
Vision loss—moderate	K8c	2.2%	0.170	0.927	0.857
COPD	M1	6.9%	0.170	0.958	0.894
Peripheral arterial disease	L8	1.5%	0.243	0.977	0.888
Cancer—medium average weight	F14–16,19,22,24	2.3%	0.255	(a)	(a)
Heart failure	L2	0.3%	0.353	(a)	(a)
Cancer—high average weight	F3,4,7,8,12	2.5%	0385	(a)	(a)
Osteoarthritis grade 3 (sympt.)	Q2	2.4%	0.420	(a)	(a)
Vision loss—severe	K8c	2.1%	0.430	(a)	(a)
Dementia	K1	5.6%	0.440	(a)	(a)
Stroke	L3	1.9%	0.540	(a)	(a)

(a) Comorbidity adjustments not made for these conditions, although they are taken into account as comorbid conditions in calculating the comorbidity adjustments for lower severity conditions

31. There is extensive epidemiological evidence that socioeconomic disadvantage is causally related to higher mortality levels (Mathers 1994a, Wilkinson and Marmot 1998). Some but not all of the mortality differentials are mediated by differences in the prevalence of lifestyle risk factors such as tobacco smoking, physical inactivity, alcohol consumption, overweight and dietary risk factors.

Chapter 3

32. If male YLL are calculated using the cohort life expectancies for females, then the male excess mortality burden rises from 26%to 43%The latter figure includes the years of life lost due to the male-female gap in projected life expectancies in Australia. If YLL are not discounted, then the male excess

mortality burden is 31% based on projected cohort life expectancies for males and females and 53% if female life expectancies are used for both males and females.

Chapter 7

33. Estimation of the proportion of current disease burden that would be prevented in the future if exposure to the risk factor were eliminated requires answers to 'what if' questions. The contribution of the risk factor can be estimated by comparing the current level and projected future levels of a summary measure of population health with the levels that would be expected for some hypothetical or 'counterfactual' distribution of risk factor exposure. Counterfactual analysis requires a model that predicts the levels of a summary measure under an alternative hypothetical scenario. Sometimes these models are extremely simple but in the case of risk factors, which can have complex time and distributional characteristics, the models can be quite complex. The validity of the estimate depends on the validity of the model used to predict the counterfactual scenarios (Murray et al. 1999).

Counterfactual analysis of summary measures has a potentially wide spectrum of uses from the assessment of specific policies or actions to more general assessments of the contribution of diseases, injuries or risk factors. Murray et al. (1999) identified four major types of counterfactual scenario that may be used for this type of assessment:

- The effect of small changes in the disease, injury or risk factor can be assessed and the results expressed as the elasticity of the summary measure with respect to changes in the disease, injury or risk factor.
- The change in a summary measure expected with complete elimination of a risk factor can be assessed for some risk factors such as tobacco or alcohol use, but not for others such as blood pressure.
- The changes in future levels of a summary measure could be assessed for elimination of the risk for one year, followed by a return to the status quo at the end of the year. The health effects that are due to one year of risk exposure would then be traced out in terms of changes in future health expectancies or future burden.
- The change in a summary measure from the appplication of an intervention can be assessed.

More generally, Murray and Lopez (1999) have developed a classification of various counterfactual risk distributions that can be used for these purposes, including the theoretical minimum risk, the plausible minimum risk, the feasible minimum risk and the cost-effective minimum risk. They used the examples of tobacco and alcohol to explore the implications of using these different types of counterfactual distributions to define attributable burden and avoidable burden.

Chapter 8

34. Wolfson (1998) has outlined a vision of a coherent and integrated statistical framework, with summary measures of population health status at the apex of a hierarchy of related measures. Such a system should include the capability to 'drill down' below the summary measure to component parts such as incidence rates, prevalence rates, severity distributions, case fatality rates, etc. It should also allow us to 'drill down' below whole of population level to examine inequalities in health and to estimate the impacts of a given intervention on various sub-groups.