Mental Disorders

Burden of disease and

cost-effectiveness

Theo Vos

Mental Disorders: Burden of Disease and Cost-Effectiveness
Vos, Theo
Thesis Erasmus MC, University Medical Center Rotterdam,
with summary in English and Dutch

Front cover: '***Outside Aunt Rosa's Window***', by Paul Best, Brisbane 2005
Back cover: '***Anonymous Girl Seeks Secure Family Home***', by Paul Best, Brisbane 2005
Cover design: Anna Bosselaar, Zoiets
Layout: Theo Vos
Printed by: PrintPartners Ipskamp, Enschede

ISBN 90-9020440-7

Mental Disorders

Burden of disease and cost-effectiveness

Psychiatrische Aandoeningen

Ziektelast en kosten-effectiviteit

Thesis

to obtain the degree of Doctor from the Erasmus University Rotterdam by command of the rector magnificus

Prof.dr. S.W.J. Lamberts

and in accordance with the decision of the Doctorate Board
The public defence shall be held on Friday 31 March 2006 at 13.30 hours

by

Eric Theo Vos

Born at Westmaas, the Netherlands

Doctoral Committee

Promotor:	Prof.dr. P.J. van der Maas
	Prof.dr. J. Ormel
Other members:	Prof.dr. M.H.C. Donker
	Prof.dr. M.W. Hengeveld
	Prof.dr. A.H. Schene

Contents

1

General introduction

CHAPTER 1

INTRODUCTION

Health care expenditure in Australia is growing faster than other sectors of the economy [1]. The ageing of the population has added further concern about the affordability of future health care provision. Recently, the Treasurer commissioned the Intergenerational Report to estimate social expenditure over the next few decades [2]. The health care expenditure projections in this report were rather crude as it used a simplistic econometric approach to extrapolate past per capita expenditure patterns by age and imposed these on the projected changes in population size. A more informative analysis is possible taking into account that patterns of disease and mortality are changing and that there are differences in the ability to intervene depending on the underlying condition [3]. Nevertheless, the Report is an indication of the growing interest at the highest level of government in using evidence to inform social and health policy.

Evidence based medicine has become a well-established policy tool in clinical medicine. This body of work has led to numerous treatment guidelines based largely on systematic reviews by the Cochrane Collaboration (http://www.cochrane.org). More recently, evidence based policy was introduced as a term to describe the collation of evidence to inform health policymaking. For instance, in the UK it is being used in the name of the Centre for Evidence Based Policy and Practice of the Economic and Social Research Council (http://www.evidencenetwork.org) and in the title of the Evidence-Based Healthcare and Public Health journal (http://www.harcourt-international.com/journals/ebhc). The Australasian Cochrane Centre runs an Evidence-Based Policy Network to support Australian health care policy makers make best use of evidence (http://www.cochrane.org.au/ebpnetwork/ ebpnpublic).

The processing and presentation of evidence to inform health policy decision making is the topic of this thesis. Burden of Disease and Cost-Effectiveness Analysis are the two major themes that are applied to mental disorders and the treatment options for depression within the context of priority setting in Australia. The two themes are closely linked as a thorough

understanding of the epidemiology of a particular disease provided by Burden of Disease is an essential building block to the development of credible models for analysing cost-effectiveness.

Chapters 2 and 3 describe methodological issues related to estimating the Burden of Disease from mental disorders in Australia [4, 5]. The next four chapters are part of a larger series of publications on the Assessing Cost-Effectiveness (ACE)–Mental Health study [6-14]. The first of these four is a description of the epidemiological and economic modelling methods [8]. Two further chapters deal with the amount of depression in Australia that could potentially be averted with episodic or longer-term maintenance intervention strategies and the cost-effectiveness of these alternative approaches [11, 12]. Chapter 7 presents an overview of the ACE-Mental Health study [14].

In this introductory chapter, a general discussion on the two major themes of this thesis, Burden of Disease and cost-effectiveness, follows. The main issues arising from chapters 2–7 are discussed in a final discussion chapter.

THEME ONE: BURDEN OF DISEASE

In the early 90s, Murray and Lopez introduced a method to comprehensively measure the health of populations using the Disability-Adjusted Life Year (DALY), a summary measure that captures loss of health from premature mortality and non-fatal health states. A first set of global Burden of Disease estimates for eight world regions in 1990 was published in the 1993 World Development Report followed by more refined 1990 estimates in a series of books published in 1996 [15-17]. A further eight volumes detailing the epidemiological data sources and methods were planned but only one of these on sexual and reproductive health has come to publication [18]. This has meant that many of the national studies that followed had information on the overall methods but could not benefit from the disease-specific epidemiology and modelling assumptions used in the initial Global Burden of Disease study. Personally, it meant that the first country Burden of

Disease studies I conducted in Mauritius and Australia did not benefit from past experiences [19-23]. While this made the task more difficult it did inspire developmental work and led to considerable methodological improvements. Subsequently, many of these improved methods have been fed back into later versions of the Global Burden of Disease study including the approach to assessing the burden of non-fatal injuries from the Mauritius study and the detailed assessment of mental disorders from the Australian studies. The latter is described in the first two publications included in this thesis.

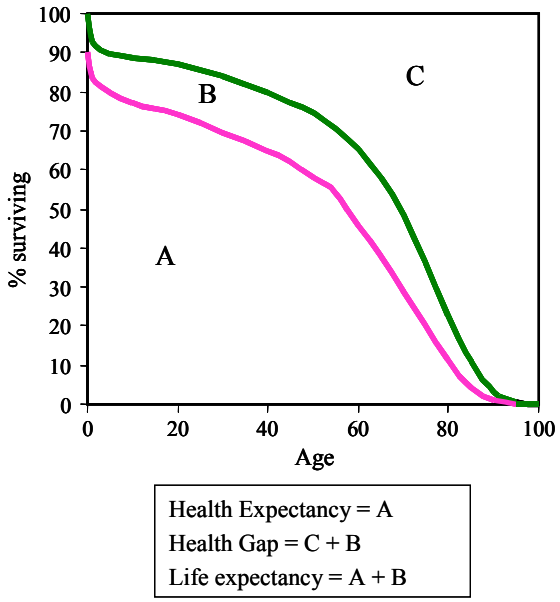

Figure 1 Life expectancy, health expectancy and health gap described in a survivorship curve.

Summary measures of population health

Burden of Disease studies quantify a population's health status using summary measures of population health (SMPH). Two classes of SMPH are distinguished: health expectancies and health gaps [24]. The difference between the two is best illustrated in a survivorship curve plotting the

proportion of a birth cohort on the Y-axis surviving to each age on the X-axis (Figure 1). The top curve represents the probability of surviving to each age in this hypothetical population. The life expectancy is the area under this curve (combined areas A and B in Figure 1).

Most commonly, life expectancy at birth is presented by calculating in a life table the hypothetical average amount of time lived if a cohort of newborns would be subjected to current age-specific mortality rates until they all have died. This is the so-called period life expectancy that appears in many annual country statistics publications such as the World Health Reports [25]. Life expectancy has intuitive appeal and is widely accepted though it is probably rarely understood that it makes the assumption of constant mortality rates and hence would underestimate true life expectancy in the majority of countries that have declining mortality trends.

Health expectancies are an adaptation of life expectancy adjusting time lived by the probability of loss of health due to non-fatal diseases. Area B in Figure 1 represents this amount of health loss and hence the health expectancy is defined as area A. There are several ways in which the amount of health loss is defined to estimate area B with the Disability-Free Life Expectancy and the Health-Adjusted Life Expectancy the most commonly used [24, 26]. Health expectancies are not discussed further as they do not feature in this thesis.

Health gaps: normative survivorship goal

The DALY, the common measure used for the work presented in this thesis, is a health gap measure. Health gaps are calculated as the difference between a population's current health status and a stated ideal that everyone in a population reaches old age free of disease. This ideal has been named the 'normative survivorship goal' [24]. In the example of Figure 1 the normative survivorship goal is set at 100 years. The mortality component of the health gap is defined by area C while the non-fatal component is the aggregate of severity-weighted health loss from all prevalent health states in the population.

The first normative survivorship goal was posited in 1947 by Dempsey introducing the concept of Potential Years of Life Lost (PYLL) to give a greater value to loss to deaths from tuberculosis at younger ages [27]. She used the 1946 life expectancy of US males and females by ethnic group in that analysis. Subsequently, researchers have used different cut-offs to determine PYLL, e.g. 60, 65 or 70 years. The drawback of this method is that it gives zero value to deaths occurring at ages above the cut-off. Unless a very high cut-off is chosen the implicit message to policy makers is that deaths in the elderly are unimportant. In reality, however, the health system in Australia provides lifesaving interventions such as influenza vaccination even to very old people and thus, implicitly gives a value to time lived in old age.

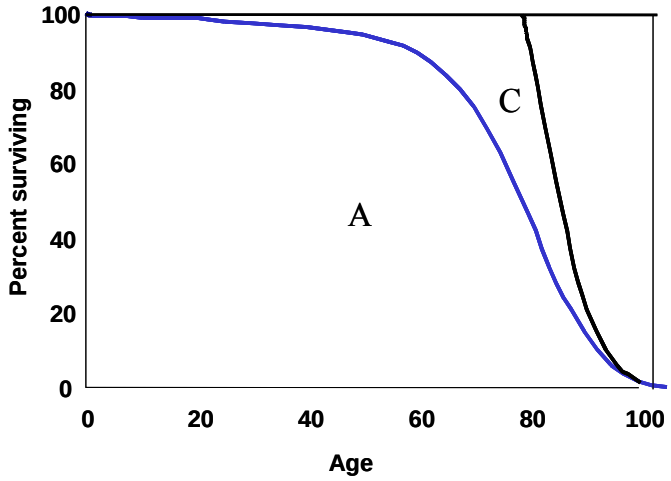

Figure 2 Survivorship curve of Australian males, 1996, and normative survivorship goal based on Coale and Demeny model life table West Level 25.

To overcome this problem, Murray used a life table approach to define a normative survivorship goal. At the time of the first global Burden of Disease study, the highest life expectancy at birth was enjoyed by Japanese women at around 82.5 years and mirrored that of the highest level 26 of the West model life tables of Coale and Demeny [28]. Arguing that the

differentials in life expectancy between males and females may in part be a reflection of a true biological difference and not all due to men's greater propensity to take health risks, the next female West Level 25 model life table with life expectancy at birth of 80 years was used to define the male normative survivorship goal. A female model life table was chosen as none of the male Coale and Demeny model life tables approached a life expectancy at birth at only a few years below the 82.5 set for females [17]. Figure 2 shows the survivorship curve for Australian males in 1996 with the normative survivorship curve based on the Global Burden of Disease study's standard life table and the 'mortality gap' defined by area C.

The advantage of the life table approach is that a death at any age is valued, as the life table will continue to give a life expectancy even at advanced ages. If the same standard life table is used in health gap comparisons between populations it has the desirable property of treating a death at a particular age equally regardless of the true life expectancy, thus meeting Murray's objective of treating 'like as like' [29]. I have argued elsewhere that there is a theoretical flaw in this method at a population level. If a population's mortality decreases, a greater proportion of people survives to older ages and their mortality gap will be estimated against a higher aggregate of age at death plus the remaining life expectancy at that age. In other words, the normative survivorship goal shifts to the right with potentially the undesirable property of increasing the health gap [30]. This violates the basic rule for a summary measure of population health that if mortality is reduced in a population a summary measure should improve. The only way to avoid this problem is to revert to using PYLL. If we also want to value a death at all ages the cut-off would have to be set at a very high age, e.g. 120 years. Such a high cut-off would give much greater emphasis to mortality at older ages relative to deaths at younger ages and also would give less emphasis to non-fatal health outcomes in comparison to mortality. It is inherent to the fact that this is a social value choice rather than an observed demographic or epidemiological parameter that there is no right or wrong choice. Rather, the decision on what method to use will depend on the emphasis one gives to theoretical soundness or the practical

consequences of the choice in terms of the resulting relative weight given to deaths at varying ages. A further argument is that the longer the current method is being used in the global and national Burden of Disease studies the more desirable it becomes not to change this fundamental assumption underlying the DALY. In practice, the shift in the normative survivorship goal with decreasing mortality in a population only has a small impact on DALY estimates and may be considered trivial. This illustrates an important principle in Burden of Disease research, i.e. concentrating efforts on issues that have the greatest bearing on the final estimates [31].

Health gaps: severity weights

A second important social value that has to be made in a health gap measure is how to value health loss from non-fatal conditions in comparison to health loss due to death. A plethora of methods exists to measure people's health-related quality of life and to elicit valuations for defined health states. This lack of consensus reflects how difficult the topic of health state valuation is. Currently most Burden of Disease studies use a combination of the original Global Burden of Disease disability weights and weights from a Dutch study [17, 32]. Both sets of weights are based on valuation panels using the person trade-off method [1] to elicit health state valuations [33]. The global Burden of Disease disability weights have been criticised (a) for having been set by a small group of international health experts; (b) for using the person trade-off method rather than the more commonly applied time trade-off or standard gamble methods; and c) for ignoring cultural differences in health state valuations [34-37]. The Global Burden of Disease valuation methods were replicated in 14 different countries representing all World Health Organisation regions. The ranking of disability severity for 17

[1] The person-trade-off is a technique used in economics and psychometry to ask people about their social values of different health care interventions. Individuals are asked to state how many outcomes of one kind they consider equivalent in social value to X outcomes of another kind [33].

health states was found to be relatively stable and highly correlated with the original disability weights [38].

Recently the World Health Organisation organised the World Health Survey in over 50 countries many of which included a health state valuation component (http://www3.who.int/whs/). The expectation is that this body of work will produce a new set of severity weights for Burden of Disease purposes but analyses of these data sets have yet to appear in publication. For comparisons of Burden of Disease across countries it is important that the same disability weights are used. For the first Australian Burden of Disease studies a conscious choice was made to concentrate on the epidemiological estimates and not to derive a set of Australian preferences for health states. At the time, that was considered the best use of limited analytical resources.

Health gaps: discounting

The Global Burden of Disease Study opted for an additional two social value choices, discounting and age weighting, that are not essential to SMPH such as the two previously discussed choices of the normative survivorship goal and disability weights. The issue of discounting is linked to the choice to base non-fatal health loss on incidence rather than prevalence. By definition, prevalence-based summary measures of disability are not discounted. Murray argues that an incidence approach for non-fatal health states fits better with the mortality component of DALYs which is also calculated as a future stream of lost health [39]. A further argument is made that it better reflects the potential for health gain and hence is better suited to inform policy making around prevention. The counterargument, of course, is that a prevalence approach is better suited to plan the health services response to current health problems. For diseases with large shifts in occurrence over time, such as an emerging epidemic of HIV or the decline in cardiovascular disease witnessed over the last few decades, incident Years Lived with Disability (YLD) estimates may vary considerably from prevalent YLD.

For instance, in 1999, in Thailand, incident YLD for HIV/AIDS were higher in women than in men while the prevalent YLD were much higher in

males [40]. This is because public health campaigns have very successfully reduced the rate of infection in young males visiting commercial sex workers but have been much less successful in curbing transmission from husband to wife. The many young men who became infected in the early stages of the epidemic are still alive and contribute to prevalence. They are also a source of infection for their wives and hence contribute to incidence in women [41]. The incident YLD would be more informative for priority setting of prevention while the prevalent estimates are more helpful in determining the current need for treatment. As both approaches have utility it is good practice to calculate both types of measures. However, in the standard reporting of Burden of Disease in DALYs the argument that incident YLD are more consistent with the way YLL are calculated should probably prevail.

Discounting of a stream of lost health is not common in epidemiology. In contrast, it is common practice in economic evaluations to discount future health outcomes as well as future costs [42, 43]. However, discounting future health gains or losses is not without controversy even among health economists and discounting is rarely used by epidemiologists and demographers for summary health measures [44]. The main argument for discounting is that people have time preference, i.e. they prefer to benefit (from income or, in this case, health) sooner rather than later. There are some specific arguments for applying discounting to the DALY in measuring population health [29]:

- to be consistent with measurement of health outcomes in cost-effectiveness analyses;
- to prevent giving excessive weight to deaths at younger ages; and
- the disease eradication/research paradox: assuming that investment in research or disease eradication has a non-zero chance of succeeding, then without discounting, all current expenditure should be shifted to such investment because the future stream of benefits is infinite.

Murray and Acharya conclude that the strongest argument for discounting is the disease eradication/research paradox [29]. They noted,

however, that the choice of a discount rate for health benefits, even if technically desirable, might result in morally unacceptable allocations between generations. Because the discount rate issue is not easily resolved, the GBD published discounted and undiscounted estimates of the global burden.

Health gaps: age weighting

A second, optional social value choice incorporated in the DALY is that of age weighting. The GBD Study weighted a year of healthy life lived at young and mid adult ages higher than at other ages. Relative to the value 1 if no age weighting were applied, greater weight is given to years of health life lost in young and middle-aged adults and lesser weight to loss of health in the very young and old (Figure 3). This choice was based on a number of studies that have indicated there is a social preference to value a year lived by a young adult more highly than a year lived by a young child or at older ages [45]. Some critics find age weights unacceptable on equity grounds (every year of life is of equal value *a priori*), others on empirical grounds (that the standard age weights do not well reflect actual social values) [35, 46-48]. Murray and Acharya have argued that age weights are not in themselves inequitable, because everyone potentially lives through every age, and that they do reflect legitimate societal priorities [29]. Interestingly, a small study among residents of a poor satellite town of Harare in Zimbabwe found age preferences closely aligned to the age weighting function of the GBD [49].

Because of the controversy about age weighting this social value choice was not incorporated in the DALYs calculated in a number of national Burden of Disease studies, e.g. in Australia, Thailand and the Netherlands [20, 40, 50]. The best argument not to include the GBD age-weighting factors is that a time-based mortality measure like YLL is an age weighting of deaths already and that adding a second age weight makes the GBD standard DALY unnecessarily complex [48].

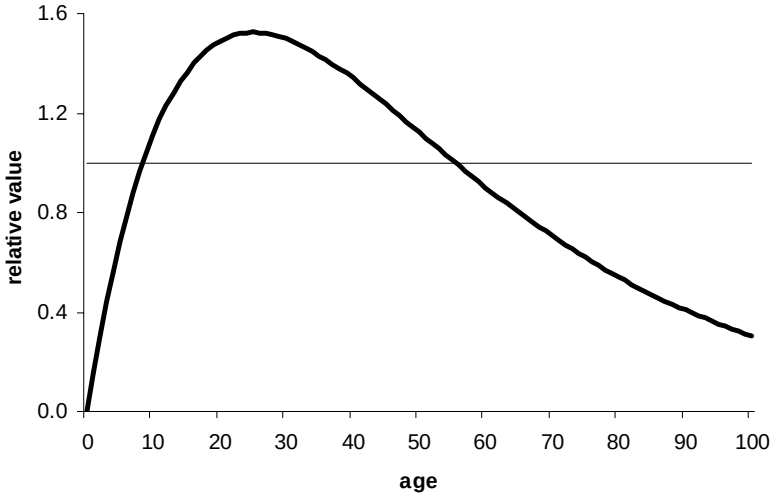

Figure 3 Age weights applied in DALYs.

Burden of Disease estimates for mental disorders

Burden of Disease measurements for mental disorders is the topic of two chapters in this thesis. While most mental disorders are associated with increased mortality risk [51], in vital registration systems there is little mortality attributed to mental disorders, apart from substance use disorders, as the underlying cause. However, mental disorders are common and often chronic and hence contribute a large proportion of YLD, the non-fatal component of DALYs. In 1990, mental disorders contributed 8% to the overall Burden of Disease in the world and 24% of all YLD [17].

When work commenced on the Burden of Disease studies in Australia and the state Victoria, all that was available from the GBD study related to its estimates for mental disorders were the disability weights and tables of incidence, prevalence, duration, YLD and DALYs. The background papers for each of the disorders prepared by expert groups assisting the GBD study have not been published to date.

Two factors facilitated the large effort to improve Burden of Disease estimates for mental disorders in Victoria. First, the National Survey of

Mental Health and Wellbeing (NSMHWB) had just been released providing up-to-date estimates of the occurrence of the main mental disorders in Australia [52]. Towards the end of the study, results from two sub-component studies of the main survey on psychotic disorders and childhood mental disorders also became available [53, 54]. Professor Gavin Andrews of the University of New South Wales was the driving force behind the main component of the survey. He provided invaluable assistance in understanding the complex questionnaire and derivation of ICD-10 and DSM-IV diagnoses.

Second, a group of mental health clinicians and epidemiologists enthusiastically responded to a request to join an expert panel to guide the difficult choices and assumptions that had to be made. In consultation with this expert panel, a number of methodological issues were queried in the GBD approach to estimating the burden of mental disorders and developed into the following research questions addressed in the first two chapters of this thesis:

1. *Which mental disorders ought to be included in a comprehensive assessment of Burden of Disease?*

   The GBD included eight mental disorders (major depression, bipolar disorder, schizophrenia, alcohol use, drug use, panic disorder, obsessive-compulsive disorder and post-traumatic stress disorder) and grouped them together with 4 neurological disorders (dementia, epilepsy, multiple sclerosis and Parkinson's disease) in one category of neuropsychiatric conditions [17]. Childhood mental disorders, eating disorders, very common anxiety disorders such as generalised anxiety disorder and social phobia as well as a lack of differentiation in the estimates of illicit drug use were identified as missing in the GBD list of conditions.

2. *Does a single disability weight per disorder suffice to describe the variation in severity of mental disorders?*

   The GBD study used a treated and untreated disability weight for each of the mental disorders. If these disability weights are then used in different countries, an implicit assumption of a constant severity distribution is

made. Given the great variability in the severity of mental disorders it was considered desirable to make this more explicit in the calculations. This was possible due to the detail of information on the level of severity in the NSMHWB and the availability of disability weights from the Dutch Burden of Disease study with three or four levels of severity for most mental disorders.

3. *How much does a co-morbidity correction affect Burden of Disease estimates for mental disorders?*

   The GBD 1990 study ignored the common co-morbidity between mental disorders found in all epidemiological surveys (e.g. [55-57]). This leads to considerable overestimation particularly if more mental disorders are included. Co-morbidity between anxiety disorders and depression is so common that some have questioned the boundaries between diagnostic categories. The fact that similar psychological and drug treatments are found to be effective in anxiety and depression adds a further argument to the hypothesis that these may actually be different expressions of the same disorder [58, 59].

4. *Are the GBD estimates of prevalence, incidence and average duration for mental disorders applicable to the Australian context?*

   Without the background papers justifying the assumptions underlying the GBD estimates for mental disorders we could only scrutinise the actual estimates of incidence, prevalence and average duration [16]. The durations estimated in the GBD for bipolar disorder (less than 1.5 years), anxiety disorders (ranging from 0.75 to 2.5 years) and alcohol use (less than two years) did not seem to fit with findings from follow-up studies for these disorders [60-62]. The question then arises how sensitive YLD results are to different assumptions of remission. An incidence-prevalence-mortality (IPM) model, DisMod, was developed for the GBD to derive missing epidemiological disease parameters and check known disease parameters for consistency [63]. DisMod makes use of the chronological sequence of disease events: incidence must precede prevalence and disease-specific mortality follows prevalence. If three of

five disease parameters (incidence, prevalence, duration, remission and excess mortality) are defined, DisMod will derive the missing two parameters. Unless case fatality for a disease is very high, remission largely determines average duration. Burden of Disease estimates for mental disorders are all based on prevalence data as incidence is rarely measured in studies representative of the whole population. Varying the remission assumption in DisMod will give different combinations of incidence and average duration: long duration and low incidence if remission is low and vice versa. YLD are the product of incidence and duration multiplied by a severity weight. Thus, even a great difference in the remission assumption causes only small variation in YLD estimates. When discounting is applied, the difference is accentuated a bit more as the combination of low incidence and long duration will be discounted more than that of high incidence and short duration. In other words, YLD estimates are not very sensitive to assumptions of remission. However, estimates of prevalence and incidence are important outputs from a Burden of Disease study in their own right used to determine health service needs and as inputs to economic evaluation models. Moreover, experts may reject the DALY estimates if they are based on implausible estimates of duration.

Also, the expert panel raised doubts about the prevalence estimates for some disorders, e.g. the high prevalence of obsessive-compulsive disorder in relation to other anxiety disorders.

THEME TWO: COST-EFFECTIVENESS ANALYSES

A number of health economists have been vocal in their criticism of Burden of Disease as a tool to inform priority setting [37, 64, 65]. A central point in their arguments is that Burden of Disease assessments are a waste of scarce intellectual resources as economic evaluation of health interventions rather than the size of health problems should inform priority setting in health. This is based on a rather selective reading of the literature. From the first publication on Burden of Disease onwards emphasis has been given to the

essential link between Burden of Disease and cost-effectiveness [15]. The strongest argument in support of Burden of Disease as an essential component of evidence-based policy is that policy makers, when presented with economic information of what the 'good buys' are amongst a choice of health interventions, will want to know the financial consequences of adopting these interventions. That information can only come from knowledge about the number of people affected by the health problems addressed by each of the interventions. Moreover, the assessment of health benefits, a key component of cost-effectiveness analysis, requires a good understanding of the epidemiology (e.g. remission, severity and case-fatality) of the diseases affected by the intervention. Of course, that is the 'bread and butter' of Burden of Disease assessments. Indeed, a counterargument to these critics is that economic evaluation without knowing the size of a health problem in the population is an inadequate use of scarce intellectual resources.

The literature provides little guidance on how best to conduct economic analyses for priority setting across multiple interventions and multiple areas of health. Most economic analyses have been undertaken to choose between two or a few alternatives addressing the same health problem. The broader the scope of a priority setting exercise, the more comparability of results becomes an essential prerequisite. The use of leagues tables of cost-effectiveness results derived from different studies using a variety of methods has rightfully been criticised [43]. However, that does not mean that it is impossible to meet the information needs of policy makers who want to set priorities across a range of interventions and disease areas. For valid comparisons between the interventions included in the league table it is paramount that the same methods have been used across all analyses. Despite the efforts by a consensus panel of US health economists and the seminal textbook on cost-effectiveness by Drummond et al. to guide cost-effectiveness studies there remains a great variety in essential components of economic evaluation methods, including [66, 67]:

- the choice of comparator (i.e. current practice or no intervention);

- the perspective (i.e. government, health sector or societal perspective);
- the horizon of implementation of the intervention and measurement of costs and benefits (ranging from the very short term to lifelong);
- the use and level of discounting (varying from 0-10%);
- the target population (implementing the intervention to a cohort of eligible recipients or modelling 'real life' implementation of the intervention over time to a population); and
- the units of measuring benefits (i.e. natural units such as cases averted, QALYs, DALYs or in dollar terms).

Traditionally, economic analyses have often been conducted in conjunction with randomised controlled trials examining the efficacy of interventions. These are expensive studies. It would be unaffordable to set up a large number of such trials for all information needs of policy makers for priority setting. Moreover, trials often exclude potentially eligible recipients of an intervention based on gender, race or co-morbidity; are conducted over short periods of time; tend to be underpowered to detect adverse effects; and often do not collect cost information. Hence, a modelling approach is advocated to make most efficient use of the evidence generated by trials and to make the most plausible interpolations and extrapolations to fill in essential data gaps [68]. For multi-intervention, multi-health problem economic evaluations it is inevitable that they are undertaken as modelling exercises making use of best available evidence on current practice, effectiveness, costs and outcomes from a variety of existing data sources. That is a practice closely resembling the principles of Burden of Disease studies which also collect, process and evaluate existing data rather than collect new information.

The WHO-CHOICE (CHOosing Interventions that are Cost-Effective) project is an example of an economic modelling exercise addressing a large number of health problems. WHO-CHOICE uses the same methods of analysing the costs and benefits for hundreds of different interventions. In contrast to common health economic practice, it advocates generalised cost-effectiveness (GCEA) rather than marginal analyses [69]. Its aim is to

formulate the most appropriate mix of cost-effective interventions to address a health problem regardless of current practice. This can only be done after back calculating a hypothetical 'clean sheet' or a 'null option' by estimating the burden that would exist if no health care interventions existed for the health problem studied. Average cost-effectiveness ratios are then modelled for the implementation of separate interventions and combinations thereof to determine the 'ideal' mix of cost-effective interventions addressing a particular health problem given fixed budget constraints.

The WHO-CHOICE GCEA approach was developed to overcome a problem of marginal analysis, i.e. that replacing cost-ineffective current practice with a more favourable intervention option can artificially make the new intervention look better than it actually is. The advantage of GCEA is that it allows reconsideration of current practice and provides an opportunity to redress any major inefficiencies. The challenge is to make credible assumptions about the null-option and to estimate the combined effect of interventions. However, the objective is not to make an accurate estimate of the null option but to use it as a starting point to estimate the costs and benefits of alternative mixes of interventions. By using the same input parameters on effectiveness, adherence and disease in both the back calculation to the null option and the forward calculations of the avoidable burden, consistency between the two approaches can be achieved. In other words, if the current mix of interventions for a health problem is modelled from the null option it should produce the current level of burden for that health problem.

Priority setting and cost-effectiveness modelling in Australia

In Australia, all publications on the Australian Burden of Disease studies have reiterated the complementarity of Burden of Disease and cost-effectiveness information to inform priority setting [4, 5, 22, 70, 71]. This was first put in practice in a priority setting exercise conducted in 2000 in collaboration with the Cancer Strategies Group of the National Health Priority Council in Australia [72]. Key features of this project were:

- an economic protocol specifically developed for a priority setting context;
- an evidence-based approach with a small research team bringing together the best available information;
- the choice of the Disability Adjusted Life Year (DALY) as the measure of health gain;
- the use of marginal cost-effectiveness analysis defining interventions as opportunities for change from current practice;
- the recognition that "benefit" is broader than just health gain; and
- adoption of a two-stage approach to the assessment of benefit involving both "technical" aspects (i.e. cost per DALY recovered; level of evidence) and "judgment" aspects (i.e. equity; size of the problem; acceptability to stakeholders; and feasibility of implementation).

Two subsequent economic modelling projects, ACE-Heart Disease and ACE-Mental Health, used and further developed the economic evaluation methods trialled in the cancer study. I was the lead investigator on both studies. Not having had formal health economic training, continued collaboration with Rob Carter was essential. While my original input largely concentrated again on the modelling methods of health outcomes, all decisions on the costing and overall economic protocol were jointly deliberated.

The ACE-Mental Health study was conceived after a presentation on the first Australian Burden of Disease study findings to Department of Health and Ageing staff in Canberra in early 2000. After the presentation key policy makers showed interest in the ACE-Heart Disease project for which I had just received funding from the National Health and Medical Research Council. I sketched the challenges ahead in terms of data requirements and development of modelling tools to carry out comparable economic evaluations even in the area of cardiovascular health for which there are good epidemiological data and a well-established evidence base of effectiveness. Casually, I mentioned that a similar undertaking addressing the efficiency of mental health care would pose a much greater challenge

due to the weaker evidence base on effectiveness; and the need for considerable developmental work on translating effectiveness of interventions into a measurement of health gain. The immediate response was an invitation to develop a proposal as the department considered this a priority area. Similar discussions with Mental Health Branch staff at the Department of Human Services in the state Victoria led to a request for joint funding from the federal and state government to establish the ACE-Mental Health study. The main study questions were:

1. *Can the diverse measures on the impact of interventions reported in the international literature be summarised in a common metric for use in an economic evaluation modelling exercise allowing meaningful comparisons of the cost-effectiveness of various interventions for the main mental disorders affecting Australians?*

   The main difference between mental disorders and most other health outcomes is that the disease burden is largely non-fatal. It is much easier to measure the impact of an intervention on mortality, as death is an unambiguous event. Quantifying the impact of an intervention on non-fatal outcomes is much harder, particularly if the main goal of intervention is to reduce severity rather than preventing complete disease episodes. Few intervention studies of mental health interventions have reported on the impact of overall quality of life. Largely, results are reported as changes in disease-specific symptom questionnaires. Therefore, the key challenge was to find a comparable way of summarising these diverse measures of impact. In systematic reviews, mental health epidemiologists tend to use the standardised mean difference as the effect size to summarise results that are described as a change in continuous symptom scores [73]. However, for many mental health interventions there are no up-to-date systematic reviews and no consistency in summarising impact across different outcome measures reported in individual trials.

2. *Can the findings from the literature on impact be translated into a change in DALYs, the measure of health outcomes chosen for ACE-Mental Health?*

   Despite the methodological developments in quantifying the severity of mental disorders described in chapters 2 and 3, the DALY remains a coarse instrument with between two and four levels of disability weights to describe the heterogeneity in the presentation of the main mental disorders. While for Burden of Disease measurements these coarse severity gradings match the limited available epidemiological data on disease occurrence by severity, the measurement of small to moderate changes in severity reported for most mental health interventions demands greater accuracy. Thus, new methods had to be developed for the ACE-Mental Health study.

3. *Focusing on interventions for depression, are there credible methods of comparing the effectiveness of drug and psychological interventions for the acute and longer-term management of depression?*

   In the Australian Burden of Disease studies, anxiety disorders, schizophrenia and bipolar disorder are modelled as chronic or chronic episodic disorders. Major depression is treated as an episodic disorder despite considerable evidence that in most individuals it is a chronic episodic illness. At the time of the Burden of Disease study key epidemiological information on the long-term course of illness was missing making it difficult to model major depression as the chronic disease it is with incidence mostly in childhood and early adulthood and frequent recurrence of episodes thereafter. As longer-term maintenance treatment strategies aim to prevent recurrence of episodes it was necessary to develop methods of modelling depression as a chronic disorder specifying the risk of recurrence of disease following an acute episode.

4. *Taking into account the uncertainty around key input variable to cost-effectiveness models can plausible differences in the cost-effectiveness of*

different treatment approaches for mental disorders in Australia be distinguished? and

5. *Can mental health policy makers and experts in Australia be meaningfully engaged in a priority setting exercise?*

   The ACE-Mental Health study aimed to combine a strong focus on the technical analyses of cost-effectiveness with involvement of policy makers and opinion leaders in framing recommendations for policy change by explicitly placing the results on cost-effectiveness in the context of other important policy considerations, the 'judgment' factors described above for the pilot study on prioritising cancer control interventions. Two main challenges were envisaged: (a) the complex financing structure of health care in Australia that is shared by state and federal governments; and (b) anecdotal evidence of the polarisation between mental health opinion leaders in the country who have a strong allegiance to research and services for either the severe spectrum mental disorders such as schizophrenia or the more common disorders such as depression and anxiety.

The four chapters on ACE-Mental Health, mainly address the first four of these study questions. The first of these describes the overall economic and epidemiological methods. Two further chapters deal with the economic evaluation of intervention strategies for major depression. Chapter 7 reports on the study's overall findings and final recommendations. It does not give much insight into the process of reaching these conclusions despite the competing interests often forcefully voiced by policy makers and experts represented in the steering committee. In the final chapter some attention will be given to this process although a complete policy analysis of the contribution of ACE-Mental Health to the priority setting debate in mental health care in Australia is beyond the scope of this thesis.

REFERENCES

1. Australian Institute of Health and Welfare. *Health system expenditure on disease and injury in Australia, 2000-01*. Canberra: AIHW (Health and Welfare Expenditure Series no. 19), 2004.
2. Treasury. *Intergenerational Report 2002-03*. Canberra: Treasury, 2002.
3. Begg S, Vos T, Goss J. *Australian Burden of Disease and Injury Study: Projected Health Care Costs*. Brisbane: University of Queensland
4. Vos T, Mathers CD. The burden of mental disorders: a comparison of methods between the Australian burden of disease studies and the Global Burden of Disease study. *Bulletin of the World Health Organization* 2000; 78:427-438.
5. Vos T, Mathers CD, Herrman H, Harvey C, Gureje O, Bui D, Watson N, Begg S. The burden of mental disorders in Victoria, 1996. *Social Psychiatry and Psychiatric Epidemiology* 2001; 36:53-62.
6. Chalamat M, Mihalopoulos C, Carter R, Vos T. Supported employment for schizophrenia. *Australian and New Zealand Journal of Psychiatry* 2005; 39:693-700.
7. Donnelly M, Haby MM, Carter R, Andrews G, Vos T. Cost-effectiveness of dexamphetamine and methylphenidate for the treatment of childhood attention deficit hyperactivity disorder. *Australian and New Zealand Journal of Psychiatry* 2004; 38:592-601.
8. Haby MM, Carter R, Mihalopoulos C, Magnus A, Sanderson K, Andrews G, Vos T. Assessing Cost-Effectiveness - Mental Health: Introduction to the study and methods. *Australian and New Zealand Journal of Psychiatry* 2004; 38:569-578.
9. Haby MM, Tonge B, Littlefield L, Carter R, Vos T. Cost-effectiveness of cognitive behavioural therapy and selective serotonin reuptake inhibitors for major depression in children and adolescents. *Australian and New Zealand Journal of Psychiatry* 2004; 38:579-91.
10. Magnus A, Carr V, Mihalopoulos C, Carter R, Vos T. Assessing cost-effectiveness of drug interventions for schizophrenia. *Australian and New Zealand Journal of Psychiatry* 2005; 39:44-54.
11. Vos T, Haby MM, Barendregt JJ, Kruyshaar ME, Corry J, Andrews G. The burden of major depression avoidable by longer-term treatment strategies. *Archives of General Psychiatry* 2004; 61:1097-1103.

12. Vos T, Corry J, Haby MM, Carter R, Andrews G. Cost-effectiveness of CBT and drug interventions for episodes of major depression. *Australian and New Zealand Journal of Psychiatry* 2005; 39:683-692.
13. Heuzenroeder L, Donnelly M, Haby MM, Mihalopoulos C, Rossell R, Carter R, Andrews G, Vos T. Cost-effectiveness of psychological and pharmacological interventions for generalized anxiety disorder and panic disorder. *Australian and New Zealand Journal of Psychiatry* 2004; 38:602-12.
14. Vos T, Haby MM, Magnus A, Mihalopoulos C, Andrews G, Carter R. Assessing Cost-Effectiveness (ACE)-Mental Health: Helping policy makers prioritise and plan health services. *Australian and New Zealand Journal of Psychiatry* 2005; 39:701-712.
15. World Bank. *World Development Report 1993: Investing in Health*. New York: Oxford University Press, 1993.
16. Murray CJL, Lopez AD. *Global Health Statistics: a compendium of incidence, prevalence and mortality estimates for over 200 conditions*. Cambridge, Mass: Harvard University Press, 1996.
17. Murray CJL, Lopez AD. *The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020*. Cambridge, Mass: Harvard University Press, 1996.
18. Murray CJM, Lopez AD. *Health Dimensions of Sex and Reproduction: the global burden of sexually transmitted diseases, HIV, maternal conditions, perinatal disorders and congenital anomalies*. Cambridge, Mass: Harvard University Press, 1996.
19. Mathers CD, Vos T, Stevenson CE. *The burden of disease and injury in Australia*. Canberra: Australian Institute of Health and Welfare, 1999.
20. Mathers CD, Vos ET, Stevenson CE, Begg SJ. The Australian Burden of Disease Study: measuring the loss of health from diseases, injuries and risk factors. *Medical Journal of Australia* 2000; 172:592-596.
21. Vos T, Begg S. *The Victorian Burden of Disease Study: mortality*. Melbourne: Public Health and Development Division, Department of Human Services, 1999.
22. Vos T, Begg S. *The Victorian Burden of Disease Study: morbidity*. Melbourne: Public Health and Development Division, Department of Human Services, 1999.
23. Vos T, Timæus I, Gareeboo H, Roussety F, Huttly S, Murray CJL. *Mauritius Health Sector Reform, National Burden of Disease Study. Final report of consultancy to Ministry of Health and Ministry of*

Economic Planning and Development, Mauritius. London: London School of Hygiene and Tropical Medicine, 1995.

24. Murray CJ, Salomon JA, Mathers C. A critical examination of summary measures of population health. *Bull World Health Organ* 2000; 78:981-94.
25. World Health Organization. *World Health Report 2004*. Geneva: WHO, 2004.
26. Robine JM, Ritchie K. Healthy life expectancy: evaluation of global indicator of change in population health. *British Medical Journal* 1991; 302:457-60.
27. Dempsey M. Decline in Tuberculosis: the death rate fails to tell the entire story. *American Reviews of Tuberculosis* 1947; 56:157-164.
28. Coale A, Demeny P, Vaughan B. *Models of mortality and age composition*. New York: Academic Press, 1983.
29. Murray CJ, Acharya AK. Understanding DALYs (disability-adjusted life years). *Journal of Health Econonomics* 1997; 16:703-30.
30. Vos T. Shifting the goal post - normative survivorship goals in health gap measures. In *Summary Measures of Population Health*. Geneva: World Health Organization, 2002.
31. Mathers CD, Vos T, Lopez AD, Salomon J, Lozano R, Ezzati M (eds): National Burden of Disease Studies: a practical guide. Edition 2.0. :. Geneva, World Health Organization
32. Stouthard MEA, Essink-Bot ML, Bonsel GJ, Barendregt JJ, Kramer PGN, van de Water HPA, Gunning-Schepers LJ, van der Maas PJ. *Disability weights for diseases in the Netherlands*. Rotterdam: Department of Public Health, Erasmus University, 1997.
33. Nord E. The person trade-off approach to valuing health care programs. *Medical Decision Making* 1995; 15:201-208.
34. Allotey P, Reidpath D, Kouame A, Cummins R. The DALY, context and the determinants of the severity of disease: an exploratory comparison of paraplegia in Australia and Cameroon. *Social Science and Medicine* 2003; 57:949-58.
35. Anand S, Hanson K. Disability-adjusted life years: a critical review. *Journal of Health Economics* 1997; 16:685-702.
36. Reidpath DD, Allotey PA, Kouame A, Cummins RA. Measuring health in a vacuum: examining the disability weight of the DALY. *Health Policy and Planning* 2003; 18:351-6.
37. Williams A. Calculating the global burden of disease: time for a strategic reappraisal? *Health Economics* 1999; 8:1-8.
38. Üstün B, Rehm J, Chatterji S. Are disability weights universal? Ranking of the disabling effects of different health conditions in 14

countries by different informants. In *Summary Measures of Population Health: concepts, ethics, measurement and applications*. Geneva: World Health Organization, 2002.

39. Murray CJL. Rethinking DALYs. In *The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020*. Cambridge, Mass: Harvard University Press, 1996.
40. Bundhamchareon K, Teerawatananon Y, Vos T, Begg S. *Burden of Disease and Injuries in Thailand.* Nonthaburi, Thailand: Ministry of Public Health, 2002.
41. The Thai working group on HIV/AIDS projection. *Projections for HIV/AIDS in Thailand: 2000-2020*. Nonthaburi, Thailand: Division of AIDS, Department of Communicable Disease Control, Ministry of Public Health, 2001.
42. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-effectiveness in Health and Medicine. *Journal of the American Medical Association* 1996; 276:1253-8.
43. Drummond MF, Torrance GW, Mason JM. Cost-effectiveness league tables: More harm than good? *Social Science and Medicine* 1993; 37:33-39.
44. Severens JL, Milne RJ. Discounting health outcomes in economic evaluation: the ongoing debate. *Value and Health* 2004; 7:397-401.
45. Murray CJ. Quantifying the burden of disease: the technical basis for disability-adjusted life years. *Bulletin of the World Health Organization* 1994; 72:429-45.
46. Arnesen T, Kapiriri L. Can the value choices in DALYs influence global priority-setting? *Health Policy* 2004; 70:137-49.
47. Tsuchiya A. Age-related preferences and age weighting health benefits. *Social Science and Medicine* 1999; 48:267-76.
48. Barendregt JJ, Bonneux L, Van der Maas PJ. DALYs: the age-weights on balance. *Bulletin of the World Health Organization* 1996; 74:439-43.
49. Jelsma J, Shumba D, Kristian H, De Weerdt W, De Cock P. Preferences of urban Zimbabweans for health and life lived at different ages. *Bulletin of the World Health Organization* 2002; 80:204-9.
50. Melse JM, Essink-Bot ML, Kramers PG, Hoeymans N. A national burden of disease calculation: Dutch disability-adjusted life-years. Dutch Burden of Disease Group. *American Journal of Public Health* 2000; 90:1241-7.

51. Harris EC, Barraclough B. Excess mortality of mental disorder. *British Journal of Psychiatry* 1998; 173:11-53.
52. Australian Bureau of Statistics. *Mental health and wellbeing: Profile of adults, Australia, 1997.* Canberra: ABS, 1998.
53. Sawyer MG, Arney FM, Baghurst PA, Clark JJ, Graetz BW, Kosky RJ. *The Mental Health of Young People in Australia*. Canberra: Mental Health and Special Programs Branch, Commonwealth Department of Health and Aged Care, 2000.
54. Jablensky A, McGrath J, Herrman H, Castle D, Gureje O, Evans M, Carr V, Morgan V, Korten A, Harvey C. Psychotic disorders in urban areas: an overview of the Study on Low Prevalence Disorders. *Australian and New Zealand Journal of Psychiatry* 2000; 34:221-36.
55. Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, Wittchen HU, Kendler KS. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Archives of General Psychiatry* 1994; 51:8-19.
56. Bijl RV, Ravelli A, van Zessen G. Prevalence of psychiatric disorder in the general population: results of The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social Psychiatry and Psychiatric Epidemiology* 1998; 33:587-95.
57. Andrews G, Henderson S, Hall W. Prevalence, comorbidity, disability and service utilisation. Overview of the Australian National Mental Health Survey. *British Journal of Psychiatry* 2001; 178:145-53.
58. Andrews G, Neilson M, Hunt C, Stewart G, Kiloh LG. Diagnosis, personality and the long-term outcome of depression. *Br J Psychiatry* 1990; 157:13-8.
59. Andrews G, Slade T, Issakidis C. Deconstructing current comorbidity: data from the Australian National Survey of Mental Health and Well-Being. *Br J Psychiatry* 2002; 181:306-14.
60. Cunningham JA, Lin E, Ross HE, Walsh GW. Factors associated with untreated remissions from alcohol abuse or dependence. *Addictive Behaviors* 2000; 25:317-21.
61. Angst J, Preizig M. Outcome of a clinical cohort of unipolar, bipolar and schizoaffective patients: Results of a prospective study from 1959 to 1985. *Schweizerische Archiv für Neurologie und Psychiatrie* 1995; 146:1-16.
62. Yonkers KA, Bruce SE, Dyck IR, Keller MB. Chronicity, relapse, and illness--course of panic disorder, social phobia, and generalized anxiety disorder: findings in men and women from 8 years of follow-up. *Depression and Anxiety* 2003; 17:173-9.

63. Barendregt JJ, Van Oortmarssen GJ, Vos T, Murray CJ. A generic model for the assessment of disease epidemiology: the computational basis of DisMod II. *Popul Health Metr* 2003; 1:4.
64. Mooney G, Wiseman V. Burden of disease and priority setting. *Health Economics* 2000; 9:369-72.
65. Mooney G, Irwig L, Leeder S. Priority setting in health care: unburdening from the burden of disease. *Australian and New Zealand Journal of Public Health* 1997; 21:680-1.
66. Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-effectiveness in health and medicine*. New York: Oxford University Press, 1996.
67. Drummond MF, O'Brien B, Stoddart GL, Torrance GT. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford Medical Publications, 1997.
68. Liew D, McNeil JJ, Peeters A, Lim SS, Vos T. Epidemiological modelling (including economic modelling) and its role in preventive drug therapy. *Med J Aust* 2002; 177:364-7.
69. Murray CJ, Evans DB, Acharya A, Baltussen RM. Development of WHO guidelines on generalized cost-effectiveness analysis. *Health Economics* 2000; 9:235-51.
70. Mathers CD, Vos ET, Stevenson CE, Begg SJ. The Australian Burden of Disease Study: measuring the loss of health from diseases, injuries and risk factors. *Medical Journal of Australia* 2000; 172:592-6.
71. Mathers CD, Vos ET, Stevenson CE, Begg SJ. The burden of disease and injury in Australia. *Bulletin of the World Health Organization* 2001; 79:1076-84.
72. Carter R, Stone C, Vos T, Hocking J, Mihalopoulos C, Peacock S, Crowley S. *Trial of Program Budgetting and Marginal Analysis (PBMA) to assist cancer control planning in Australia*. Canberra: Commonwealth Department of Health and Aged Care, 2000.
73. Deeks JJ, Altman DG, Bradburn MJ. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In Systematic reviews in health care: meta-analysis in context. London: BMJ Publishing Group, 2001.

THEME A

The burden of mental disorders in Australia

2

The burden of mental disorders in Victoria, 1996

Vos T, Mathers C, Herrman H, Harvey C, Gureje O, Bui D, Watson N, Begg S (2001). The burden of mental disorders in Victoria, 1996. *Social Psychiatry and Psychiatric Epidemiology,* 36:53–62.

SUMMARY

Background – Between 1998 and 1999, a burden of disease assessment was carried out in Victoria, Australia applying and improving on the methods of the Global Burden of Disease Study. This paper describes the methods and results of the calculations of the burden due to 22 mental disorders adding 14 conditions not included in previous burden of disease estimates.

Methods – The National Survey of Mental Health and Wellbeing provided recent data on the occurrence of the major adult mental disorders in Australia. Data from international studies and expert advice further contributed to the construction of disease models describing each condition in terms of incidence, average duration and level of severity with adjustments for co-morbidity with other mental disorders. Disability weights for the time spent in different states of mental ill health were borrowed mainly from a study in the Netherlands supplemented by weights derived in a local extrapolation exercise.

Results – Mental disorders were the third largest group of conditions contributing to the burden of disease in Victoria ranking behind cancers and cardiovascular diseases. Depression was the greatest cause of disability in both men and women. Eight other mental disorders in men and seven in women ranked among the top twenty causes of disability.

Conclusions – Insufficient information on the natural history of many of the mental disorders, the limited information on the validity of mental disorder diagnoses in community surveys and considerable differences between ICD–10 and DSM–IV defined diagnoses were the main concerns about the accuracy of the estimates. Similar and often greater concerns were raised during the estimation of the burden from common non-fatal physical conditions such as asthma, diabetes and osteoarthritis. In comparison, psychiatric epidemiology can boast greater scientific rigour in setting standards for population surveys.

BACKGROUND

Over the last two decades, Australian Health Ministers have identified mental health as a priority area with the potential for significant health gain at the population level. Initially the focus was on the development of a set of goals and targets for improving health and reducing inequalities [1, 2]. In 1994, the focus shifted to four major areas for action one of which was mental health [3]. In a report on progress in all priority areas in 1996, the only indicator identified in the area of mental health was mortality from suicide [4]. While mortality and health facility data to a large extent describe the disease burden in other priority areas such as injuries, cardiovascular disease and cancer, this is not the case in mental health. The disease burden of mental disorders is largely due to non-fatal conditions and a large proportion of people with a mental disorder never presents to a health facility. Thus, routine data collection systems do not contribute much to analysis of the disease burden due to mental disorders. The National Survey of Mental Health and Wellbeing (SMHWB) of 1997, on the other hand, provides a wealth of information on the occurrence of mental disorders in Australia [5]. The Australian Burden of Disease Studies relied heavily on this source of information for the majority of mental disorders [6-9]. This paper describes the methods and results of these analyses.

METHODS

Two burden of disease studies were carried out in Australia between 1998 and 1999: a national study conducted at the Australian Institute of Health and Welfare (AIHW) and a separate study for the state of Victoria by the Department of Human Services (DHS). The two studies used identical methods and shared a considerable proportion of the workload. The Victorian team was responsible for the calculation of the mental health burden and was guided by a panel of local mental health experts. The results presented in this paper are for the State Victoria. Australian estimates of the

burden of disease can be found at the AIHW website (www.aihw.gov.au).

The burden of disease was calculated in disability-adjusted life years (DALY), a summary measure of population health developed by Murray and Lopez for the Global Burden of Disease (GBD) study [10]. The DALY combines information on mortality and non-fatal health outcomes to describe population health in a single number. The burden of disease in a population in a given year is the sum of the years of life lost due to premature deaths (YLL) and an estimate of the future years lost due to disability (YLD) for new cases of disease or injury, weighted for severity. The DALY is a health gap measure that extends the concept of potential years of life lost (PYLL) so as to also include years of 'healthy' life lost because a proportion of the population lives in states of less than full health [11]. The DALY was designed to allow estimates of health impact to be mapped to causes, whether in terms of disease and injury, or risk factors and broader social determinants, and to use common values and health standards for all regions of the world. The Australian burden of disease studies depart from the general GBD methodology in a number of key areas:

1. The inclusion of a wider range of disease and injury categories than the GBD and a more detailed age breakdown of the burden of disease.
2. The use of the 1996 Australian cohort life expectancy (a method that takes declining mortality trends into account) rather than the standard life table used by the GBD to calculate the years of life lost for a death at each age. This results in a life expectancy at birth of 85.7 years in women and 81.5 years in men compared to the GBD's standard life table with a life expectancy at birth of 82.5 years in women and 80 years in men.
3. The application of uniform age-weighting rather than the greater weight given in the GBD to loss of health in young and mid adult life to reflect a greater social responsibility at those ages to care for the young and old in society. We decided not to use age weighting because of its controversial nature (see, for instance [12-14]) and because GBD results were reported not to be sensitive to the age-weighting assumptions [10].
4. The use of disability weights developed in the Netherlands [15] in addition to those generated for the GBD. The Dutch weights were

preferred because of their detail and their focus on health states that are common in Australia.

5. Adjustments for the effects of co-morbidity between highly prevalent physical conditions, between mental disorders and between injuries.

The methods used to estimate the burden of mental disorders in Australia have been described in considerable detail in a previous paper comparing the Australian and GBD methods [16]. A summary is presented here with additional details on the mental disorders examined in the Victorian and Australian studies but not the GBD study.

Mortality

The Australian Bureau of Statistics provided data on deaths occurring within Australia to people, whose usual place of residence was Victoria. We chose 1996 as the baseline year, as this was the most recent year of mortality data at the start of the study. The number of deaths in Australia attributed to mental disorders was small with the exception of drug overdose deaths and to a lesser extent deaths due to alcohol dependence. We included 85 heroin deaths coded under the injury chapter of the International Classification of Diseases version nine (ICD–9) as accidental poisoning due to opioids because they had a similar age and sex profile to the 81 heroin deaths directly coded under the Mental Disorders chapter.

Disability

The estimation of the disability associated with mental disorders required information on the incidence, average duration and severity for each disease and its associated health states. The SMHWB was the main source of information to determine the burden in adults for six anxiety disorders, depression, most substance use disorders and borderline personality disorder. The survey was conducted on a representative sample of 10,641 Australian adults with a response rate of 78% [5]. Interviewers used a computerised version of the Composite International Diagnostic Interview (CIDI) [17]. The unit record data of the survey provided information on the prevalence of mental disorders by ICD–10 and the Diagnostic and Statistical Manual of

Mental Disorders (DSM–IV) categories [18, 19]. We chose the ICD–10 diagnoses for consistency with our other disease categories that were largely based on ICD–9 categories.

The ICD–10 criteria for post-traumatic stress disorder (PTSD) are less strict than the DSM–IV stipulates [20]. To avoid overestimation we used the DSM–IV diagnosis of PTSD, which resulted in a 60% lower prevalence of PTSD. Our initial estimates of psychotic disorders were based on prevalence data from a US study [21]. These were subsequently confirmed by the recently released prevalence figures (4.7 per 1,000 adults) of the Low Prevalence Disorders sub-study of the SMHWB [22]. We estimated a prevalence of heroin dependence of five per 1,000 persons 15–44 years old from the number of heroin users enrolled in methadone programmes adjusted by estimates of the proportion of regular heroin users reached by these programmes. This estimate was consistent with the finding of a national drug household survey after accounting for the likelihood of both underreporting of heroin use and users not being included in a household sample [23]. The prevalence of eating disorders, depression in childhood and childhood disorders such as autism and Attention Deficit and Hyperactivity Disorder (ADHD) were derived from international epidemiological studies for lack of relevant Australian data [24-30]. We classified dementia under the category of neurological and sense disorders and not under mental disorders as is the convention in ICD.

Disease models

For each of the twenty-two mental disorders included in our analyses, we aimed to model the course of the disease in the average incident case. The incidence of mental disorders is rarely measured and cross-sectional surveys tend to report one-year period prevalence. To derive incidence we made extensive use of the DisMod software package developed by Harvard University that allows the modelling of internally consistent relationships between estimates of incidence, prevalence, duration, remission and mortality [31]. The disease models for anxiety disorders and bipolar disorder were based on evidence from the literature which describes them as chronic

diseases with periods of remission and relapse [32-35]. This differs from the approach adopted in the GBD 1990 study in which short durations of between 0.75 and 2.5 years for all anxiety disorders, substance use disorders and bipolar disorder were assumed. We estimated the average duration for anxiety and bipolar disorders in the Australian studies basing our disease models on rates of disease remission reported in the literature [32, 33, 36, 37]. Lacking relevant published information, we estimated the duration of substance use disorders, borderline personality disorder and dysthymia from DISMOD models that gave the best fit between the observed age pattern of prevalence and assumptions of incidence and remission. Disease experts were consulted to check the estimates for plausibility. To account for asymptomatic periods during the course of the chronic conditions for which we derived prevalence estimates from the SMHWB, we adjusted the final YLD estimates by the ratio of two-week prevalence and one-year prevalence.

There is evidence to suggest that dysthymia—a long-term condition with recurrent periods of low mood— and major depressive disorder are different expressions of the same chronic disease entity with periods of greater, lesser or no symptoms [38-41]. However, the great diversity in the course of illness made it difficult to find a credible, 'average', chronic disease model for depression. Some people have only one episode in a lifetime; some have chronic unremitting disease and the majority experiences multiple episodes over a lifetime with no or low-grade symptoms in between. We therefore decided to model major depression as episodes and dysthymia separately but to present the results under a single category of depression.

We estimated a 4–6 year average duration for borderline personality disorder. This may seem short for a personality disorder but there was no alternative DISMOD model with credible estimates of incidence and remission to match the prevalence observed in the SMHWB. Psychoses were modelled as a life long condition that generally starts in young adulthood and has an elevated risk of dying based on standardised mortality ratios of 154 in males and 162 in females in the UK [42]. The estimated average

duration for anorexia and bulimia was 8 and 5 years, respectively, based on overseas follow-up studies reporting 24% persistence of anorexia after 10–15 years and 20% persistence of bulimia after 5–10 years [43, 44]. We assumed an annual mortality risk of 0.58 per cent in anorexia and no elevated mortality in bulimia [45]. An estimate of 50 per cent remission after five years of follow-up led to an estimate of the average duration for ADHD of almost 7 years [46]. The duration for autism was extrapolated from the average life expectancy reported for moderate intellectual disability in California [47].

Severity

For the main mental disorders we used the Dutch disability weights (DWs). The DWs for drug dependence disorders, manic episodes in bipolar disorder and borderline personality disorder were extrapolated by panels of local experts in a manner similar to the way in which the GBD expert panel derived weights for most conditions after having determined weights for 22 indicator conditions by person trade-off methods. Our weight for psychoses is a compound of 30 percent the untreated weight and 70 per cent the treated weight from the GBD study based on an average time spent in psychosis reported from a number of industrialised countries that took part in the International Pilot Study of Schizophrenia [48] (Table 1).

The Dutch DWs typically describe the mental disorders to which they refer by two or three levels of severity. However, it was not so easy to find the epidemiological data by level of severity. When that level of detail was available, as for instance was the case with three levels of severity of ICD–10 diagnoses of depression and panic disorder in the SMHWB, local mental health experts doubted its validity. Instead, for conditions derived from the SMHWB we found it necessary to develop a new classification of severity from six questions on mental well being in the SF–12. This method has been described previously in detail [16].

Table 1. Disability weights (DW) for mental disorders used in Australian burden of disease studies.

Disease category	Disability weight				Origin of DWs
		Mild	Moderate	Severe	
Substance use disorders					
Alcohol harmful use	0.11				Dutch weights
Alcohol dependence		0.11	0.33	0.55	
Heroin dependence	0.27				Locally
Marihuana dependence	0.11				derived
Sedative dependence	0.18				weights
Stimulant dependence	0.11				
Anxiety disorders					
Panic disorder		0.16	0.16	0.69	Dutch weights
Obsessive-compulsive disorder		0.17	0.17	0.60	
Post-traumatic stress disorder		0.13	0.13	0.51	
Agoraphobia		0.11	0.11	0.55	
Social phobia		0.17	0.17	0.59	
Generalised anxiety disorder		0.17	0.17	0.60	
Separation anxiety disorder		0.11	0.11	0.55	
Affective disorders					
Major depressive episodes		0.14	0.35	0.76	Dutch weights
Dysthymia		0.14	0.35	0.76	
Bipolar disorder	0.18	Combined DW: 0.50 for mania (locally derived weight), 0.34 for depressive episodes (Dutch weight for moderate depression) and 0.14 for time in between episodes (Dutch weight for mild depression)			
Psychoses	0.43	30% untreated GBD DW (0.63) and 70% treated GBD DW (0.43)			
Borderline personality disorder	0.54				Locally derived weight
Eating disorders (anorexia and bulaemia)	0.28				Dutch weight
Attention-deficit hyperactivity disorder		0.02	0.15	0.15	Dutch weight
Autism	0.55				Dutch weight

In comparison to the ICD–10 categories of mild, moderate and severe depression, our severity classification resulted in larger differentials for the average disability score (and in the expected direction) than measured by each of the other disability instruments used in the SMHWB, including the

mental component score of SF–12, the General Health Questionnaire, the Brief Disability Questionnaire and the Kessler psychological distress scale. This suggested that it was better than ICD–10 at discerning different levels of severity in depression and lent support to its use in other conditions derived from the SMHWB. Because of the large overall size of the mental health burden and the broad spectrum of severity experienced in mental disorders, we chose to use this unvalidated classification scheme and apply the graded Dutch DWs rather than the single GBD weights, which imply an undefined and potentially inappropriate distribution of severity.

Co-morbidity

Co-morbidity between mental disorders identified in the SMHWB was very common. The prevalence of people in Australia with a mental disorder was 17.8%, 35% of whom qualified for two or more diagnoses. At the level of individual diagnoses the proportion of persons with co-morbid conditions was even higher. For instance, of all people with a current diagnosis of major depression, 61% had at least one other concurrent diagnosis. Co-morbidity with anxiety disorders was common, occurring in a third of people with depression. Co-morbidity in people with borderline personality disorder was even more frequent (94%): depression, anxiety disorder and substance use disorder occurred in 62%, 48% and 52%, respectively. There were people identified with up to eight different mental disorders. Calculating the disability for each co-morbid disorders as if it belonged to a separate person could result in attributing an amount of non-fatal burden in one person in excess of a disability weight of one or the equivalent of being dead. As this is obviously not credible, we decided to split the number of prevalent cases of anxiety disorders, affective disorders, borderline personality disorder and substance dependence categories equally between concurrent diagnoses. By capturing the level of severity separately, as described above, we allowed for the fact that people with multiple diagnoses were likely to experience more severe disease than those with a single diagnosis only. An increase in the number of co-morbid diagnoses was associated with a higher proportion of people reporting moderate and severe disability (Table 2).

Table 2. Reported level of severity from SF-12 questions on mental well-being by number of concurrent mental diagnoses, National Survey of Mental Health and Wellbeing 1997.

No. of diagnoses	Reported level of severity (%)				
	none	mild	moderate	severe	Total
0	71	22	6	0	100
1	46	29	21	3	100
2	27	29	36	8	100
3 or more	14	24	46	16	100

Depression as a risk factor

The contribution of certain diseases to the total burden of disease is not well captured by the mutually exclusive disease categories used in this study. This is because in addition to the direct sequelae from these conditions, there is also an increased risk for other diseases or injuries. This 'excess' attributable burden can be estimated using similar methods to those used for estimating the burden attributable to risk factors. We made a separate more inclusive calculation of the burden attributable to depression including all depression without adjustment for co-morbidity and part of the burden of suicide, self-inflicted harm and ischaemic heart disease. Estimation of the fraction of the burden from these conditions that is attributable to depression is based on the two-week prevalence of depression observed in the SMHWB and relative risks of 30 for suicide [49] and 2.3 for ischaemic heart disease [50].

RESULTS

Mental illness was responsible for about one-seventh of the total disease burden in Victoria in 1996, or 40,776 DALYs in men and 41,451 in women (Figure 1).

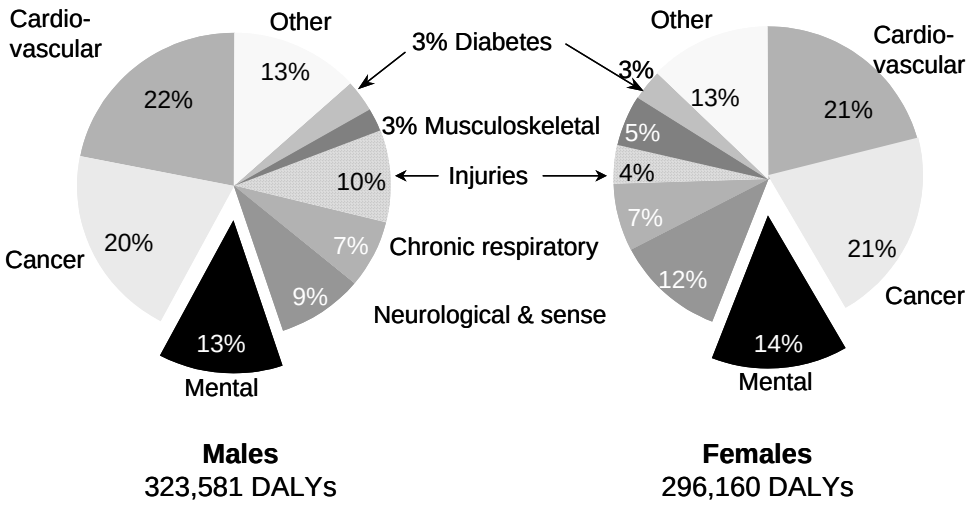

Figure 1. The burden of disease by disorder and sex, Victoria 1996.

Only 6% of this burden was due to mortality, most of which can be accounted for by fatal outcomes associated with substance use disorders. The picture is dominated by substance use, affective and anxiety disorders, which together constitute four-fifths of the overall burden attributable to mental illness (Figure 2). There were marked sex differentials in the distribution of the mental illness burden due to particular causes. The contribution from affective disorders was 50% higher in women than in men, while the anxiety burden was three-fifths higher. Conversely, the male burden from substance abuse was two and a third times higher compared with women. Eating disorders occurred mainly in women, with men having only 5% of the female burden attributable to these disorders. Childhood conditions (such as

autism and ADHD) were predominantly found in boys who had a three times higher burden from these conditions than girls.

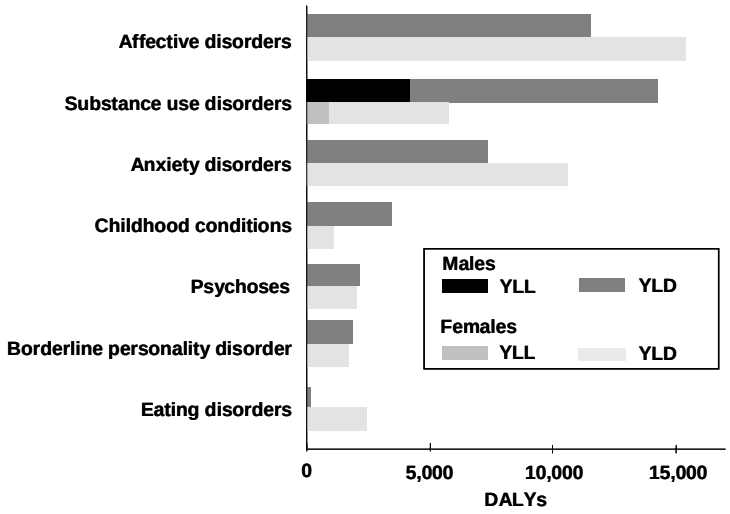

Figure 2. The burden of mental illness (YLL, YLD and DALYs) by disorder and sex, Victoria 1996

In men, depression was the eighth leading cause of overall burden, while alcohol and drug use disorders were the 13th and 15th, respectively. In women, depression was the fifth leading cause of overall burden, while generalised anxiety disorder and alcohol use disorder were 12th and 17th, respectively (Table 3).

A more inclusive calculation of depression as a risk factor for suicide and ischaemic heart disease and without adjustment for co-morbidity with other mental disorders increased the share of depression as a proportion of total disease burden making it the third largest condition in males and the top condition in females.

Table 3. Top twenty causes of burden of disease in DALYs by sex, Victoria 1996

	Males	% of total		Females	% of total
1	Ischaemic heart disease	13.1%	1	Ischaemic heart disease	10.7%
2	Stroke	4.8%	2	Stroke	6.1%
3	Lung cancer	4.7%	3	Breast cancer	5.4%
4	COPD[1]	4.4%	4	Dementia	4.9%
5	Diabetes mellitus	3.3%	5	Depression[2]	4.7%
6	Bowel cancer	2.9%	6	Osteoarthritis	3.3%
7	Depression[2]	2.9%	7	COPD	3.3%
8	Prostate cancer	2.7%	8	Diabetes mellitus	3.2%
9	Suicide	2.7%	9	Asthma	2.9%
10	Road traffic accidents	2.7%	10	Bowel cancer	2.8%
11	Dementia	2.6%	11	Lung cancer	2.7%
12	Hearing loss	2.6%	12	Generalised anxiety disorder	1.7%
13	Alcohol abuse/dependence	2.3%	13	Other vision disorders[3]	1.5%
14	Asthma	2.3%	14	Hearing loss	1.3%
15	Drug abuse/dependence	2.1%	15	Ovary cancer	1.2%
16	Osteoarthritis	1.8%	16	Parkinson's	1.2%
17	Benign prostatic hypertrophy	1.1%	17	Alcohol abuse/dependence	1.1%
18	Inflammatory heart disease	1.0%	18	Road traffic accidents	1.1%
19	HIV/AIDS	1.0%	19	Peripheral vascular disease	1.0%
20	Lymphoma	0.9%	20	Lymphoma	1.0%

[1] Chronic obstructive pulmonary disease

[2] Ignoring co-morbidity and adding an attributable fraction for suicide and ischaemic heart disease increases the share of the burden for depression in males to 5.0% (rank 3) and 11.2% in females (rank 1)

[3] Excludes cataract, glaucoma and diabetic retinopathy and mainly concerns macula degeneration

Mental disorders accounted for the largest proportion of non-fatal burden of any group of diseases or injuries, representing 25.5% of disability

in males and 27.5% in females. Depression, alcohol dependence, generalised anxiety disorder, psychotic disorders, bipolar disorder and social phobia ranked among the top 20 causes of disability in males and females. Also ranked in the top 20 were borderline personality disorder and attention-deficit hyperactivity disorder in males and eating disorders in females (Table 4).

Table 4. Top twenty causes of disability in YLD by sex, Victoria 1996

	Males	% of total		Females	% of total
1	Depression	6.5%	1	Depression	10.7%
2	Hearing loss	5.8%	2	Dementia	6.1%
3	Alcohol abuse/dependence	4.9%	3	Osteoarthritis	5.4%
4	Dementia	4.5%	4	Asthma	4.9%
5	Asthma	4.5%	5	Diabetes mellitus	4.7%
6	COPD	4.3%	6	Generalised anxiety disorder	3.3%
7	Diabetes mellitus	4.0%	7	Breast cancer	3.3%
8	Osteoarthritis	3.9%	8	Other vision disorders	3.2%
9	Ischaemic heart disease	3.7%	9	Stroke	2.9%
10	Stroke	3.5%	10	COPD	2.8%
11	Benign prostatic hypertrophy	2.4%	11	Hearing loss	2.7%
12	Prostate cancer	2.2%	12	Ischaemic heart disease	1.7%
13	Generalised anxiety disorder	2.0%	13	Alcohol abuse/dependence	1.5%
14	Borderline personality disorder	1.6%	14	Parkinson's	1.3%
15	Attention-deficit disorder	1.3%	15	Eating disorders	1.2%
16	Bipolar disorder	1.5%	16	Social phobia	1.2%
17	Psychoses	1.5%	17	Infertility	1.1%
18	Parkinson's	1.5%	18	Bipolar disorder	1.1%
19	Social phobia	1.4%	19	Psychoses	1.0%
20	Bowel cancer	1.0%	20	Rheumatoid arthritis	1.0%

The usual way of calculating DALYs for non-fatal outcomes is to multiply incident cases of disease with the average duration of disease and the severity weight. Thus calculated, the per capita incident non-fatal burden attributable to mental illness was far greater in early adulthood than at any other age (Figure 3). This is largely due to the peak in new cases of chronic mental illnesses at this life stage, the disability of which is experienced for many years into the future.

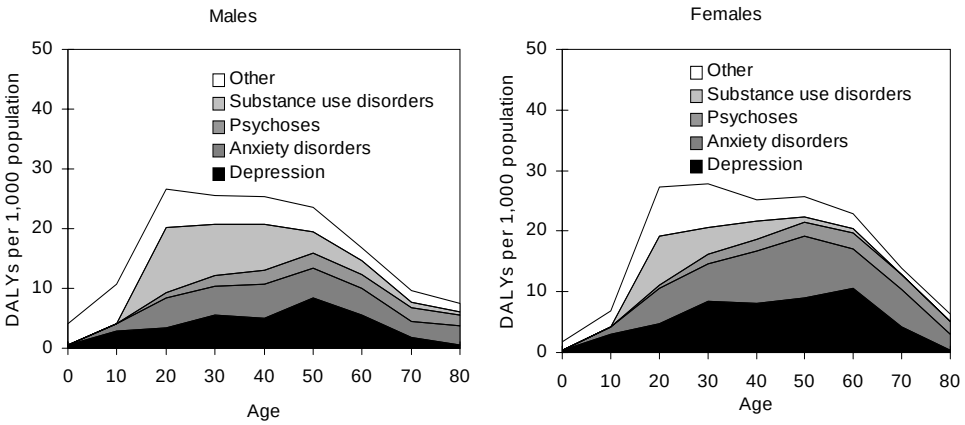

Figure 3. Incident YLD rates per 1,000 population by mental disorder, age and sex, Victoria 1996

Using an incidence approach, this disability is attributed back to age at onset. Boys experienced incident non-fatal burden from childhood disorders at twice the rate of girls. Sex differences in the age distribution of incident burden rates, however, were the most pronounced for depression, which we modelled as an episodic condition. In men, this burden steadily rises to half the male mental illness burden at middle age, after which it diminishes quickly. In women, the burden from depression increases rapidly to the highest rate for any mental illness at any age by early adulthood, after which it declines slightly to three-quarters the female mental illness burden at retirement age.

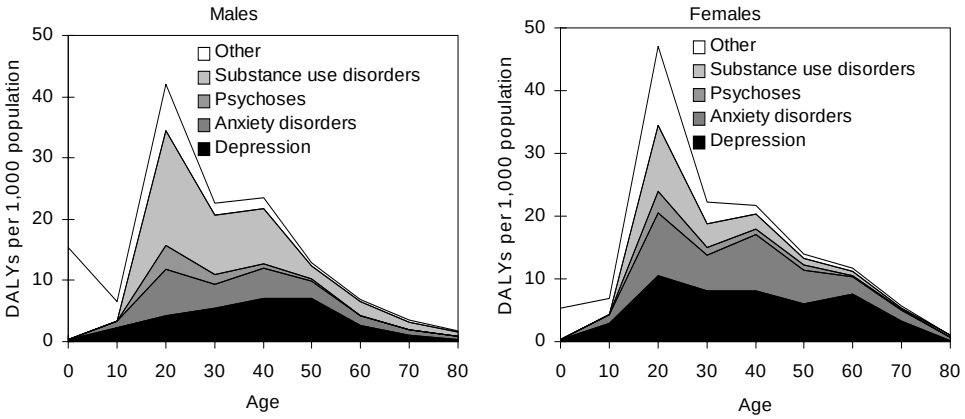

Figure 4. Prevalent YLD rates per 1,000 population by mental disorder, age and sex, Victoria 1996

Another way of presenting the burden of disease is by calculating DALYs for prevalent cases of disease during the one-year duration of the year of study again weighted for severity. This per capita prevalent non-fatal burden attributable to mental illness provides a substantially different picture to an incidence approach (Figure 4). The large burden from high incidence mental disorders of early adulthood illustrated in Figure 3 is reduced by more than a third and appears instead throughout older ages until death, this being the period during which chronic conditions (i.e. anxiety disorders and psychoses) remain prevalent. Only a proportion of substance use disorders in early adulthood leads to chronic mental illness, which explains the larger reduction in prevalent burden after this life stage for substance abuse than for anxiety disorders and psychoses. The difference in the picture for depression between an incidence and prevalence approach is not nearly as apparent because we modelled this condition as an episodic disorder. A useful way of interpreting the prevalent burden is as a measure of the proportion of the population experiencing disability at a particular point in time. Viewed in these terms, four-fifths of the burden from mental illness was experienced at ages between 15 and 64 in both men and women.

DISCUSSION

This paper presents results from the first attempt at a comprehensive assessment of the burden of mental disorders in Australia. The loss of health due to 22 mental disorders was quantified in DALYs. The DALY is a summary measure of population health that allows the quantification of mortality and morbidity in one measure and thus allows comparisons between diseases of a very different nature. Mental disorders rank third in size of burden behind cardiovascular disease and cancer. We have reported elsewhere that mental disorders would have ranked first with an increase in its share of the total burden from 13.3% to 20.7% if age-weighting were applied [16]. In other low-mortality countries with most of the burden of physical disorders occurring at old age and a high prevalence of mental disorders in young and mid adult ages, burden of disease results will be similarly sensitive to the GBD age weighting function. This is an undesirable outcome for a social value in a summary measure of population health that does not enjoy universal support.

Because mental disorders largely run a non-fatal course and are often not presented or recognised as such during contacts with health services [5], routine data collection systems are unable to capture the information that is needed to quantify the burden of mental disorders. It was fortunate that data from the SMHWB became available during the course of the national and Victorian burden of disease studies. Mental health epidemiologists have made considerable progress in setting standards on how to obtain and classify mental disorders in population surveys [51]. The CIDI performs well in terms of inter-rater and test-retest reliability [52]. The validity of mental diagnoses derived from population surveys is less well established. Most validity studies have examined a few selected aspects of validity and have been conducted on small selected clinical samples [53]. Validity studies are difficult because of the lack of a well-defined gold standard of psychiatric diagnoses and the existence of two major psychiatric classification systems. Discordance between ICD–10 and DSM–IV diagnostic categories in a study from New South Wales was 32% for all mental disorders examined [54].

Analysing the SMHWB in a similar fashion shows an even greater lack of agreement between ICD–10 and DSM–IV diagnoses (42%). However, when using survey data in a burden of disease analysis it is more important to compare the prevalence estimates and the subsequent DALY estimates of mental disorders between the two classification systems. In the SMHWB, the prevalence of ICD–10 defined mental disorders was 14% higher than that of DSM–IV diagnoses. The biggest outliers were PTSD (59% more common in ICD–10), dysthymia (55% more common in ICD–10), OCD (51% less common in ICD–10), social phobia (49% more common in ICD–10) and agoraphobia (42% more common in ICD–10). The calculations in the Australian burden of disease studies were based on ICD–10 diagnoses with the exception of PTSD for which we opted for the DSM–IV diagnosis following expert advice.

For this paper, we recalculated the DALYs for 14 mental disorders that were derived from the SMHWB using the DSM–IV diagnoses but otherwise the same methods. In comparison, the sum of DALYs for the 14 DSM–IV defined mental disorders is 7% greater. Under DSM–IV, a greater burden of major depression is partially compensated by a much smaller estimate for dysthymia. Large differences in size between individual anxiety disorders cancel each other out when added together. DSM–IV also gives higher estimates for alcohol dependence, harmful use of alcohol and cannabis dependence and lower estimates of sedative dependence (Table 5). On balance the differences between ICD–10 and DSM–IV diagnoses are not great although the large differences in estimates for social phobia, agoraphobia and OCD challenge the accuracy of estimates for individual anxiety disorders. Concern about the validity of survey results is not restricted to mental health. In the course of collecting information for the burden of disease studies we were concerned about the validity of self-reported health status measures from population surveys for common physical conditions such as diabetes, asthma and osteoarthritis. In comparison, psychiatric epidemiology can boast a much greater scientific rigour in setting standards for interview-based population surveys.

Our estimates for bipolar disorder, childhood mental disorders and

eating disorders relied on overseas data and should be considered provisional until local data become available. This was also the case for psychotic disorders although an Australian estimate of prevalence obtained prior to finalising our results confirmed our initial estimate.

Table 5. The prevalence and burden in Years Lived with Disability (YLD) for 14 mental disorders estimated from the National Survey of Mental Health and Wellbeing, calculated by ICD–10 and DSM–IV diagnostic categories.

Mental disorder	Prevalence		YLD		
	ICD-10	DSM-IV	ICD-10	DSM-IV	Difference[1]
Depression			**23,330**	**26,145**	**-12%**
Major depression	5.2%	6.3%	17,288	23,261	-35%
Dysthymia	1.1%	0.5%	6,042	2,884	52%
Anxiety disorders			**17,984**	**17,098**	**5%**
Panic disorder	1.1%	1.0%	1,497	1,441	4%
Agoraphobia	0.9%	0.5%	1,149	756	34%
Social phobia	2.7%	1.4%	4,548	3,175	30%
Generalised anxiety disorder	3.1%	2.6%	7,990	7,806	2%
Obsessive-compulsive disorder	0.4%	0.7%	1,191	2,312	-94%
Post-traumatic stress disorder	3.3%	1.3%	1,609	1,609	-26%
Substance use disorders			**12,073**	**14,236**	**-18%**
Harmful alcohol use	3.0%	1.9%	1,243	1,807	-45%
Alcohol dependence	3.5%	4.1%	8,912	9,375	-5%
Sedative dependence	0.4%	0.4%	744	419	44%
Cannabis dependence	1.6%	1.5%	1,097	1,332	-21%
Stimulant dependence	0.2%	0.2%	78	73	6%
Borderline personality disorder	0.4%		3,653	3,596	2%
Total for 14 mental disorders			**57,040**	**61,075**	**-7%**

[1] Difference expressed as % of ICD-10 YLD

Co-morbidity between mental disorders is very common. For an estimate of the burden of disease in a population, estimating each disorder independently would lead to considerable overestimation. We were able to adjust for co-morbidity between the disorders derived from the SMHWB. This was the first time that adjustment for co-morbidity has been attempted in a burden of disease study. As we included a greater number of mental disorders than previously had been included in the GBD, it became more difficult to ignore the issue. With the results in table 2 we show that people with more than one mental disorder had more severe disease. Because their greater severity is translated into higher disability weights, people with co-morbid mental disorders contribute more per individual to the burden of disease than people with a single mental disorder, even though we apportion them equally between conditions. We were unable to adjust for co-morbidity arising from mental disorders that were not derived from the SMHWB, such as the co-morbidity between psychoses and depression or between eating disorders and depression. This could have led to some degree of overestimation of the mental burden. We decided not to adjust for co-morbidity between physical and mental disorders even though the SMHWB collected self-reported data on physical conditions because of the concerns we had in general about the accuracy of self-reported health status and the difficulty of attributing self-reported disability categories (such as 'heart trouble', 'kidney disease' and 'liver trouble') to specific causes. For the majority of mental disorders, the age at which they are most prevalent is between young and mid-adulthood when physical disorders are less common. Therefore, ignoring co-morbidity between mental and physical disorders may not lead to much overestimation of the burden of disease.

Declining prevalence of depressive and anxiety disorders with increasing age is a consistent finding of population surveys of mental health. Four types of error in case ascertainment of depression have been suggested as potential explanations:

1. Exclusion of the institutionalised and homeless,
2. Failure of the elderly to respond adequately to the complexity of the interview schedule,

3. Failure to detect depression in the presence of chronic physical illness, and
4. A lesser likelihood of elderly persons acknowledging symptoms of depression [54-56].

The balance of arguments, however, leans towards accepting the validity of lower prevalence of depression and anxiety in the elderly. "Psychological immunisation", that is increasing resistance to depression after prolonged exposure to adversity over a lifetime, has been suggested as an explanation for this phenomenon [57, 58].

Depression in old age is commonly associated with chronic physical conditions such as Parkinson's, dementia and stroke [59]. The Dutch disability weight study used descriptions, which included some or considerable depression and/or anxiety at all severity levels for these three diseases [15]. Thus our DALY calculations for these diseases include a component of depression. Not all of these depressive symptoms would qualify as major depression or dysthymia, the two categories of depression estimated in this study. In fact, it is reported that survey instruments and diagnostic classification systems erroneously downplay the importance of depression associated with physical illness [59].

For the Australian burden of disease studies we made a conscious choice to concentrate on the epidemiological inputs and not to derive separate DWs for Australia. Instead, we adopted DWs developed in the Netherlands that covered the main sources of disability in Australia. Dutch DWs were available for varying levels of severity for the major mental disorders. This had the advantage over the single DWs used in the GBD study that it allowed the specification of a local distribution of severity. However, there were no readily available data on the distribution by severity of mental disorders in Australia to match the description of the health states valued in the Dutch study. For conditions derived from SMHWB we devised a classification of severity based on answers to six questions on mental health status of the SF–12. Our assumption that the same classification system was valid when applied to the other conditions measured in the SMHWB, however, is an example of the need to improvise when

undertaking a burden of disease study. More accurate information about the natural history of mental disorders including duration and patterns of severity would require longitudinal studies of people identified with a mental disorder in community samples. The fact that such a study was not planned to follow-up a sample SMHWB respondents with a mental disorder represents a lost opportunity. Follow-up studies of clinical samples can only provide partial information as they report on those patients who are in contact with health services and therefore probably have more severe disease.

CONCLUSIONS

While there is room for improvement in our estimates of the burden due to mental disorders in Australia, the results published here represent a major step forward in our understanding of the importance of mental disorders as a population health problem. As a health gap measure, DALYs describe the unfinished health agenda, that is the amount of ill health in a population that currently is not addressed by health service efforts. As such, it suggests where opportunities for further health gain may exist. Complementary analyses of the cost-effectiveness of interventions are required to identify the mix of mental health interventions that gives greatest value for money. The Department of Human Services in Victoria is planning to conduct cost-effectiveness analyses in mental health using the DALY as the measure of health outcome. The disease models developed for the burden of disease study will be a starting point to assess the scope for change and the resources required to reduce the size of the burden due to mental disorders.

REFERENCES

1. Nutbeam D, Wise M, Bauman A, Harris E, Leeder S. *Goals and targets for Australia's health in the year 2000 and beyond*. Canberra: Australian Government Printing Service, 1993.
2. Health Targets and Implementation (Health for All) Committee. *Health for all Australians: Report to the Australian Health Ministers*

Conference. Canberra: Australian Government Printing Service, 1988.

3. Commonwealth Department of Health and Family Services (DHFS). *Better health outcomes for Australians*. Canberra: Australian Government Printing Service, 1994.
4. Australian Institute of Health and Welfare (AIHW) & Commonwealth Department of Health and Family Services (DHFS). *First report on national health priority areas 1996*. Canberra: AIHW & DHFS, 1997.
5. Australian Bureau of Statistics. *Mental health and wellbeing: Profile of adults, Australia, 1997*. Canberra: ABS, 1998.
6. Mathers CD, Vos T, Stevenson CE. *The burden of disease and injury in Australia*. Canberra: Australian Institute of Health and Welfare, 1999.
7. Mathers CD, Vos ET, Stevenson CE, Begg SJ. The Australian Burden of Disease Study: measuring the loss of health from diseases, injuries and risk factors. *Medical Journal of Australia* 2000; 172:592-6.
8. Vos T, Begg S. *The Victorian Burden of Disease Study: mortality*. Melbourne: Public Health and Development Division, Department of Human Services, 1999.
9. Vos T, Begg S. *The Victorian Burden of Disease Study: morbidity*. Melbourne: Public Health and Development Division, Department of Human Services, 1999.
10. Murray CJL, Lopez AD. *The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020*. Cambridge, Mass: Harvard University Press, 1996.
11. Murray CJ, Salomon JA, Mathers C. A critical examination of summary measures of population health. *Bull World Health Organ* 2000; 78:981-94.
12. Williams A. Calculating the global burden of disease: time for a strategic reappraisal? *Health Economics* 1999; 8:1-8.
13. Barendregt JJ, Bonneux L, Van der Maas PJ. DALYs: the age-weights on balance. *Bulletin of the World Health Organization* 1996; 74:439-43.
14. Anand S, Hanson K. Disability-adjusted life years: a critical review. *Journal of Health Economics* 1997; 16:685-702.
15. Stouthard MEA, Essink-Bot ML, Bonsel GJ, Barendregt JJ, Kramer PGN, van de Water HPA, Gunning-Schepers LJ, van der Maas PJ. *Disability weights for diseases in the Netherlands*. Rotterdam: Department of Public Health, Erasmus University, 1997.

16. Vos T, Mathers CD. The burden of mental disorders: a comparison of methods between the Australian burden of disease studies and the Global Burden of Disease study. *Bulletin of the World Health Organization* 2000; 78:427-438.
17. World Health Organization. *Composite International Diagnostic Interview - version 2.1*. Geneva: WHO, 1997.
18. World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders - Diagnostic criteria for research*. Geneva: WHO, 1993.
19. American Psychiatric Association. *Diagnostic and Statistical Manual of mental disorders, 4th edn: DSM-IV*. Washington, DC: APA, 1994.
20. Peters L, Slade T, Andrews G. A comparison of ICD10 and DSM-IV criteria for posttraumatic stress disorder. *J Trauma Stress* 1999; 12:335-43.
21. Kendler KS, Gallagher TJ, Abelson JM, Kessler RC. Lifetime prevalence, demographic risk factors, and diagnostic validity of nonaffective psychosis as assessed in a US community sample. The National Comorbidity Survey. *Arch Gen Psychiatry* 1996; 53:1022-31.
22. Jablensky A, McGrath J, Herrman H, Castle D, Gureje O, Evans M, Carr V, Morgan V, Korten A, Harvey C. Psychotic disorders in urban areas: an overview of the Study on Low Prevalence Disorders. *Australian and New Zealand Journal of Psychiatry* 2000; 34:221-36.
23. Australian Institute of Health and Welfare. *1998 National drug strategy household survey: first results*. Canberra: AIHW, 1999.
24. Herder GA. [Infantile autism among children in the county of Nordland. Prevalence and etiology]. *Tidsskr Nor Laegeforen* 1993; 113:2247-9.
25. Fombonne E, Du Mazaubrun C, Cans C, Grandjean H. Autism and associated medical disorders in a French epidemiological survey. *J Am Acad Child Adolesc Psychiatry* 1997; 36:1561-9.
26. Gilchrist PN, Ben-Tovim DI, Hay PJ, Kalucy RS, Walker MK. Eating disorders revisited. I: Anorexia nervosa. *Med J Aust* 1998; 169:438-41.
27. Angold A, Costello E. The epidemiology of depression in children and adolescents. In *The depressed child and adolescent: developmental and clinical perspectives*. Cambridget: Cambridge University Press, 1995.
28. Nordin V, Gillberg C. Autism spectrum disorders in children with physical or mental disability or both. I: Clinical and epidemiological aspects. *Dev Med Child Neurol* 1996; 38:297-313.

29. Steinhausen HC, Winkler C, Meier M. Eating disorders in adolescence in a Swiss epidemiological study. *Int J Eat Disord* 1997; 22:147-51.
30. Wolraich ML, Hannah JN, Pinnock TY, Baumgaertel A, Brown J. Comparison of diagnostic criteria for attention-deficit hyperactivity disorder in a county-wide sample. *J Am Acad Child Adolesc Psychiatry* 1996; 35:319-24.
31. DisMod v1.0, Burden of Disease Unit, Harvard University, 1994.
32. Angst J, Vollrath M. The natural history of anxiety disorders. *Acta Psychiatr Scand* 1991; 84:446-52.
33. Allsopp M, Verduyn C. A follow-up of adolescents with obsessive-compulsive disorder. *Br J Psychiatry* 1989; 154:829-34.
34. Davidson JR, Hughes D, Blazer DG, George LK. Post-traumatic stress disorder in the community: an epidemiological study. *Psychol Med* 1991; 21:713-21.
35. Murphy JM. Trends in depression and anxiety: men and women. *Acta Psychiatr Scand* 1986; 73:113-27.
36. Keller MB, Hanks DL. Course and outcome in panic disorder. *Prog Neuropsychopharmacol Biol Psychiatry* 1993; 17:551-70.
37. Pollack MH, Smoller JW. The longitudinal course and outcome of panic disorder. *Psychiatr Clin North Am* 1995; 18:785-801.
38. Angst J, Wicki W. The Zurich Study. XI. Is dysthymia a separate form of depression? Results of the Zurich Cohort Study. *Eur Arch Psychiatry Clin Neurosci* 1991; 240:349-54.
39. Judd LL. The clinical course of unipolar major depressive disorders. *Arch Gen Psychiatry* 1997; 54:989-91.
40. Keller MB. Dysthymia in clinical practice:course, outcome and impact on the community. *Acta Psychiatr Scand Suppl* 1994; 383:24-34.
41. Andrews G, Sanderson K, Slade T, Issakidis C. Why does the burden of disease persist? Relating the burden of anxiety and depression to effectiveness of treatment. *Bull World Health Organ* 2000; 78:446-54.
42. Kelly C, McCreadie RG, MacEwan T, Carey S. Nithsdale schizophrenia surveys. 17. Fifteen year review. *Br J Psychiatry* 1998; 172:513-7.
43. Keel PK, Mitchell JE. Outcome in bulimia nervosa. *Am J Psychiatry* 1997; 154:313-21.
44. Strober M, Freeman R, Morrell W. The long-term course of severe anorexia nervosa in adolescents: survival analysis of recovery, relapse, and outcome predictors over 10-15 years in a prospective study. *Int J Eat Disord* 1997; 22:339-60.

45. Nielsen S, Moller-Madsen S, Isager T, Jorgensen J, Pagsberg K, Theander S. Standardized mortality in eating disorders--a quantitative summary of previously published and new evidence. *J Psychosom Res* 1998; 44:413-34.
46. Hill JC, Schoener EP. Age-dependent decline of attention deficit hyperactivity disorder. *Am J Psychiatry* 1996; 153:1143-6.
47. Strauss D, Kastner TA. Comparative mortality of people with mental retardation in institutions and the community. *Am J Ment Retard* 1996; 101:26-40.
48. Leff J, Sartorius N, Jablensky A, Korten A, Ernberg G. The International Pilot Study of Schizophrenia: five-year follow-up findings. *Psychol Med* 1992; 22:131-45.
49. Bertschy G, Vandel S. [The link between suicide and depression. Epidemiologic aspects]. *Encephale* 1991; 17:33-6.
50. Australian Institute of Health and Welfare. *Heart, stroke and vascular diseases, Australian facts*. Canberra: AIHW and the Heart Foundation of Australia, 1999.
51. Anthony JC, Eaton WW, Henderson AS. Looking to the future in psychiatric epidemiology. *Epidemiol Rev* 1995; 17:240-2.
52. Andrews G, Peters L. The psychometric properties of the Composite International Diagnostic Interview. *Soc Psychiatry Psychiatr Epidemiol* 1998; 33:80-8.
53. Wittchen HU. Reliability and validity studies of the WHO--Composite International Diagnostic Interview (CIDI): a critical review. *J Psychiatr Res* 1994; 28:57-84.
54. Andrews G, Slade T, Peters L. Classification in psychiatry: ICD-10 versus DSM-IV. *Br J Psychiatry* 1999; 174:3-5.
55. Heithoff K. Does the ECA underestimate the prevalence of late-life depression? *J Am Geriatr Soc* 1995; 43:2-6.
56. Snowdon J, Draper B, Chiu E, AMes D, Brodaty H. Surveys of mental health and wellbeing: critical comments. *Australas Psychiatry* 1998; 6:246-247.
57. Henderson AS. Does ageing protect against depression? *Soc Psychiatry Psychiatr Epidemiol* 1994; 29:107-9.
58. Henderson AS, Jorm AF, Korten AE, Jacomb P, Christensen H, Rodgers B. Symptoms of depression and anxiety during adult life: evidence for a decline in prevalence with age. *Psychol Med* 1998; 28:1321-8.
59. Snowdon J. Epidemiologic questions on mood disorders in old age. *Clin Neurosci* 1997; 4:3-7.

3

The burden of mental disorders: a comparison of methods between the Australian burden of disease studies and the Global Burden of Disease Study

Vos T, Mathers C (2000). The burden of mental disorders: a comparison of methods between the Australian burden of disease studies and the Global Burden of Disease Study. *Bulletin of the World Health Organization*, 78:427–38.

ABSTRACT

The national and Victorian burden of disease studies in Australia set out to critically examine the methods used in the Global Burden of Disease study to estimate the burden of mental disorders. The main differences include the use of a different set of disability weights allowing estimates in greater detail by level of severity, adjustments for co-morbidity between mental disorders, a greater number of mental disorders measured and modelling of substance use disorders, anxiety disorders and bipolar disorder as chronic conditions. Uniform age-weighting in the Australian studies produces considerably lower estimates of the burden due to mental disorders in comparison to age-weighted disability-adjusted life years. A lack of follow-up data of people identified in cross-sectional surveys with a mental disorder poses the greatest challenge in determining the burden of mental disorders more accurately.

CHAPTER 3

BACKGROUND

The prominent position of mental disorders, particularly depression as a cause of disease burden is a widely quoted result of the Global Burden of Disease (GBD) study [1-5]. However, there has been little discussion of the methods used to estimate the mental health burden at either the global or national level. The general debate about the use of the disability-adjusted life year (DALY) as a summary measure of population health has largely concentrated on the underlying assumptions of age-weighting and discounting and the relevance of burden of disease measurements to policy-making [6-11].

One report has challenged the severity weights for mental disorders used in the GBD study. Findings from a small community sample in New South Wales, Australia, led to a cautious conclusion that the GBD study may have overestimated the disability weights for depression and substance disorders, while underestimating the level of disability associated with anxiety disorders [12]. However, the epidemiological assumptions that fed into the GBD study's calculation of the burden of mental disorders have not been scrutinised in the peer-reviewed literature. Only three of the intended ten volumes in the *Global Burden of Disease and Injury Series* have been published to date [13-15]; and a description of methods used in calculating the burden of mental disorders has not yet appeared. The only epidemiological information is in vol. 2 of the series, detailing for each disease and world region the age-specific and sex-specific values of incidence, prevalence, average duration and mortality.

As part of the two recent burden of disease studies in Australia, an effort was made to examine critically the GBD estimates for mental disorders, to improve the methods and to apply them to the most appropriate information on the epidemiology of mental disorders in the country. The results of the national Australian study conducted by the Australian Institute of Health and Welfare and of an analysis of the burden of disease in Victoria carried out by the state's Department of Human Services are available as

printed reports and on the Internet [16-19]. The two project teams worked closely together and shared methods and analyses.

The methods used to estimate the burden of mental disorders in Australia are discussed below, and departures from those of the GBD study are identified and justified. The consequences are described and discussed of the methodological changes on the estimates for the state of Victoria and the results are compared with those of the GBD estimates for the Established Market Economy (EME) region. Although burden was estimated for dementia and other neurological conditions in the Australian studies, these conditions were included, together with Alzheimer's disease, in a separate category for nervous system disorders and they are not discussed here.

METHODS

The disability-adjusted life year

Summary measures of population health combine information on mortality and non-fatal health outcomes to describe population health in a single number. The DALY was designed to provide a common measure for fatal and non-fatal health outcomes, to allow estimates of health impact to be mapped to causes, and to enable common values and health standards to be applied to all regions of the world [14].

DALYs for a disease are the sum of the years of life lost because of premature mortality in the population and the years lost because of disability for incident cases of the health condition in question. The DALY is a *health gap* measure that extends the concept of potential years of life lost because of premature death to include equivalent years of healthy life lost in states of less than full health, broadly termed disability [20].

The Australian burden of disease studies depart from the general methodology used in the GBD study in the following key areas:

- The GBD study used a standard life table (West level 26) with a life expectancy at birth of 82.5 years for women and 80 years for men, whereas the Australian studies use the Australian cohort life expectancy (taking declining mortality trends into account) for 1996.

This resulted in a life expectancy at birth of 85.7 years for women and 81.5 years for men.

- In the GBD study, DALYs were discounted at 3% and age-weighted. In the Australian studies, DALYs were discounted but were not age-weighted. Age-weighting is intended to capture a greater social responsibility in young and mid adult life for the very young and old. Age-weighting was not used in the Australian studies because it is perceived as inequitable by some people and becasue the GBD sensitivity analyses showed that it did not essentially change the overall estimates of burden [14].
- In addition to DWs developed for the GBD stduy, the Australian studies used those developed by Dutch researchers [21, 22] for many conditions because of their greater detail and their focus on the most common disabilities found in countries of low-mortality.
- The GBD did not account for the occurrence of co-morbid health states whereas the Australian studies make adjustments for the effects of co-morbidity between highly prevalent physical conditions, between mental disorders and between injuries
- The Australian studies included a wider range of disease and injury categories than the GBD study and provided a more detailed age breakdown of the burden of disease.

Data sources

Apart from deaths associated with drug overdoses and, to a lesser extent, deaths due to alcohol dependence, the number of deaths in Australia attributed to mental disorders is small. We included as heroin deaths a substantial number of deaths coded under accidental poisoning due to opioids in the International Classification of Diseases, ninth revision (ICD–9).

The estimation of the disability associated with mental disorders requires information on the incidence, average duration and severity for each disease and its sequelae. The incidence of mental disorders is rarely measured; surveys tend to report one-year prevalence (the number of people

who experienced relevant symptoms at any time during the past 12 months). To derive the incidence we made extensive use of the DisMod software package developed by Harvard University to examine the consistency between estimates of incidence, prevalence, duration and mortality [23].

Table 1. Data sources used in Australian burden of disease studies to estimate the incidence and/or prevalence of mental disorders.

Data source	Mental disorder
National Mental Health and Wellbeing Survey, 1997 [24]	Anxiety disorders (panic disorder, agoraphobia, social phobia, generalised anxiety disorder, obsessive-compulsive disorder) Depression (major depressive episode and dysthymia) Schizophrenia [28] Most substance abuse (alcohol, cannabis, sedative and stimulant drug dependence or harmful use) Borderline personality disorder
National Drug Strategy Household Survey 1998 [29]	Heroin and residual 'other drugs' category
Reviews of international epidemiological studies	Schizophrenia [30] Bipolar disorder [27] Eating disorders (anorexia and bulimia) [31, 32] Childhood disorders (attention-deficit hyperactivity disorder, autism and Asperger's syndrome) [33-35] Depression and separation anxiety disorder in childhood [36, 37]

The data sources are summarised in Table 1. The main source of prevalence data for adults was the National Mental Health and Wellbeing Survey (MHS) of 1997 [24], in which information was collected on symptoms experienced in the last 12 months, 1 month and 2 weeks for a representative sample of 10,560 adults. A computerised version of the Composite International Diagnostic Interview was used in this work. Interviews were completed for 78% of the individuals approached. The unit record data of the survey contained information on the prevalence of mental disorders by ICD-10 and DSM-IV (Diagnostic and Statistical Manual of

Mental Disorders) categories as well as a number of measures of disability, namely the abbreviated Short Form (SF-12), the General Health Questionnaire (GHQ), the Brief Disability Questionnaire (BDQ) and the Kessler psychological distress scale. We used the ICD-10 diagnoses for consistency with our other disease categories which were largely based on ICD-9 categories. The only exception was post-traumatic stress disorder (PTSD) for which we used the DSM-IV diagnosis, because the ICD-10 criteria were too broad and would have resulted in overestimation due to misclassification of other anxiety disorders [25]. One of the modules of the Composite International Diagnostic Interview on mania was omitted from the survey, and this meant that estimates of bipolar disorder were inaccurate. Instead, we relied on estimates from international epidemiological studies [26].

Initially, we built our estimates of schizophrenia on overseas data as well [27]. The results of the Low-Prevalence Disorders sub-study of the MHS showed a prevalence of psychotic disorders, which was similar to our initial estimates [28]. The exclusion of institutionalised patients from the MHS sample is only important for these low-prevalence conditions, for which we relied on other data sources.

We based our estimates of heroin dependence and harmful use (5 per 1,000 15–44 year olds) on the numbers of heroin users enrolled in methadone programs together with expert estimates of the proportion of dependent users reached by these programmes. We checked the resulting estimates for consistency with prevalence data from the National Drug Strategy Household Survey 1998, making allowance for underreporting and for heroin users who would not be captured in a household sample [29].

Severity

The Dutch disability weights study [21, 22] provided weights for three levels of severity for most of the mental disorders. Because of the large size of the mental health burden we felt it desirable to use Australian information on severity distributions together with these weights to obtain more accurate estimates of burden. In using a single overall DW for the spectrum of

severity of a disease, as in the GBD study, a severity distribution is assumed and judgement of that implicit distribution becomes part of the DW estimation task.

The use of the Dutch weights required matching of Australian epidemiological data to the severity levels defined in the Dutch study. Where level of severity was available, as with the three levels of severity of ICD-10 diagnoses of depression and panic disorder in the National Mental Health and Wellbeing Survey (MHS), local mental health experts doubted whether the diagnostic categories were representative of the actual level of severity of disability.

We used information from the SF-12 for each respondent in order to classify the conditions in the MHS in the mild, moderate and severe categories defined in the Dutch DW study, which DWs using EuroQol 5D+ (EQ5D+) descriptions for each of the health states valued. The EQ5D+ is an extended version of the EuroQol health measurement instrument with an additional domain of cognitive functioning. The EQ5D+ descriptions of the six anxiety disorders in adults distinguish mild/moderate from severe mostly in the third domain of usual activities and the fifth domain of anxiety/depression with a score of 2 for mild/moderate and 3 for severe. Similarly, harmful alcohol use (score of 2) differs from alcohol dependence (score of 3) in the usual activities domain. The difference between mild, moderate and severe depression involves further domains but it has the same split between mild/moderate and severe on the usual activities and anxiety/depression domains.

We thus mapped responses to three questions (B6, B7 and B12 relating to usual activities) in the SF-12 to the three levels of the EuroQol usual activities domain. Similarly, we mapped responses to three other questions (B9, B10 an B11 relating to anxiety, depression and energy level) in the SF-12 to the three levels of the EuroQol anxiety/depression domain. This enabled us to categorise MHS respondents in accordance with the EuroQol levels for usual activities and anxiety/depression and hence with the severity categories for the DWs.

In order to validate this method of determining severity, we examined how it compared with the way ICD-10 classifies depression into mild, moderate and severe levels. This involved comparing the scores of all other instruments used in the MHS each of which, in its limited way, describe severity of disability, as follows: the mental and physical component score of SF-12 (MCS and PCS, with the latter not expected to be much affected by mental disorders); the overall score and Likert scores of the GHQ; reported days out of role from the BDQ; the main BDQ score; the Mental Outcome Study score of BDQ; the WHO score of BDQ; and the Kessler psychological distress scale. As these instruments refer to the preceding four weeks, we based our comparisons of scores on those reporting symptoms of depression in this period only. As shown in Table 2, the classification system based on SF-12 resulted in larger differentials for the average disability score (in the expected direction) as measured by each of the other disability instruments. This suggested that our classification system was better at discerning different levels of severity in depression and supports our use of this classification system across the conditions for which we derived estimates from the MHS.

Co-morbidity

We derived prevalence estimates from the MHS for six anxiety disorders, six substance use disorders, borderline personality disorder, major depressive episodes and dysthymia. Co-morbidity between these mental disorders was very common; the prevalence in Australia of people with one of them being 17.8%, 35% of whom had two or more diagnoses. At the level of individual diagnoses, the proportion of persons with co-morbid conditions was even higher. For example, of the people with a current diagnosis of major depression, 61% had at least one other concurrent diagnosis. Co-morbidity with anxiety disorders was common, occurring in a third of people with depression. Co-morbidity with borderline personality disorder was even more frequent (94%). Depression, anxiety disorder and substance use disorder, occur in 62%, 48% and 52%, respectively, of people with borderline personality disorder.

Table 2. Comparison of ICD-10 and SF-12 based classification of severity of depression with scores from multiple instruments.

	ICD-10 classification			SF-12 based classification		
Instrument	**mild**	**moderate**	**severe**	**mild**	**moderate**	**severe**
SF-12 MCS	46.0	34.1	29.8	46.4	30.9	21.6
SF-12 PCS	47.9	46.5	44.2	50.3	44.7	42.1
GHQ score	3.0	5.2	5.5	2.9	5.2	7.0
GHQ Likert score	13.5	17.4	18.1	13.6	17.4	20.9
BDQ - days out of role	4.8	4.8	9.1	1.6	6.8	12.8
BDQ main score	1.7	1.8	2.2	1.4	2.0	2.4
BDQ MOS score	4.2	4.5	6.1	2.9	5.7	6.8
BDQ WHO score	5.9	6.3	9.1	4.0	8.3	10.2
Kessler psychological distress scale	40.1	35.3	30.4	39.0	33.9	27.3

Counting each of these disorders as a separate episode could result in attributing disability in one person in excess of a disability weight of 1 — the equivalent of being dead! In order to avoid overestimation of burden, we shared co-morbidity between anxiety disorders, affective disorders, borderline personality disorder and alcohol and drug dependence equally so that a person with 2 or more disorders was partially counted in each category. On the advice of mental health experts, 75% of co-morbidity with harmful alcohol use was attributed to the relevant other mental disorder and 25% was attributed to harmful alcohol use, whereas co-morbidity between anxiety and affective disorders was attributed equally between categories.

Because we captured the level of severity separately, as described above, our estimates took into account that people with multiple diagnoses were likely to have greater severity than people with a single diagnosis.

Duration

It is argued that dysthymia and major depressive disorder are part of the same disease entity as people with dysthymia frequently develop superimposed major depressive episodes and the symptomatic course of major depression commonly changes between levels of severity [38-40]. Modelling major depression and dysthymia as one disease proves very difficult because of the heterogeneity of the course of the disease: some people have only one episode, some people are continuously depressed for a long time and the majority have multiple episodes that vary in frequency. We therefore decided to model major depression as episodes and to model dysthymia separately. We added the years lost due to disability (YLD) of major depression and dysthymia to give total YLD for depression.

We derived estimates of the average duration of major depressive episodes from the one-year prevalence (symptomatic at some time during the preceding 12 months) and point prevalence (symptomatic in the preceding two weeks) as follows:

$$1\text{-year prevalence} = \text{point prevalence} \times \frac{52 \text{ weeks} + \text{average duration}}{\text{average duration}}$$

This gave an average durations of 38.2 weeks and 24.9 weeks for males and females, respectively. The literature suggests that the mean duration of of both initial and recurrent episodes of major depression is consistent at 20 weeks [26, 41]. Our higher estimate of duration may be attributed to recall bias of those with symptoms of major depression in the last year who did not have current symptoms. Having no evidence to support or reject this hypothesis we used the estimates of duration calculated above.

The overall YLD estimates are not very sensitive to the assumed duration as a shorter estimate of duration leads to a higher estimate of incidence derived from the same observed prevalence, and the total YLD values are proportional to incidence multiplied by average duration. In fact, a 20-week duration for depressive episodes resulted in almost identical estimates of YLD although there was a moderate shift in the age distribution from younger adults to older age groups. We used a remission rate of 0.124

in our DISMOD model for dysthymia based on a reported remission in 58% during seven years of follow-up [38]. From young adults to older ages the modelled average duration ranges from 8.0 to 3.8 years.

The literature on anxiety disorders [42, 43], obsessive-compulsive disorder [44] and PTSD [45, 46] suggests that these disorders generally run a chronic course with periods of remission and relapse. This contrasts with the relatively short duration estimated in the GBD study of 0.75, 1.6 and 2.5 years, for panic disorder, OCD and PTSD, respectively. To account for asymptomatic periods during the course of these chronic conditions we adjusted our durations by the ratio of point prevalence and one-year prevalence.

In our DisMod models for substance use disorders, we assumed age-specific remission rates and incidences that were consistent with the observed large age differences in prevalence. This leads to duration estimates of 2–5 years for harmful alcohol use, 3–7 years for alcohol dependence, 5–17 years for heroin dependence, and 3–4 years for marijuana, sedative and stimulant dependence. The DisMod models for alcohol dependence and harmful use included a twofold increased risk of death resulting in a number of deaths amounting to about 60% of all deaths attributed to alcohol in our risk factor calculations. We accepted that as plausible, since not all deaths attributed to alcohol occur in people who qualify for a diagnosis of substance use disorder. For harmful alcohol use, marijuana, sedative and stimulant dependence we used the same method of adjustment to account for asymptomatic periods as described above for anxiety disorders. This same method cannot be applied to alcohol dependence as the definition assumes a chronic unremitting state. However, local alcohol experts did not believe that the large numbers of people identified with alcohol dependence, particularly at younger ages, all experienced disability for the duration of the condition. We therefore adjusted our YLD calculations, assuming that the 47–60% per cent in males and 16–53% in females, depending on age, who reported higher scores on the SF–12 than the population average represented an equivalent to the asymptomatic periods estimated for anxiety disorders.

Table 3. Comparison of disability weights (DW) for mental disorders used in the Australian burden of disease studies and in the GBD

Disease category	GBD DW [(a)]	Australian DW [(b)]	Comments on Australian DWs
Alcohol dependence	0.18	0.07-0.18	Dutch weights (0.11-0.55)
Drug use	0.25		
Heroin dependence		0.27	Locally derived weights
Marijuana dependence		0.11	
Sedative dependence		0.18	
Stimulant dependence		0.11	
Anxiety disorders			Range of Dutch weights[(c)]
Panic disorder	0.15	0.21–0.27	0.11–0.69
Obsessive-compulsive disorder	0.12	0.20–0.28	0.17–0.60
Post-traumatic stress disorder	0.11	0.14–0.15	0.13–0.51
Agoraphobia		0.14–0.16	0.11–0.55
Social phobia		0.18–0.21	0.17–0.59
Generalised anxiety disorder		0.22–0.23	0.17–0.60
Affective disorders			Range of Dutch weights
Major depressive episodes	0.50	0.37–0.41	0.14–0.76
Dysthymia		0.33–0.38	
Bipolar disorder	0.51	0.18	Composite of local weight for mania (0.50), moderate Dutch weight for depressive episodes (0.34) and mild Dutch depression weight in between episodes (0.14)
Schizophrenia	0.41	0.43	Composite of 30% untreated and 70% treated GBD DW
Borderline personality disorder		0.54	Locally derived weight
Eating disorders		0.28	Dutch weight
ADHD		0.02–0.15	Dutch weights
Autism		0.55	Dutch weight

[(a)] Calculated from treated and untreated DW together with proportion assumed treated [14].
[(b)] Range reflects variations by age and sex in the severity distribution.
[(c)] For further details, refer to Dutch disability weights study [21, 22].

We estimated durations of 4–6 years for borderline personality disorder. This may seem short for a personality disorder but there is no

alternative credible set of incidence and remission estimates that can be matched to the observed MHS prevalence figures using DisMod.

Schizophrenia was modelled as a life long condition that generally starts in young adulthood and has an elevated risk of dying based on a meta-analysis reporting standardised mortality ratios of 154 in males and 162 in females [47].

Disability weights

For most disorders we used the Dutch DWs (Table 3). The DWs for drug dependence disorders, manic episodes in bipolar disorder, and borderline personality disorder were extrapolated by panels of local experts in the same way as the GBD study's expert panel derived weights for most conditions after anchoring weights for 22 indicator conditions. The weight for schizophrenia is a compound of 30% for the untreated weight and 70% for the treated weight from the GBD study, reflecting the average time spent in psychosis reported from a number of industrialised countries in the International Pilot Study of Schizophrenia [30].

Comparisons with GBD estimates

In order to make meaningful comparisons with the results of the GBD study for the EME region, we recalculated our results with age-weighting and applied the same standard model life table for years of life lost (YLL). All comparisons are presented as rates of DALYs per 1,000 population age-standardised to the 1990 EME population.

RESULTS

According to Australian reports based on non-age-weighted DALYs, mental illness contributed 13.2% of the total disease burden in Victoria and 13.3% of that in the country as a whole in 1996 [16, 19]. Only 6% of the mental illness burden was attributable to mortality, mostly involving fatal outcomes associated with substance use disorders. Affective, anxiety and substance use disorders together accounted for four-fifths of the overall burden attributable

to mental illness (Figure 1). For men, depression was the eighth leading cause of overall burden, while alcohol and drug use disorders were the thirteenth and fifteenth, respectively. For women, depression was the fifth leading cause of overall burden, while generalised anxiety disorder and alcohol use disorder were the twelfth and seventeenth, respectively.

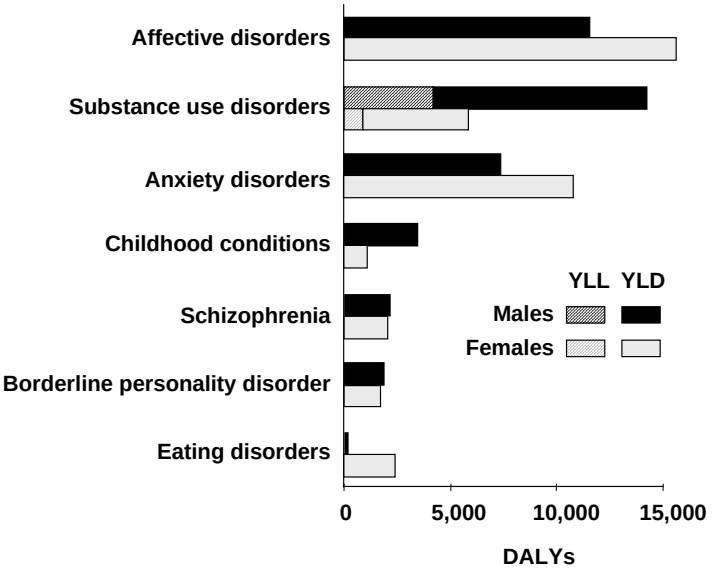

Figure 1. The burden of mental illness (YLL, YLD and DALYs) by disorder and sex, Victoria 1996

Adding the age-weighting of the GBD study to the results for Victoria gave considerably more prominence to mental disorders. They thus became the largest group of conditions, contributing 20.7% to the burden of disease. This was similar to the estimate of the GBD for EMEs, in which mental disorders other than dementia accounted for 19.5% of the total burden of disease. With age-weighting, depression became the leading cause of DALYs for women and the second cause for men in Victoria. Alcohol dependence in men and women and drug dependence in men entered the top ten while generalised anxiety disorder, social phobia and bipolar disorder ranked among the top twenty conditions for both men and women; eating disorders and drug dependence also entered the top twenty for women.

Table 4. DALY rates per 1,000 population[(a)] for mental disorders, EME region 1990 and Victoria 1996.

	Males		Females	
	GBD	**Victoria**	**GBD**	**Victoria**
All mental disorders	25.9	20.1	21.4	21.2
Affective disorders	8.2	6.7	12.7	9.0
Depression	6.0	5.3	10.7	7.7
Bipolar disorder	2.2	1.4	2.1	1.3
Anxiety disorders	2.4	4.3	3.6	6.0
Post-traumatic stress disorder	0.3	0.5	0.4	0.5
Obsessive-compulsive disorder	1.6	0.4	2.0	0.3
Panic disorder	0.6	0.2	1.1	0.6
Other anxiety disorders	0.0	3.2	0.0	4.6
Substance abuse	12.4	6.4	2.6	3.1
Alcohol dependence/harmful use	9.7	4.3	1.7	1.9
Drug dependence/harmful use	2.7	2.1	0.9	1.2
Schizophrenia	2.9	1.2	2.6	1.0
Other mental disorders	0.0	1.5	0.0	2.1

[(a)] Age-weighted and discounted DALYs, age-standardised to EME population 1990.

Notwithstanding the close agreement on the total size of the mental health burden, there were substantial differences between the Victorian estimates and those of the GBD study for EME countries in respect of particular disorders (Table 4). These differences are the result of variations in estimates for one or more of the YLD parameters (i.e. incidence, duration and disability weights) for most mental illnesses (Table 5).

Table 5. Estimates of prevalence, incidence and duration of mental disorders in GBD 1990 and Victoria (Vic) 1996

Disorder	Prevalence per 1,000 [a]		Incidence per 1,000 [a]		Duration in years[b]	
	GBD	Vic	GBD	Vic	GBD	Vic
Males						
Major depression	0.9	1.0	16.3	10.9	0.56	0.73
Bipolar disorder	0.3	0.7	2.7	0.3	1.4–1.5	21–24
Panic disorder	0.3	0.3	3.8	0.2	0.75	10–16
OCD	1	0.1	6.2	0.2	1.6	4–7
PTSD	0.2	0.4	0.8	0.9	2.5	5
Schizophrenia	0.9	0.3	0.2	0.1	52–54	21–57
Alcohol dependence	3.9	2.7	24.0	12.9	1.5–1.7	4–5
Drug use[c]	0.8	2.0	7.7	4.3	1	3–13
Females						
Major depression	1.7	1.5	29.8	27.3	0.56	0.48
Bipolar disorder	0.3	0.7	2.6	0.3	1.4–1.5	21–24
Panic disorder	0.6	1.1	7.5	0.7	0.75	10–16
OCD	1.3	0.2	8.2	0.3	1.6	4–7
PTSD	0.3	0.6	1.3	1.1	2.5	5
Schizophrenia	0.9	0.3	0.2	0.1	52–54	21–57
Alcohol dependence	0.7	0.9	4.3	5.0	1.5–1.7	4–5
Drug use[b]	0.2	0.8	2.5	1.7	1	3–13

[a] Prevalence and incidence figures include an adjustment for comorbidity with other mental disorders

[b] Ranges indicate variations in duration by age and sex

[c] The GBD estimated prevalence of 'dysfunctional and harmful drug use' as one category

Two main factors explain most of the differences. Firstly, the occurrence of the disease, i.e. its prevalence, may have differed between Australia and the EMEs. It should be noted that almost all estimates were derived from prevalence data. The prevalence estimates for Victoria (Table

5) included the downward adjustment for co-morbidity between mental disorders. Secondly, the disease models may have differed, particularly in respect of the assumptions on the distribution of severity, DWs and average duration.

Our estimates of the burden of alcohol dependence and harmful use in men are less than half those for the EMEs. To some extent this reflected a lower prevalence in Victoria (partly attributable to by the adjustment for co-morbidity) but the main influence was the downward adjustment of the average level of disability, on the assumption that the proportion of survey respondents not reporting disability on six key questions of the SF–12 reflected symptom-free periods during the course of illness.

Despite taking a similar prevalence of major depression as a starting point and including dysthymia, we estimated a lower burden from depression in Victoria. This was largely because of the use of lower DWs. Our YLD estimates for bipolar disorder were based on a prevalence more than double that estimated in the GBD study. Nevertheless, the Victorian estimate of the associated burden was much lower because we use different DWs. The GBD study used a uniform weight of 0.51 for the EME region while in Victoria we model manic episodes at a weight of 0.50, depressive episodes at 0.34 and the time in between episodes at 0.14, resulting in an average weight of 0.18 for the overall course of illness. The very different estimates of the duration of bipolar disorder between the two studies cannot explain the large differences in burden estimates because the incidence and duration are modelled on prevalence data. Combinations of high incidence and short duration or low incidence and long duration that are consistent with the same prevalence figures lead to marginal differences in burden estimates as YLD are determined by the product of incidence and duration.

Australian figures indicate higher prevalence of PTSD in both sexes and of panic disorder in women but much lower prevalence of OCD. The latter was probably overestimated in the GBD, which may have relied on survey results in which lay interviewers used too broad criteria to diagnose this disorder, as occurred in a Canadian study [48]. In contrast to the short durations for anxiety disorders in the GBD study, we modelled them as

chronic conditions with periods of remission and relapses. This leads to very different YLD estimates, which are also affected by the higher DWs. Our estimate of the burden due to PTSD was higher, mostly because of the higher prevalence. Our much lower estimate of YLD for OCD was mostly influenced by the sixfold lower prevalence estimate, the higher DW estimates and the MHS-based estimates indicating that men and women with OCD have 11% and 41%, respectively, of the time symptom-free during the course of their illness. Higher average DWs and an estimated 74% symptom-free periods during the course of illness, and a higher observed prevalence, resulted in YLD estimates for panic disorder that were considerably lower in men and women in comparison with the estimates in the GBD study. With the inclusion of three additional anxiety disorders in adults (social phobia, generalised anxiety disorder and agoraphobia) and separation anxiety disorder in childhood, the estimates in the Victoria study for anxiety disorders are two-thirds higher than in the GBD study.

The YLD estimates for schizophrenia in Victoria were 60% lower than in the GBD study and reflected lower estimates of prevalence as other assumptions of DWs and average duration were almost identical.

DISCUSSION

The epidemiological assumptions underlying the calculations of the burden of disease caused by mental disorders have not previously been examined. The Australian burden of disease studies developed new models for each of the mental disorders that were included in the GBD study and made estimates for ten new disease categories. Major differences in the underlying assumptions are outlined below.

The use of Dutch DWs in the Australian studies allowed the construction of disease models with details by level of severity. In principle it is a great improvement if estimates can be made explicitly by level of severity, as each of the mental disorders shows considerable variation in severity. The use of a single DW for a disease implicitly assumes an average distribution of severity, but does not give the flexibility to adapt the weight

to populations with different severity distributions. The big difficulty in our approach lay in matching epidemiological data to the severity categories for which DWs have been derived.

We decided not to derive disability weights separately for Australia in order to concentrate on the epidemiological inputs to the burden of disease calculations. This was made easier by the availability of the Dutch weights covering the main sources of disability in Australia. However, this left us with the task of finding Australian data on mental disorders to fit the Dutch severity levels as defined by the six domains of the EQ5D+. Because MHS did not use the EuroQol, we mapped six questions of the SF-12 to two domains of the EuroQol in order to distinguish between levels of severity. This was not a tested method but in comparison to ICD-10 diagnoses it showed greater distinction between levels of mild, moderate and severe depression on all other measures of disability in the survey. We then assumed that the same method could distinguish between levels of severity for the other disorders included in the same survey. The very large contribution of mental disorders to the overall burden of disease means that there is a need for longitudinal studies of people identified with a mental disorder in community samples. This would allow a better understanding to be obtained of the level of disability suffered during the course of illness.

Our revised disease models for anxiety disorders and bipolar disorder as chronic conditions were more in line with the literature than the durations estimated in the GBD study. Follow-up studies of people with anxiety disorders justified our choice of disease models of long duration with periods of remission and relapses [43, 44, 49, 50]. As we had only cross-sectional data on mental disorders in Australia we used the ratio of one-year prevalence and point prevalence to represent the proportion of time during the chronic course of illness which was spent with symptoms. Because the one-year prevalence figures may be influenced by recall bias, it would be helpful to examine whether our estimates of the proportion of time symptomatic can be confirmed in longitudinal follow-up studies.

Similarly, there is uncertainty about the amount of disability associated with alcohol dependence and harmful use. On the advice of local

alcohol experts, we reduced YLD estimates for alcohol dependence by assuming no disability for the proportion of respondents not reporting disability on the SF-12. That was an arbitrary decision made in relation to alcohol dependence but not in relation to the other mental disorders. Again this identifies a critical need for follow-up data on people identified in community surveys of mental disorders.

Ignoring the common co-morbidity between mental disorders can lead to significant double counting and overestimation of the burden of disease. This is particularly important if a large number of mental disorders is included. We were able to assess the occurrence of co-morbidity between five substance use disorders, six anxiety disorders, two affective disorders and borderline personality disorder from the results of a representative Australian mental health survey. The proportion of co-morbidity ranged from 49% in people identified with social phobia to 94% in people with borderline personality disorder. The latter figure raised the question as to whether borderline personality disorder should be excluded from the burden of disease list of mutually exclusive conditions and, instead, valued separately as a risk factor for other mental disorders. Our method of adjusting for co-morbidity assumed that each diagnosis contributes equally to the overall level of disability. By capturing severity separately from the SF-12 we are able to allow for the fact that people with more than one diagnosis were likely to experience greater disability than people with only one of the co-morbid conditions. It is possible that in this method some individuals with two or more diagnoses contribute less YLD than a person with the severest of these conditions. A laborious procedure would be required in order to analyse this on a case-by-case basis for each possible combination of co-morbid conditions. We decided against this because we did not think it would make a great difference to our overall results.

Contrary to the sensitivity analyses for the GBD study, which showed only a marginal difference between results that were age-weighted and those that were not, in Victoria age-weighting increased the contribution of mental disorders to overall DALYs by 57%. Thus mental disorders surpassed cardiovascular diseases and cancer as a leading cause of burden. Although

the authors of the GBD study mention that age-weighting preferentially gave more weight to mental disorders, their overall conclusion that results were insensitive to assumptions such as age-weighting did not hold for the Victorian study. In the debate about age-weighting in summary population health measures it is important to note that its impact may not be as slight as is commonly assumed.

CONCLUSIONS

Using different detailed methods and data sources, the Australian burden of disease studies have confirmed that mental disorders are a leading cause of burden of disease in developed countries. Partly as a result of methodological differences (including use of different weights, analysis of severity distributions and adjustments for co-morbidity) the Australian estimates show marked differences in comparison to the results of the GBD study for particular disorders. Lower severity weighting for alcohol dependence in younger men together with differences in prevalence resulted in lower burden in Australia. Lower estimates of burden for major depression were predominantly due to lower estimated severity. Despite using higher DWs, our estimates for OCD are less than a quarter of the estimates of the GBD study because of much lower prevalence estimates. Much lower estimates of the burden for panic disorder and higher estimates for PTSD were derived from disease models with higher DWs, higher prevalence estimates and adjustments for symptom-free periods during the chronic course of illness. Estimates of the total burden due to anxiety disorders in Australia were almost twice as high as the GBD results because four additional conditions were included in the Australian work.

In conclusion, we hope that our efforts to improve disease modelling for mental disorders make a useful contribution to burden of disease methodology. We believe that our models are an improvement over those used in the GBD study. There is still much scope for further developments, particularly in the measurement of severity over the course of mental disorders and in dealing with co-morbidity. Full details of our calculations

are available from the Internet (http://www.health.vic.gov.au/healthstatus/bod/daly.htm).

ACKNOWLEDGEMENTS

We thank Derrick Bui for his extensive and systematic review of the literature, which laid the foundation for the calculation of the burden attributable to mental disorders in Australia. We thank Nicole Watson for assistance with the analysis of data from the National Mental Health and Wellbeing Survey. We are grateful for generous and valuable support given by Helen Herrman, Carol Harvey, Oye Gureje, Mike Salzberg, Bruce Singh and Fiona Judd in discussing the plausibility of our disease models for mental disorders.

REFERENCES

1. Neugebauer R. Mind matters: the importance of mental disorders in public health's 21st century mission. *Am J Public Health* 1999; 89:1309-11.
2. Ustun TB. The global burden of mental disorders. *Am J Public Health* 1999; 89:1315-8.
3. Jenkins R. Reducing the burden of mental illness. *Lancet* 1997; 349:1340.
4. Appleby L, Shaw J, Amos T, Dennehy J. Global burden of disease. *Lancet* 1997; 350:143.
5. Eisenberg L. Global burden of disease. *Lancet* 1997; 350:143.
6. Barendregt JJ, Bonneux L, Van der Maas PJ. DALYs: the age-weights on balance. *Bulletin of the World Health Organization* 1996; 74:439-43.
7. Anand S, Hanson K. Disability-adjusted life years: a critical review. *Journal of Health Economics* 1997; 16:685-702.
8. Sayers BM, Fliedner TM. The critique of DALYs: a counter-reply. *Bull World Health Organ* 1997; 75:383-4.
9. Barker C, Green A. Opening the debate on DALYs (disability-adjusted life years). *Health Policy Plan* 1996; 11:179-83.
10. Ugalde A, Jackson JT. The World Bank and international health policy: a critical review. *J Int Dev* 1995; 7:525-41.

11. Williams A. Calculating the global burden of disease: time for a strategic reappraisal? *Health Economics* 1999; 8:1-8.
12. Andrews G, Sanderson K, Beard J. Burden of disease. Methods of calculating disability from mental disorder. *Br J Psychiatry* 1998; 173:123-31.
13. Murray CJL, Lopez AD. *Global Health Statistics: a compendium of incidence, prevalence and mortality estimates for over 200 conditions*. Cambridge, Mass: Harvard University Press, 1996.
14. Murray CJL, Lopez AD. *The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020*. Cambridge, Mass: Harvard University Press, 1996.
15. Murray CJL, Lopez AD. *Health Dimensions of Sex and Reproduction: the global burden of sexually transmitted diseases, HIV, maternal conditions, perinatal disorders and congenital anomalies*. Cambridge, Mass: Harvard University Press, 1996.
16. Mathers CD, Vos T, Stevenson CE. *The burden of disease and injury in Australia*. Canberra: Australian Institute of Health and Welfare, 1999.
17. Mathers C, Vos T, Stevenson CE. *The burden of disease and injury in Australia - summary report*. Canberra: Australian Institute of Health and Welfare, 1999.
18. Vos T, Begg S. *The Victorian Burden of Disease Study: mortality*. Melbourne: Public Health and Development Division, Department of Human Services, 1999.
19. Vos T, Begg S. *The Victorian Burden of Disease Study: morbidity*. Melbourne: Public Health and Development Division, Department of Human Services, 1999.
20. Murray CJ, Salomon JA, Mathers C. A critical examination of summary measures of population health. *Bull World Health Organ* 2000; 78:981-94.
21. Stouthard MEA, Essink-Bot ML, Bonsel GJ. Methodology. Disability weights for diseases: a modified protocol and results for a Western European region. *European Journal of Public Health* 2000; 10:24-30.
22. Stouthard MEA, Essink-Bot ML, Bonsel GJ, Barendregt JJ, Kramer PGN, van de Water HPA, Gunning-Schepers LJ, van der Maas PJ. *Disability weights for diseases in the Netherlands*. Rotterdam: Department of Public Health, Erasmus University, 1997.
23. DisMod v1.0, Burden of Disease Unit, Harvard University, 1994.
24. Australian Bureau of Statistics. *Mental health and wellbeing: Profile of adults, Australia, 1997*. Canberra: ABS, 1998.

25. Peters L, Slade T, Andrews G. A comparison of ICD10 and DSM-IV criteria for posttraumatic stress disorder. *J Trauma Stress* 1999; 12:335-43.
26. Angst J, Preisig M. Course of a clinical cohort of unipolar, bipolar and schizoaffective patients. Results of a prospective study from 1959 to 1985. *Schweiz Arch Neurol Psychiatr* 1995; 146:5-16.
27. Kendler KS, Gallagher TJ, Abelson JM, Kessler RC. Lifetime prevalence, demographic risk factors, and diagnostic validity of nonaffective psychosis as assessed in a US community sample. The National Comorbidity Survey. *Arch Gen Psychiatry* 1996; 53:1022-31.
28. Jablensky A, McGrath J, Herrman H, Castle D, Gureje O, Evans M, Carr V, Morgan V, Korten A, Harvey C. Psychotic disorders in urban areas: an overview of the Study on Low Prevalence Disorders. *Australian and New Zealand Journal of Psychiatry* 2000; 34:221-36.
29. Australian Bureau of Statistics. *1998 National drug strategy household survey: first results*. Canberra: Australian Bureau of Statistics, 1999.
30. Leff J, Sartorius N, Jablensky A, Korten A, Ernberg G. The International Pilot Study of Schizophrenia: five-year follow-up findings. *Psychol Med* 1992; 22:131-45.
31. Gilchrist PN, Ben-Tovim DI, Hay PJ, Kalucy RS, Walker MK. Eating disorders revisited. I: Anorexia nervosa. *Med J Aust* 1998; 169:438-41.
32. Steinhausen HC, Winkler C, Meier M. Eating disorders in adolescence in a Swiss epidemiological study. *Int J Eat Disord* 1997; 22:147-51.
33. Fombonne E, Du Mazaubrun C, Cans C, Grandjean H. Autism and associated medical disorders in a French epidemiological survey. *J Am Acad Child Adolesc Psychiatry* 1997; 36:1561-9.
34. Wolraich ML, Hannah JN, Pinnock TY, Baumgaertel A, Brown J. Comparison of diagnostic criteria for attention-deficit hyperactivity disorder in a county-wide sample. *J Am Acad Child Adolesc Psychiatry* 1996; 35:319-24.
35. Herder GA. [Infantile autism among children in the county of Nordland. Prevalence and etiology]. *Tidsskr Nor Laegeforen* 1993; 113:2247-9.
36. Bowen RC, Offord DR, Boyle MH. The prevalence of overanxious disorder and separation anxiety disorder: results from the Ontario Child Health Study. *J Am Acad Child Adolesc Psychiatry* 1990; 29:753-8.
37. Angold A, Costello E. The epidemiology of depression in children and adolescents. In *The depressed child and adolescent: developmental and clinical perspectives*. Cambridget: Cambridge University Press, 1995.

38. Keller MB. Dysthymia in clinical practice:course, outcome and impact on the community. *Acta Psychiatr Scand Suppl* 1994; 383:24-34.
39. Judd LL. The clinical course of unipolar major depressive disorders. *Arch Gen Psychiatry* 1997; 54:989-91.
40. Angst J, Wicki W. The Zurich Study. XI. Is dysthymia a separate form of depression? Results of the Zurich Cohort Study. *Eur Arch Psychiatry Clin Neurosci* 1991; 240:349-54.
41. Solomon DA, Keller MB, Leon AC, Mueller TI, Shea MT, Warshaw M, Maser JD, Coryell W, Endicott J. Recovery from major depression. A 10-year prospective follow-up across multiple episodes. *Arch Gen Psychiatry* 1997; 54:1001-6.
42. Murphy JM. Trends in depression and anxiety: men and women. *Acta Psychiatr Scand* 1986; 73:113-27.
43. Angst J, Vollrath M. The natural history of anxiety disorders. *Acta Psychiatr Scand* 1991; 84:446-52.
44. Allsopp M, Verduyn C. A follow-up of adolescents with obsessive-compulsive disorder. *Br J Psychiatry* 1989; 154:829-34.
45. Davidson JR, Hughes D, Blazer DG, George LK. Post-traumatic stress disorder in the community: an epidemiological study. *Psychol Med* 1991; 21:713-21.
46. Kessler RC, Sonnega A, Bromet E, Hughes M, Nelson CB. Posttraumatic stress disorder in the National Comorbidity Survey. *Arch Gen Psychiatry* 1995; 52:1048-60.
47. Kelly C, McCreadie RG, MacEwan T, Carey S. Nithsdale schizophrenia surveys. 17. Fifteen year review. *Br J Psychiatry* 1998; 172:513-7.
48. Stein MB, Forde DR, Anderson G, Walker JR. Obsessive-compulsive disorder in the community: an epidemiologic survey with clinical reappraisal. *Am J Psychiatry* 1997; 154:1120-6.
49. Keller MB, Hanks DL. Course and outcome in panic disorder. *Prog Neuropsychopharmacol Biol Psychiatry* 1993; 17:551-70.
50. Pollack MH, Smoller JW. The longitudinal course and outcome of panic disorder. *Psychiatr Clin North Am* 1995; 18:785-801.

THEME B

The cost-effectiveness of treating major depression in Australia

4

Assessing Cost-Effectiveness (ACE) – Mental Health: Introduction to the Study and Methods

Haby M, Carter R, Mihalopoulos C, Magnus A, Andrews G, Vos T (2004). Assessing Cost-Effectiveness (ACE) – Mental Health: Introduction to the Study and Methods. Australian and New Zealand Journal of Psychiatry, 38:569–578.

ABSTRACT

Objective: The Assessing Cost-Effectiveness (ACE)-Mental Health Study aims to assess from a health sector perspective, whether there are options for change that could improve the effectiveness and efficiency of Australia's current mental health services by directing available resources towards 'best practice' cost-effective services.

Method: The use of standardised evaluation methods addressed the reservations expressed by many economists about the simplistic use of League Tables based on economic studies confounded by differences in methods, context and setting. The cost-effectiveness ratio for each intervention is calculated using economic and epidemiological data.This includes systematic reviews and randomised controlled trials for efficacy, the Australian Surveys of Mental Health and Wellbeing for current practice and a combination of trials and longitudinal studies for adherence. The cost-effectiveness ratios are presented as cost (A$) per disability adjusted life year (DALY) saved with a 95% uncertainty interval based on Monte Carlo simulation modelling. An assessment of interventions on 'second filter' criteria ('equity', 'strength of evidence', 'feasibility' and 'acceptability to stakeholders') allows broader concepts of 'benefit' to be taken into account, as well as factors that might influence policy judgements in addition to cost-effectiveness ratios.

Conclusions: The main limitation of the study is in the translation of the effect size from trials into a change in the DALY disability weight, which required the use of newly developed methods. While comparisons within disorders are valid, comparisons across disorders should be made with caution. A series of articles is planned to present the results.

BACKGROUND

Scarcity of funds dedicated to health care means that choices of what to fund are inevitable. Currently choices are driven largely by historical patterns of expenditure and the influence of professional, industrial and community interest groups. Cost-effectiveness analyses provide additional information that can help decision-makers set funding priorities that will improve the effectiveness and efficiency of health services [1].

In the national and Victorian Burden of Disease studies, mental disorders ranked third behind cardiovascular disease and cancer, accounting for 13% of the overall burden of disease [2, 3]. This raises the question as to why such a large burden persists while there is subsidised access to potentially effective treatments for most disorders [4, 5]. This question prompted the Mental Health Branches of the Australian Department of Health and Ageing and the Department of Human Services Victoria to fund the Assessing Cost-Effectiveness – Mental Health (ACE–MH) study. The research question for the ACE–MH study is to assess from a health sector perspective, whether there are options for change that could improve the effectiveness and efficiency of Australia's current mental health services by directing available resources towards 'best practice' cost-effective services.

OVERVIEW OF METHODS

The ACE–MH study uses a macro economic evaluation approach to aid decision-making in mental health care in Australia [6]. This approach aims to give policy makers and health experts in Australia a greater involvement in both the study design and conclusions, as recommended by the Panel on Cost Effectiveness in Health and Medicine [1]. Recently the Centre for Health Program Evaluation and the Department of Human Services, in a collaborative effort, trialled a similar approach in the analysis of cancer control priorities in Australia [7]. Concurrently, a further study, ACE–Heart Disease, is being conducted using the same methods.

The key characteristics of the ACE–MH approach are: a clear rationale for selection of options for change; an evidence-based approach; adoption of a two-stage approach to the assessment of benefit, involving both health gain (i.e. cost per Disability Adjusted Life Year [DALY] saved) and "judgement" aspects which are included as second stage filters ('equity', 'strength of evidence', 'feasibility' and 'acceptability to stakeholders'); use of an economic protocol specifically developed for the study, which ensures transparency and consistency of methods; and extensive uncertainty testing.

The DALY was chosen as the measure of health gain for this study because it captures both morbidity and mortality effects across a wide range of disorders and intervention types and because baseline information on health status was available using the DALY for Australia [2, 3].

ACE–MH was guided by a Steering Committee of mental health experts, policy-makers and representatives of community mental health organisations, working to the following terms of reference:

- To select up to 30 major interventions for mental disorders, based on: the size of the problem addressed; importance in terms of current expenditure; relevance to current policy decision-making; availability of evidence to support analyses; indications that additional expenditure would lead to significant health gain or conversely, that decreased expenditure would lead to little or no reduction in health outcomes; and ability to specify the intervention in clear concrete terms.
- To define 'benefit' and the associated criteria by which the interventions will be judged.
- To critically examine the evidence and analyses presented by the researchers.
- To formulate conclusions based on the presented evidence.

The interventions chosen for analysis are shown in Table 1. The comparator to the interventions selected as options for change in the ACE studies is 'current practice'. To determine 'current practice' we utilise the Australian National Surveys of Mental Health and Wellbeing: Adult

Component [8], the Child and Adolescent Component [9], the Low Prevalence Component [10] and expert advice. Current practice includes a mixture of no treatment, treatment with evidence-based medicine (EBM) and treatment with non-EBM. For most interventions in the ACE–MH study we model the effect of switching those currently on non-EBM onto an evidence-based intervention. An intervention was classed as evidence-based if it was currently recommended in clinical practice guidelines and supported by randomised controlled trial evidence. Current contact with EBM was defined from self-reported receipt of interventions.

The perspective in ACE–MH is that of the health sector. This includes the government as health service provider, as well as the impact on patients and their families/carers. Discounting at 3% is applied to both costs and benefits to match the rate chosen in the Australian BOD studies [2, 3].

The target audience for the interventions is Australians with prevalent disease in the year 2000. The time horizon, both for the provision of the interventions and for tracking the associated costs and consequences, is specified in the individual reports. Due to the heterogeneity in the illness/intervention combinations, it is not possible to have a uniform time horizon. Time horizons were chosen to realistically reflect how the interventions would be applied in real life; and to ensure tracking of all relevant costs and benefits. Regardless of the time horizon, it is assumed that all interventions are in 'steady-state', that is they work at their full potential and that trained practitioners are available. Thus, the analyses do not address implementation and 'learning curve' issues or costs.

In addition, scoping reports were conducted for interventions for anorexia nervosa, dual diagnosis (drug dependence and psychoses), and assertive community treatment of psychoses, which showed that there was insufficient evidence of effectiveness for a full cost-effectiveness analysis.

Table 1. Interventions selected for analysis in ACE–Mental Health

Major depression in adults
Individual cognitive behavioural therapy (CBT)
Group Cognitive Behavioural Therapy (CBT)
Selective serotonin reuptake inhibitors (SSRIs)
Tricyclic antidepressants (TCAs)
Bibliotherapy
Major depression in children and adolescents
Individual Cognitive Behavioural Therapy (CBT)
Selective serotonin reuptake inhibitors (SSRIs)
Schizophrenia
Stand-alone early psychosis intervention vs generic service
Enhanced standard care for early psychosis vs generic service
Behavioural Family Management (BFM)
Behavioural Intervention in Families (BIF)
Multiple Family Groups (MFG)
Supported employment (Individual placement and support)
Risperidone vs typicals
Olanzapine vs typicals
Olanzapine vs risperidone
Clozapine vs typicals
Risperidone vs low dose typicals
Generalised Anxiety Disorder
Individual Cognitive Behavioural Therapy (CBT)
Venlafaxine (and other antidepressants)
Panic disorder / agoraphobia
Individual Cognitive Behavioural Therapy (CBT)
Selective serotonin reuptake inhibitors (SSRIs)
Tricyclic antidepressants (TCAs)
ADHD
Dexamphetamine
Methylphenidate

ASSESSMENT OF BENEFIT

Benefits are calculated by a two-stage process. The first stage involves the estimation of the health gain that could be attributed to each intervention using the DALY. The second stage involves the assessment of issues that either influence the degree of confidence that can be placed in the cost-effectiveness ratios (such as the level of available evidence), or broader issues that need to be taken into account in decision-making about resource allocation (such as equity and acceptability to stakeholders).

Stage One: Measurement of the health gain

In comparison to other health areas such as cancer and heart disease, the emphasis in mental health services is largely on improvement in quality of life of those with the mental disorder, as the current repertoire of interventions provides limited opportunities for primary prevention and improved survival. This puts a heavy emphasis on the ability to translate measures of impact described in the literature into a change in the DALY measure, and in particular, the Years Lived with Disability (YLD) component. YLD are determined by the incidence, duration and severity of disease. Severity is measured by the DALY disability weights on a scale from 0 to 1, with zero being full health and one being death. In the ACE–MH study we use the Dutch disability weights (DWs) [11], which are also used in the Australian burden of disease studies [2, 3]. Most interventions included in ACE–MH impact on disease severity, with prevention or delay of relapse also being measured in some interventions (e.g. maintenance treatment of depression and family interventions for schizophrenia).

The first step in measuring the health gain is to determine the efficacy of an intervention, i.e. the impact of the intervention on severity, duration or risk of relapse. When the intervention affects the severity of the mental disorder we use the effect size as our main measure of efficacy. The second step is to transfer the measure(s) of effect into a change in the DALY DW. These two steps are described below.

Effect size

Effect sizes are used to measure the benefit of an intervention in relation to a placebo or other control group. In the psychiatric literature outcomes are most commonly expressed as continuous measures but many different scales are used. Thus, the most relevant measure of effect size for mental health interventions is the standardised mean difference. It quantifies the magnitude of the difference between two groups in a metric-free unit, by expressing the mean difference in standard deviation units. We use 'Hedges' g [12] because it includes an adjustment to correct for small sample bias and is used in Cochrane Collaboration systematic reviews.

If there is a systematic review or meta-analysis of the intervention the effect size can often be obtained from this. However, for many of the interventions we have conducted our own meta-analysis because systematic reviews are not available, do not include all relevant studies or present the results in a form that is different to that required for ACE–MH. Wherever possible we use randomised controlled trials to calculate the effect size, as this is the best methodology for determining treatment efficacy [13].

We pool effect sizes using the random effects method of DerSimonian and Laird [12]. Firstly, an effect size is calculated for each study by averaging across selected outcome measures within the study. This differs from Cochrane systematic reviews but is consistent with meta-analyses of the psychiatric literature [14, 15] and allows inclusion of all relevant outcome measures that reflect health related quality of life in calculating the effect size. In practice, much of the efficacy data presented for mental health interventions is limited to symptom specific measures which can affect various aspects of health related quality of life, to varying extents. Following meta-analysis, any heterogeneity between trials is investigated according to established methods [16].

A 'placebo effect' is not included in our modelling of effectiveness in ACE–MH due to controversy regarding its existence [17, 18]. Further, its inclusion has little or no impact on the relative effectiveness or cost-effectiveness of treatments because patients in the non-evidence based

treatment group or the non-adherent group also get a placebo effect (but no treatment effect).

Transferring the effect size into a change in the disability weight

Translating the effect size into a change in the DALY DW is the most difficult methodological issue in ACE–MH, as there is no well-established and accepted method available. From our exploration of available methods we settled on the 'conversion factor method' and the 'survey severity method'. The results of both methods are used as a range to determine the total Years Lived with Disability, costs and cost-effectiveness ratios for the interventions. Uncertainty from each of the methods is included in the results.

The 'conversion factor method' uses a DW 'conversion factor' to translate an effect size into a change in the DW [19]. We multiply the effect size by the DW conversion factor for the particular mental disorder. This conversion factor is an average change in DALY DWs for the equivalent of a standard deviation change in severity for the particular mental disorder.

An assumption of this method is that the degree of change in effect size units in clinical trials reflects the degree of change in disability weights from application of the translation factor. While the relationship between disability weight and effect size change was defined for symptoms and the overall level of disability, the effect size from meta-analyses predominantly summarises changes in symptoms. While, there is a close correspondence between symptoms and disability in anxiety and depression [20], and in schizophrenia greater severity elicits less favourable preference values [21], it is not known if the magnitude of the change is comparable.

For the 'survey severity method' a health status measure from the mental health survey is used to classify the severity of respondents to calculate an average DW. The effect size is then applied directly to this health status measure, severity of respondents is reclassified and a new average DW calculated. The difference in average DWs is the anticipated change due to treatment.

Table 2. The survey severity method: specific steps to calculate a change in the disability weight due to an intervention

1. We classify cases in the NSMHW into 'severe' (>2.5 SD below mean; i.e. <25), 'moderate' (>1.5 to 2.5 SD below mean; i.e. 25-34.9), 'mild' (>0.5 to 1.5 SD below mean; i.e. 35-44.9) and 'no disability' (a score of ≥ 45). This is done in the group eligible for the intervention, i.e. those having the diagnosis, and consulting, but currently not receiving evidence based treatment.
2. The proportion of cases in each severity category is multiplied by the appropriate DW for the category to get a weighted average DW for the eligible group.
3. We then multiply the effect size from trial data by the standard deviation of the MCS/PSS in the group with the disorder to determine the average increase in the MCS/PSS with the intervention.
4. We apply this increase in the MCS/PSS to those respondents in the survey eligible for the intervention.
5. These respondents are reclassified into the categories 'none', 'mild', 'moderate' and 'severe' and a new weighted average DW is calculated for the group.
6. The change in average DW before and after applying the effect size is then the DW change due to the treatment.
7. Steps 3 to 6 are repeated for both the lower and upper limits of the 95% confidence interval of the effect size. This gives upper and lower limits around the DW change due to treatment. These are used for uncertainty analyses.

DW = disability weight; MCS = mental component score; NSMHW = National Survey of Mental Health and Wellbeing; PSS = Psychosocial Summary Score

For anxiety and depression we use the Mental Component Score of the SF-12 in the adult component of the National Survey of Mental Health and Wellbeing [8] and for ADHD we use the Psychosocial Summary Score from the Child Health Questionnaire in the child and adolescent component of the Survey [9]. Both measures are designed to have a mean population value of 50 and a SD of 10. The specific steps are illustrated in Table 2. A similar approach is used for schizophrenia but more assumptions are required (see Appendix).

An assumption of the 'survey severity method' is that the effect size calculated from clinical trials can be applied directly to a general health status measure such as the Mental Component Score or Psychosocial Summary Score. This may be a significant issue when the effect size from

meta-analyses is calculated from a predominance of symptom measures, as these are usually more sensitive to outcome than generic measures.

Non-adherence

The non-adherence rate is important to the cost-effectiveness ratio because patients who do not adhere to treatment would be expected to incur costs at no or reduced health benefit. For ACE–MH, we use both the dropout rate from trials and from longitudinal studies to determine the adherence rate. We use the uncertainty analysis to incorporate possible differences between trial results and 'real-life' by using a uniform distribution between the observations from trials and longitudinal studies. In the absence of longitudinal studies we assumed a lower value of 50%.

Table 3. The Second Stage Filter Criteria

Strength of evidence: We used the NHMRC 'Designation of levels of evidence' but also highlight potential threats to the quality of the evidence by identifying possible biases and issues that affect the generalisability of the evidence. Based on the level and quality of evidence we categorised the strength of the evidence of effectiveness (or ineffectiveness) as: 'sufficient', 'limited' or 'inconclusive' (Table 4). This criterion is also addressed in the uncertainty modelling. For example, if the evidence on effectiveness is weak, large uncertainty is put around the size of the impact measure.
Equity: This criterion addressed the capacity of the intervention to affect inequity in the distribution of the mental disorder, as well as highlighting issues that may affect access to, or utilisation of, the intervention. The special needs groups considered include those with a lower socio-economic status, non English speaking background, Aboriginal and Torres Strait Islanders, or rural/remote residence.
Feasibility: This criterion addressed issues such as the availability of an appropriate workforce with appropriate training to conduct the intervention, whether the intervention could be implemented under current institutional arrangements, and ease of implementation.
Acceptability to stakeholders: This criterion referred to the anticipated acceptability of the proposed interventions to the various stakeholders affected by the intervention. Stakeholders include consumers and their family/carers, clinicians, policy makers, the general community and third-party funders.

Table 4. Classifying the Strength of the Evidence [6]

Strength category	Strength of the evidence
'Sufficient evidence' of effectiveness (or ineffectiveness): Effectiveness (or ineffectiveness) is demonstrated by sufficient evidence from well-designed research.	• The effect is unlikely to be due to chance (e.g. P is < 0.05) and • The effect is unlikely to be due to bias (e.g. evidence from: ▪ a level I study design; ▪ several good quality level II studies; or ▪ several high quality level III-1 or III-2 studies from which effects of bias and confounding can be reasonably excluded on the basis of the design and analysis)
'Limited evidence' of effectiveness (or ineffectiveness): Effectiveness (or ineffectiveness) is demonstrated by limited evidence from studies of varying quality	• The effect is probably not due to chance (e.g. P is < 0.05) but • Bias, while not certainly an explanation for the effect, cannot be excluded as a possible explanation (e.g., evidence from: ▪ one level II study of uncertain or indifferent quality; ▪ evidence from one level III-1 or III-2 study of high quality; ▪ evidence from several level III-1 or III-2 studies of insufficiently high quality to rule out bias as a possible explanation; or ▪ evidence from a sizeable number of level III-3 studies which are of good quality and consistent in suggesting an effect).
'Inconclusive evidence of effectiveness (or ineffectiveness)': Inadequate evidence due to insufficient or inadequate quality research.	• No position could be reached on the presence or absence of an effect of the intervention (eg. no evidence from level I or level II studies and level III studies are available, but they are few and of poor quality, or only level IV studies are available.)

Stage Two: the Second Filter criteria

While evidence on cost-effectiveness is the main focus of activity in ACE–MH, recognition is also given to broader aspects where decisions rest heavily on judgement and notions of 'due process'. These additional criteria function as a 'second filter' by which each of the interventions is judged before recommending allocation of more or less resources. One of the roles of the Steering Committee was to select and apply the second filter criteria.. The filters chosen are described in Table 3. The main outcome of the second

stage analysis was a table for each intervention in which the issues were flagged and a qualitative judgement made explicit about each of the criterion and its impact.

ASSESSMENT OF COSTS

A common convention [22] in costing is to describe the analysis in three steps:

- Identification – which activities/services and which cost impacts are included in the analysis?
- Measurement – what is the extent of resource usage associated with these activities/services? and
- Valuation – what is the monetary value of this resource usage?

For 'identification' the health sector perspective means that costs (and cost offsets) impacting on government, together with costs (and cost offsets) impacting on patients and their families/carers are included. Costs to government are essentially the resources involved in organising and operating the services. Costs to patients and their families/carers are primarily identified as out-of-pocket expenses associated with visits to health professionals and associated treatments. The steering committee also expressed an interest in time costs (i.e. travelling time; waiting time; treatment time). These costs are calculated, where available and presented separately. With the exception of supported employment for people with schizophrenia, we do not include production losses/gains associated with ill health itself (i.e. premature death; absence due to morbidity; reduced productivity whilst at work), nor do we include non-health care goods and services. Pathway analysis is used to identify component activities of the various interventions and their current practice comparator, based on the published literature and supplemented with expert advice.

The 'measurement' of resource usage is also facilitated by the pathway analysis. Component activities are identified and quantities estimated. For example, a contact for CBT therapy usually entails a standard

60-minute consultation with a psychologist. The number and type of medical visits, drugs etc are estimated during this measurement step, with references documented.

When attempting to cost non-adherence to treatment, we could find no information on the likely subsequent health seeking behaviour and associated costs. Thus, we assume that the non-adherers have the same health seeking behaviour, and the same costs, as those currently not receiving evidence based treatment. When estimating the cost of non-evidence based treatment, we have utilised the National Surveys of Mental Health and Wellbeing to determine the average number of visits made to different health professionals in the previous 12 months (in the case of the adult component) or 6 months (in the case of the child and adolescent component). We have not included the cost of non-evidence based drugs or natural therapies in the cost of non-evidence based treatment. The conservative approach to the estimation of non-evidence based care will produce conservative estimates of the economic merit of options for change.

In the 'valuation' step, a unit price for each of the activities, together with the data source, is specified. Costs to the government and to the patient (including family/carers) are reported separately. Costs are measured in real prices for the reference year (2000). Where necessary, the AIHW health sector deflators [23] are used to adjust prices to the reference year. Unit costs for doctor visits (GP, paediatrician and psychiatrists) and drugs are obtained from the Australian Department of Health and Ageing. For doctor visits costs to government (i.e. average benefit paid) and to patient (i.e. average difference between fee charged and benefit paid) are calculated from Medicare services that were processed in the 1999/2000 financial year. Both patient billed and bulk billed services are included but services provided in hospital are excluded, due to unreliability of the fee charged data**.** For drugs, costs are for scripts filled under the Pharmaceutical Benefits Scheme and the Repatriation Pharmaceutical Benefits Scheme for the 1999/2000 financial year. The cost to the patient is averaged over 'general' and 'concession' patients and 'safety net' and 'non safety net' patients, in proportion to the number of prescriptions for each item (i.e. brand, form and quantity of the

drug) within a class. Cost per script is converted to an average cost per day for each quantity and form using the recommended dose per day. When costing a class of drugs we weight the cost by current usage of the individual brands, rather than using the cheapest brand in the class.

INCREMENTAL COST-EFFECTIVENESS

The cost-effectiveness ratio is determined as the incremental cost of the intervention divided by the incremental benefit and presented as cost (A$) per DALY saved. The incremental cost is defined as the cost of the intervention minus the cost of current practice. Likewise, the incremental benefit is the benefit of the intervention minus the benefit associated with current practice.

UNCERTAINTY ANALYSIS

Simulation-modelling techniques (with Monte Carlo sampling) using @Risk software [24] are used to allow the presentation of an uncertainty range around the benefits, costs, and cost-effectiveness ratios. This approach is recommended by the Canadian Coordinating Office for Health Technology Assessment [25] and also mentioned as one of a number of methods of exploring uncertainty in the 1996 US Consensus Panel on Cost-Effectiveness in Health and Medicine [1].

The probability distributions around the input variables are based on standard errors quoted in, or calculated from, the literature; the range of parameter values quoted in, or calculated from, the literature; and from expert advice on the likely scenarios under Australian conditions. In addition to the uncertainty range, the @RISK analysis can show the input parameters with the greatest influence on the final results and hence is an indication of research priorities if greater accuracy of results is desired

STRENGTHS AND LIMITATIONS OF THE METHODS

Strengths of ACE–MH include the use of a common economic protocol to ensure comparability of the results; extensive uncertainty testing; interpretation of cost-effectiveness ratios within a broader decision-making framework that includes consideration of second filter criteria; and use of Australian data for demography, health system costs and offsets, disease incidence/prevalence, risk factors, and disease burden.

The main limitation in ACE–MH methods is the measurement of the health benefit due to the intervention. The measurement of the effect size is based on accepted methods and uses the best study design for measuring efficacy, that is systematic reviews of randomised controlled trials [13] (Table 4). However adjustment needs to be made to transfer trial results into real-life settings (effectiveness) where results may vary due to the motivation of clinicians and patients, the availability of skilled clinicians and the capacity to vary the intervention to suit the needs of the patient, among other things. We attempt to take these factors into account using the adherence rate. However, this may not explain all of the variation likely to be seen in real-life practise, nor does it fully address the issue of generalisability of trial results, which are often conducted in a very selected patient group.

Translation of the effect size into a change in the DALY disability weight is another key issue. This required the use of new methods – the 'conversion factor method' and 'survey severity method' – each of which has its own strengths and limitations. The advantage of the 'conversion factor method' is that it allows measurement of health gain in a comparable manner for most mental disorders. The disadvantages of the method are that: the health state descriptions of severity states for psychosis were less formally derived than those for depression and anxiety; there is no DW conversion factor for attention deficit hyperactivity disorder (ADHD); and it cannot be used when outcomes from the trials are primarily dichotomous.

Like the 'conversion factor method', the 'survey severity method' also has the advantage of allowing measurement of health benefits in a

comparable manner for most mental disorders. An important advantage is that it allows dichotomous outcome data to be incorporated when only data on relapse are available. This is done by calculating the odds ratio or relative risk of relapse and converting it to a risk difference [26], which is then used to change the proportions in each of two or more severity states. The disadvantages of the method are that: for most disorders only three or four severity states are specified by Dutch disability weights (this reduces the precision of continuous measurements of impact, although applying the effect size to a large number of survey respondents to some degree 'evens out' the discontinuous steps); and it relies on the accuracy of the survey instruments for measuring severity, such as the SF-12, DIP-DIS [27] (for schizophrenia) and Child Health Questionnaire. Both methods have the disadvantage that they do not allow measurement of the impact of side effects, which is particularly important for drug interventions for psychosis.

ACE–MH was also limited by the lack of health-related quality of life outcome data from trials and the availability of only two to four DALY disability weights, which coarsely describe severity for each mental disorder. Future work would be enhanced by the use, in trials, of a combination of utility-based economic instruments that facilitate comparisons across quite different interventions and disorders (a multi-attribute utility instrument) and which are sensitive to small changes (a disease-specific instrument) [28]. The accuracy of translating change in severity into DALY units would benefit from having disability weights specified for a greater number of different health states within each disorder the valuation study from which the DALY 'conversion factor' by severity level is derived for mental disorders [19] would need to be replicated to increase confidence in its validity.

The limitation in the methods for translating an effect size into a change in the DALY disability weight mainly influences comparisons of interventions across disorders. However, it is important to note that the main driver of cost-effectiveness ratios is the effect size. Further, the effect sizes for different interventions for a particular disorder are comparable as the

outcome measures are consistent. Thus, comparisons of cost-effectiveness ratios within disorders are more valid than across disorders.

CONCLUSIONS

Despite the limitations in methods the ACE–MH study is providing useful information for policy-makers. To our knowledge there have been no similar attempts at determining cost-effectiveness of interventions across a wide range of mental health interventions in a comparable manner. Despite considerable uncertainty around key input variables, clear distinctions in cost-effectiveness between mental health interventions (particularly within disorders) are apparent. Nevertheless, until there is greater consensus on how to quantify health benefits in mental disorders our estimates should be considered provisional, though indicative of the relative magnitude of the health gain. Results of the study will be published in upcoming papers in this and other journals. We hope the publication of our results will both encourage debate about future directions for mental health policy, and encourage further research to clarify those issues where current knowledge is lacking.

CAVEAT

The ACE-Mental Health project was jointly funded by the Australian Department of Health and Ageing, Mental Health and Suicide Prevention Branch and the Department of Human Services, Mental Health Branch, Victoria in recognition of the importance of research into the cost-effectiveness of interventions in mental health treatment and care. This work draws upon, but is also limited by the available research and the assumptions necessary to complete the work.

The results of the analyses provide valuable material, likely to contribute to future policy deliberations by all service providers. Conclusions drawn from the economic evaluations should be considered within the context of the second stage filter process, which qualifies the results taking into account issues of equity, feasibility, strength of evidence,

and acceptability to stakeholders. This second stage filter process addresses some of the practical considerations required for changes in actual service practice.

ACKNOWLEDGEMENTS

Principal Investigators for the project are: Dr Theo Vos, A/Professor Rob Carter and Professor Gavin Andrews. We thank members of the ACE–Mental Health Steering Committee for their input into the project: Dr David Barton, Professor Graham Burrows (Chair), Sue Caleo, Professor Vaughan Carr, Dermot Casey, Joy Easton, Dr William Hart, Professor Helen Herrman, Barbara Hocking, Professor Assen Jablensky, Professor Anthony Jorm, Dr Lyn Littlefield, Professor Patrick McGorry, John McGrath, Dr Paul Morgan, Lorna Payne, Deb Podbury, Dr Kristy Sanderson, Suzy Saw, Professor Bruce Singh, Professor Bruce Tonge, Dr Ruth Vine, Professor Harvey Whiteford. We also thank the other researchers who have worked on the project: Maturot Chalamat, Marie Donnelly, Justine Corry, Louise Heuzenroeder and Ruth Rossell.

REFERENCES

1. Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-effectiveness in health and medicine*. New York: Oxford University Press, 1996.
2. Mathers CD, Vos T, Stevenson CE. *The burden of disease and injury in Australia*. Canberra: Australian Institute of Health and Welfare, 1999.
3. Vos T, Begg S. *The Victorian Burden of Disease Study: morbidity*. Melbourne: Public Health and Development Division, Department of Human Services, 1999.
4. Nathan PE, Gorman JM. *A guide to treatments that work*. New York: Oxford University Press, 1998.
5. Andrews G, Sanderson K, Slade T, Issakidis C. Why does the burden of disease persist? Relating the burden of anxiety and depression to effectiveness of treatment. *Bull World Health Organ* 2000; 78:446-54.

6. Carter R: The Macro Economic Evaluation Model (MEEM): an approach to priority setting in the health sector, in Faculty of Business and Economics. Melbourne, Monash University, 2001
7. Carter R, Stone C, Vos T, Hocking J, Mihalopoulos C, Peacock S, Crowley S. *Trial of Program Budgetting and Marginal Analysis (PBMA) to assist cancer control planning in Australia*. Canberra: Commonwealth Department of Health and Aged Care, 2000.
8. Australian Bureau of Statistics. *Mental health and wellbeing: Profile of adults, Australia, 1997*. Canberra: ABS, 1998.
9. Sawyer MG, Arney FM, Baghurst PA, Clark JJ, Graetz BW, Kosky RJ. *The Mental Health of Young People in Australia*. Canberra: Mental Health and Special Programs Branch, Commonwealth Department of Health and Aged Care, 2000.
10. Jablensky A, McGrath J, Herrman H, Castle D, Gureje O, Evans M, Carr V, Morgan V, Korten A, Harvey C. Psychotic disorders in urban areas: an overview of the Study on Low Prevalence Disorders. *Australian and New Zealand Journal of Psychiatry* 2000; 34:221-36.
11. Stouthard MEA, Essink-Bot ML, Bonsel GJ, Barendregt JJ, Kramer PGN, van de Water HPA, Gunning-Schepers LJ, van der Maas PJ. *Disability weights for diseases in the Netherlands*. Rotterdam: Department of Public Health, Erasmus University, 1997.
12. Deeks JJ, Altman DG, Bradburn MJ. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In *Systematic reviews in health care: meta-analysis in context*. London: BMJ Publishing Group, 2001.
13. National Health and Medical Research Council. *How to use the evidence: assessment and application of scientific evidence*. Canberra: NHMRC, 2000.
14. Gould RA, Clum GA. A meta-analysis of self-help treatment approaches. *Clinical Psychology Review* 1993; 13:169-186.
15. Rosenthal R. *Meta-analytic procedures for social research*. Newbury Park: SAGE Publications, 1991.
16. Egger M, Smith GD, Altman DG (eds): Systematic reviews in health care: meta-analysis in context. London, BMJ Publishing Group, 2001
17. Hrobjartsson A, Gotzsche PC. Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. [see comments]. [Review] [134 refs]. *New England Journal of Medicine* 2001; 344:1594-1602.
18. The Quality Assurance Project. A treatment outline for depressive disorders. *Australian and New Zealand Journal of Psychiatry* 1983; 17:129-146.

19. Sanderson K, Andrews G, Corry J, Lapsley H. Using the effect size to model change in preference values from descriptive health status. *Qual Life Res* 2004; 13:1255-64.
20. Ormel J, Von Korff M, Van den Brink W, Katon W, Brilman E, Oldehinkel T. Depression, anxiety, and social disability show synchrony of change in primary care patients. *Am J Public Health* 1993; 83:385-90.
21. Chouinard G, Albright PS. Economic and health state utility determinations for schizophrenic patients treated with risperidone or haloperidol. *J Clin Psychopharmacol* 1997; 17:298-307.
22. Drummond MF, O'Brien B, Stoddart GL, Torrance GT. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford Medical Publications, 1997.
23. Australian Institute of Health and Welfare. *Health system expenditure on disease and injury in Australia, 2000-01.* Canberra: AIHW (Health and Welfare Expenditure Series no. 14), 2002.
24. Palisade Corporation: @RISK Risk analysis add-in for Micorsoft Excel, Version 4. Newfield, NY, Palisade corporation, 2001
25. Canadian Coordinating Office for Health Technology Assessment. *Guidelines for economic evaluation of pharmaceuticals: Canada, 2nd edn.* Ottawa: Canadian Coordinating Office for Health Technology Assessment, 1997.
26. Deeks JJ. Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. *Stat Med* 2002; 21:1575-600.
27. Gureje O, Herrman H, Harvey C, Trauer T, Jablensky A. Defining disability in psychosis: performance of the diagnostic interview for psychosis-disability module (DIP-DIS) in the Australian National Survey of Psychotic Disorders. *Aust N Z J Psychiatry* 2001; 35:846-51.
28. Richardson J, Olsen J, Hawthorne G, Mortimer D, Smith R. *The measurement and valuation of utility based quality of life: recommendations from a review of the literature. CHPE Working Paper*. Melbourne: Centre for Health Program Evaluation, 1999.

APPENDIX

Application of the Survey Severity Method to schizophrenia

The Low Prevalence Component of the NSMHW [10] did not use a general measure of quality of life with comparable psychometric validity as the SF-12 or Child Health Questionnaire that were used for depression and anxiety. However, Gureje et al. [27] proposed a severity measure in the survey combining six DIP-DIS questions – a scale which attempts to capture overall functioning in the previous year in the areas of socialising, social withdrawal, intimate relationships, self-care, interest & information, participation in household activities, perceived deterioration in interpersonal relationships – and two questions on satisfaction with own independence and life as a whole. The scores on this measure range from 0 ('well') to 16 ('very poor').

This summary DIP-DIS score discriminates – and in the right direction – between four categories describing the course of the disorder in the respondents in the survey (Table A). The distribution of DIP-DIS scores is skewed to the right with the largest concentration in the scores of 0–5. After log-transformation of the DIP-DIS score the higher values are normally distributed but not the lower values. The Dutch DWs for schizophrenia are for the full course of the disorder and as such are not suited to measure change in severity over time (or due to an intervention). However, we assume that the 0.98 weight for 'several psychotic episodes, severe and increasing permanent impairments' reflects the most severe state in schizophrenia and that the 0.21 weight for "one psychotic episode, no permanent impairments" is the best possible health state in schizophrenia. We then assume that the 0.21 weight corresponds with a score of 0 on the DIP-DIS and the 0.98 with a DIP-DIS score of 16 and that the DW changes linearly with a change in DIP-DIS score. This allows us to calculate a DW for each individual in the survey and to calculate mean values for different subgroups. Table A shows the mean DW by the variable course of disorder in the survey.

The calculated DWs distinguish between interviewer-rated course of the disorder categories, albeit that the difference between the least and most severe categories is only 0.15. It is not clear if this is due to a weakness in the DIP-DIS score or an insensitive rating on course of disorder. The effect size is then applied to the SD of the DIP-DIS score (with a preference for using the log transformed score as it approximates a normal distribution for most of the DIP-DIS score values). An effect size of 1 translates into a DW change of 0.113.

Table A. Mean DIP-DIS score for Schizophrenia and calculated DW by course of disorder

Course of disorder	**N**	**Mean DIP-DIS score (95% CI) ***	**Mean DW (95% CI)**
Multiple episodes, good recovery †	222	2.3 (2.0, 2.6)	0.32 (0.31, 0.33)
Multiple episodes, partial recovery	224	3.4 (3.1, 3.8)	0.38 (0.36, 0.39)
Chronic, little deterioration	237	3.7 (3.4, 4.1)	0.39 (0.37, 0.41)
Chronic, clear deterioration	283	5.4 (5.0, 5.8)	0.47 (0.45, 0.49)
Total	966	3.8 (3.6, 4.0)	0.39 (0.38, 0.40)

* Calculated from the Low Prevalence Component of the NSMHW [10]
† Also includes single episodes, good or unknown recovery

5

The burden of major depression avoidable by longer-term treatment strategies

Vos T, Haby MM, Barendregt JJ, Kruyshaar ME, Corry J, Andrews G (2004). The burden of major depression avoidable by longer-term treatment strategies. Archives of General Psychiatry, 61:1097–1103.

ABSTRACT

***Background**: Major depression is the largest single cause of non-fatal disease burden in Australia. Effective drug and psychological treatments exist, yet are underutilised.*

Objective: To quantify the burden of disease currently averted in people seeking care for major depression and the amount of disease burden that could be averted in these people under optimal episodic and maintenance treatment strategies.

Design: Modelling impact of current and optimal treatment strategies based on secondary analysis of mental health survey data, studies of the natural history of major depression and meta-analyses of effectiveness data. Monte Carlo simulation of uncertainty in the model.

Setting: The cohort of Australian Adults experiencing an episode of major depression in 2000 are modelled through 'what if' scenarios of no treatment, current treatment and optimal treatment strategies with cognitive behavioural therapy or anti-depressant drug treatment.

Main outcome measure: Disability-Adjusted Life Year

Results: Current episodic treatment averts 9% (95% uncertainty interval 6%-12%) of the disease burden of major depression in Australian adults. Optimal episodic treatment with cognitive behavioural therapy could avert 28% (19%-39%) of this disease burden and with drugs 24% (19%-30%) could be averted. Over five years following an episode of major depression, current episodic treatment patterns would avert 13% (10%-17%) of DALYs while maintenance drug treatment could avert 50% (40%-60%) and maintenance cognitive behavioural therapy could avert 52% (42%-64%) even if adherence of around 60% is taken into account.

Conclusions: Longer-term maintenance drug or psychological treatment strategies are required to make significant inroads into the large disease burden associated with major depression in the Australian population.

CHAPTER 5

INTRODUCTION

As in other industrialised countries, depression is the commonest mental disorder in Australia [1]. It is the largest single cause of disability responsible for 6.2% and 9.8% of Years Lived with Disability, in men and women, respectively. It is the third largest cause of disease burden in Disability-Adjusted Life Years (DALY) in women and ranked eighth in men [2]. Treatment guidelines recommend antidepressant (AD) drugs and/or a specific, effective psychological therapy for major depressive disorder [3, 4]. Cognitive-Behavioural Therapy (CBT) and interpersonal therapy are the psychological treatments that have the best-documented efficacy [4]. However, only 59% of respondents identified with major depression in the 1997 National Survey of Mental Health and Wellbeing (NSMHWB) had sought any help for their problem and 35% reported consulting a relevant health professional and receiving medication or psychological treatment [1].

The high burden of depression and the poor utilisation of efficacious treatments mean that there is considerable potential for health gain. A previous analysis quantified the cost-effectiveness of an ideal mix of interventions for affective disorders in Australia and concluded that evidence-based health care is supported on grounds of efficacy as well as cost-effectiveness [5]. However, it did not report the scope of health gain as a proportion of the overall disease burden of major depression that can be achieved by offering effective treatment to all people with depression who seek care from health services. Further limitations were that no separate conclusion could be made about the impact of different treatment options and that by limiting analyses to a one-year period the impact of long-term treatment strategies could not be evaluated. This paper quantifies the impact of treatment strategies on the disease burden due to major depression. In particular, it answers the following questions: (1) what is the proportion of the depression burden averted by current treatment; (2) what is the potential of episodic drug and psychological treatment options to further reduce this

burden; and (3) what is the potential of longer-term maintenance drug and psychological treatment options to further reduce this burden?

METHODS

The impact of evidence-based psychological and drug treatment strategies is modelled as a change in DALY. Separate estimates are presented for short-term treatments directed at episodes including a short continuation phase and longer-term maintenance treatments over five years of follow-up. Data are derived from existing surveys and routine health information collection systems in Australia as well as findings on the epidemiology of depression and its treatments in the international literature. The analysis starts with a description of the epidemiology of depression and current health service utilisation patterns in Australia. The next step is an evaluation of the impact of effective treatments by translating outcome measures from meta-analyses of trials into a change in DALY. The main comparisons are between the amount of depression experienced under current and expanded treatment options versus the hypothetical disease burden in the absence of treatment. Our analysis applies to Australian adults who experienced an episode of major depression in the year 2000 and sought care from health services.

Epidemiology of depression

We derive parameters on the prevalence of major depression and treatment patterns from the 1997 National Survey of Mental Health and Wellbeing (NSMHWB) [1] and apply these to 2000 population figures. The main outcome of the survey was the one-year prevalence, i.e. people qualifying for a diagnosis of major depression in the twelve months prior to survey. An additional question on the recency of symptoms allows identification of respondents who are currently prevalent, i.e. having had symptoms in the last two weeks, the minimum duration of an episode. Of the survey respondents identified as having major depression as defined by the *International Classification of Diseases, 10th Revision (ICD–10) [6]*, 58.9% had consulted a psychologist, psychiatrist and/or general practitioner for a

mental health problem while 35.1% fulfilled our criteria for potentially having received evidence based treatment (EBM): consulting at least three times and having had medication and/or CBT ("learning how to change thoughts, behaviours and emotions") [5].

We grade the severity of prevalent cases of depression from the NSMHWB by the number of standard deviations from the mean Mental Component Score of the 12-item Short-Form Health Survey (SF-12) [7] into normal (45+) mild (35–44.9), moderate (25–34.9) and severe (<25). Disability weights (DW) for mild (0.14), moderate (0.35) and severe (0.76) depression which were used in the Australian Burden of Disease Study [2] and derived from a Dutch study [8] are assumed to apply to these categories.

Natural history

Next, we use data from international follow-up studies on the natural history of major depression to mathematically describe the variation in duration of episodes and time to next episode. While there are many naturalistic studies of the duration of MDD episodes in clinical samples, there are few follow-up studies of MDD in community samples [9-13]. The four US studies show a similar pattern of recovery over time after the start of an episode. The median time to recovery in the four studies ranged from 8 to 12 weeks while at one year between 3 and 11% of cases had not yet recovered. The Kendler figures [10] have been adjusted for the 7% of excluded cases with onset more than 1 year prior to study. The fifth study from the Netherlands [11] reports a considerably longer duration of episodes. Inclusion of sub-syndromal depression and dysthymia in life chart histories is a possible explanation for this higher estimate. From the data reported in the US studies [9, 10, 12, 13] we fit a lognormal distribution [14] that has the lowest sum of squared differences between modelled and observed time to recovery starting from a minimum duration of 2 weeks specified in the definition of an episode of major depression (Figure 1).

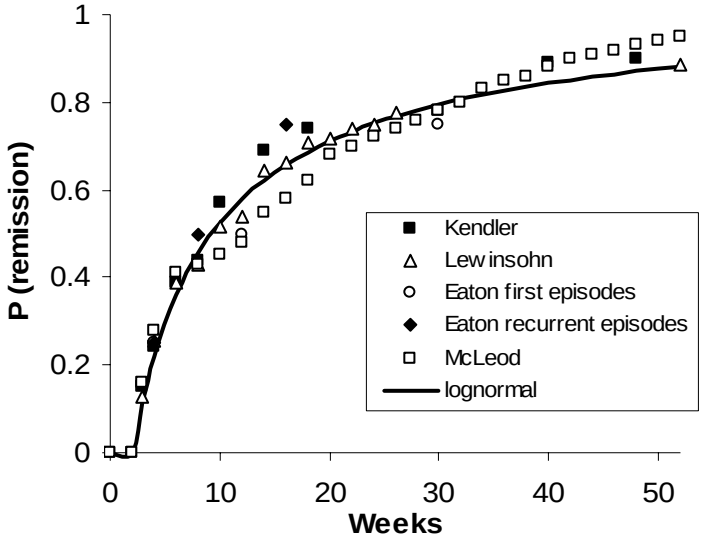

Figure 1. Time to recovery in episodes of major depression from four community samples [9, 10, 12, 13].

Major depression is a chronic episodic disorder and hence for our modelling purposes, it is important to describe the pattern of time to a next episode following a previous episode. Over a few decades of follow-up major depression is reported as a recurrent disorder in 80% of cases [15]. We assume that over a lifetime at least 90% of cases experiences a recurrence. Six naturalistic follow-up studies [16-21] report on the risk of relapse during periods varying between six months and two years after cessation of drug treatment for an acute episode of major depression. We fit lognormal and Weibull [22] distributions that give the best fit as determined by the lowest sum of squared differences between modelled and observed data points (Figure 2). We decide to use the lognormal distribution as it gives a slightly better fit. In a Monte Carlo simulation model we use the lognormal distributions describing the length of episodes and the time between episodes to estimate the mean number of episodes and the mean time depressed over a six-month and a five-year period following an episode.

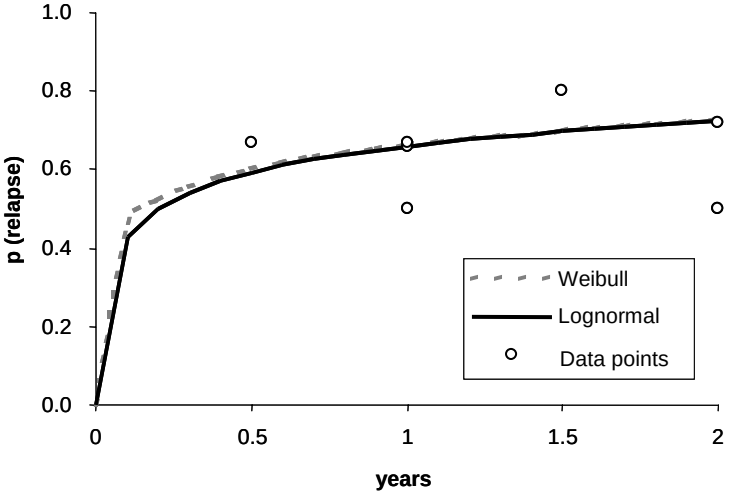

Figure 2. Weibull and Lognormal distributions fitted to time to relapse after an episode of major depression from six naturalistic follow-up studies [16-21].

Impact of interventions

We separately evaluate drug treatment for episodes of major depression plus a continuation phase after remission of symptoms and maintenance treatment of five years following remission of an episode; CBT treatment of major depressive episodes; and a maintenance variant of CBT with booster sessions over a period of five years.

A meta-analysis reporting on 48 trials estimated an effect size (ES) of 0.55 (95%CI 0.40-0.70) for selective serotonin re-uptake inhibitors (SSRIs) over placebo [23]. No differences were found between four different SSRIs. Meta-analyses examining the efficacy of SSRIs and tricyclic anti-depressants (TCAs) consistently show no significant differences between the two drug classes [23-26]. Therefore, we assume the same efficacy for all anti-depressants.

From the figures presented in a recent meta-analysis [27] of the odds ratios of relapse in 26 maintenance drug studies and 7 continuation drug

studies we derive a pooled relative risk (RR) of 0.416 (0.312 - 0.555) for relapse with continuation AD drug treatment and 0.437 (0.394 - 0.485) for maintenance AD drug treatment.

A meta-analysis of cognitive therapy reports a pooled ES of 0.82 from 48 studies [28]. On closer inspection, several studies included in this systematic review do not fit the stated inclusion criteria. Our own meta-analysis of CBT interventions including many of the same studies as well as a few additional studies, gives a random effects ES of 0.77 (0.44-1.10) close to the Gloaguen point estimate but with wider confidence intervals. We use these figures in our main analyses. Excluding two outlier studies (by the same author) with particularly high ES estimates reduces the Q statistic for heterogeneity from 50.8 (df=16, P<0.001) to 22.3 (df=13, P=0.051). In a separate sensitivity analysis we recalculate the model using the ES (0.54; 0.29-0.79) calculated after excluding these two outliers.

While the effect of AD drugs ceases when treatment is stopped, there is evidence for a prolonged effect of CBT beyond the treatment period. From a review of naturalistic longer-term follow-up studies (ranging from 1.5 to 4 years) after RCTs[17, 20, 29-31] that were set up to compare CBT with ADs in the acute phase we calculate a lower risk of relapse following CBT (RR 0.64; 95% CI 0.51–0.79).

Maintenance CBT is described in two trials. The first compares CBT maintenance with AD drug maintenance and over one year follow-up found no difference in relapse [32]. The other reports on a trial of maintenance CBT following acute CBT [21]. At two year follow-up the groups who had maintenance CBT had 25% relapse (5/20) compared to 80% (16/20) in the group getting case management only after CBT (AD drugs were tapered off and discontinued in both groups). The scanty evidence from these two trials suggests similar impact of maintenance strategies with AD drugs or CBT.

Adherence

Several meta-analyses with a large overlap in the included studies report discontinuation rates of between 27% and 39% with 3-6% lower rates for SSRIs in comparison to TCAs [23, 24, 26, 33]. However, as most trials are

of short duration representing what is possible with motivated patients and doctors, adherence rates may be lower than reported in the controlled trial literature. Adherence in four studies of primary care ranges from 50 to 66% [34-37]. We decide to model drug adherence ranging uniformly between the recorded adherence level in trials and an estimated lower level of 50% adherence in community settings.

We have found one community study of the attrition rate of CBT for depression where volunteers were recruited via the local media for a 12-week course of CBT. Total dropouts were 47% with almost half of those dropping out in the first three weeks [38]. As with AD drugs, we model adherence ranging between the estimate reported in trials (81%) [39] and a lower estimate of 50% in community settings.

Translating treatment impact in DALYs

The health benefit of interventions is measured in DALYs which is the sum of a non-fatal component (YLD) determined by the severity-weighted time lived with depression and a fatal component, Years of Life Lost (YLL), calculated as the stream of life lost due to suicide.

As described elsewhere [40], we use two methods to translate effect sizes from trial literature into a reduction in the DW. Briefly, the first method relies on an estimate of DW change for each standard deviation change in severity of depression which we call the conversion factor [41]. As the effect size quantifies the impact of an intervention in SD units, health gain in DALY units can be calculated as the product of the effect size, the conversion factor and the duration spent in the health state. The second (survey severity) method applies the effect size to the MCS of the SF-12 across eligible respondents in the mental health survey after which the difference in average DW with and without treatment is calculated. Results from both methods are incorporated in our uncertainty analyses and hence broaden the uncertainty ranges around the results presented.

Reductions in the DW are only applied to the time from the commencement of the intervention i.e. taking into account that there is a lag to treatment seeking after onset of symptoms. A UK study found a median

10-week interval between onset and seeking care for patients with an affective disorder [42]. We cannot assume a similar lag as the proportion of cases with a duration shorter than 10 weeks in a community sample is greater than the total proportion not seeking care in NSMHWB. Instead, based on expert consultation we decide to model a lag varying between 2 and 6 weeks.

As ESs are calculated from continuous measures and are not calculated on an intention-to-treat basis we apply the full non-adherence rate as a reduction in impact. For cases not adherent with treatment no reduction in DW is modelled.

From the point prevalence of depression in the NSMHWB, a UK estimate of the RR for suicide of 20.4 [43] and observed suicide deaths in Australia in 2000, we derive suicide deaths attributable to depression by age and sex. We assume a RR of 1.8 from Swedish routine data collection systems [44] applies to time lived with depression while not effectively on AD drugs. In the absence of long-term studies, we assume that suicide rates are similar in patients on CBT as in those on AD drugs.

From these estimates we derive suicide rates in those currently on effective treatment and those ineffectively treated. As in the Australian Burden of Disease Study, the YLL associated with a death are calculated as the cohort life expectancy for each age and sex category. We then divide the sum of YLL for suicide in treated and untreated depression by the person-years of depression in 2000.

The size of the burden averted by current treatment strategies requires a back-calculation of the burden if no treatment were given. This is done by applying the ES estimates for CBT and AD drugs to the mean disability weight of respondents in the NSMHWB on these treatments taking into account the estimated lag to treatment and level of adherence.

Table 1. Model input parameter values and sources of information

Parameter		Value (uncertainty range)	Source
RR suicide in prevalent depression		triangular distribution (18.2, 20.4, 22.6)	[43]
RR suicide on treatment vs. not on treatment		triangular distribution (1.6, 1.8, 2.0)	[44]
Effect size	AD drugs	triangular distribution (0.4, 0.55, 0.7)	[23]
	CBT	triangular distribution (0.44, 0.77, 1.10)	Own meta-analysis
RR relapse during 6 months continuation AD drug treatment		triangular distribution (0.312, 0.416, 0.555)	Own meta-analysis of 7 studies reported in [27]
RR relapse maintenance treatment		triangular distribution (0.394, 0.437, 0.485)	Own meta-analysis of placebo arms of trials evaluated in [27]
RR relapse in 18 months following CBT		triangular distribution (0.514, 0. 636, 0. 787)	Own meta-analysis of follow-up studies after CBT
Average duration episodes		29.9 weeks (0.57 years)	Based on fitted lognormal distribution (μ2.0, σ1.6) and two weeks minimum duration
Time depressed:			Based on fitted
6 months following episode		19.5%	lognormal (μ2.4, σ3.9)
5 years following episode		20.8%	distribution of time to
Average number of episodes over 5 years following episode		2.4	next episode and fitted lognormal distribution of duration of episodes
Disability weight conversion factor		uniform distribution (0.139–0.172)	[41]
Adherence with	AD drugs	50% – 73%	Upper values from [24]
	CBT	50% – 81%	and [39]
Lag to treatment		uniform distribution (2–6)	Reduced from 10 week estimate by [42]
Proportion of cases seeking care		triangular distribution (0.541, 0. 589, 0. 637)	Mental health survey [1]
Proportion of cases on evidence-based treatments		59.5% of those seeking care	Mental health survey [1]

Uncertainty

We use simulation-modelling techniques and present uncertainty ranges instead of point estimates that reflect all the main sources of uncertainty in the calculations. Details of the parameters and distributions for the uncertainty assumptions are shown in Table 1. The probability distributions around the input variables are based on (1) standard errors quoted in, or calculated from, the literature; (2) a range of parameter values quoted in, or calculated from, the literature; or (3) expert advice. We use the @RISK software (Palisade Corporation, Newfield, NY), which allows multiple recalculations of a spreadsheet each time choosing a value from uncertainty distributions defined for input variables. We run a Monte-Carlo simulation and calculate 95% uncertainty ranges for our output variables (bounded by the 2.5 and 97.5 percentiles of the 4000 values generated).

To identify the main sources of uncertainty affecting our results we regress the values of each of the input variables against results in each of the iterations of our simulation modelling. We report on input variables with a regression coefficient greater than 0.2 or less than –0.2. All results are presented to 2 significant digits only.

RESULTS

The fitted lognormal distribution for the duration of episodes (corresponding to a normal distribution with a mean of 2.049 and SD of 1.599) has a mean of 27.9 weeks resulting in an average duration of episodes of 29.9 weeks after adding the minimum 2 weeks of duration. In combination with the fitted lognormal distribution of time to next episode (corresponding to a normal distribution with a mean of 2.353 and SD of 3.876) the modelled mean number of episodes during five years of follow-up after an episode is 2.4 and the mean time spent in major depression is 20.8%. The mean time spent with depression in six months following an episode is 19.5%.

The mean DW for mental health survey respondents on evidence-based treatment (0.429), those consulting but not receiving evidence-based treatment (0.364) and those not consulting (0.282) indicate that those with

more severe disease are more likely to seek care and to be offered potentially effective treatments. We attribute a reduction in DW from 0.490 (the hypothetical level of severity without treatment) to 0.429 to current treatment strategies.

In the year 2000, we estimate that 555 male and 198 female suicide deaths are attributable to major depression in Australia (or 30% of all suicides). Per person-year lived with major depression the suicide risk is 0.8% in males and 0.3% in females. For both sexes combined, the risk of suicide in those on medication or CBT is 0.26% and in those not treated 0.47%. This translates on average across all ages into an annual loss of 0.093 YLL if treated and 0.167 YLL without treatment, i.e. a net health gain of 0.074 YLL that we attribute to treatment per year lived with depression.

During the first year following onset of an episode of major depression, current treatment strategies avert 10% (95% uncertainty interval 6%-12%) of the burden experienced by those in contact with health services. Treatment during the episode with an additional 6 months continuation treatment can raise this proportion to 28% (95% uncertainty interval 19%-39%) with CBT and 24% (95% uncertainty interval 19%-30%) with AD drugs (Figure 3a).

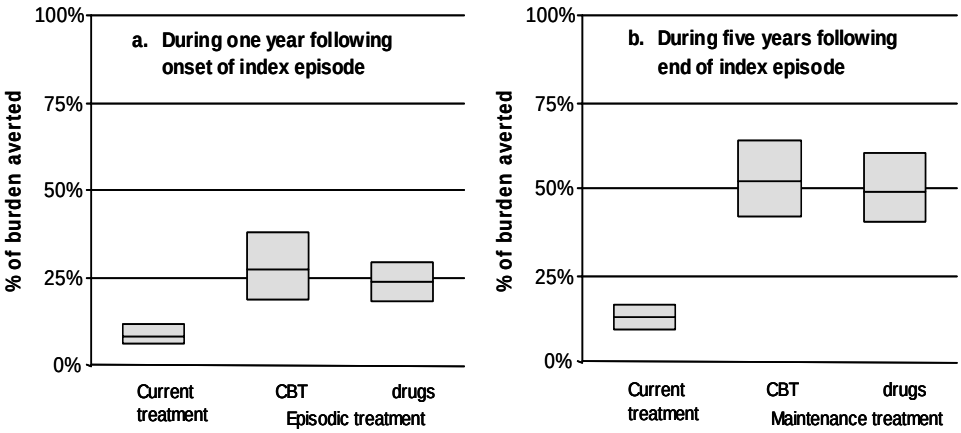

Figure 3. Proportion of time in depression averted (95% uncertainty interval) for episodic (a) and maintenance (b) treatment strategies in those in contact with health services

If all those seeking care for an episode are offered five years of maintenance treatment 52% (95% uncertainty interval 42%-64%) of the burden can be averted with CBT and 50% (95% uncertainty interval 40%-60%) with AD drugs compared to 13% (95% uncertainty interval 10%-17%) under a scenario in which episodic treatment continues (Figure 3b).

The results for CBT are not very sensitive to the choice of ES for CBT (0.54 vs 0.77). The lower ES estimates brings the estimates of burden averted by CBT down by less than 2 percentage points. Similarly, the proportion of burden averted by AD drugs is only modestly sensitive to the assumed effect size. Altering the effect size for AD drugs by 25% alters results by 3 percentage points for episodic treatment and less than one percentage point for maintenance treatment.

The main sources of uncertainty in the model are the assumed treatment discontinuation rates; the method of calculating a reduction in DW; and, to a lesser extent, the ESs. Prevention of suicide contributes to almost a third of the amount of health gain in DALYs for each of the 4 intervention scenarios in comparison to no treatment. In the episodic treatment scenarios reduction in severity is the main impact of treatments while in maintenance treatment the impact on preventing relapse contributes more to overall health gain than reduction of severity while depressed (Table 2).

Table 2. Proportion of health gain attributed to reduced severity, relapse prevention and suicide prevention due to episodic or maintenance treatments with AD drugs or CBT

	Prevention of relapse	Reduced severity while in episode	Prevention of suicide
AD drugs in acute episode & 6 months continuation	24%	46%	31%
CBT including impact on relapse over one year period	18%	53%	29%
Maintenance AD drugs	43%	28%	30%
Maintenance CBT	39%	33%	28%

COMMENT

Our results strongly support longer-term treatment strategies for depression. Despite assuming rates of adherence to treatment of around 60% we estimate that half of depression experienced in five years following an episode of major depression can be averted. The main reason for this favourable outcome is that maintenance treatment prevents relapses and that relapses that do occur are being treated from the start rather than being implemented after a lag time to seeking appropriate care. As the vast majority of people with depression experience multiple episodes over a lifetime and are particularly prone to relapses shortly after an index episode there are convincing arguments to treat all depression as a chronic disorder and not just those presenting with recurrent or more severe episodes as recommended in current treatment guidelines [3, 4].

We have made a conscious choice to simplify our modelling by using averages, e.g. for severity of episodes, the duration of the index episode and by modelling all ages and both sexes together. Some of these decisions do not do justice to the great complexity and variation in the presentation of depression. However, each added complexity requires more epidemiological input data with associated uncertainty and is limited by the lack of efficacy data for different durations, severities, gender and age. We believe we have struck a reasonable balance. The model takes enough of the complexities into account but still is simple enough for others to scrutinise and apply in other situations.

Elsewhere, we discuss the difficulties we encountered in translating trial findings into a health benefit in DALY terms [40]. To some extent we were able to incorporate this into our uncertainty analysis by using the range of results between two different methods of determining health benefit. The difference in burden averted between current practice and alternative treatment options is less affected as the same imperfect method is used for each treatment scenario. More accurate measurements of change in health status that can be attributed to interventions requires further developmental

work such as the use of general quality of life outcome measures in trials and more sensitive disability weights in DALY.

Our analyses are enhanced by the use of local epidemiological information. We had to rely on the 1997 National Survey of Mental Health and Wellbeing as the only and most recent source for much of the epidemiology of depression in Australia. Regular updates of the survey are needed to sustain this kind of analysis in the future. As this has been the only community prevalence study in Australia we are unable to incorporate temporal trends in the occurrence of depression. However, the time horizon over which we calculate our results is 5 years at most and hence, results are not much affected by the assumption of stable incidence of major depressive episodes. It would be very useful if a future survey identifying people with depression in the community endeavoured to follow-up people over time to examine if our modelled assumptions of duration, time to next episode and proportion of time with depression can be replicated in the Australian context.

The studies from which we derived our mathematical descriptions of the average duration of episodes and time to next episode are few and of relatively small size. Our 20% estimate of the average time with major depression over five years of follow-up is higher than that from a clinical study in the US which found that 15% of time was spent with depressive symptoms at the level of major depression over nine years of follow-up [45]. Our results are rather insensitive to this finding as the treatment impact measures applied to a 15% or 20% amount of depression over follow-up gives similar estimates of the proportion of depression burden averted.

We have limited our analyses to major depression ignoring that over follow-up time many people will spend time with sub-syndromal symptoms or dysthymia [45]. Assuming treatments are also effective for these types of depression this means that we have underestimated the true impact of treatments.

The measures of efficacy of maintenance treatment strategies are derived from studies of people who responded to treatment during an episode and hence it is not evident that this would apply equally to all people

with depression as we have modelled. However, our results are not very sensitive to the estimates of ES for AD drugs or CBT and hence our conclusions would not alter even if the effectiveness of treatment in primary care cases is estimated to be as much as 25% higher or lower. The information we used from two European studies [43, 44] to determine the risk of suicide is not so strong. However, the inclusion of YLL from suicide in the analyses is important because it constitutes almost a third of the overall health benefits. Our Australian estimates of suicide are high in comparison to a US estimate of suicide risk in people followed up after a diagnosis of depression [46]. However, if we take into account that suicide rates in young adults are 30% higher in Australia (based on analysis of deaths reported to WHO; available at http://www3.who.int/whosis/mort) and that we estimated risk of suicide only while depressed and not for all follow-up time, our estimates are only marginally higher (by 12% in males and 20% in females) than the US estimates.

Despite the limitations associated with lack of data on the course of depression and the impact of treatments, our results suggest that only by treating depression as a chronic episodic disorder with longer-term treatment strategies is it possible to make a meaningful reduction in the large burden of depression in Australia. Similarities in community survey findings on the epidemiology of major depression in the US [47, 48] and Australia [1] and the predominantly US studies on the impact of treatments used in our model, make it likely that our results also have relevance to depression in the US.

Psychological and drug treatments have similar impact on reducing the depression burden giving clinicians a choice of treatments. Additional information on cost-effectiveness is needed to complement these results in informing priority setting.

Funding/Support: This study was conducted as part of the Assessing Cost Effectiveness (ACE)–Mental Health project funded by the Australian Department of Health & Ageing, Canberra, Australia and the Victorian Department of Human Services, Melbourne, Australia.

Acknowledgments: Unit record data of the National Survey of Mental Health and Wellbeing were obtained from the Australian Bureau of Statistics and Gavin Andrews, MD, provided a revised scoring algorithm to determine *ICD–10* and *DSM–IV* diagnoses.

REFERENCES

1. Australian Bureau of Statistics. *Mental health and wellbeing: Profile of adults, Australia, 1997.* Canberra: ABS, 1998.
2. Mathers CD, Vos T, Stevenson CE. *The burden of disease and injury in Australia*. Canberra: Australian Institute of Health and Welfare, 1999.
3. American Psychiatric Association: Practice guideline for the treatment of patients with major depression, 2000
4. Ellis PM, Hickie IB, Smith DAR. Summary of guideline for the treatment of depression. *Australasian Psychiatry* 2003; 11:34-38.
5. Sanderson K, Andrews G, Corry J, Lapsley H. Reducing the burden of affective disorders: is evidence-based health care affordable? *J Affect Disord* 2003; 77:109-25.
6. World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders - Diagnostic criteria for research*. Geneva: WHO, 1993.
7. Ware J, Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996; 34:220-33.
8. Stouthard M, Essink-Bot M, Bonsel G, Barendregt JJ, P K. *Disability weights for diseases in the Netherlands.* Rotterdam: Department of Public Health, Erasmus University, 1997.
9. Lewinsohn PM, Clarke G, Seeley JR, Rohde P. Major depression in community adolescents; age at onset, episode duration and time to recurrence. *Journal of the American Academy of Child and Adolescent Psychiatry* 1994; 33:809-818.
10. Kendler KS, Walters EE, Kessler RC. The prediction of length of major depressive episode: results from an epidemiological sample of female twins. *Psychological Medicine* 1997; 27:110-117.
11. Spijker J, de Graaf R, Bijl RV, Beekman AT, Ormel J, Nolen WA. Duration of major depressive episodes in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *British Journal of Psychiatry* 2002; 181:208-13.

12. Eaton WW, Anthony JC, Gallo J, Cai G, Tien A, Romanoski A, Lyketsos C, Chen L-S. Natural history of diagnostic interview schedule/DSM-IV major depression. *Archives of General Psychiatry* 1997; 54:993-999.
13. McLeod JD, Kessler RC, Landis KR. Speed of recovery from major depressive episodes in a community sample of married men and women. *Journal of Abnormal Psychology* 1992; 101:277-286.
14. Ustun TB, Kessler RC. Global burden of depressive disorders: the issue of duration. *British Journal of Psychiatry* 2002; 181:181-3.
15. Judd LL. The clinical course of unipolar major depressive disorders. *Archives of General Psychiatry* 1997; 54:989-991.
16. Simons AD, Murphy GE, Levine JL, Wetzel RD. Cognitive therapy and pharmacotherapy for depression. *Archives of General Psychiatry* 1986; 43:43-48.
17. Shea MT, Elkin I, Imber SD, Sotsky SM, Watkins JT, Collins JF, Pikonis PA, Beckham E, Glass DR, Dolan RT, Parloff MB. Course of depressive symptoms over follow-up: Findings from the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Archives of General Psychiatry* 1992; 49:782-787.
18. McLean PD, Hakstian AR. Relative endurance of unipolar depression treatment effects: Longitudinal follow-up. *Journal of Consulting and Clinical Psychology* 1990; 58:482-488.
19. Kovacs M, Rush AJ, Beck AT, Hollon SD. Depressed outpatients treated with cognitive therapy or pharmacotherapy. A one-year follow up. *Archives of General Psychiatry* 1981; 38:33-39.
20. Evans MD, Hollon SD, DeRubeis RJ, Piasecki JM, Grove WM, Garvey MJ, Tuason VB. Differential relapse following cognitive therapy and pharmacotherapy for depression. *Archives of General Psychiatry* 1992; 49:802-808.
21. Fava GA, Rafanelli C, Grandi S, Conti S, Belluardo P. Prevention of recurrent depression with cognitive behavioral therapy: preliminary findings. *Archives of General Psychiatry* 1998; 55:816-20.
22. Weibull W. A statistical distribution of wide applicability. *Journal of Applied Mechanics* 1951; 18:193-297.
23. Trindade E, Menon D. *Selective serotonin reuptake inhibitors (SSRIs) for major depression. Part 1. Evaluation of the clinical literature.* Ontario, Canada: Canadian Coordinating Office for Health Technology Assessment, 1997.
24. Anderson IM. Selective serotonin reuptake inhibitors versus tricyclic antidepressants: a meta-analysis of efficacy and tolerability. *Journal of Affective Disorders* 2000; 58:19-36.

25. Geddes JR, Freemantle N, Mason J, Eccles MP, Boynton J. SSRIs versus other antidepressants for depressive disorder. *The Cochrane Library* 2000; Update 2000
26. Williams JW, Jr., Mulrow CD, Chiquette E, Noel PH, Aguilar C, Cornell J. A systematic review of newer pharmacotherapies for depression in adults: evidence report summary. *Annals of Internal Medicine* 2000; 132:743-756.
27. Geddes J, Carney SM, Davies C, Furukawa TA, Kupfer D, Frank E, Goodwin GM. Relapse prevention with antidepressant drug treatment in depressive disorders: a systematic review. *The Lancet* 2003; 361:653-61.
28. Gloaguen V, Cottraux J, Cucherat M, Blackburn I-M. A meta-analysis of the effects of cognitive therapy in depressed patients. *Journal of Affective Disorders* 1998; 49:59-72.
29. Hollon SD, Evans MD, DeRubeis RJ. Cognitive mediation of relapse prevention following treatment for depression: Implications of differential risk. In *Contemporary psycholgical approaches to depression*. New York: Guilford Press, 1990.
30. Fava GA, Grandi S, SZielezny M, Rafanelli C, Canestrari R. Four-year outcome for cognitive behavoiural treatment of residual symptoms in major depression. *American Journal of Psychiatry* 1996; 153:945-947.
31. Paykel ES, Scott J, Teasdale JD, Johnson AL, Garland A, Moore RG, Jenaway A, Cornwall PL, Hayhurst H, Abbott R, Pope M. Prevention of relapse in residual depression by cognitive therapy: a controlled trial. *Archives of General Psychiatry* 1999; 56:829-835.
32. Blackburn IM, Moore RG. Controlled acute and follow-up trial of cognitive therapy and pharmacotherapy in out-patients with recurrent depression. *British Journal of Psychiatry* 1997; 171:328-334.
33. Barbui C, Hotopf M, Freemantle N, Boynton J, Churchill R, Eccles MP, Geddes JR, Hardy R, Lewis G, Mason JM. Selective serotonin reuptake inhibitors versus tricyclic and heterocyclic antidepressants: comparison of drug adherence (Cochrane review). In *The Cochrane Library, issue 1*. Oxford: Update software, 2001.
34. Simon GE, Von Korff M, Rutter CM, Peterson DA. Treatment process and outcomes for managed care patients receiving new antidepressant prescriptions from psychiatrists and primary care physicians. *Archives of General Psychiatry* 2001; 58:395-401.
35. Lin EH, Von Korff M, Katon W, Bush T, Simon GE, Walker E, Robinson P. The role of the primary care physician in patients' adherence to antidepressant therapy. *Medical Care* 1995; 33:67-74.

36. Katon W, Robinson P, Von Korff M, Lin E, Bush T, Ludman E, Simon G, Walker E. A multifaceted intervention to improve treatment of depression in primary care. *Archives of General Psychiatry* 1996; 53:924-932.
37. Katon W, Rutter C, Ludman EJ, Von Korff M, Lin E, Simon G, Bush T, Walker E, Unutzer J. A randomized trial of relapse prevention of depression in primary care. *Archives of General Psychiatry* 2001; 58:241-247.
38. Oei TP, Kazmierczak T. Factors associated with dropout in a group cognitive behaviour therapy for mood disorders. *Behavioral Research and Therapy* 1997; 35:1025-30.
39. Antonuccio DO, Thomas M, Danton WG. A cost-effectiveness analysis of cognitive behavior therapy and fluoxetine (prozac) in the treatment of depression. *Behavior Therapy* 1997; 28:187-210.
40. Haby MM, Carter R, Mihalopoulos C, Magnus A, Sanderson K, Andrews G, Vos T. Assessing Cost-Effectiveness - Mental Health: Introduction to the study and methods. *Australian and New Zealand Journal of Psychiatry* 2004
41. Sanderson K, Andrews G, Corry J, Lapsley H. Modelling changes in preference values from descriptive health status using the effect size. *Quality of Life Research* 2003; (in press)
42. Gater R, Goldberg D. Pathways to psychiatric care in South Manchester. *British Journal of Psychiatry* 1991; 159:90-96.
43. Harris EC, Barraclough B. Suicide as an outcome for mental disorders: A meta-analysis. *British Journal of Psychiatry* 1997; 170:205-228.
44. Isacsson G, Bergman U, Rich CL. Epidemiological data suggest antidepressants reduce suicide risk among depressives. *Journal of Affective Disorders* 1996; 41:1-8.
45. Judd LL, Akiskal HS, Maser JD, Zeller PJ, Endicott J, Coryell W, Paulus M, Kunovac JL, Leon AC, Mueller TI, Rice JA, Keller. MB. A prospective 12-year study of subsyndromal and syndromal depressive symptoms in unipolar major depressive disorders. *Archives of General Psychiatry* 1998; 55:694-700.
46. Simon GE, VonKorff M. Suicide mortality among patients treated for depression in an insured population. *Am J Epidemiol* 1998; 147:155-60.
47. Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, Wittchen HU, Kendler KS. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Arch Gen Psychiatry* 1994; 51:8-19.

48. Blazer DG, Kessler RC, McGonagle KA, Swartz MS. The prevalence and distribution of major depression in a national community sample: the National Comorbidity Survey. *Am J Psychiatry* 1994; 151:979-86.

6

Cost-effectiveness of CBT and drug interventions for major depression

Vos T, Corry J, Haby MM, Carter R, Andrews G (2005). Cost-effectiveness of CBT and drug interventions for major depression. Australian and New Zealand Journal of Psychiatry, 39:683–692.

ABSTRACT

Objective: Antidepressant drugs and cognitive–behavioural therapy (CBT) are effective treatment options for depression and recommended by clinical practice guidelines. As part of the Assessing Cost-Effectiveness (ACE)–Mental Health project we evaluate the available evidence on costs and benefits of CBT and drugs in the episodic and maintenance treatment of major depression.

Method: The cost-effectiveness is modelled from a health care perspective as the cost per disability-adjusted life year. Interventions are targeted at people with major depression who currently seek care but receive non evidence-based treatment. Uncertainty in model inputs is tested using Monte Carlo simulation methods.

Results: All interventions for major depression examined have a favourable incremental cost-effectiveness ratio under Australian health service conditions. Bibliotherapy, group CBT, individual CBT by a psychologist on a public salary and tricyclic antidepressants (TCAs) are very cost-effective treatment options falling below A$10,000 per disability-adjusted life year (DALY) even when taking the upper limit of the uncertainty interval into account. Maintenance treatment with selective serotonin re-uptake inhibitors (SSRIs) is the most expensive option (ranging from A$17,000 to A$20,000 per DALY) but still well below A$50,000 considered the affordable threshold.

Conclusions: A range of cost-effective interventions for episodes of major depression exists and is currently under-utilised. Maintenance treatment strategies are required to significantly reduce the burden of depression but the cost of long-term drug treatment for the large number of depressed people is high if SSRIs are the drug of choice. Key policy issues with regards to expanded provision of CBT concern the availability of suitably trained providers and the funding mechanisms for therapy in primary care.

CHAPTER 6

INTRODUCTION

Major depression is the commonest mental disorder in Australia [1] and causes more disability than any other disease [2]. Effective treatment options with antidepressant drugs and/or a specific, effective psychotherapy exist and are recommended by treatment guidelines [3, 4]. Among the psychotherapeutic approaches cognitive-behavioural therapy (CBT) and interpersonal therapy have the best-documented efficacy. In addition, there are self-help variants of CBT using books ("bibliotherapy") and/or the Internet. However, a significant proportion of people identified with depression in a community survey reported not to have sought any care and those seeking care often did not get adequate treatment [1]. In a previous analysis, we have indicated that a meaningful reduction of the burden of depression in Australia would only take place if depression were treated as a chronic episodic condition requiring maintenance rather than episodic treatment [5]. In this paper, we present the cost-effectiveness of different treatment options for depression.

There is little comparable information on the cost-effectiveness of the different treatment options in the literature which consists predominantly of drug comparison studies of tricyclic (TCAs) and newer antidepressants such as the selective serotonin reuptake inhibitors (SSRIs). There are a number of methodological problems with most of these studies and none are directly transferable to the Australian health care system due to differences in costs of these treatments.

To overcome these shortcomings, as part of the Assessing Cost-Effectiveness–Mental Health (ACE–MH) project and using a common economic protocol [6] we decided to evaluate the available evidence on costs and benefits of CBT and antidepressant drugs in the treatment of episodes of major depressive disorder and as maintenance treatment in the prevention of recurrence of depressive episodes.

METHODS

The general methods of ACE-MH are reported elsewhere in detail [6]. In brief, interventions are chosen as options for change compared to current practice in the year 2000. Costs accruing in the Australian health sector, i.e. including all health service costs and patient out-of-pocket costs, are estimated and discounted at 3% back to year 2000 values.

A Steering Committee of mental health experts, policy makers and representatives of community organisations selected the interventions for analysis, scrutinised the assumptions made in modelling cost-effectiveness and assisted in defining and applying a set of 'second filter' criteria (quality of evidence, equity, feasibility and acceptability to stakeholders) which potentially can influence conclusions drawn based only on cost-effectiveness considerations. The selected interventions for adult depression include:

1. drug treatment with TCAs for acute major depressive episodes plus a six month continuation phase after remission of symptoms;
2. drug treatment with SSRIs (including other 'new' drugs such as venlafaxine, nefazodone and mianserin) for acute episodes and six months continuation;
3. CBT treatment of acute major depressive episodes consisting of 12 sessions analysed separately whether provided by a psychologist or a psychiatrist in public or private service and whether provided to individuals or in a group;
4. bibliotherapy for acute episodes;
5. maintenance treatment with TCAs for five years following an acute episode;
6. maintenance treatment with SSRIs for five years following an acute episode; and
7. a maintenance variant of CBT with booster sessions over a period of five years.

We model the impact of the acute episode interventions (options 1–4) in Australians experiencing at least one episode of major depression in the year 2000 (513,000) who sought care (302,000) but did not receive an evidence-

based treatment (122,000). These estimates are based on the 1997 Survey of Mental Health and Well Being (SMHWB) [1]. The criterion for seeking care was consulting for a mental health problem with a psychologist, psychiatrist, GP, surgeon, social worker, mental health team worker or an admission to hospital. We defined evidence-based treatment as a minimum of three such consultations and having received medication and/or CBT. We assume that only 15% of patients are likely to take up bibliotherapy. We model the impact of maintenance interventions (options 5–7) in all people seeking care for depression in the year 2000 assuming maintenance treatment is not part of current practice.

Health benefit

The health benefit of each intervention is measured in DALYs including a non-fatal component, years lived with disability (YLD), determined by the severity-weighted time lived with depression and a smaller fatal component, years of life lost (YLL), calculated as the discounted stream of life lost due to suicide.

Modelling health benefits requires data on the epidemiology of depression (duration of episodes, time to next episode and severity) and the impact of interventions (on severity and duration of episodes, on the risk of recurrent episodes and the risk of suicide while depressed). A companion paper focussing on the impact of treatment options on the size of the depression burden in Australia [5] provides detail of the methods of estimating health benefits summarised below and in Table 1.

We estimate the average duration of episodes and the proportion of time spent in depression following an episode from data published on community follow-up studies. We grade the severity of prevalent cases of depression from the SMHWB by the number of standard deviations (SDs) from the mean Mental Component Score (MCS) of the SF-12 into normal (>45) mild (35.1–45), moderate (25–34.9) and severe (<25) and apply disability weights (DW) for mild (0.14), moderate (0.35) and severe (0.76) depression [24] that were also used in the Australian Burden of Disease Study [2]. From the international literature of randomised controlled trials

we derive impact measures that alter the severity of depression during an episode or alter the duration to the next episode. The pooled effect size for SSRIs over placebo is 0.55 (95% CI 0.40–0.70) and no difference is reported between four different SSRI drugs [9]. Meta-analyses also consistently show no difference in efficacy between SSRIs and TCAs [21, 25, 26].

From the published figures of a recent meta-analysis [19] we derive a pooled reduction in the risk of relapse for continuation and maintenance antidepressant drug treatment (Table 1).

Table 1. Model input parameter values and sources of information

Parameter	**Value (uncertainty range)**	**Source**
RR suicide in prevalent depression	20.4 (SE 1.1)	[7]
RR suicide on treatment vs. not on treatment	1.8 (SE 0.1)	[8]
Effect size antidepressant drugs	0.55 (SE 0.075)	[9]
CBT	0.77 (0.44–1.10)	Own meta-analysis
Bibliotherapy	0.98 (0.62–1.35)	Own meta-analysis of 8 studies [10-18]
RR relapse during 6 months continuation drug treatment	0.416 (0.312–0.555)	Own meta-analysis of 7 studies reported in [19]
RR relapse maintenance treatment	0.437 (0.394–0.485)	Own meta-analysis of trials evaluated in [19]
RR relapse in 18 months following CBT	0.636 (0.514–0.787)	Own meta-analysis of follow-up studies after CBT
Disability weight change per SD change in severity of depression	0.139–0.172	[20]
Adherence with TCAs	50% – 69% (SE 3.3%)	Upper values from [21] and [22]
SSRIs	50% – 73% (SE 3.7%)	
CBT	50% – 81% (SE 4.4%)	
Lag to treatment	6 weeks (2 to 4 weeks)	Reduced from 10 week estimate in clinical sample [23] based on expert advice

Parameter	Value (uncertainty range)	Source
% of cases seeking care	58.9% (SE 2.4)	[1]
% cases on evidence-based treatments	59.5% of those seeking care	[1]
Types of antidepressant drugs		
SSRIs	77% (SE 0.9%)	BEACH
TCAs	15% (SE 0.7%)	
MAO-inhibitors	6% (SE 0.5%)	
Other	2% (SE 0.3%)	
- Group CBT sessions per day	2–3	Assumptions based on expert advice
- Participants group CBT	6–8	
- Individual CBT sessions/day for psychologist	5–7	
- CBT sessions per half day public psychiatrist	2–4	
Uncertainty fees health professionals not covered by Medicare	± 10%	Assumption; no uncertainty modelled around Medicare fees

From our own update of the Gloaguen meta-analysis of cognitive therapy [27] we estimate a pooled effect size of 0.77 (95% CI 0.44–1.10). Two outlier studies by the same author determine much of the heterogeneity between results. Excluding these studies reduces the effect size to 0.54 (95% CI 0.29–0.79) We base our main results on the first estimate but test the sensitivity of results if using the lower estimate. Our review of naturalistic longer-term follow-up studies following trials comparing CBT with antidepressant drugs in acute episodes [28-32] indicates a prolonged effect during 18 months after the treatment period. Two small trials of maintenance CBT suggest similar impact of maintenance strategies with antidepressant drugs or CBT [31, 33].

Adherence

Several recent meta-analyses with a large overlap in the studies included report 3-6% lower discontinuation rates for SSRIs in comparison to TCAs

[9, 21, 25, 34]. As most trials are of short duration representing what is possible with motivated patients and doctors, adherence rates for SSRIs and TCAs may be more divergent than what has been found in the controlled trial literature. Four studies of primary care report patient adherence ranging from 50 to 66% but do not make mention of the type of antidepressant [35-38]. While trials of CBT on average report 81% adherence [22], a community study recruiting volunteers via the local media for a 12-week course of CBT reported 47% dropping out mostly in the first three weeks [39]. Consistent with other analyses in the ACE-MH project we model adherence to range between the recorded adherence level in trials and an estimated lower level of 50% adherence in community settings.

Calculation of benefits

As described elsewhere [6], we use two methods to translate the effect size into a health benefit. Briefly, the first method relies on an Australian valuation study which quantified a mean DW change for each SD change in severity of depression [20]. As the effect size quantifies the impact of an intervention in SD units, the product of the effect size and the mean DW change per SD represents the health gain. The second method applies the effect size to the MCS of the SF-12 across eligible respondents in the mental health survey after which the difference in average DW with and without treatment is calculated. Both methods are incorporated in our uncertainty analyses. This results in a broadening of the uncertainty ranges around the cost-effectiveness ratios reflecting the absence of an established method for translating effect sizes into a change in DALY DWs.

Reductions in the DW are only applied to the time from the commencement of the intervention taking into account a lag to treatment seeking. We model that those not adhering to treatment gain no health benefit although they can incur some costs.

During the continuation and maintenance treatment periods following an acute episode we assume that the reduction in risk of recurrence or relapse can be applied to the average time spent in depression. The comparator in the analysis for each intervention is continuation of current

treatment seeking patterns including a lag to treatment seeking for each new episode. For bibliotherapy we model health gain during the acute episode only. This may well be an underestimate as it is not unlikely that this self-help variant of CBT also has benefits beyond the acute episode. We make this decision because we have no direct evidence for a sustained benefit and because the cost-effectiveness ratio is very favourable anyway due to the very low cost of the intervention.

From the point prevalence of depression in the SMHWB, an estimate of the RR for suicide of 22.4 from a meta-analysis [7] and observed suicide deaths in Australia in 2000, we derive suicide deaths attributable to depression by age and sex. We assume a RR of 1.8 from Swedish routine data collection systems [8] applies to those who are not on effective treatment. No information is available that suggests that TCAs or SSRIs differ in their ability to prevent suicide and therefore we use the same estimate for both drug classes. In the absence of long-term studies, we also assume that suicide rates are similar in patients on CBT as in those on antidepressant drugs. As in the Australian Burden of Disease Study, the YLL associated with a death are calculated as the cohort life expectancy for each age and sex category. We then divide the total YLL for suicide in treated and untreated depression by the person-years of depression in 2000.

Health Service Costs

In our cost estimates we assume the same mix of providers as reported in the SMHWB, prescribing and supervising drug treatment (56% general practitioner (GP) only, 10% psychiatrist only and 33% GP plus psychiatrist). We take into account weekly visits in the first month, monthly visits for two months and then every three months for review. For those seeing both a GP and a psychiatrist we include one initial GP visit for referral, weekly sessions with a psychiatrist for one month and then fortnightly for another month before being referred back to the GP for further monitoring and management once every 1-2 months. We assume that cases not adherent with drug treatment accrue similar costs as those on non-evidence based treatments.

Table 2. *Summary of unit cost information, data sources and assumptions*

Element costed	**Cost to govt (A$)**	**Cost to patient (A$)***	**Source**	*Assumptions*
1 month SSRI	$73.54	$18.70	PBS	
1 month TCA	$10.94	$4.53	PBS	
1 GP visit of 20-40 mins	$39.51	$1.87	MBS	MBS item 36 – for initial visit and/or for referral/diagnosis
1 GP visit of < 20 mins	$21.88	$2.21	MBS	MBS item 23 – for second and subsequent visits
1 psychiatrist visit of 45-75 minutes	$117.02	$16.47	MBS	MBS item 306
1 psychiatrist visit of 15-30 minutes	$56.38	$5.95	MBS	MBS item 302
1 session of 46-60 mins with a private psychologist	$0	$115.00	Australian Psychological Society	The Australian Psychological Society recommended fee is $161 (as of July 1 2001) but personal communication with the Executive Director suggests the fee most commonly charged is $110-$120.
1 session of 60 mins with a public psychologist	$47.05	$0	Base salary from Victorian Hospitals' Industrial Association	Grade 3, Year 2 psychologist (PL2). Salary effective from 1 July 2000: $1085.80 per week. On-costs of 30% have been added. The psychologist has 6 patient contacts per day with the remainder of the working day used for preparation, administration, professional development etc.
1 session of 60 mins with a public psychiatrist	$129.64	$0	Base salary from the Austin & Repatriation Medical Centre Human Resources	The psychiatrist is paid as a visiting medical officer (VMO). An average salary per session has been used from the range: Specialist to Senior Specialist. Salary effective from 1 July 2000: $260.76 to $337.60 per session. The psychiatrist has 3 patient contacts per session (of 3.5 hours). On-costs of 30% have been added.

* Patients costs obtained from Department of Health and Ageing; PBS – Pharmaceutical Benefits Scheme; MBS - Medical Benefits Scheme

Drug costs are derived from the Health Insurance Commission for scripts filled under the Pharmaceutical Benefits Scheme (PBS) for the 1990/2000 financial year. The average dosage modelled for TCAs is 150 mg and for the group of SSRIs and related drugs 40mg (fluoxetine, paroxetine or citalopram, mianserin), 150mg (sertraline, venlafaxine), 200 mg (fluvoxamine) or 400 mg (nefazodone). The cost to the patient is averaged over general and concession patients and safety net and non-safety net patients. A summary monthly drug cost for SSRIs is calculated by weighting the cost for each drug by the proportion of prescriptions for each SSRI or TCA derived from the Bettering the Evaluation and Care of Health (BEACH) GP data collection system for the year 2000 [40].

For CBT, we cost 12 one-hour sessions for five scenarios: individual therapy sessions by a clinical psychologist or psychiatrist in public service or private practice as well as group therapy by a publicly funded psychologist. A GP visit for referral is included in all scenarios. Bibliotherapy is modelled as one long and two short visits to a GP and the purchase of a book (Table 2).

Depressed patients who receive non-evidence based treatment under current practice are costed for the mean number of visits to a GP (4.2), psychiatrist (0.3) and psychologist (2.2) as observed in SMHWB.

Uncertainty

In determining the benefits and costs associated with each intervention, we have assembled the best evidence available. However, there is uncertainty associated with all cost and outcome estimates. We use Monte Carlo simulation modelling and present uncertainty ranges as well as point estimates for benefits, costs, and cost-effectiveness ratios that reflect all the main sources of uncertainty in the calculations.

Details of the parameters and distributions for the uncertainty assumptions used are shown in Table 1. The probability distributions around the input variables are based on i) standard errors quoted in, or calculated from, the literature; ii) a range of parameter values quoted in, or calculated from, the literature; or iii) expert advice. We use the @RISK software

which basically is a macro allowing multiple recalculations of an Excel spreadsheet each time choosing a value from uncertainty distributions defined for input variables. We calculate 95% uncertainty ranges for our output variables (bounded by the 2.5th and 97.5th percentiles of the 4000 values generated). Results are presented to 2 significant digits only.

RESULTS

All interventions for major depression examined have a favourable incremental cost-effectiveness ratio (ICER) under Australian health service conditions. Bibliotherapy, group CBT, individual CBT by a psychologist on a public salary and TCAs are very cost-effective treatment options falling below A$10,000 per DALY even when taking the upper limit of the uncertainty interval into account (Table 3).

The ICER of maintenance treatment with CBT is very favourable as the costs per individual are relatively low in relation to considerable benefits. Maintenance and episodic treatment with TCAs have similar ICERs while maintenance treatment with SSRIs is less favourable (ranging from A$17,000 to A$20,000 per DALY) but still well below the threshold of A$50,000 we use in the ACE-Mental Health studies as 'affordable'.

Maintenance treatment strategies are associated with the largest amount of overall health gain. The total cost of treating all people with depression who present for care (taking an average 62% adherence into account) with 5 years of maintenance treatment with SSRIs is considerable: close to two billion A$. The costs are less than a third of that if TCAs are used and CBT is again a cheaper option than drug interventions.

In a sensitivity analysis, we examine using a lower ES for CBT (0.56 instead of 0.82), the pooled estimate of the impact of CBT trials for depression after exclusion of two outlier studies with very high effect sizes. This increases the ICER of CBT for acute episodes by 35% and by 6% for maintenance CBT. However, the overall conclusions remain the same: CBT is more cost-effective than drugs if provided by a publicly funded psychologist and particularly if done in groups; maintenance CBT is a very

cost-effective option; CBT by private providers has a similar ICER as TCAs; and SSRIs are less cost-effective, but still affordable.

Second filter considerations are presented in Table 4. The evidence base for both CBT and antidepressant drugs is sufficient although there are some concerns that almost all the evidence on CBT is in Anglo-Saxon populations and delivered by psychologists. There are moderate concerns about equity particularly if it would be largely available on a 'user pays' basis by private providers. Expanding access to CBT is challenging in terms of workforce and funding arrangements. The main issue with antidepressant drugs is that it is not very likely that many clinicians will go back to prescribing TCAs despite the fact that SSRIs give only marginal added benefits at much higher cost and hence the less favourable ICERs.

Table 3. Summary of point estimates and 95% uncertainty ranges of benefits, costs and incremental cost-effectiveness ratios of interventions for major depression.

Intervention	Target group	Benefit ('000s)[#] YLD	Benefit ('000s)[#] YLL	Costs (A$ Millions)[#] Intervention	Costs (A$ Millions)[#] Comparator	ICER ('000 A$ per DALY)
1. Antidepressant R/ for acute episode and 6/12 continuation R/						
SSRIs	Non-EBM	6.0 (3.8–8.9)	2.1 (1.5–2.9)	120 (99–140)	7 (4–11)	14 (11–18)
TCAs		5.9 (3.8–8.6)	2.1 (1.5–2.8)	51 (44–58)		5.5 (4.2–7.2)
2. CBT including extended benefit for 18 months following episode						
Individual CBT public psychologist	Non-EBM	9.0 (5.6–14)	3.0 (2.0–4.1)	50 (36–71)		3.5 (2.3–5.4)
Individual CBT private psychologist	Non-EBM			120 (87–150)		8.9 (6.7–12)
Individual CBT public psychiatrist	Non-EBM			130 (94–170)		10 (7.4–14)
Individual CBT private psychiatrist	Non-EBM			130 (100–160)		10 (8.1–14)
Group CBT public psychologist	Non-EBM			21 (15–31)	8 (4–12)	1.1 (0.5–2.0)
3. Bibiliotherapy	15% of non-EBM	1.1 (0.6–1.8)	0.3 (0.2–0.4)	1.4 (1.0– 1.9)	1.2 (0.7–1.9)	0.1 (dominant*–0.4)
4. Maintenance R/	All seeking					
SSRI	care	73 (58–89)	22 (16–28)	1,900 (1,600–2,300)	220 (180–260)	18 (17–20)
TCA		71 (58–85)	21 (16–27)	640 (530–740)		4.5 (4.1–4.9)

Table 3. Summary of point estimates and 95% uncertainty ranges of benefits, costs and incremental cost-effectiveness ratios of interventions for major depression (continued).

Intervention	Target group	Benefit ('000s)[#] YLD	Benefit ('000s)[#] YLL	Costs (A$ Millions)[#] Intervention	Costs (A$ Millions)[#] Comparator	ICER ('000 A$ per DALY)
5. Maintenance CBT						
Individual CBT public psychologist	All seeking care	84 (62–110)	23 (16–31)	190 (130–280)	220 (180–260)	dominant* (UL 0.2)
Individual CBT private psychologist				470 (340–600)		2.2 (1.7–2.8)
Individual CBT public psychiatrist				520 (370–700)		2.8 (2.0–3.8)
Individual CBT private psychiatrist				540 (420–660)		3.0 (2.6–3.4)
Group CBT public psychologist				67 (40–110)		dominant*

[#] Costs and benefits based on offering treatment for acute episodes (intervention options 1–3) to 122,000 people with an incident episode of major depression who sought care in the year 2000 but did not receive an evidence-based treatment (non-EBM). Maintenance treatment (intervention options 4 and 5) is modelled over 5 years in 302,000 people who sought care for an incident episode of major depression in the year 2000.

* An intervention is dominant if it produces net health benefit and net cost saving.

Table 4. Second stage filter considerations for cognitive-behavioural therapy (CBT) and anti-depressants.

Filter	CBT	Anti-depressants
Evidence	Sufficient evidence of adequate quality, noting however that there are: • few trials in people with non-English speaking background (NESB) • few trials for providers other than psychologists • no trials amongst the indigenous population	Sufficient evidence of adequate quality.
Equity	Moderate equity concerns require attention, i.e. • appropriateness for minority groups (e.g. NESB, indigenous) • access for rural / remote consumers and in outer metropolitan areas • inequity in access if 'user pays' (e.g. for private providers)	No important equity issues
Feasibility	Possible but challenging to implement in short term. Issues: • ensuring an adequate workforce, i.e. appropriately trained and accredited providers; adequate distribution • ability of health funding to enable adequate access via primary care	Feasible within current workforce and institutional arrangements

Table 4. Second stage filter considerations for cognitive-behavioural therapy (CBT) and anti-depressants (continued).

Filter	CBT	Anti-depressants
Acceptability	Some issues that require resolution: • cost to consumers if private providers • acceptance of treatment by clinicians and consumers • acceptance of a shift towards non-pharmacological treatments	Issues that require resolution: • Worries of consumers about side effects particularly of long-term drug treatment strategies • Reluctance by clinicians to prescribe TCAs due to the perception of greater side effects (while trial data indicate only small differences in drop-out rates due to side effects) • Pressure on clinicians to prescribe newer drugs (with a new generation of drugs already under development) • For depression - policy makers may be reluctant to endorse long-term drug treatment strategies particularly if the more expensive SSRIs are likely to be prescribed as there are already great pressures on the PBS budget

DISCUSSION

A range of cost-effective interventions for episodes of major depression exists. Bibliotherapy is the cheapest option with net cost savings if offered instead of ineffective treatment options. However, bibliotherapy may be a treatment option for only a small proportion of people with depression. CBT when delivered by psychologists on a 'public' salary is the next best option, followed by TCAs, CBT by private providers and SSRIs.

Maintenance treatment with CBT is a very favourable option but maintenance treatment with TCAs is also very cost-effective. However, the total cost for CBT would be a lot less than that for drugs. With a ICER around A$18,000 maintenance treatment with SSRIs is still a cost-effective option but the overall cost over five years is considerable if maintenance treatment is offered to all cases of major depression currently seeking care. The annual cost of maintenance treatment would be higher still as we have not included incident cases in the years following our baseline year of 2000.

Our results strongly favour longer-term treatment strategies for depression. As the vast majority of people with depression experience multiple episodes over a lifetime and are particularly prone to relapses shortly after an index episode there are strong arguments to treat all depression as a chronic episodic disorder and not just those with three or more episodes as recommended in current treatment guidelines [3]. Unlike drugs, CBT conveys a longer-lasting impact beyond the time of treatment. It therefore seems advisable to make CBT available to all people experiencing depression rather than to resort to large numbers of people taking psychotropic drugs as first-line care over long periods. CBT using books should be strongly advocated for those willing and able to use it.

Key policy issues with regards to CBT concern the availability of suitably trained providers and the funding thereof. The bulk of the evidence on CBT is for psychologists as providers. While there may be capacity to increase provision of CBT services above current levels it is unlikely that there are enough trained psychologists to extend services to all those with

major depression let alone including those with anxiety disorders who would benefit from CBT as well [41]. There is no explicit evidence that other providers such as nurses and social workers can deliver CBT with similar effectiveness. However, the fact that bibliotherapy has similar effectiveness indicates that the type of provider may not be the most critical element of CBT. Alternative modes of delivery of computer administered CBT (e.g. via the Internet or at GP practices) are currently being explored. Such methods would in particular facilitate access to CBT in rural and outer metropolitan areas where there are few therapists.

Funding is the other main policy issue. Current mental health services concentrate on treating the more severe and complex mental disorders. Large-scale provision of CBT for depression would mean a major change in focus of these services. Expanding services in primary care in collaboration with GPs and community health centres is a more feasible option. However, the main challenge is how to fund it. A recent initiative by the Australian government to improve access to psychologists through GPs (as being piloted in the "Better outcomes in mental health care" initiative) shows that funding mechanisms outside of Medicare are feasible.

More widespread implementation of CBT could potentially lead to cost offsets: (i) for the PBS due to a reduction in prescription of antidepressant drugs and (ii) for the health system in general due to a decrease in resource usage resulting from a reduction in relapse and severity of depression. These have not been considered in the analyses but would only have made the findings more favourable towards CBT.

The cost-effectiveness analyses assume steady state operation, i.e. assuming interventions are fully implemented and operate in accordance with its efficacy potential. In practise, resources will need to be allocated to implementation of the intervention, e.g. in administering a new delivery system and in ensuring adequate training and accreditation of providers.

Based on cost-effectiveness TCAs should be the drug of choice for depression unless not well tolerated. There is some worry about the toxicity of TCAs when taken in overdose but there is insufficient evidence that prescribing TCAs leads to higher suicide rates [42]. While the ICERs for

giving SSRIs to those currently not treated are still within range of what is considered 'affordable' in Australia, a comparison of changing from TCAs to SSRIs has very high ICERs (exceeding A$250,000 per DALY). Reverting to widespread use of TCAs is unlikely to find favour with clinicians and patients despite these findings. However, the cost mainly to government through the PBS is very high due to the large number of people with depression eligible for treatment.

Although directly observed economic data on the cost-effectiveness of psychotherapies is limited and based on short-term follow-up [43], modelling the available evidence provides very strong support for CBT under Australian health service conditions. The generalisability of this finding to other industrialised countries is largely dependent on the cost of delivering the interventions, as there are similarities in the epidemiology of depression between countries.

Our finding that TCAs are more cost-effective than SSRIs is contrary to the conclusions of published drug comparison studies. There are three reasons for this difference. First, we model no difference in impact between the drugs. In contrast, a study conducted by the non-profit Canadian Organisation for Health Technology Assessment [44] 'borrowed' utility weights from another study [45] giving greater utility while depressed on SSRIs than while depressed on TCAs despite acknowledging the equivalence between the two drugs in reducing symptoms of depression. Second, we use modest differentials in adherence rates as consistently reported from meta-analyses. Third, most drug comparison studies estimate cost offsets from reduced health service use that compensate for the higher cost of SSRIs. We decided not to include a change in hospital costs for depression in our analyses. For the analyses of intervention strategies for acute episodes and a continuation phase this would make little difference as we model the changes in the group currently not receiving evidence based care when consulting a health professional. In the 1997 survey, they reported an average number of 0.05 days in hospital which is so small that omission of these costs could not have a sizeable effect on the results. The average number of hospital days reported by all cases seeking any health care for

their depression in the SMHWB is 1.9 days. This is our target group for maintenance treatments. However, we have no trial data to determine what proportion of hospital costs can be saved. Moreover, evidence from elsewhere may not apply to the health care system we have in Australia. Therefore, we decide to ignore any impact on hospital costs. This may have led to some underestimation of the true cost savings associated with maintenance treatment.

Elsewhere, we discuss that translating trial findings into a health benefit in DALY terms is problematic [6]. Further developmental work is recommended such as the use of general quality of life outcome measures in trials and more sensitive disability weights in DALYs. Nevertheless, we argue that comparisons within depression are valid as the same imperfect methods are used to translate trial findings derived by similar means and the health benefit is largely driven by the effect size. Comparisons of these results with those for other mental or physical disorders are more affected by these methodological limitations.

Our estimates of the 20% average time with major depression over five years of follow-up is somewhat higher than reported from a clinical study in the US which found that 15% of time was spent with depressive symptoms at the level of major depression over nine years of follow-up [46]. Using the lower estimate would increase the cost-effectiveness ratios for maintenance treatment options by 38% without altering the ranking order of interventions by their cost-effectiveness ratios.

Our ability to model depression as a chronic episodic disorder rather than episode by episode allows the evaluation of longer-term treatment strategies. This is an important new development as we find such strategies to be cost-effective and potentially much more effective in curbing the burden of depression [5] even if we take ‘realistic’ adherence rates into account.

CAVEAT

The ACE–Mental Health project was jointly funded by the Australian Department of Health and Ageing, Mental Health and Suicide Prevention

Branch and the Department of Human Services, Mental Health Branch, Victoria in recognition of the importance of research into the cost-effectiveness of interventions in mental health treatment and care. This work draws upon, but is also limited by the available research and the assumptions necessary to complete the work.

The results of the analyses provide valuable material, likely to contribute to future policy deliberations by all service providers. Conclusions drawn from the economic evaluations should be considered within the context of the second stage filter process, which qualifies the results taking into account issues of equity, feasibility, strength of evidence, and acceptability to stakeholders. This second stage filter process addresses some of the practical considerations required for changes in actual service practice.

ACKNOWLEDGMENTS

Principal Investigators for the project are: A/Professor Theo Vos, A/Professor Rob Carter and Professor Gavin Andrews. We thank members of the ACE–Mental Health Steering Committee for their input into the project: Dr David Barton, Professor Graham Burrows (Chair), Sue Caleo, Professor Vaughan Carr, Dermot Casey, Joy Easton, Dr William Hart, Professor Helen Herrman, Barbara Hocking, Professor Assen Jablensky, Professor Anthony Jorm, Dr Lyn Littlefield, Professor Patrick McGorry, John McGrath, Dr Paul Morgan, Lorna Payne, Deb Podbury, Dr Kristy Sanderson, Suzy Saw, Professor Bruce Singh, Professor Bruce Tonge, Dr Ruth Vine, Professor Harvey Whiteford.

The unit record data for the Survey of Mental Health and Wellbeing were obtained from ABS and a revised scoring algorithm to determine ICD–10 and DSM–IV diagnoses was provided by Gavin Andrews. Data on current prescribing behaviour by GPs was obtained from the Bettering the Evaluation and Care of Health data collection system run by the Australian Institute of Health and Welfare and the General Practice Statistics and Classification Unit of the University of Sydney. The average cost of various types of medical attendances and the various forms of SSRI and TCA were

obtained from Medicare Benefits Schedule and Pharmaceutical Benefits Scheme data from the Department of Health and Ageing.

REFERENCES

1. Australian Bureau of Statistics. *Mental health and wellbeing: Profile of adults, Australia, 1997.* Canberra: ABS, 1998.
2. Mathers CD, Vos ET, Stevenson CE, Begg SJ. The Australian Burden of Disease Study: measuring the loss of health from diseases, injuries and risk factors. *Medical Journal of Australia* 2000; 172:592-596.
3. Ellis P. Australian and New Zealand clinical practice guidelines for the treatment of depression. *Aust N Z J Psychiatry* 2004; 38:389-407.
4. American Psychiatric Association. Practice guideline for the treatment of patients with major depressive disorder (revision). *American Journal of Psychiatry* 2000; 157:1-45.
5. Vos T, Haby MM, Barendregt JJ, Kruyshaar ME, Corry J, Andrews G. The burden of major depression avoidable by longer-term treatment strategies. *Archives of General Psychiatry* 2004; 61:1097-1103.
6. Haby MM, Carter R, Mihalopoulos C, Magnus A, Sanderson K, Andrews G, Vos T. Assessing Cost-Effectiveness - Mental Health: Introduction to the study and methods. *Australian and New Zealand Journal of Psychiatry* 2004; 38:569-578.
7. Harris EC, Barraclough B. Suicide as an outcome for mental disorders: A meta-analysis. *British Journal of Psychiatry* 1997; 170:205-228.
8. Isacsson G, Bergman U, Rich CL. Epidemiological data suggest antidepressants reduce suicide risk among depressives. *Journal of Affective Disorders* 1996; 41:1-8.
9. Trindade E, Menon D. *Selective serotonin reuptake inhibitors (SSRIs) for major depression. Part 1. Evaluation of the clinical literature.* Ontario, Canada: Canadian Coordinating Office for Health Technology Assessment, 1997.
10. Ackerson J, Scogin F, McKendree-Smith N, Lyman RD. Cognitive bibliotherapy for mild and moderate adolescent depressive symptomatology. *Journal of Consulting and Clinical Psychology* 1998; 66:685-690.
11. Brown RA, Lewinsohn PM. A psychoeducational approach to the treatment of depression: comparison of group, individual, and minimal contact procedures. *Journal of Consulting and Clinical Psychology* 1984; 52:774-783.

12. Jamison C, Scogin F. The outcome of cognitive bibliotherapy with depressed adults. *Journal of Consulting and Clinical Psychology* 1995; 63:644-650.
13. Selmi PM, Klein MH, Greist JH, Sorrell SP, Erdman HP. Computer-administered cognitive-behavioral therapy for depression. *American Journal of Psychiatry* 1990; 147:51-56.
14. Schmidt MM, Miller WR. Amount of therapist contact and outcome in a multidimensional treatment program. *Acta Psychiatrica Scandinavica* 1983; 67:319-332.
15. Scogin F, Hamblin D, Beutler L. Bibliotherapy for depressed older adults: a self-help alternative. *Gerontologist* 1987; 27:383-387.
16. Scogin F, Jamison C, Gochneaur K. Comparative efficacy of cognitive and behavioral bibliotherapy for mildly and moderately depressed older adults. *Journal of Consulting and Clinical Psychology* 1989; 57:403-407.
17. Scogin F, Jamison C, Davis N. Two-year follow-up of bibliotherapy for depression in older adults. *Journal of Consulting and Clinical Psychology* 1990; 58:665-667.
18. Wollersheim JP, Wilson GL. Group treatment of unipolar depression: a comparison of coping, supportive, bibliotherapy, and delayed treatment groups. *Professional Psychology: Research and Practice* 1991; 22:496-502.
19. Geddes J, Carney SM, Davies C, Furukawa TA, Kupfer D, Frank E, Goodwin GM. Relapse prevention with antidepressant drug treatment in depressive disorders: a systematic review. *The Lancet* 2003; 361:653-661.
20. Sanderson K, Andrews G, Corry J, Lapsley HM. Modelling change in preference values from descriptive health status using the effect size. *Quality of Life Research* 2004; 13:1255-1264.
21. Anderson IM. Selective serotonin reuptake inhibitors versus tricyclic antidepressants: a meta-analysis of efficacy and tolerability. *Journal of Affective Disorders* 2000; 58:19-36.
22. Antonuccio DO, Thomas M, Danton WG. A cost-effectiveness analysis of cognitive behavior therapy and fluoxetine (prozac) in the treatment of depression. *Behavior Therapy* 1997; 28:187-210.
23. Gater R, Goldberg D. Pathways to psychiatric care in South Manchester. *British Journal of Psychiatry* 1991; 159:90-96.
24. Stouthard MEA, Essink-Bot ML, Bonsel GJ, Barendregt JJ, Kramer PGN, van de Water HPA, Gunning-Schepers LJ, van der Maas PJ. *Disability weights for diseases in the Netherlands*. Rotterdam: Department of Public Health, Erasmus University, 1997.

25. Williams JW, Jr., Mulrow CD, Chiquette E, Noel PH, Aguilar C, Cornell J. A systematic review of newer pharmacotherapies for depression in adults: evidence report summary. *Annals of Internal Medicine* 2000; 132:743-756.
26. Geddes JR, Freemantle N, Mason J, Eccles MP, Boynton J. SSRIs versus other antidepressants for depressive disorder (Cochrane Review). In *The Cochrane Library*, Issue 1. Oxford: Update Software, 2000.
27. Gloaguen V, Cottraux J, Cucherat M, Blackburn I. A meta-analysis of the effects of cognitive therapy in depressed patients. *Journal of Affective Disorders* 1998; 49:59-72.
28. Evans MD, Hollon SD, DeRubeis RJ, Piasecki JM, Grove WM, Garvey MJ, Tuason VB. Differential relapse following cognitive therapy and pharmacotherapy for depression. *Archives of General Psychiatry* 1992; 49:802-808.
29. Shea MT, Elkin I, Imber SD, Sotsky SM, Watkins JT, Collins JF, Pilkonis PA, Beckham E, Glass DR, Dolan RT. Course of depressive symptoms over follow-up. Findings from the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Archives of General Psychiatry* 1992; 49:782-787.
30. Hollon SD, Evans MD, DeRubeis RJ. Cognitive mediation of relapse prevention following treatment for depression: Implications of differential risk. In *Contemporary psycholgical approaches to depression*. New York: Guilford Press, 1990.
31. Fava GA, Grandi S, SZielezny M, Rafanelli C, Canestrari R. Four-year outcome for cognitive behavoiural treatment of residual symptoms in major depression. *American Journal of Psychiatry* 1996; 153:945-947.
32. Paykel ES, Scott J, Teasdale JD, Johnson AL, Garland A, Moore R, Jenaway A, Cornwall PL, Hayhurst H, Abbott R, Pope M. Prevention of relapse in residual depression by cognitive therapy: a controlled trial. *Archives of General Psychiatry* 1999; 56:829-835.
33. Blackburn IM, Moore RG. Controlled acute and follow-up trial of cognitive therapy and pharmacotherapy in out-patients with recurrent depression. *British Journal of Psychiatry* 1997; 171:328-334.
34. Barbui C, Hotopf M, Freemantle N, Boynton J, Churchill R, Eccles MP, Geddes JR, Hardy R, Lewis G, Mason JM. Selective serotonin reuptake inhibitors versus tricyclic and heterocyclic antidepressants: comparison of drug adherence (Cochrane Review). In *The Cochrane Library*, Issue 1. Oxford: Update Software, 2001.
35. Simon GE, Von Korff M, Rutter CM, Peterson DA. Treatment process and outcomes for managed care patients receiving new antidepressant

prescriptions from psychiatrists and primary care physicians. *Archives of General Psychiatry* 2001; 58:395-401.

36. Lin EH, Von Korff M, Katon W, Bush T, Simon GE, Walker E, Robinson P. The role of the primary care physician in patients' adherence to antidepressant therapy. *Medical Care* 1995; 33:67-74.
37. Katon W, Rutter C, Ludman EJ, Von Korff M, Lin E, Simon G, Bush T, Walker E, Unutzer J. A randomized trial of relapse prevention of depression in primary care. *Archives of General Psychiatry* 2001; 58:241-247.
38. Katon W, Robinson P, Von Korff M, Lin E, Bush T, Ludman E, Simon G, Walker E. A multifaceted intervention to improve treatment of depression in primary care. *Archives of General Psychiatry* 1996; 53:924-932.
39. Oei TP, Kazmierczak T. Factors associated with dropout in a group cognitive behaviour therapy for mood disorders. *Behaviour Research and Therapy* 1997; 35:1025-1030.
40. Britt H, Miller G, Charles J, Knox S, Sayer G, Valenti L, Henderson J, Kelly Z. *General practice activity in Australia 1999–00. AIHW Cat. No. GEP 5*. Canberra: Australian Institute of Health and Welfare (General Practice Series no.5), 2000.
41. Heuzenroeder L, Donnelly M, Haby MM, Mihalopoulos C, Rossell R, Carter R, Andrews G, Vos T. Cost-effectiveness of psychological and pharmacological interventions for generalized anxiety disorder and panic disorder. *Australian and New Zealand Journal of Psychiatry* 2004; 38:602-12.
42. Jick SS, Dean AD, Jick H. Antidepressants and suicide. *British Medical Journal* 1995; 310:215-218.
43. Churchill R, Hunot V, Corney R, Knapp M, McGuire H, Tylee A. A systematic review of controlled trials of the effectiveness and cost-effectiveness of brief psychological treatments for depression. *Health Technology Assessment Reports* 2001; 5.
44. Drummond M, McGuire A. *Economic Evaluation in Health Care. Merging theory with practice*. Oxford: Oxford University Press, 2001.
45. Revicki DA, Wood M. Patient-assigned health state utilities for depression-related outcomes: differences by depression severity and antidepressant medications. *Journal of Affective Disorders* 1998; 48:25-36.
46. Judd LL, Akiskal HS, Maser JD, Zeller PJ, Endicott J, Coryell W, Paulus M, Kunovac JL, Leon AC, Mueller TI, Rice JA, Keller. MB. A prospective 12-year study of subsyndromal and syndromal depressive symptoms in unipolar major depressive disorders. *Archives of General Psychiatry* 1998; 55:694-700.

7

Assessing Cost-Effectiveness (ACE)–Mental Health: Helping policy makers prioritise and plan health services

Vos T, Haby MM, Magnus A, Mihalopoulos C, Andrews G, Carter R (2005). Assessing Cost-Effectiveness (ACE)–Mental Health: Helping policy makers prioritise and plan health services. Australian and New Zealand Journal of Psychiatry, 39:701–712.

ABSTRACT

Objective: We assessed, from a health sector perspective, options for change that could improve the efficiency of Australia's current mental health services by directing available resources towards 'best practice' cost-effective services

Method: We summarise cost-effectiveness results of a range of interventions for depression, schizophrenia, ADHD and anxiety disorders that have been presented in previous papers in this journal. Recommendations for change are formulated after taking into account 'second filter criteria' of equity, feasibility of implementing change, acceptability to stakeholders and the strength of the evidence. In addition, we estimate the impact on total expenditure if the recommended mental health interventions for depression and schizophrenia are to be implemented in Australia.

Results: There are cost-effective treatment options for mental disorders that are currently underutilised (e.g. cognitive-behavioural therapy for depression and anxiety, bibliotherapy for depression, family interventions for schizophrenia and clozapine for the worst course of schizophrenia). There are also less cost-effective treatments in current practice (e.g. widespread use of olanzapine and risperidone in the treatment of established schizophrenia and, within those atypicals, a preference for olanzapine over risperidone). Feasibility of funding mechanisms and training of staff are the main second filter issues for CBT and family interventions. Acceptability to various stakeholders is the main barrier to implementation of more cost-effective drug treatment regimens. More efficient drug intervention options identified for schizophrenia would cost A$68 Million less than current practice. These savings would more than cover the estimated A$36M annual cost of delivering family interventions to the 51% of people with schizophrenia whom we estimated to be eligible and this would lead to an estimated 12% improvement in their health status. Implementing recommended strategies for depression would cost A$121 Million annually for the 24% of people with depression who seek care currently but do not receive an evidence-based treatment.

Conclusions: Despite considerable methodological problems, a range of cost-effective and less cost-effective interventions for major mental disorders can be discerned. The biggest hurdle to implementation of more efficient mental health services is that this change would require reallocation of funds between interventions, between disorders and between service providers with different funding mechanisms.

CHAPTER 7

BACKGROUND

The 'ACE approach' to priority setting

Although there is nothing new about the task of making difficult choices in health care, policy-makers are now discussing the issue of priority setting with renewed interest. Three main reasons are given in a growing international literature:

- growing evidence that the deployment of current resources is far from optimal [1, 2];
- continued growth in health care expenditure, both in absolute terms and as a percentage of GDP [3, 4]; and
- the wish to avoid 'government failure' in resource allocation decisions following government intervention in health care to achieve social justice objectives or to correct for 'market failure' [5, 6].

Although the importance and need for priority setting is clearly established, the central question of how priority setting is to be achieved remains contested. There are a variety of approaches and models available, offered from a range of disciplines. There are models offered by behavioural scientists based on achieving consensus; by epidemiologists/ clinicians based on needs assessment; by philosophers based on notions of social justice; and of course, there are a variety of approaches offered by economists based on efficiency [7]. An important step in assessing these various approaches is to reflect on the question of what constitutes an 'ideal' approach to priority setting.

The logical starting point to address this question is to develop a checklist based on a clearly specified rationale. The literature provides few guidelines that focus on priority setting across multiple interventions (as opposed to evaluation of individual projects) or that take a suitably broad-based approach. Carter [7] advises that a checklist be based on four general considerations:

- guidance from a review of economic theory, both applied economic evaluation methods, as well as their theoretical foundations in normative economics;
- lessons from empirical experience with priority setting;
- guidance from a review of literature on the role of ethics and social justice in priority setting; and
- reflection on the pragmatic needs of decision-makers, evident from the empirical experience and key issues of setting and context.

A central issue is the extent to which priority setting approaches focus on 'technical analysis' or 'due process' for their legitimacy. The 'technical school' is characterised by a reliance on rational decision rules and the development of technical frameworks in which they are applied. This school has in large part been the preserve of health economists (pursuing the goal of efficiency) and epidemiologists/clinicians (pursuing evidence-based medicine guidelines and/or needs-based equity). In the technical school, decisions are based on applying the correct rules whether efficiency, effectiveness and/or equity focused. Provided one accepts these principles results should give guidance to decision-makers on how services should be ranked.

In contrast, advocates of the 'due process' school question the assumption that it is possible to devise 'rational' decision rules and see the technical approaches as based on a simplistic view of the health care system. For the due process school the task is less to refine the technical basis of decision-making than to construct a process that enables proper debate and discussion to occur. This does not mean implicit rationing, but instead a system whereby decisions are made explicitly and the reasoning behind specific judgements is clearly explained.

However, the debate between the two schools may be drawn too starkly in the literature. There is no inherent conflict between provision of information on the costs, outcomes and evidence base for different interventions, and strengthening of the processes for debating that information and arriving at judgements on priorities. The reality is that neither option alone is likely to fulfil the theoretical and practical

requirements of an ideal approach to explicit priority setting. Technical methods alone will never be able to deal with the complexity and contested nature of priority setting, but 'due process' should ideally utilise the sort of information on needs, efficacy/effectiveness, efficiency, and equity provided by technical approaches. Both elements need to be involved in any approach to priority setting that is seeking strong theoretical foundations and empirical validity. The 'ACE Approach' reflects our endeavours to develop such a joint approach. It involves close attention to technical rigour in its economic and epidemiological analyses, but also seeks to ensure due process by involving stakeholders in a steering committee; it takes into account broader considerations that are less amenable to quantification; and seeks consensus decisions after informed discussion.

On the technical side the ACE methodology applies the key economic concepts of 'opportunity cost'; 'marginal analysis' and a 'clear concept of benefit' [8, 9] using standardised evaluation methods clearly documented in an evaluation protocol. Undertaking the evaluations in this way as part of the priority setting exercise, addresses the reservations expressed by economists [10] about the simplistic use of league tables, where economic studies are assembled from the literature with little regard to differences in methods, context and setting. The key technical features of the ACE approach are:

- the rationale for the selection of interventions is clearly explained and consistently applied;
- the evaluation methods are standardised, documented and open to scrutiny;
- the setting and context is common to all interventions (i.e. target populations within Australia);
- Australian data have been used for demography, health system costs, disease incidence/prevalence, and the description of 'current practice';
- information is assembled by a multi-disciplinary research team, preparing briefing papers to a standardised format agreed by a steering committee with policy makers, experts and representatives of community organisations as members;

- a range of results is reported (not just point estimates) reflecting explicitly the uncertainty and sensitivity of cost, outcome and value estimates; and
- the incremental costs per Disability Adjusted Life Year (DALY) ratios are placed within a broader decision-making framework that includes 'strength of evidence', 'equity', 'feasibility' and 'acceptability', which we have called 'second stage filter' criteria.

The steering committee in ACE studies has an important role in achieving balance between the technical analyses and achieving due process. On the technical side members contribute in areas of their expertise and discuss issues of method and evidence. On the due process side, members ensure stakeholder interests and views are articulated; assist with 'value' judgement aspects of the analysis; assist with sensible interpretation of the technical analysis; and assist in ensuring transparency and appropriate use of relevant data.

'ACE' applied to mental health

Previously, the ACE approach has been applied to cancer and heart disease [11, 12] and further studies have commenced on obesity and prevention of non-communicable disease in general. The choice of mental health reflected a number of considerations. First, mental disorders are the largest contributor to the non-fatal burden of disease in Australia and rank third behind cardiovascular disease and cancer in terms of overall burden of disease [13, 14]. Second, there are efficacious interventions for the main mental disorders [15] but, until now, there was little evidence of their cost-effectiveness under routine Australian health service conditions. The aim of the ACE–Mental Health study was to assess from a health sector perspective, whether there are options for change that could improve the effectiveness and efficiency of Australia's current mental health services by directing available resources towards cost-effective services. The overall methods and results of intervention-specific analyses have been published in a series of papers in this journal [16-23].

CHAPTER 7

In this final paper we present an overview of the results and conclusions from the study. We add one new analytical component by estimating the impact of the recommended options for change on overall expenditure for depression and schizophrenia which have been the main focus of our analyses. The first step in doing this was to estimate the annual cost of current practice using the same assumptions of health service use, drug costs and adherence as in the cost-effectiveness models for depression and schizophrenia for the baseline year 2000. This 'bottom-up' way of assembling total costs based on unit costs for the various elements of an intervention follows usual practice in economic analysis [24].Thereafter we estimate the change in annual expenditure that would occur if our major recommendations for a shift towards more cost-effective practice were implemented.

OVERVIEW OF RESULTS

The cost-effectiveness results

The results of the economic evaluations conducted as part of the ACE–Mental Health study can be summarised in a 'league table' ranking the interventions in order of their economic merit (Table 1). This is possible because the methodology is sound, as consistent as possible across the various studies (with any variations clearly specified) and applied to a common setting [18].

The results of the ACE–Mental Health study show that there are cost-effective treatment options for mental disorders that are currently under-utilised (e.g. cognitive behavioural therapy for depression and anxiety, bibliotherapy for depression, family interventions for schizophrenia and clozapine for the worst course of schizophrenia). There are also less cost-effective treatments in current practice (e.g. use of olanzapine and risperidone in the treatment of established schizophrenia and, within those atypicals, a preference for olanzapine over risperidone).

Table 1. Interventions analysed in ACE–Mental Health: cost-effectiveness and second filter issues that require attention.

Disorder and Intervention	Cost-effectiveness ratio in AUD per DALY (95% uncertainty range)	Second filter issues that require attention
Schizophrenia		**Drugs:**
Behavioural intervention in families (Tarrier)	8,000 (4,000 – 18,000)	• Quality of the evidence: trials suffer from absence of patient rated quality of life measures, high attrition rates, short duration, frequent use of excessively high dose levels of the typical comparators (thus biasing the results towards a better outcome for risperidone and olanzapine) • Return to the widespread use of typicals not likely to happen • The apparent tolerability advantages of olanzapine and risperidone in the short term make them more acceptable to patients and clinicians than typical neuroleptics
Risperidone vs typicals mix for people with troublesome side-effects	20,000 (15,000 – 27,000)	
Multiple family groups (McFarlane)	21,000 (10,000 – 53,000)	
Clozapine vs typicals for those with chronic disease and clear deterioration	23,000 (17,000 – 47,000)	
Behavioural family management (Falloon)	28,000 (13,000 – 64,000)	
Olanzapine (15mg/day) vs typicals mix for people with troublesome side-effects	38,000 (28,000 – 50,000)	**Family interventions:** • Quality of evidence: i) mostly small trials, ii) only relapse rate reported, not continuous measures of symptoms / quality of life • Rural and remote access • Implementation of Behavioural Interventions in Families is more feasible due to the short nature (9 months) of this intervention • Appropriateness for NESB and indigenous families not known • Acceptability to clinicians – although effectiveness known for many years, intervention not widely provided
Clozapine vs typicals for those with chronic disease and little deterioration	42,000 (31,000 – 62,000)	
Risperidone vs typicals	48,000 (27,000 – 110,000)	
Risperidone vs low dose typicals	80,000 (36,000 – infinity)	
Olanzapine (15mg/day) vs typicals mix	92,000 (53,000 – 170,000)	
Olanzapine (15mg/day) vs risperidone	160,000 (44,000 – infinity)	

Disorder and Intervention	Cost-effectiveness ratio in AUD per DALY (95% uncertainty range)	Second filter issues that require attention
Generalised Anxiety Disorder		**CBT:**
CBT by psychologist (public)	6,900 (4,000 – 12,000)	• Generalisability of results to minority groups and all providers • Equity in access if 'user pays', access in rural/remote and outer metropolitan areas • Ensuring an adequate workforce • Development of implementation arrangements • Acceptability to stakeholders
CBT by psychologist (private)	20,000 (12,000 – 33,000)	
CBT by psychiatrist (private)	23,000 (15,000 – 38,000)	
CBT by psychiatrist (public)	23,000 (14,000 – 38,000)	
Venlafaxine	23,000 (16,000 – 40,000)	**Venlafaxine:** • Not available on the PBS for GAD • Consumer concern about side-effects
Panic Disorder		**CBT:**
CBT by psychologist (public)	6,800 (2,900 – 14,000)	• Same as for Generalised Anxiety Disorder
TCAs	17,000 (9,700 – 42,000)	**TCAs and SSRIs:**
CBT by psychologist (private)	26,000 (15,000 – 45,000)	• Antidepressants less effective than CBT • Consumer concern about side-effects • Reluctance by clinicians to prescribe TCAs due to the perception of greater side-effects
CBT by psychiatrist (private)	27,000 (19,000 – 48,000)	
CBT by psychiatrist (public)	30,000 (18,000 – 55,000)	
SSRIs	38,000 (27,000 – 89,000)	

Disorder and Intervention	Cost-effectiveness ratio in AUD per DALY (95% uncertainty range)	Second filter issues that require attention
Major Depression (adults)		**CBT:**
Group CBT by psychologist (public) – ***maintenance*** treatment	Dominant	• Same as for Generalised Anxiety Disorder • Long-term treatment more effective and cost-effective than treatment of the acute episode **Bibliotherapy:** • Only suitable for some patients **TCAs and SSRIs:** • Antidepressants less effective than CBT, particularly for longer-term treatment strategies • Consumer concern about side-effects, particularly for long-term treatment • Reluctance by clinicians to prescribe TCAs due to the perception of greater side-effects • Reluctance by policy makers to endorse long-term treatment due to cost pressures on PBS
CBT by psychologist (public) – ***maintenance*** treatment	Dominant (dominant – 230)	
Bibliotherapy	160 (dominant – 460)	
Group CBT by psychologist (public) – ***acute episode***	1,300 (600 – 2,300)	
CBT by psychologist (private) – ***maintenance*** treatment	2,200 (1,700 – 2,800)	
CBT by psychiatrist (public) – ***maintenance*** treatment	2,800 (2,000 – 3,800)	
CBT by psychiatrist (private) – ***maintenance*** treatment	2,900 (2,600 – 3,400)	
CBT by psychologist (public) – ***acute episode***	4,100 (2,700 – 6,200)	
TCAs – ***maintenance*** treatment	4,500 (4,100 – 4,900)	
TCAs – ***acute episode***	5,700 (4,400 – 7,500)	
CBT by psychologist (private) – ***acute episode***	10,000 (7,600 – 14,000)	
CBT by psychiatrist (public) – ***acute episode***	12,000 (8,300 – 16,000)	
CBT by psychiatrist (private) – ***acute episode***	12,000 (9,200 – 16,000)	
SSRIs for ***acute episode***	14,000 (11,000 – 19,000)	
SSRIs for ***maintenance*** treatment	18,000 (17,000 – 19,000)	

Disorder and Intervention	Cost-effectiveness ratio in AUD per DALY (95% uncertainty range)	Second filter issues that require attention
Major depression (children and adolescents)		
CBT by psychologist (public)	9,000 (39,000 – 24,000)	**CBT:**
SSRIs – 1st line treatment	23,000 (13,000 – 53,000)	• Same as for Generalised Anxiety Disorder
SSRIs – 2nd line treatment	23,000 (13,000 – 54,000)	• Only 1 trial in children < 13 years
CBT by psychologist (private)	28,000 (13,000 – 70,000)	**SSRIs:**
CBT by psychiatrist (public)	34,000 (16,000 – 82,000)	• Parental concerns about using drugs in children and adolescents
CBT by psychiatrist (private)	34,000 (16,000 – 82,000)	• Worries about increased suicide risk associated with SSRIs • Less effective than CBT • Ethical concerns about drugs (which have side effects) as the first line treatment while more effective treatment (i.e. CBT) is available with no side effects
Attention Deficit Hyperactivity Disorder		
Dexamphetamine	4,100 (dominant - 14000)	• Short duration of trials (average 4 weeks)
Methylphenidate	15,000 (9,100 – 22,000)	• Access to specialist treatment in remote or rural areas • Methylphenidate not available on the PBS – which is an issue for those children who do not respond to dexamphetamine • Concern about over-prescription

AUD – Australian dollars; CBT – cognitive behavioural therapy (individual unless otherwise specified); GAD – Generalised Anxiety Disorder; DALY – disability adjusted life year; NESB – non-English speaking background; PBS – pharmaceutical benefits scheme; SSRIs – selective serotonin reuptake inhibitors; TCAs – tricyclic antidepressants

Impact on annual expenditure for schizophrenia and depression

We estimate A$119M was spent in 2000 on pharmaceuticals for schizophrenia based on the proportions of people with schizophrenia using different types of neuroleptic drugs from the 1998 Low Prevalence Disorder Study and adjusting for the change in volume of scripts for each drug under the Pharmaceutical Benefits Scheme since 1998. This estimate is close to the A$110M reported by the Australian Institute of Health and Welfare (AIHW) as total expenditure on pharmaceuticals for schizophrenia for the financial year 2000/2001 [25] using a method referred to as 'top-down' costing or 'gross-costing' [24]. These estimates include costs to government as well as patient contributions.

Changing treatment in those currently on olanzapine to risperidone would result in A$27M cost saving. Implementation of all our recommended drug treatment strategies [21], i.e. giving (i) clozapine to the 15% of patients with the worst course of schizophrenia; (ii) risperidone to those who experience moderate to severe side effects on typicals; and (iii) low dose typicals to all others, would cost A$42M or a cost saving compared to current practice of A$68M with virtually no change in health status. These savings would more than cover the estimated A$36M annual cost of delivering family interventions to the 51% of people with schizophrenia whom we estimated to be eligible and this would lead to an estimated 12% improvement in their health status [22].

We estimate A$153M was spent in 2000 on out-of-hospital medical services (largely general practitioner and specialist care) and A$190M on pharmaceuticals for depression taking into account a change in prescribing patterns since 1997 (the year of the National Survey of Mental Health and Wellbeing [26]) as indicated by the change in volume of scripts for each drug under the Pharmaceutical Benefits Scheme. These amounts are considerably lower than the A$276M for out-of-hospital medical services and A$340M for pharmaceuticals AIHW attributed to depression in 2000/01 [25]. This difference in cost estimates can be explained by two reasons. First, our estimates are based on a cross-sectional survey asking respondents about

health service use in the last 12 months and a sizeable proportion of respondents identified with depression would not yet have completed their episode Thus we underestimate the true costs of health care for depression. Second, it is not easy to apportion total expenditure on anti-depressant drugs to individual disorders as they are also widely used for other disorders, such as anxiety disorders, dysthymia and subsyndromal depression not included in our estimates. This would lead to overestimation in the AIHW costs for depression.

Our cost-effectiveness analyses of anti-depressant drugs and CBT targeted the 24% of people with depression who seek care from a health professional but do not receive an evidence-based treatment. We estimate that in 2000, these people accrued A$ 27M out-of-hospital medical costs. Providing them with the same mix of therapy and/or drugs by the same mix of providers as those currently receiving evidence-based care the annual cost would be at least A$172M based on our low estimate of total costs. This amount would reduce to A$121M if: (i) GPs supervise drug treatment; (ii) psychologists in public service provide CBT; and (iii) prescribed anti-depressants were equally split between tricyclic antidepressants and newer drugs such as selective serotonin-reuptake inhibitors (rather than the 36% – 61% split recorded in 2000-01 which has become a 81% – 17% split in 2003/04). All these estimates are considerably higher than the current expenditure on non-evidence based treatments and would be higher still if we take into account that some of the difference between our total cost estimates and AIHW estimates for depression is due to the fact that we underestimate costs in respondents who had not yet completed treatment at the time of measurement in the cross-sectional survey. Moreover, it is not clear how much of current expenditure on non-evidence based treatment would cease if evidence-based treatment were offered instead.

The total amount of expenditure estimated for funding evidence-based care in just a proportion of people with depression exceeds expenditure on medical and pharmaceutical services for people with schizophrenia because of the much higher prevalence of depression. While our recommendations for depression are costly, it would result in a 20% reduction in disease

burden for treatment of depressive episodes including six months continuation of drug treatment after the episode. Providing all those seeking care for their depression in the baseline year 2000 with drug maintenance treatment following an episode would cost A$312M over five years and A$160M if maintenance CBT were offered instead. Both types of maintenance treatment are estimated to reduce the disease burden of depression by around 50% [27]

A shift towards our recommended strategy to improve health services for people with depression and schizophrenia is not easy because it relies on: (i) changing the practice of clinicians; (ii) flexibility in financing mental health care between different types of health care and between different health care funders; and (iii) addressing opinions of patients, care givers, health care workers and advocacy groups about the relative merits of different treatments in an environment influenced by intensive marketing efforts of pharmaceutical companies. We have tried to address these issues by shaping our recommendations in discussion with the steering committee taking into account four 'second stage filter' criteria.

The second stage filters

To date, ACE steering committees have endorsed the value judgement that in the health sector, it is appropriate to maximise 'health' (defined using DALYs), subject to acceptable performance on other criteria (such as the health gain being fairly distributed). While there are different techniques for taking such broader considerations into account, steering committees have so far chosen to adopt what we have called the 'second stage filter' process. Under this process the broader considerations are selected, clearly defined and presented as qualifying statements together with the evidence of cost-effectiveness for each intervention [18].

Second stage filter issues in ACE–Mental Health that require attention for each of the interventions are summarised in Table 1. For example, provision of CBT to all eligible patients with depression and anxiety requires the development of implementation arrangements. Resources would need to be committed to ensuring providers have adequate training and are available

in sufficient numbers and this would be a greater challenge in rural and outer metropolitan areas. Also, funding mechanisms for psychological care would need to be extended to ensure equitable access in primary care. For schizophrenia, major issues are the acceptability of the side effects of typical neuroleptics to various stakeholders and overcoming the inertia in mental health services failing to implement family therapy despite two decades of evidence of efficacy.

Selection of the interventions

In addition to the interventions included in Table 1, five other interventions were selected for analysis by the steering committee. However, we found insufficient evidence of efficacy to justify a full cost-effectiveness analysis. First, we found limited evidence for specialist versus generalist services for eating disorders. Two randomised controlled trials of in-patient versus out-patient care [28-30] indicated no clear advantage of either.

Second, integrated treatment for clients with the dual diagnosis of severe mental illness and substance abuse showed a greater availability of evidence at the level of controlled trials, including 7 randomised and 6 non-randomised controlled trials, summarised in one systematic review [31] and one narrative review of the literature [32]. However, these studies gave no clear support for integrated programs over standard care due to the lack of differential impact on health-related quality of life.

Third, a review of the evidence for assertive community treatment (ACT) of psychoses including 30 studies, 25 of which are randomised controlled trials found no clear health benefit with ACT nor sufficient evidence to suggest overall health system cost savings associated with ACT. However, greater satisfaction was reported by clients of ACT and this may make ACT a preferred option as has been reported previously by advocates [33].

Fourth, a cost-comparison analysis was conducted for supported employment in schizophrenia due to the lack of evidence for a clinical benefit [16].This analysis showed that, largely due to disincentives in welfare legislation to taking up full-time employment, supported

employment costs more than it saves from both a patient and government perspective.

Finally, we investigated a stand-alone early psychosis intervention and enhanced standard care for early psychosis versus standard care. The evidence for the effectiveness of the two interventions from a non-randomised controlled trial was inconclusive, thus preventing a full cost-effectiveness analysis. However, a comparison of average first year costs showed that the two interventions were not significantly more expensive than standard care [34].

DISCUSSION

Despite considerable uncertainty around key input variables, clear distinctions in cost-effectiveness between mental health interventions (particularly within disorders) are apparent. The results suggest that substantial opportunities exist to improve efficiency within our current mental health resources, if resources were shifted towards more cost-effective interventions. We have also identified a significant amount of under-treatment, particularly for depression and anxiety, which would require considerable additional resources but which would lead to sizeable improvements in health outcomes. A previous study found similar order of magnitude findings of the cost-effectiveness of interventions for schizophrenia and depression and draws the same conclusions about the potential for greater efficiency in mental health services [35-37]. Results between the two studies are not directly comparable because of some differences in methodology (we use a longer time horizon, include mortality outcomes and have not included hospital costs) and presentation of results (we present cost-effectiveness ratios for individual interventions while the other study presents the cost-effectiveness of the current and an optimal mix of interventions per disorder). However, the main underlying data sources describing the epidemiology of mental disorders in Australia and the effectiveness of interventions are the same and hence the similarity of results is not surprising.

The brief analysis of the impact of our recommendations for more efficient services for depression and schizophrenia shows the value of combining information on cost-effectiveness with the size of health problems, thus refuting the criticism of some health economists that burden of disease studies are a waste of valuable resources [38-40]. Clearly, the quantification of the potential impact on mental health budgets of implementing change is essential to policy makers and this requires basic burden of disease information on prevalence, incidence and duration of disease.

However, the cost-effectiveness results of this study need to be considered in the context of the second stage filter criteria and the limitations of data and methods.

Limitations in the selection of disorders and interventions

The ACE–Mental Health study is limited to interventions for major depression, panic disorder, generalised anxiety disorder, schizophrenia and related conditions, and ADHD. Clearly some important mental disorders with significant health burden were not included, such as substance use disorders, bipolar disorder, dysthymia, borderline personality disorder, eating disorders, post-traumatic stress disorder, social phobia, obsessive-compulsive disorder and conduct disorder. Although the steering committee expressed interest in these disorders, given the two-year time constraint and limited project budget it was simply not possible to cover all mental disorders. Preference was given to disorders for which the disease burden was greatest, current practice could be quantified and for which evidence-based interventions could be specified. Thus, the ACE–Mental Health study did not intend to provide a comprehensive strategic plan for mental health, but rather to suggest options for directing available resources towards cost-effective services.

Within the disorders included in the study, not all possible interventions were analysed. Important omissions include physical exercise [41] and electroconvulsive treatment [42] for depression, cognitive

behavioural therapy for schizophrenia [43] and benzodiazepines for anxiety disorders.

We failed to complete full analyses for a number of interventions with limited and inconclusive evidence of effectiveness. This raises the point of selection bias in economic analyses towards interventions with a stronger evidence base and often these include pharmacological and surgical interventions driven by commercial interests.

An important simplification made in the ACE–Mental Health study was to ignore comorbidity with other physical and mental disorders because evidence of efficacy for interventions for comorbid disorders is lacking. Randomised controlled trials (RCTs) often exclude subjects with comorbid conditions and/or do not present the results stratified by comorbidity. This raises doubts about the applicability of trial findings to patients in routine health services as comorbidity between mental disorders is very common

Limitations in the cost per DALY results

As mentioned in the individual papers, there were limitations in the methods used to calculate the cost per DALY, the most significant being the measurement of health benefits. While the measurement of the effect size is based on accepted methods [44, 45], the translation into a change in the DALY disability weight required the development of new methods [18]. The absence of an established method of translating effect size into health gain in DALY units is further compounded by the lack of health-related quality of life outcome data from trials and survey data and the availability of only two to four DALY disability weights, which coarsely describe severity for each mental disorder. However, it is important to recognise that the effect size is usually the main driver of the cost-effectiveness ratio and that, within disorders, the method for translating the effect size into a change in the DALY disability weight is the same. Thus, comparisons within disorders have greatest validity but it is more problematic to compare results between mental disorders. Furthermore, comparison of mental health interventions where health gains are dominated by non-fatal outcomes, with interventions for heart disease and cancer where extension of life dominates health

outcomes, are less certain even though the studies have been conducted with comparable methods [12, 46].

Second stage filters process

The second stage filter process involved the assessment of issues that either influence the degree of confidence that can be placed in the cost-effectiveness ratios (such as the level of available evidence), or broader issues that need to be taken into account in decision-making about resource allocation (such as equity and acceptability to stakeholders). The second stage filter process was facilitated by the involvement of policy-makers in the steering committee. This helped ensure that the research was conducted from within a policy context, thus having 'real-life' applicability. However, this also challenged specific interests. For example, the recommendation that patients with established schizophrenia be switched to low dose typicals (rather than olanzapine or risperidone) goes against current practice and was not supported by all members of the steering committee. It would also not be supported by pharmaceutical companies which are gradually withdrawing typicals from the market and continuing to introduce newer and higher cost atypicals. Thus, for some interventions and for some criteria, the steering committee could not reach complete consensus. Where this was an issue, the researchers were guided by the weight of evidence and by the majority view.

A significant limitation of the process was the limited consumer input into the project. The Steering Committee included a representative from SANE Australia (a community advocacy organisation) and a representative from the Mental Health Council but no consumers. Thus the view of consumers is incorporated indirectly only, from data presented in trials on side-effects and drop-outs, from consumer perceptions recorded in survey data [26, 47] and from the perception of the mental health experts and representatives from community organisations.

Overall, the cost-effectiveness analyses conducted in the ACE–Mental Health study are all based on Level I evidence of efficacy, i.e. systematic reviews of all relevant randomised controlled trials [48]. The number of

patients/trials included for each intervention was good for CBT and drug interventions but modest for family interventions for schizophrenia. An issue relevant for all intervention trials is the lack of patient-rated quality of life measures. The quality of trials was reasonable, but there are worries about the validity of some of the trial results. This was a particular problem with drug interventions for schizophrenia, which are mostly funded by drug companies and suffer from multiple reporting, which is not always obvious to the reader [49, 50]. A more insidious issue is the use of high dose levels of the typical comparators, which bias the results to a better outcome for atypicals [51]. Not only does this inflate the apparent benefit of atypicals on extrapyramidal side effects, it may also inflate the apparent symptom benefits. In the analysis of the risperidone versus haloperidol RCTs from a Cochrane systematic review [52] and supplemented with a more recent RCT [53], we found an effect size (based predominantly on symptom measures) of 0.35 (95% CI 0.08 to 0.62) for trials in which risperidone was compared to 12mg or greater of haloperidol and a lower, non-significant effect size of 0.13 (95% CI -0.11 to 0.36) for trials which used mean haloperidol doses of less than 12mg. This finding is consistent with that found for atypicals in general in the systematic review by Geddes et al. [54].

As well as the second filter criteria of 'strength of evidence', 'equity', 'feasibility' and 'acceptability', the steering committee proposed a filter based on 'severity of the disorder'. Such a criterion would have to apply similarly to all interventions for the same disorder unlike the other second filter criteria which are intervention-specific. As the greatest validity of comparisons is for intervention options addressing the same disorder we decided not to incorporate severity as an additional second filter. However, while conventional economic analysis does not accord any weight to the severity of the illness per se, there is evidence [55-58] that people often do wish to give greater priority to treatment of those who are worse off, above and beyond the priority it already has within a cost-effectiveness analysis as captured by the incremental health gain. For example, given the greater disability associated with schizophrenia, policy-makers may decide to accept a higher willingness-to-pay threshold for interventions for schizophrenia.

Alternatively, greater weight may be given to minor health gains associated with severe conditions, thus increasing the health benefit associated with any particular effect size change. In fact, there is evidence to suggest that even when more severely ill patients are clearly receiving less benefit than patients with less severe illnesses, there remains a distinct preference for treating the more severe illnesses [59].

What next?

The results of these cost-effectiveness analyses provide valuable material likely to contribute to future policy deliberations by all service providers in mental health. The recommendations should not be seen as 'one size fits all'. Not all patients respond to any one treatment and patients (and doctors) have their own preferences for treatment, which will inevitably impact on its effectiveness. For example, some people with depression may prefer antidepressants while others prefer psychological approaches to treatment. Further, CBT may not be appropriate for all people who prefer psychological therapies. Bibliotherapy (a self-help form of CBT) is also cost-effective and suitable for some patients [23]. Interpersonal therapy may also be an efficacious treatment [60] that may better suit some patients. A combination of antidepressant and psychological therapies may be appropriate for others. Thus, the results of the study should not be seen as prescriptive. However, current practice patterns in mental health services indicate that there are inadequate incentives to promote efficiency. The intention of the ACE-Mental Health project was to raise awareness about these inefficiencies. We realise that presenting these results and recommendations will not automatically translate into changed practice. With our discussion of second filter criteria we have identified potential barriers to implementing change.

Policy could benefit from a more comprehensive analysis of mental health care including analyses of the efficiency of services for other mental disorders, for prevention and for high-cost services (e.g. hospital care), which we have not included in this study. The value of these analyses will be further enhanced by comparison with interventions in other areas of health if the same methods are used. That will indicate if there are cost-effectiveness

arguments to change the share of total health funding dedicated to mental health services.

Caveat

The ACE–Mental Health project was jointly funded by the Australian Department of Health and Ageing, Mental Health and Suicide Prevention Branch and the Department of Human Services, Mental Health Branch, Victoria in recognition of the importance of research into the cost-effectiveness of interventions in mental health treatment and care. This work draws upon, but is also limited by the available research and the assumptions necessary to complete the work.

The results of the analyses provide valuable material, likely to contribute to future policy deliberations by all service providers. Conclusions drawn from the economic evaluations should be considered within the context of the second stage filter process, which qualifies the results taking into account issues of equity, feasibility, strength of evidence, and acceptability to stakeholders. This second stage filter process addresses some of the practical considerations required for changes in actual service practice.

Acknowledgements

Principal Investigators for the project are: A/Professor Theo Vos, A/Professor Rob Carter and Professor Gavin Andrews. We thank members of the ACE–Mental Health Steering Committee for their input into the project: Dr David Barton, Professor Graham Burrows (Chair), Sue Caleo, Professor Vaughan Carr, Dermot Casey, Joy Easton, Dr William Hart, Professor Helen Herrman, Barbara Hocking, Professor Assen Jablensky, Professor Anthony Jorm, Dr Lyn Littlefield, Professor Patrick McGorry, John McGrath, Dr Paul Morgan, Lorna Payne, Deb Podbury, Dr Kristy Sanderson, Suzy Saw, Professor Bruce Singh, Professor Bruce Tonge, Dr Ruth Vine, Professor Harvey Whiteford. We also thank the other researchers who have worked on the project: Justine Corry, Ruth Rossell, Maturot Chalamat, Marie Donnelly and Louise Heuzenroeder.

REFERENCES

1. McPherson K. International differences in medical care practice. In *Health Care Systems in Transition: The Search for Efficiency*. Paris: OECD, 1990.
2. Richardson J, Robertson I. *Variation in procedure rates across statistical local areas in Victoria*. Melbourne: Centre for Health Program Evaluation, 1998.
3. Richardson J. The health care financing debate. In *Economics and Australian Health Policy*. Sydney: Allen & Unwin, 1998.
4. Klein R, Day P. *Managing scarcity: priority setting and rationing of medical services*. New York: Open University Press, 1996.
5. Hurley J. An Overview of the Normative Foundations of the Health Sector. In *Handbook of Health Economics*. Amsterdam: North-Holland, 2000.
6. Rice T. *The Economics of Health Reconsidered*. Chicago: Health Administration Press, 2003.
7. Carter R: The Macro Economic Evaluation Model (MEEM): an approach to priority setting in the health sector, in Faculty of Business and Economics. Melbourne, Monash University, 2001
8. Drummond MF, O'Brien B, Stoddart GL, Torrance GT. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford Medical Publications, 1997.
9. Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-effectiveness in health and medicine*. New York: Oxford University Press, 1996.
10. Drummond MF, Torrance GW, Mason JM. Cost-effectiveness league tables: More harm than good? *Social Science and Medicine* 1993; 37:33-39.
11. Cancer Strategies Group. *Priorities for Action in Cancer Control: 2001-2003.* Canberra: Commonwealth of Australia, 2001.
12. Liew D, Lim SS, Bertram M, McNeil JJ, Vos T. A model for undertaking effectiveness and cost-effectiveness analyses of primary preventive strategies in cardiovascular disease. *European Journal of Cardiovascular Prevention and Rehabilitation* (submitted).
13. Mathers CD, Vos T, Stevenson C. *The burden of disease and injury in Australia.* Canberra: Australian Institute of Health and Welfare, 1999.
14. Vos T, Begg S. *The Victorian Burden of Disease Study: morbidity*. Melbourne: Public Health and Development Division, Department of Human Services, 1999.
15. Nathan PE, Gorman JM. *A guide to treatments that work*. New York: Oxford University Press, 1998.

16. Chalamat M, Mihalopoulos C, Carter R, Vos T. Supported employment for schizophrenia. *Australian and New Zealand Journal of Psychiatry* (submitted).
17. Donnelly M, Haby MM, Carter R, Andrews G, Vos T. Cost-effectiveness of dexamphetamine and methylphenidate for the treatment of childhood attention deficit hyperactivity disorder. *Australian and New Zealand Journal of Psychiatry* 2004; 38:592-601.
18. Haby MM, Carter R, Mihalopoulos C, Magnus A, Sanderson K, Andrews G, Vos T. Assessing Cost-Effectiveness - Mental Health: Introduction to the study and methods. *Australian and New Zealand Journal of Psychiatry* 2004; 38:569-578.
19. Haby MM, Tonge B, Littlefield L, Carter R, Vos T. Cost-effectiveness of cognitive behavioural therapy and selective serotonin reuptake inhibitors for major depression in children and adolescents. *Australian and New Zealand Journal of Psychiatry* 2004; 38:579-91.
20. Heuzenroeder L, Donnelly M, Haby MM, Mihalopoulos C, Rossell R, Carter R, Andrews G, Vos T. Cost-effectiveness of psychological and pharmacological interventions for generalized anxiety disorder and panic disorder. *Australian and New Zealand Journal of Psychiatry* 2004; 38:602-12.
21. Magnus A, Carr V, Mihalopoulos C, Carter R, Vos T. Assessing cost-effectiveness of drug interventions for schizophrenia. *Australian and New Zealand Journal of Psychiatry* (accepted).
22. Mihalopoulos C, Magnus A, Carter R, Vos T. Assessing cost-effectiveness in mental health: family interventions for schizophrenia and related conditions. *Australian and New Zealand Journal of Psychiatry* 2004; 38:511-9.
23. Vos T, Corry J, Haby MM, Carter R, Andrews G. Cost-effectiveness of CBT and drug interventions for episodes of major depression. *Australian and New Zealand Journal of Psychiatry* (submitted).
24. Brouwer W, Rutten F, Koopmanschap M. Costing in economic evaluations. In *Economic evaluation in health care*. Oxford: Oxford University Press, 2001.
25. Australian Institute of Health and Welfare. *Health system expenditure on disease and injury in Australia, 2000-01*. Canberra: AIHW (Health and Welfare Expenditure Series no. 19), 2004.
26. Australian Bureau of Statistics. *Mental health and wellbeing: Profile of adults, Australia, 1997*. Canberra: ABS, 1998.
27. Vos T, Haby MM, Barendregt JJ, Kruyshaar ME, Corry J, Andrews G. The burden of major depression avoidable by longer-term treatment strategies. *Archives of General Psychiatry* 2004; 61:1097-1103.

28. Freeman C. Day patient treatment for anorexia nervosa. *British Review of Bulimia and Anorexia Nervosa* 1992; 6:3-8.
29. Gowers S, Norton K, Halek C, Crisp AH. Outcome of outpatient psychotherapy in a random allocation treatment study of anorexia nervosa. *International Journal of Eating Disorders* 1994; 15:165-77.
30. Crisp AH, Norton K, Gowers S, Halek C, Bowyer C, Yeldham D, Levett G, Bhat A. A controlled study of the effect of therapies aimed at adolescent and family psychopathology in anorexia nervosa. *British Journal of Psychiatry* 1991; 159:325-33.
31. Ley A, Jeffery DP, McLaren S, Siegfried N. Treatment programmes for people with both severe mental illness and substance misuse. *Cochrane Database of Systematic Reviews [computer file]* 2000:CD001088.
32. Drake RE, Mercer-McFadden C, Mueser KT, McHugo GJ, Bond GR. Review of integrated mental health and substance abuse treatment for patients with dual disorders. *Schizophrenia Bulletin* 1998; 24:589-608.
33. Mihalopoulos C, Carter R, Vos T. *Assertive Community Treatment (ACT): The evidence re-visited from an economic perspective*. Melbourne: Department of Human Services, 2003.
34. Mihalopoulos C. *ACE-Mental Health: Early psychosis interventions*. Melbourne: Department of Human Services, 2002.
35. Sanderson K, Andrews G, Corry J, Lapsley H. Reducing the burden of affective disorders: is evidence-based health care affordable? *Journal of Affective Disorders* 2003; 77:109-25.
36. Andrews G, Issakidis C, Sanderson K, Corry J, Lapsley H. Utilising survey data to inform public policy: comparison of the cost-effectiveness of treatment of ten mental disorders. *British Journal of Psychiatry* 2004; 184:526-33.
37. Andrews G, Sanderson K, Corry J, Issakidis C, Lapsley H. Cost-effectiveness of current and optimal treatment for schizophrenia. *British Journal of Psychiatry* 2003; 183:427-35.
38. Mooney G, Irwig L, Leeder S. Priority setting in health care: unburdening from the burden of disease. *Australian and New Zealand Journal of Public Health* 1997; 21:680-1.
39. Mooney G, Wiseman V. Burden of disease and priority setting. *Health Economics* 2000; 9:369-72.
40. Williams A. Calculating the global burden of disease: time for a strategic reappraisal? *Health Economics* 1999; 8:1-8.
41. Jorm AF, Christensen H, Griffiths KM, Rodgers B. Effectiveness of complementary and self-help treatments for depression. *Medical Journal of Australia* 2002; 176 Suppl:S84-96.

42. The UK ECT Review Group. Efficacy and safety of electroconvulsive therapy in depressive disorders: a systematic review and meta-analysis. *Lancet* 2003; 361:799-808.
43. Cormac I, Jones C, Campbell C, Silveira da Mota Neto J. Cognitive behaviour therapy for schizophrenia (Cochrane Review). In *The Cochrane Library, Issue 2*. Oxford: Update Software, 2003.
44. Deeks JJ, Altman DG, Bradburn MJ. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In *Systematic reviews in health care: meta-analysis in context*. London: BMJ Publishing Group, 2001.
45. Rosenthal R. *Meta-analytic procedures for social research*. Newbury Park: SAGE Publications, 1991.
46. Carter R, Stone C, Vos T, Hocking J, Mihalopoulos C, Peacock S, Crowley S. *Trial of Program Budgetting and Marginal Analysis (PBMA) to assist cancer control planning in Australia*. Canberra: Commonwealth Department of Health and Aged Care, 2000.
47. Jablensky A, McGrath J, Herrman H, Castle D, Gureje O, Morgan V, Korten A. *People living with psychotic illness: an Australian study 1997-98*. Canberra: Department of Health and Aged Care, 1999.
48. National Health and Medical Research Council. *How to use the evidence: assessment and application of scientific evidence*. Canberra: NHMRC, 2000.
49. Duggan L, Fenton M, Dardennes RM, El-Dosky A, Indran S. Olanzapine for schizophrenia (Cochrane Review). In *The Cochrane Library, Issue 2*. Oxford: Update Software, 2003.
50. Huston P, Moher D. Redundancy, disaggregation, and the integrity of medical research. *Lancet* 1996; 347:1024-6.
51. Safer DJ. Design and reporting modifications in industry-sponsored comparative psychopharmacology trials. *Journal of Nervous and Mental Disease* 2002; 190:583-92.
52. Kennedy E, Song F, Hunter R, Clarke A, Gilbody S. Risperidone versus typical antipsychotic medication for schizophrenia (Cochrane Review). In *The Cochrane Library, Issue 2*. Oxford: Update Software, 2003.
53. Csernansky JG, Mahmoud R, Brenner R. A comparison of risperidone and haloperidol for the prevention of relapse in patients with schizophrenia. *New England Journal of Medicine* 2002; 346:16-22.
54. Geddes J, Freemantle N, Harrison P, Bebbington P. Atypical antipsychotics in the treatment of schizophrenia: systematic overview and meta-regression analysis. *British Medical Journal* 2000; 321:1371-6.

55. Core Services Committee (New Zealand). *Core Services for 1995/1996*. Wellington: New Zealand Ministry of Health, 1994.
56. Government Committee on Choices in Health Care (The Netherlands). *Choices in Health Care*. The Hague: The Netherlands Ministry of Welfare, Health, and Cultural Affairs, 1992.
57. Nord E. The trade-off between severity of illness and treatment effect in cost-value analysis of health care. *Health Policy* 1993; 24:227-38.
58. Menzel P, Gold MR, Nord E, Pinto-Prades JL, Richardson J, Ubel P. Toward a broader view of values in cost-effectiveness analysis of health. *Hastings Center Report* 1999; 29:7-15.
59. Haddorn D. Setting health care priorities in Oregon: Cost-effectiveness analysis meets the rule of resuce. *JAMA* 1991; 265:2218-2225.
60. Churchill R, Hunot V, Corney R, Knapp M, McGuire H, Tylee A. A systematic review of controlled trials of the effectiveness and cost-effectiveness of brief psychological treatments for depression. *Health Technology Assessment Reports* 2001; 5.

8

General Discussion

CHAPTER 8

INTRODUCTION

Each of the chapters 2–7 in this thesis have been published as stand-alone papers in peer-reviewed journals. There is no need to repeat the discussion and conclusions already presented. However, what remains is to evaluate the combined body of work. The emphasis of this last chapter is on the technical feasibility of assembling the evidence on the size of mental health problems in Australia and the efficiency of the mental health services response to major depression. The main questions to be answered are:

1. to what extent have the methodological questions raised in the introductory chapter been addressed?
2. what have been the scientific achievements?
3. what are the main deficiencies in methods and available data? and
4. what are the opportunities to improve the quality and accuracy of the evidence?

In the following sections these questions are addressed separately for the two main topics, i.e. burden of mental disorders and cost-effectiveness of mental health interventions.

While these technical aspects are the focus of this thesis, this body of work has been much more than an academic endeavour. The explicit aim was to provide relevant evidence to Australian policy makers and service providers to improve the health services response to mental disorders. While this thesis was not set up as a formal health policy analysis it is appropriate to finish with some statements about the process of collating the evidence and the impact so far on influencing mental health policy and practice in Australia.

BURDEN OF MENTAL DISORDERS

The Victorian and Australian burden of disease studies have made multiple contributions to the methods of measuring the size of mental disorders in populations. Many of these developments have subsequently been

incorporated into revisions of the Global Burden of Disease (GBD) study (facilitated by the fact that Colin Mathers who was the lead researcher on the Australian study took up a post to work on the GBD at WHO) and a host of national burden of disease studies facilitated by the fact that we provided transparency of methods and data sources in publications and by posting all Excel workbooks on the world wide web.

Survey data

Access to the National Survey of Mental Health and Wellbeing (NSMHWB) was essential [1]. Without such a survey neither burden of disease nor cost-effectiveness modelling would be possible. As many mental disorders remain untreated, health service data collection systems fail to capture the true occurrence and characteristics of mental illness in the population. This is not only the case for common mental disorders such as depression, anxiety and ADHD for which the survey found that between 40% and 60% of cases had not sought any health care for their mental health problem. It also applies to schizophrenia. The Low Prevalence Disorder study, a sub-study of the NSMHWB, identified people with psychotic disorder by contacting a range of mental health service providers in four selected urban settings [2]. Due to the poor response of general practitioners contacted to take part in the survey, a very low proportion of cases identified under their care but not other mental health services. For the ACE-Mental Health study we identified a study in Newcastle, New South Wales, that found a third of people with schizophrenia under the sole care of their GP [3]. Assuming that this pattern of care is representative of the country, it means that the survey underestimates the true prevalence.

While the diagnosis of psychotic disorders in the Low Prevalence Study is that reported by mental health practitioners, there is no such equivalent for most other mental disorders. Hence, the NSMHWB relies on responses to a diagnostic survey instrument to identify people in the population who meet the criteria for a range of mental disorders. Experience from burden of disease studies has led to considerable scepticism about self-reported health status. A common finding is that self-report survey data

suffer from low sensitivity (e.g. type 2 diabetes), low specificity (e.g. asthma, angina), systematic bias (e.g. height and weight) or a combination these problems [4, 5]. For many physical disorders, measurements can provide a more accurate diagnosis in surveys. This is not the case for mental disorders. Inevitably the diagnosis rests on self-report. Mental health epidemiologists have a long history of developing and validating survey instruments. The Composite Diagnostic Interview Schedule developed by WHO [6] has been used in several large mental health surveys, including the NSMWHB, and has been shown to have variable validity when compared with diagnosis by a psychiatrist depending on which mental disorder is examined [7-12]. All of these validation studies were conducted on clinical samples rather than community samples as in surveys. It is therefore not clear how valid these population estimates are. However, surveys are the only way of collecting population data on mental disorders, as routine health service data collection systems do not capture the large proportion of people with a mental disorder who never seek care.

Problematic is also that there are two major classification systems for mental disorders, the ICD-10 by WHO and DSM-IV by the American Psychiatric Association [13, 14]. Responses to the CIDI allow diagnosis by either classification. In the NSMHWB there was lack of agreement between the ICD-10 and DSM-IV in 42% of diagnoses. At the population level, differences in prevalence are smaller. We estimated a small difference of 7% in the overall size of the mental health burden in Victoria if we would have chosen the DSM-IV rather than the ICD-10 diagnostic categories. For individual mental disorders, the differences were much greater [15]. Until a consensus is reached on a single classification system it is important that burden of disease studies clearly specify on what diagnostic system results are based.

Choice of mental disorders

Twenty mental disorders were included, 12 more than in the global burden of disease study. New conditions included childhood disorders, eating disorders and a greater number of substance abuse and anxiety disorder

categories. These conditions represented 36% of the mental health burden in Victoria in 1996. This is an important finding as it indicates that the selection of conditions for inclusion in a burden of disease study may have more than a trivial impact on the ranking of mental disorders relative to other disease categories.

There is no such problem with the mortality component of DALYs as there is a set total amount of deaths recorded in the Australian vital registration system. While there can be arguments about the attribution of deaths by cause, the total amount of YLL is fixed by the 'envelope' of age-and-sex specific counts of death. For non-fatal outcomes there is no such 'envelope' and hence it is important to carefully examine if the most important potential contributors to YLD are included. For some disease categories it is possible to define a rest category and thus avoid underestimating YLD. For instance, cancer registers provide information on all cancers allowing an estimation of YLD for a rest category 'all other cancers'. For other disease categories, an approximation of YLD for a rest category has been advocated by applying a YLD to YLL ratio to deaths within a larger disease category but not assigned to a specific disease.

This method does not work for largely non-fatal disease categories such as mental disorders [16]. The aim of a burden of disease study is to be comprehensive in its measurement of mortality and disability. Hence, if there is no recognised disability envelope and no measurable estimate or approximation of the disability due to excluded disorder categories, the choice of which disorders to include in the study is critical. The choice for 20 conditions comes close to being comprehensive. Future studies could still contemplate including conduct disorder in childhood; other personality disorders apart from borderline personality disorder; and cocaine and ecstasy as substance use categories. While information on prevalence is available for each of these conditions in the 1997 adult or child mental health surveys [1, 17] and/or the more recent national drug and alcohol household survey [18], their inclusion in a future burden of disease study would require additional data on remission or average duration as well as the development of disability weights for the associated health states.

Co-morbidity corrections

The addition of many new mental disorders to the burden of disease list made it imperative to deal with the very common co-morbidity between mental disorders. Previous burden of disease studies had ignored co-morbidity, implicitly assuming an additive model of disability. If individuals have several co-morbid conditions, this can mean that the total amount of disability in a particular individual adds to more than 1 or more than the equivalent of being dead. While there may be some arguments for the existence of health states that some people might consider worse than being dead, these are probably rare and limited in duration.

The correction for co-morbidity was restricted to the mental disorders identified in the adult mental health survey, i.e. anxiety disorders, depression, substance use disorders and borderline personality disorders. At the time of analysis the survey estimates for bipolar disorder were considered inaccurate, as one module on mania had been excluded from the interview schedule. Subsequently, a further analysis of the data indicated that estimates are possible for bipolar disorder type 1 and hence also its inclusion in corrections for co-morbidity[19]. A choice was made to make simple adjustments allocating prevalence equally between co-morbid states. This decision followed consultations with mental health experts failing to reach agreement on the likely direction of causality between co-morbid mental disorders. For instance, anxiety experts thought that those with anxiety and alcohol dependence would have started drinking to alleviate their anxiety, while alcohol experts indicated that their clients often develop anxiety or depression while already dependent on alcohol.

As the allocation of co-morbid prevalence between disorders was done at the unit record level of the survey it implicitly dealt with the relative frequency of combinations of co-morbid mental disorders. In contrast, the corrections for common physical disorders in the elderly assumed independence between the occurrence of co-morbid conditions. Taking the dependence between co-morbid physical disabilities into account was not possible due to time constraints and was flagged as a potential improvement in methods for further burden of disease studies.

Another advantage of making the adjustment at the level of prevalent cases was that we could take into account the potentially greater severity of co-morbid states as we derived severity independently of the ICD-10 diagnoses from questions on the Medical Outcomes Study short form (SF-12) questionnaire [20] included in the survey [15].

The disadvantages of the chosen approach are that (a) it artificially reduces the prevalence of mental disorders (and prevalence is an important output from a burden of disease study in its own right); and (b) that it deviates from the chosen method of adjusting disability weights rather than prevalence or incidence for co-morbidity between physical disorders. Further shortcomings are that no corrections were made for co-morbidity with mental disorders not included in the adult survey (such as childhood disorders, eating disorders and psychotic disorders) and that co-morbidity between mental and physical disorders was ignored. In future analyses, a number of improvements in a more integrated approach to corrections for co-morbidity can be made, including:

1. using uniform methods of correction for co-morbidity between mental disorders, and between mental and physical disabilities with a preference for adjusting disability weights rather than prevalence or incidence;
2. adding corrections for co-morbidity for mental disorders currently not included, particularly for co-morbidity between schizophrenia and depression or substance use disorders making use of data from the Low Prevalence Disorders survey [2]; and
3. taking into account the dependence between disabling conditions.

In the absence of surveys that capture co-morbidity between all major contributors to the non-fatal disease burden, it is inevitable that information on the probability of disabilities being co-morbid will have to come from disparate data sources. It also is unlikely that disability weights can easily be derived for the multitude of potential co-morbid states. Instead, we introduced a multiplicative model to adjust the level of disability avoiding

the problem of awarding a combined disability weight greater than 1 to some individuals[4, 5] and that still seems the most reasonable approach.

Choice of disability weights

At the time of study two sets of disability weights were available: (a) from the GBD study with a treated and untreated weight for each condition and an assumption of the proportion of cases under treatment [21]; and (b) from a Dutch study with between 2 and 4 severity levels per mental disorder [22]. We chose the Dutch DWs because (a) epidemiological data on the severity distribution of mental disorders showed great variability in severity between individuals; and (b) weights were available for a greater number of mental disorders. A major problem with the GBD approach of using treated and untreated weights is best illustrated with an example from the adult mental health survey. The Dutch DWs are 0.14, 0.35 and 0.76 for mild, moderate and severe depression. The average Dutch DW calculated for individuals who consulted a health professional for their depression (0.403) is higher than the DW for those with depression who did not consult (0.282). Furthermore, the DW for those reporting having had CBT or drugs prescribed (0.429) is higher than the DW for those consulting a health professional but not receiving treatment with an evidence base (0.364). If one believes that treatment is effective –or at least not harmful–, the conclusion must be that the more severe cases of depression are more likely to seek treatment and receive treatment with an evidence base. Clearly, the GBD DWs for treated (0.302) and untreated depression (0.600) would not have described this pattern of severity in the Australian population very well.

For burden of disease analysis one can argue that this does not matter as long as the DW reflects the average case in the population. Using the Dutch DWs the average DW for depression was 0.41 in men and 0.37 in women and would have been 0.50 assuming the treated weight applies to the 35% receiving an evidence-based treatment or 0.42 if the treated weight applies to the full 58% reporting to have consulted a health professional for their depression. There is no correct DW level as this is one of the social value choices that underpin the DALY. However, if over time the underlying

severity of depression would remain constant in the population but access to treatment would improve, use of the GBD weights would implicitly assume that the additional cases seeking treatment on average had moved from the untreated to the treated DW level. This does not reflect the finding that those previously not seeking treatment have less severe disease. The Dutch DWs with three levels of severity can better capture the change in severity after such a shift in treatment access.

For burden of disease analyses it may be sufficient to describe the average severity of depression in the population even though they are just three values across a wide spectrum. Cost-effectiveness models of interventions that ameliorate disabling health states demand greater accuracy in measuring the difference in severity with and without treatment as discussed in the papers on the ACE-Mental Health study [23, 24].

Conclusions

In conclusion, the Victorian and Australian burden of disease studies made significant improvements in the methods of estimating the size of mental health problems. As indicated above, a number of further improvements are possible including better corrections for co-morbidity and inclusion of a few more disorders. However, the main issue is to have up-to-date data from a representative mental health survey. There currently are plans to do another mental health survey in Australia in the year 2007, ten years after the first survey. Repeats of the child and psychotic disorder components of the first survey are unlikely to be included. For a group of conditions with such a dominant share of disease burden where surveys are the only reliable data source, a frequency of one survey per decade is not enough. Another recommendation for the future survey is to design a follow-up component to re-interview respondents to get data on episode duration, relapse and time spent in symptomatic disease. These epidemiological parameters are crucial to the current estimates and are derived from isolated international studies that often are conducted on clinical samples rather than community samples. Lastly, given the equivocal results of validation studies of the main survey instrument (the CIDI) in clinical samples, validating diagnoses in a sub-

sample of the survey against diagnosis obtained from a formal psychiatric interview would provide important information of the accuracy of current estimates in Australia.

ECONOMIC EVALUATION OF INTERVENTIONS FOR MAJOR DEPRESSION

Estimating effectiveness

The main outcome measures of randomised controlled trials of mental health interventions are responses to symptom-based questionnaires. Very rarely, quality of life measures are used. Recently, a few randomised controlled trials for depression have emerged that measured overall quality of life [25-27] but the vast majority of trials report on limited symptoms only. Hence, a major assumption is that differences on the symptom-based measures reflect the magnitude of change in overall quality of life, which is what we want to capture in the DALY measure. There is some supportive evidence that disorder-specific symptoms correlate with overall psycho-social disability [28, 29].

The next issue is that disparate rating scales are being used between trials with the Beck's Depression Inventory and the Hamilton Depression Rating Scale most commonly used for depression [30, 31]. In meta-analyses of mental health interventions, the effect size based on the standardised mean difference is the recommended measure to pool study results if the outcome measure is continuous [32, 33]. The method relies on the assumption that the change in a symptom rating scale between an intervention and a control group can be expressed in standard deviation units. This allows pooling of results even if studies used different scales as long as the scales measured the same construct. Not all statisticians accept this method [34, 35] and the Cochrane systematic reviews only report pooled results for the same rating scale. Adhering strictly to this position would have ruled out a considerable amount of evidence and would have necessitated a difficult choice of what rating scale to include or exclude. We

argued that exclusion of a substantial amount of evidence would lead to a greater potential for error than making the decision to pool results obtained from different rating scales.

For many of the interventions we evaluated there were existing meta-analyses. We endeavoured to update these results if significant new studies had become available since. We followed established methods in meta-analysis to (a) examine the data for heterogeneity; (b) apply the random effects method to estimate the effect size and its confidence limits; and (c) do a meta-regression if the data shows considerable heterogeneity [32]. For example, we did a meta-regression of studies on CBT for depression, panic disorder and generalised anxiety disorder [36]. The main finding was that CBT is less effective in patients with severe disease.

Translating effectiveness into a change in DALY

While there are well-established methods to pool data from trials of mental health interventions that have used disparate measures of outcome into an effect size, the translation of an effect size into a change in DALY disability weight units is more problematic. The main problem is that there are disability weights for three severity levels in major depression while the effect size is a continuous measure of change in severity. We used two approaches to make this translation. The first relies on a study by Sanderson et al. that used time trade-off methods to estimate the amount of change in disability weight that is equivalent to one standard deviation change in severity [37]. As the effect size is also expressed in standard deviation units health gain in DALY terms is simply the product of effect size and that estimated change in disability per standard deviation change in severity. Thus it avoids having to make the transition from a continuous to a categorical measure. When we discussed this method with the ACE-Mental Health Steering Committee, the Sanderson study had not yet been published. Committee members felt uneasy about using a new but not peer-reviewed method and encouraged the researchers to explore another method.

We developed an alternative method estimating for each individual in the Australian mental health survey an average reduction in severity due to

an intervention by altering the Mental Component Score (MCS) of the Medical Outcomes Study Short Form (SF–12) [38]. This average reduction in severity was determined by the product of the effect size of the intervention and the standard deviation of the MCS. The next step was to recategorise survey respondents into mild, moderate and severe depression and calculate an average disability weight in those modelled to receive the treatment. We called this the 'survey severity method' as opposed to the initial 'conversion factor method'.

In the absence of an established method and in consultation with the steering committee it was decided to incorporate both methods into uncertainty analyses, thus widening the uncertainty interval around our cost-effectiveness ratios by using either method in half of all simulations. The advantages of the conversion factor method are that (a) it is easier to use; (b) it avoids having to translate a continuous measure into change in categorically defined disability weights; and (c) conversion factors for most common mental disorders are available, thus increasing the comparability of cost-effectiveness results across mental disorders. However, the conversion factors were derived from a small pilot study with a convenience sample of general practitioners as the valuators and it used a different health state valuation method from the person trade off method used to derive the DALY disability weights [37]. Confirmation of its findings in a larger study would give more credence to the method.

The drawback of the alternative 'survey severity method' is that it required a change from a continuous measure (the effect size) into a categorical measure (the disability weights). At the level of individual survey respondents, it meant that after applying the effect size only some would make the transition into a less severe category. We then argued that if this were done over enough cases in the survey it would approximate the true average impact of the intervention. In hindsight, we could have refined the method by interpolating disability weight levels between the three anchoring points for mild moderate and severe depression (which we defined as 0.5, 1.5 and 2.5 standard deviation levels below the population mean of the MCS). That would have allowed a continuum of disability weights across

the full spectrum of severity as measured by MCS in the survey rather than for three categories only.

These methodological difficulties are less influential when comparing cost-effectiveness ratios between different interventions for the same disorder as the same method of translating change into DALYs is consistently applied. The conversion factor method is consistently derived for different mental disorders and hence increases the validity of comparisons of cost-effectiveness results between mental disorders. Comparisons are more problematic between depression and schizophrenia for the severity survey method as there were no common measures of severity in the main mental health survey and the low prevalence disorder survey.

Comparisons between ACE-Mental Health results and cost-effectiveness results of interventions for physical disorders where extension of life is the dominant health outcome are more problematic. If making such comparisons it is important not to over-interpret small differences in cost-effectiveness but to concentrate on large difference, e.g. by classifying interventions into a few broad categories such as 'very cost-effective', 'probably cost-effective', 'probably not cost-effective' and 'certainly not cost-effective'.

Modelling cost- effectiveness of interventions for depression

There are two main treatment options for major depression: psychological and drug interventions. Treatment guidelines for general practitioners in Australia recommend treatment to continue for at least one year for first episodes and for at least two years for repeat episodes [39]. Specialist guidelines recommend the same for first episodes but extend the maintenance treatment period to three years [40]. These guidelines were mainly based on evidence of effectiveness rather than economic arguments.

The ACE-Mental Health project has provided such additional evidence of cost-effectiveness on (a) the choice of psychological or drug treatment; and (b) the merits of episodic versus maintenance treatment options. This required the development of a model of depression as a chronic

episodic condition rather than the more simplistic modelling of depression, episode by episode. Thus, information was needed on the duration of episodes and the probability of repeat episodes occurring after an initial episode. There were only a few studies that had followed up people identified with depression from a community sample. Lognormal distributions best described the data points on duration of episodes and on risk of recurrence of an episode. There were only seven data points from six studies on the risk of a recurrent episode over a period of two years (see figure 2 in [41]). It clearly is desirable to have more data and hence, the recommendation made earlier that a next mental health survey in Australia include a follow-up component. Assuming these mathematical functions are a valid description of the course of depression over time, we were able to estimate the average number of episodes and the average amount of time with depression over a follow-up period in a micro-simulation model. Having defined the course of the depression over time following an episode we were able to examine the impact of various treatment options. The Archives of General Psychiatry paper [41] compared the impact of episodic and maintenance treatment on the amount of disease burden from depression experienced over a 5 year period following an episode. It demonstrated that maintenance treatment could avoid up to half of all depression burden while episodic treatment would only avoid a quarter. There was not a lot of information on a maintenance variant of the main psychological intervention, Cognitive Behavioural Therapy (CBT), for which there is best evidence of effectiveness during or following an episode. Thus, our conclusions about maintenance CBT are more speculative. Also, the evidence base for the impact of maintenance drug treatments is rather slim and more long-term and independent trials are needed to verify the impact.

In the cost-effectiveness analyses that followed we found CBT to be more cost-effective than drug treatments particularly if provided by a psychologist in public service and, more so, if delivered as group therapy. Of the two main drug treatment options the cheaper tricyclic agents (TCAs) were more cost-effective than the newer SSRIs and related drugs. All these conclusions held after examining the uncertainty ranges around each result

derived from our multivariate sensitivity analyses. The cost-effectiveness ratios for episodic and maintenance treatment were similar but the total costs and benefits were much greater. Maintenance CBT had an even more favourable cost-effectiveness ratio than CBT during an episode due to the relatively lower cost of only a few booster sessions per year.

The cost-effectiveness ratios of all treatment options fell well below the A$50,000 threshold commonly used in Australia to demarcate cost-effective from cost-ineffective interventions. However, the incremental analysis of SSRIs as an alternative to TCAs gave a cost-effectiveness ratio of more than four times this threshold. This is because SSRIs are six times more expensive than TCAs while there is no difference in effectiveness on depressive symptoms and only a small documented difference in adherence due to side effects. This example illustrates that simply analysing the cost-effectiveness of individual interventions and sticking them in a league table can mask the inefficiency of an intervention option if there is a more cost-effective alternative. This is the argument strongly put forward by the WHO-CHOICE project that the cost-effectiveness results of individual intervention options for a particular health problem need to be put along an intervention pathway that will determine the most cost-effective mix of interventions for a particular budget level [42]. We did not perform such analyses within the ACE-Mental Health project, as we did not yet have the expertise to do so. In future analyses, this is the strongly recommended strategy.

A further refinement of the cost-effectiveness modelling would be to use micro-simulation techniques rather than modelling the 'average case'. That would do better justice to the wide variation in presentation of depression and the differences in treatment response between individuals. For instance, we did not take into account the lesser impact of CBT in patients with severe depression [36] nor did we take into treatment-resistant depression account [43].

Engagement of policy makers and experts

From the start of ACE-Mental Health we engaged mental health policy makers and mental health experts in a Steering Committee. The mental

health experts came from around the country and included many prominent mental health researchers. There was also representation from the mental health lobby groups SANE and the Mental Health Council. The main tasks of the Steering Committee were to guide the choice of interventions for analysis, scrutiny of the methods and applying a set of 'second filter criteria' to help formulate policy recommendations. The latter are policy relevant criteria other than cost-effectiveness that were deemed necessary to take into account before formulating policy recommendations. Four second filter criteria were selected: equity, feasibility, acceptability to stakeholders and an overall judgment of the quality and strength of the evidence. Meetings were held twice a year over the three-year project period.

At the onset of the project it was anticipated that there would be tension between a number of different groups on the Steering Committee. First, the policy makers from the federal Department of Health and Ageing (DoHA) and Victorian Department of Human Services (DHS) were likely to have different interests in the project as the States fund hospital and community mental health services (largely dealing with more severe psychotic disorders) while the federal government directly funds pharmaceuticals and doctor visits. Initially, there was great enthusiasm in the Mental Health Branch of DHS but this subsided after the director left who had provided half of the project funds from her budget. It was only a year later that a successor gave any attention to the project and, as a true bureaucrat, first looked at the potential risks to the Branch's interests from the project. The new director frankly stated that she would never have part-funded such a study with the federal government. The ACE-Mental Health project was located in another part of the same department and therefore could not as easily be 'neutralised' as would have been the case with research commissioned from a university or private company. After many heated discussions a semblance of cooperation was found after we agreed to accompany all our publications with a caveat around the interpretation of results. The policy makers from the federal DoHA were much keener participants. This reflected their policy interests which included concern about the high expenditure on anti-depressant drugs and a pilot initiative to

allow financing of psychological treatments in primary care by non-physicians (there are no provisions for reimbursing the cost of visits to allied health workers in primary care in Australia).

Second, we expected tension between the mental health experts who mainly research psychotic disorders and those with an interest in depression and anxiety. Gavin Andrews as a proponent of the latter group had seriously ruffled feathers of the former group by presenting his (at the time not yet formally published) work comparing the cost-effectiveness of treatment options for various mental disorders [44]. His conclusions were there are a lot of under-utilised cost-effective treatment options for depression and anxiety while there are few cost-effective options in the treatment of schizophrenia. The ACE-Mental Health results, using more sophisticated analyses but largely the same underlying data sources, confirmed Gavin Andrews' findings. This set the scene for many heated discussions during each of the steering committee meetings. My impression is that with time the worst critics came to (possibly reluctantly) accept that these findings were based on currently available evidence. The second filter discussions allowed the critics to voice some of their concerns. In particular, the acceptability criterion was at times invoked to voice concerns about advocating cheap and old drugs over new and expensive drugs. There was some suspicion that part of the expert criticism of our results could have been influenced by the fact that most mental health research in Australia is heavily reliant on funds from pharmaceutical companies.

It is hard to judge how much influence the ACE-Mental Health project has had on policy decision-making and this thesis did not intend to answer that question. Nevertheless, the nature of burden of disease and cost-effectiveness work is that it is intended to inform policy. Within the Department of Human Services in Victoria, it is unlikely that the Mental Health Branch policy makers have made much direct use of the results given their reluctant participation in the project. Elsewhere in the department, the work has received much greater recognition and led to the development of another cost-effectiveness modelling project, ACE-Obesity. At the level of the federal government the favourable results for CBT gave support to the

new initiatives to find alternative funding streams for brief psychological interventions. For drugs to be listed under the Pharmaceutical Benefits Scheme there are strict criteria including cost-effectiveness that need to demonstrate the merits of a new drug. After listing, however, there is no established mechanism to restrict use of less cost-effective drugs or to promote use of cheaper drug alternatives. Hence, there has been no immediate impact on policy regarding drug options for the treatment of mental disorders. In other parts of the DoHA there is growing recognition for the need to support resource allocation decisions with evidence on the economic credentials of alternative health service options. There is considerable interest in the cost-effectiveness of prevention in the hope that strengthening cost-effective prevention option can lead to containment of the large increases in health expenditure projected over the next few decades due to the ageing of the population and the introduction of new technologies [45]. This strong interest has been influential in obtaining a large 5-year grant from the National Health and Medical Research Council (NHMRC) to evaluate the cost-effectiveness of prevention of non-communicable disease in Australia that has now been called the ACE-Prevention project. Having also secured 5-year funding from the UK Wellcome Trust and NHMRC for a similar project in Thailand is further recognition that the work on burden of disease and cost-effectiveness described in this thesis is perceived as useful and informative.

REFERENCES

1. Australian Bureau of Statistics, *Mental health and wellbeing: Profile of adults, Australia, 1997.*, in *Cat No. 4326.0*. 1998, ABS: Canberra.
2. Jablensky, A., et al., *Psychotic disorders in urban areas: an overview of the Study on Low Prevalence Disorders.* Australian and New Zealand Journal of Psychiatry, 2000. **34**(2): p. 221-36.
3. Lewin, T.J. and V.J. Carr, *Rates of treatment of schizophrenia by general practitioners. A pilot study.* Med J Aust, 1998. **168**(4): p. 166-9.
4. Vos, T. and S. Begg, *The Victorian Burden of Disease Study: morbidity*. 1999, Melbourne: Public Health and Development Division, Department of Human Services.

5. Mathers, C.D., T. Vos, and C.E. Stevenson, *The burden of disease and injury in Australia*. 1999, Australian Institute of Health and Welfare: Canberra.
6. World Health Organization, *Composite International Diagnostic Interview - version 2.1*. 1997, Geneva: WHO.
7. Wittchen, H.U., *Reliability and validity studies of the WHO--Composite International Diagnostic Interview (CIDI): a critical review.* J Psychiatr Res, 1994. **28**(1): p. 57-84.
8. Wittchen, H.U., et al., *Reliability and clinical validity of UM-CIDI DSM-III-R generalized anxiety disorder.* J Psychiatr Res, 1995. **29**(2): p. 95-110.
9. Peters, L. and G. Andrews, *Procedural validity of the computerized version of the Composite International Diagnostic Interview (CIDI-Auto) in the anxiety disorders.* Psychol Med, 1995. **25**(6): p. 1269-80.
10. Wittchen, H.U., et al., *Reliability and procedural validity of UM-CIDI DSM-III-R phobic disorders.* Psychol Med, 1996. **26**(6): p. 1169-77.
11. Rosenman, S.J., A.E. Korten, and C.T. Levings, *Computerised diagnosis in acute psychiatry: validity of CIDI-Auto against routine clinical diagnosis.* J Psychiatr Res, 1997. **31**(5): p. 581-92.
12. Cooper, L., L. Peters, and G. Andrews, *Validity of the Composite International Diagnostic Interview (CIDI) psychosis module in a psychiatric setting.* J Psychiatr Res, 1998. **32**(6): p. 361-8.
13. American Psychiatric Association, *Diagnostic and Statistical Manual of mental disorders, 4th edn: DSM-IV*. 1994, Washington, DC: APA.
14. World Health Organization, *The ICD-10 Classification of Mental and Behavioural Disorders - Diagnostic criteria for research*. 1993, Geneva: WHO.
15. Vos, T., et al., *The burden of mental disorders in Victoria, 1996.* Social Psychiatry and Psychiatric Epidemiology, 2001. **36**: p. 53-62.
16. Vos, T., *Specific approaches to the calculation of YLD*, in *National Burden of Disease Studies: a practical guide. Edition 2.0.*, C. Mathers, et al., Editors. 2001, World Health Organization: Geneva. p. 86-109.
17. Sawyer, M.G., et al., *The mental health of young people in Australia: key findings from the child and adolescent component of the national survey of mental health and well-being.* Aust N Z J Psychiatry, 2001. **35**(6): p. 806-14.
18. Australian Institute of Health and Welfare, *2001 National drug strategy household survey*. 2001, AIHW: Canberra.
19. Mitchell, P.B., T. Slade, and G. Andrews, *Twelve-month prevalence and disability of DSM-IV bipolar disorder in an Australian general population survey.* Psychol Med, 2004. **34**(5): p. 777-85.

20. Sanderson, K. and G. Andrews, *The SF-12 in the Australian population: cross-validation of item selection.* Aust N Z J Public Health, 2002. **26**(4): p. 343-5.
21. Murray, C.J.L. and A.D. Lopez, *The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020*. The Global Burden of Disease and Injury, ed. C.J.M. Murray and A.D. Lopez. 1996, Cambridge, Mass: Harvard University Press.
22. Stouthard, M.E.A., et al., *Disability weights for diseases in the Netherlands*. 1997, Department of Public Health, Erasmus University: Rotterdam.
23. Vos, T., et al., *Assessing Cost-Effectiveness (ACE)-Mental Health: Helping policy makers prioritise and plan health services.* Australian and New Zealand Journal of Psychiatry, 2005. **39**: p. 701-712.
24. Haby, M.M., et al., *Assessing Cost-Effectiveness - Mental Health: Introduction to the study and methods.* Australian and New Zealand Journal of Psychiatry, 2004. **38**: p. 569-578.
25. Bower, P., et al., *Randomised controlled trial of non-directive counselling, cognitive-behaviour therapy, and usual general practitioner care for patients with depression. II: cost effectiveness.* Bmj, 2000. **321**(7273): p. 1389-92.
26. Allsup, S.J. and M.A. Gosney, *Anxiety and depression in an older research population and their impact on clinical outcomes in a randomised controlled trial.* Postgrad Med J, 2002. **78**(925): p. 674-7.
27. Thomas, H.V., et al., *Computerised patient-specific guidelines for management of common mental disorders in primary care: a randomised controlled trial.* Br J Gen Pract, 2004. **54**(508): p. 832-7.
28. Ormel, J., et al., *Depression, anxiety, and social disability show synchrony of change in primary care patients.* Am J Public Health, 1993. **83**(3): p. 385-90.
29. Ormel, J., *Synchrony of change in depression and disability: what next?* Arch Gen Psychiatry, 2000. **57**(4): p. 381-2.
30. Beck, A.T. and R.A. Steer, *Internal consistencies of the original and revised Beck Depression Inventory.* J Clin Psychol, 1984. **40**(6): p. 1365-7.
31. Mowbray, R.M., *The Hamilton Rating Scale for depression: a factor analysis.* Psychol Med, 1972. **2**(3): p. 272-80.
32. Egger, M., G.D. Smith, and D.G. Altman, eds. *Systematic reviews in health care: meta-analysis in context*. 2001, BMJ Publishing Group: London.
33. Rosenthal, R., *Meta-analytic procedures for social research*. 1991, Newbury Park: SAGE Publications.

34. Petitti, D.B., *Meta-analysis, decision analysis, and cost-effectiveness analysis: Methods for quantitative synthesis in medicine*. Monographs in Epidemiology and Biostatistics, ed. J.L. Kelsey, et al. 2000, New York: Oxford University Press.
35. Greenland, S., J.J. Schlesselman, and M.H. Criqui, *The fallacy of employing standardized regression coefficients and correlations as measures of effect.* Am J Epidemiol, 1986. **123**(2): p. 203-8.
36. Haby, M.M., et al., *Cognitive behavioural therapy for depression, panic disorder and generalised anxiety disorder - a meta-regression of factors that may predict outcome.* Aust N Z J Psychiatry, 2005. **(in press)**.
37. Sanderson, K., et al., *Using the effect size to model change in preference values from descriptive health status.* Qual Life Res, 2004. **13**(7): p. 1255-64.
38. Jenkinson, C. and R. Layte, *Development and testing of the UK SF-12 (short form health survey).* J Health Serv Res Policy, 1997. **2**(1): p. 14-8.
39. Ellis, P.M. and D.A. Smith, *Treating depression: the beyondblue guidelines for treating depression in primary care. "Not so much what you do but that you keep doing it".* Med J Aust, 2002. **176 Suppl**: p. S77-83.
40. Ellis, P., *Australian and New Zealand clinical practice guidelines for the treatment of depression.* Aust N Z J Psychiatry, 2004. **38**(6): p. 389-407.
41. Vos, T., et al., *The burden of major depression avoidable by longer-term treatment strategies.* Archives of General Psychiatry, 2004. **61**: p. 1097-1103.
42. Hutubessy, R.C., et al., *Generalised cost-effectiveness analysis: an aid to decision making in health.* Appl Health Econ Health Policy, 2002. **1**(2): p. 89-95.
43. Fava, M. and K.G. Davidson, *Definition and epidemiology of treatment-resistant depression.* Psychiatr Clin North Am, 1996. **19**(2): p. 179-200.
44. Andrews, G., et al., *Utilising survey data to inform public policy: comparison of the cost-effectiveness of treatment of ten mental disorders.* Br J Psychiatry, 2004. **184**: p. 526-33.
45. Treasury, *Intergenerational Report 2002-03*. 2002, Treasury: Canberra. p. 30.

Summary

SUMMARY

The cost of providing health care is rising faster than other sectors of the economy in Australia. For this reason, government, as the main providers of health services, is under increasing pressure to justify its health expenditure. This also applies to the provision of health services for people suffering mental disorders.

In the nineties, the Global Burden of Disease study identified mental disorders as a large, hitherto often neglected, health problem. The use of the Disability-Adjusted Life Year (DALY) – a summary measure of population health, i.e. a combined measure of mortality and morbidity – revealed the large non-fatal health burden associated with mental disorders. For the first time in the late nineties, the Burden of Disease methods were applied to the Australian context. The mental health component of this study is the topic of chapters two and three.

As the methods of measuring the burden of mental disorders were still rather crude, the study put a particular emphasis on development of new methods. The main improvements were a) the inclusion of a greater number of mental disorders; b) more accurate disease models reflecting the chronic course of most mental disorders; c) greater detail in describing the distribution by severity; and d) a correction for the very common co-morbidity between different mental disorders.

Mental disorders were the third largest disease category, ranking behind cardiovascular disease and cancer. Depression was the largest cause of disability in men and women with another seven mental disorders ranking among the top-twenty causes of disability.

Recent mental health survey data provided good information on the prevalence and severity of disease. As the DALY is an incidence-based measure, it required further information from the scarce follow-up studies of community samples of people with a mental disorder. DALY estimates are based on a generalisation of the 'average' case of disease in the population, and hence, cannot do full justice to the great variety in duration and severity which characterises most mental disorders. A further problem is that there is lack of consensus in the definition of mental disorders. For instance, we found considerable differences in the size of mental disorders between the

International Classification of Disease version 10 (ICD–10) and the Diagnostic and Statistical Manual version 4 (DSM–IV) defined diagnoses. Moreover, the common co-morbidity between disorders and the similarities in treatment approaches put in doubt the validity of the diagnostic distinction between e.g. depression and anxiety disorder. Psychiatric diagnoses in community surveys by definition are based on self-report. For physical disorders, there is a strong preference for measured health status data as self-reported health information is often a poor proxy for true disease. However, it must be emphasised that mental health epidemiologists have put greater scientific rigour in setting standards for population surveys in comparison to epidemiologists studying the non-fatal physical conditions.

Large health problems are not necessarily priorities for health intervention. Prioritising health service action requires further evidence on the effectiveness and costs of different intervention options. Following on from the Burden of Disease studies the federal government and the state government of Victoria commissioned an economic evaluation study, the Assessing Cost-Effectiveness (ACE) – Mental Health study, to identify opportunities to address the large mental health burden with cost-effective interventions. Over a period of two years, the main treatment options for depression, anxiety, schizophrenia and Attention-Deficit and Hyperactivity Disorder were analysed.

Most health economic evaluation studies compare a limited number of alternative intervention options addressing the same health problem. Often these studies are tagged onto randomised controlled trials during which costing information has been collected. If the aim is to evaluate multiple interventions for different disorders, as was the case with ACE-Mental Health study, modelling of available epidemiological and economic data sources is the only option. Superficially, modelling may seem a less accurate evaluation method than using observed data from trials. However, most trials collect data on selected patient groups that are not representative of the distribution by age, gender and severity of the disease in the whole population. A modelling approach thus can better reflect the reality of the context in which decisions are to be made. Moreover, if multiple

comparisons are to be made, comparing results from existing economic evaluation studies in a 'league table' often is not valid as methods between studies can vary widely.

Chapter four describes the standard methods applied to all interventions evaluated in ACE-Mental Health. The main modelling principles include the use of a) systematic reviews and meta-analysis to estimate the effect size; b) Australian epidemiological survey data to describe disease and treatment patterns; c) appropriate costing data from routine government collection systems and patient surveys; d) extensive uncertainty testing around the cost per DALY results using simulation techniques; e) intensive consultations with policy makers, disease experts and community representatives throughout the study; and f) other policy-relevant criteria (equity, strength of evidence, feasibility and acceptability to stakeholders) to formulate policy recommendations.

Results of the evaluation of interventions for major depression are presented in Chapters five and six. The first of these compares the impact of the more commonly implemented short-term episodic treatment approaches with maintenance pharmacological or psychological treatments. This analysis demonstrates the value of combining burden of disease analyses with economic evaluations countering the arguments of some health economists who consider burden of disease studies a waste of analytical resources. Current treatment averts only 9% of the total burden of depression in Australia. Optimal provision of episodic treatment could avert a quarter of disease burden while maintenance treatment options potentially could halve the disease burden even if adherence of only 60% is factored into the analyses. Drug and brief psychological intervention by Cognitive-Behavioural Therapy (CBT) are equally effective. The main conclusion of the superiority of maintenance treatment remains after taking all uncertainty around input variables into account. Results suggest that to make significant inroads into the large burden of depression health service providers need to recognise and treat depression as a chronic disorder. However, the evidence base for the impact of maintenance treatments is rather slim and more long-term trials are needed to verify the impact.

The analyses are taken to full economic evaluation in Chapter six. Specifically, we chose to examine the cost-effectiveness of targeting intervention options to those with major depression who seek care but currently are not receiving evidence-based treatment. We made this choice because we did not have any evidence for interventions addressing depression in people who do not present to, or are not recognised as depressed by health services.

All intervention options for depression were found to have a favourable Incremental Cost-Effectiveness Ratio (ICER) in comparison to the current inadequate treatment options received. The most cost-effective options were bibliotherapy (CBT through a book), group CBT, individual CBT provided by psychologists in government service and the tricyclic anti-depressants (TCAs) with an ICER well below A$10,000 per DALY. Maintenance treatment with the more expensive Selective Serotonin-Uptake Reinhibitors (SSRIs) was the most expensive option with an ICER of around A$20,000 per DALY in comparison to non-evidence based treatment. However, if compared to TCAs (rather than non-evidence based treatment) SSRIs have a very high ICER (>$200,000 per DALY) reflecting the high price, lack of differential effectiveness and only modestly better adherence rates that are ascribed to a better side effect profile.

The main conclusions are that a range of cost-effective treatment options exists for major depression and that these options are currently under-utilised. Maintenance and episodic treatment options have similar economic credentials. The total cost associated with maintenance treatment is high particularly if SSRIs are the drug of choice.

During the 'second filter' criteria discussion, key policy issues were identified with regards to the expanded provision of CBT concerning the availability of suitably trained providers and the funding mechanism for therapy in primary care. Given the proven effectiveness of bibliotherapy in some patients new delivery forms of CBT such as via the Internet need to be evaluated.

Chapter seven is a discussion of the overall ACE-Mental Health project placing the results for major depression together with those of other

mental disorders. While for schizophrenia a more efficient drug treatment strategy could pay for the introduction of family therapy, recommended treatment strategies for anxiety and depression would require significant new resources. A caution was made that the study results should not be seen as prescriptive as not all treatment options were evaluated and the analyses are based on 'averages' ignoring the diversity of patients and treatment responses in clinical practice. However, current practice patterns in mental health services indicate that there are inadequate incentives to promote efficiency.

Measurement of health benefits in DALYs is identified as the main methodological problem. While the estimation of the effect size is based on accepted methods, the translation of effect size into DALY units required the development of new methods basically forcing a continuous outcome (effect size) into a small number of categorical values for different levels of severity of disease. This makes comparisons between interventions for the same disorder most valid as the same translation –regardless of how inadequate – applies to all results. Comparisons of interventions for different disorders are more difficult due to differences in available epidemiological data and hence variations in the translation methods from effect size to DALYs. Furthermore, comparison of ACE–Mental Health interventions where health gains are dominated by non-fatal outcomes with interventions for heart disease and stroke where extension of life dominates health outcomes, are less certain even though these were analysed in a companion project, ACE–Heart Disease, using comparable evaluation methods.

Chapter eight wraps up the discussion on these two large research endeavours that aimed to provide an evidence base for policy making of mental health services in Australia. Major achievements are evident in the development of credible methods of measuring disease burden and cost-effectiveness in the area of mental health. Nevertheless, further improvements are possible and recommendations for future work are made. The chapter concludes with a brief analysis of the impact of this body of work on mental health policy.

Samenvatting

SAMENVATTING

De kosten van de gezondheidszorg stijgen sneller dan die van andere sektoren in de economie in Australië. Als gevolg hiervan neemt de druk toe op de overheid om de verdeling van het gezondheidsbudget te rechtvaardigen. Dit is zeker ook het geval als het gaat om de geestelijke gezondheidszorg.

De Global Burden of Disease studie trok in de negentiger jaren de aandacht voor de tot dan toe vaak genegeerde, maar belangrijke rol van psychiatrische aandoeningen. De toepassing van de Disability-Adjusted Life Year (DALY) – een samengestelde volksgezondheidsmaat waarin mortaliteit en morbiditeit gecombineerd worden – bracht aan het licht dat psychiatrische aandoeningen een grote hoeveelheid morbiditeit bijdragen aan de totale ziektelast in de wereld. In de tweede helft van de negentiger jaren werd de ziektelast voor het eerst in Australië gemeten. De geestelijke ziektelast is het onderwerp van hoofdstuk twee en drie.

Indertijd waren de methoden om de ziektelast van psychiatrische aandoeningen te meten nogal rudimentair. Daarom werd speciale nadruk gelegd op het ontwikkelen van een nieuwe aanpak. De belangrijkste verbeteringen waren: a) de beschrijving van een groter aantal psychiatrische aandoeningen; b) ziektemodellen die beter het chronische verloop van de meeste psychiatrische aandoeningen weergeven; c) meer aandacht voor de verdeling naar ernst van ziekte; en d) een correctie voor de zeer vaak voorkomende comorbiditeit tussen psychiatrische aandoeningen.

Psychiatrische aandoeningen vormden de op twee na grootste ziektecategorie, volgend op hart- en vaatziektes en kanker. Depressie was de grootste oorzaak van de morbiditeitscomponent van de ziektelast bij mannen en vrouwen en zeven andere psychiatrische aandoeningen staan in de top twintiglijst van morbiditeit.

Een recente geestelijke gezondheidsenquête leverde goede informatie over de prevalentie en ernst van ziekte. Omdat de DALY gebaseerd is op incidentie, was er behoefte aan aanvullende informatie uit de weinige beschikbare vervolgstudies van mensen bij wie in een enquête een geestelijke aandoening was vastgesteld. Voor de DALY schattingen wordt een generalisatie gemaakt voor het gemiddelde geval in de bevolking. Het is

daarom onvermijdelijk dat geen rekening gehouden kan worden met alle verscheidenheid in duur en ernst van psychiatrische aandoeningen. De definitie vershillen van psychiatrische diagnoses is een bijkomend probleem. Wij vonden, bijvoorbeeld, aanzienlijke verschillen in de ziektelastschatting als we de diagnose van psychiatrische aandoeningen in de enquête baseren op de International Classification of Disease versie 10 (ICD–10) of de Diagnostic and Statistical Manual versie 4 (DSM–IV). Bovendien zijn er twijfels of er een diagnostisch verschil bestaat tussen depressie en angststoornissen, omdat er zoveel overlap bestaat tussen deze aandoeningen en de behandelingsstrategieën grotendeels hetzelfde zijn. Nog een probleem is dat psychiatrische aandoeningen in bevolkingsenquêtes worden gemeten op grond van zelf-gerapporteerde symptomen. Voor lichamelijke ziektes is er een sterke voorkeur voor meer objectieve gezondheidsmaten omdat de ervaring leert dat zelf-gerapporteerde symptomen vaak een slechte aanduiding zijn van de echte ziekte. Het moet echter gezegd worden dat over het algemeen psychiatrische epidemiologen meer aandacht hebben besteed aan het opzetten van wetenschappelijke grondregels voor enquêtes dan hun collegas die zich bezig houden met het onderzoeken van niet-fatale lichamelijke aandoeningen.

Grote gezondheidsproblemen zijn niet noodzakelijk prioriteiten voor gezondheidszorginterventie. Daarvoor is meer informatie over de effectiviteit en de kosten van verschillende interventieopties nodig. Als vervolg op de ziektelaststudies financierden de federale regering en de staatsregering in Victoria een economische evaluatiestudie, de Assessing Cost-Effectiveness (ACE) – Mental Health studie, om te onderzoeken wat de kosteneffectieve mogelijkheden zijn om de grote ziektelast van psychiatrische aandoeningen aan te pakken. In een project van twee jaar werden de belangrijkste behandelingsmethodes voor depressie, angststoornissen, schizophrenie en Attention-Deficit and Hyperactivity Disorder (ADHD) onderzocht.

De meeste economische evaluaties in de gezondheidszorg vergelijken maar een beperkt aantal interventies voor eenzelfde gezondheidsprobleem. Vaak worden deze studies uitgevoerd als onderdeel van een randomised

controlled trial waarbij extra informatie over kosten wordt verzameld. Als het doel is, zoals in de ACE–Mental Health studie, om meerdere interventies voor verschillende aandoeningen te bestuderen dan is de enige optie om te modelleren met beschikbare epidemiologische en economische informatiebronnen. Oppervlakkig gezien lijkt modelleren een minder nauwkeurige manier van evalueren dan evaluatie op grond van data uit een trial. De meeste trials, echter, verzamelen data in geselecteerde patiëntengroepen die niet representatief zijn voor leeftijd, geslacht en de ernst van ziekte in de hele bevolking. Via modelleren kunnen de werkelijke omstandigheden van de bevolking waarvoor beslissingen genomen moeten worden beter worden weergegeven. Het bijelkaar zetten van gepubliceerde evaluatiestudies in een ranglijst is geen goed alternatief omdat de methodes erg verschillen van studie tot studie.

De gestandardiseerde evaluatie methodes in ACE-Mental Health zijn het onderwerp van hoofdstuk vier. De belangrijkste basisregels voor de modellen zijn het gebruik van a) systematisch literatuuronderzoek en meta-analyse om de effectgrootte te bepalen; b) Australische epidemiologische enquêtedata om het ziekteverloop en behandelingstrajecten te beschrijven; c) kosteninformatie uit routinematig verzamelde informatie en patiëntenenquêtes; d) uitgebreide onzekerheidsanalyses rond de kosten per DALY uitkomsten met behulp van simulatietechnieken; e) intensieve raadpleging van beleidsmakers, ziektenexperts en vertegenwoordigers van patientenorganisaties gedurende de hele studie; en f) andere criteria van belang voor het formuleren van aanbevelingen (rechtvaardige verdeling, sterkte van het bewijsmateriaal, uitvoerbaarheid en de mate waarin beleidsaanbevelingen acceptabel zijn voor verschillende belangengroepen).

De resultaten van de evaluatie van interventies voor depressie worden besproken in de hoofdstukken vijf en zes. In hoofdstuk vijf wordt een vergelijking gemaakt tussen de gebruikelijke kortetermijn behandelingsaanpak en een langer durende onderhoudsbehandeling met medicijnen of cognitieve gedragstherapie. Deze analyse toont de waarde aan van het combineren van ziektelast en economische evaluatie in tegenstelling tot de beweringen van sommige gezondheidseconomen dat

ziektelastmetingen een verspilling van analytische capaciteit zijn.De huidige behandelingsaanpak die gericht is op het behandelen van patienten die zich melden met een episode van depressie voorkomt slechts 9% van de totale ziektelast van depressie in Australië. Optimaal gebruik van de episodische behandeling zou ongeveer een kwart van de ziektelast kunnen wegnemen terwijl onderhoudsbehandeling potentieel de helft kan voorkomen, zelfs wanneer aangenomen wordt dat maar 60% van patienten de behandeling voortzet. Medicijnen en kortdurende psychologische behandeling met cognitieve gedragstherapie hebben eenzelfde uitkomst. De belangrijkste conclusie dat onderhoudsbehandeling meer ziektelast voorkomt blijft overeind ook als alle onzekerheid over de ingevoerde variabelen meegerekend wordt. Deze resultaten suggereren dat een aanzienlijke vermindering in de ziektelast van depressie alleen kan gebeuren als behandelaars het chronische verloop herkennen en de ziekte als zodanig behandelen. Een caveat is dat het bewijsmateriaal voor de uitkomsten van onderhoudsbehandeling nogal zwak is. Er zijn meer en langere trials nodig om de resultaten te ondersteunen.

De analyses worden tot een volledige economische evaluatie uitgewerkt in hoofdstuk zes. We besloten om met name de kosten-effectiviteit te onderzoeken van verschillende interventiekeuzes gericht op diegenen met depressie die wel hulp zoeken en als depressief herkend worden, maar geen bewezen effectieve behandeling ontvangen. Deze keuze werd bepaald door het feit dat we geen informatie hadden over interventies voor depressieve mensen die geen hulp zoeken bij de gezondheidszorg of daar niet als zodanig herkend worden. .

Alle behandelingsopties voor depressie hadden een gunstige kosteneffectiviteitsratio in vergelijking met de huidige inadequate behandeling. Bibliotherapie (cognitieve gedragstherapie uit een boek), cognitieve gedragstherapie in een groep of individueel door een psycholoog in overheidsdienst en de tricyclische antidepressiva (TCAs) waren het meest kosteneffectief met een ratio onder de A$10,000 per DALY. Onderhoudsbehandeling met de duurdere Selective Serotonin-Uptake Reinhibitors (SSRIs) was de duurste optie met een ratio van rond de

A$20,000 per DALY in vergelijking met de huidige inadequate behandeling. Als we echter een directe vergelijking maken met TCAs vinden we een erg hoge ratio voor SSRIs (>A$200,000 per DALY) als gevolg van de hoge kosten van SSRIs, het ontbreken van een betere uitkomst voor depressie en een slechts bescheiden voordeel in termen van behandelingstrouw vanwege een iets mindere bijwerkingen.

De belangrijkste conclusies zijn dat er een aantal kosteneffectieve behandelingsopties bestaat voor depressie en dat deze behandelingen op dit moment te weinig worden gebruikt. In termen van kosteneffectiveit zijn onderhoudsbehandeling en episodische behandeling gelijkwaardig. De totale kosten van onderhoudsbehandeling zijn echter veel groter, vooral als voor SSRIs gekozen wordt.

Na de economische analyses was het op uitgebreide schaal invoeren van cognitieve gedragstherapie het belangrijkste discussiepunt, met name het gebrek aan opgeleide therapeuten en het ontbreken van een vergoedingensyteem voor psychotherapie door niet-dokters in de eerstelijnszorg. Omdat cognitieve gedragstherapie uit een boek ook effectief is bij sommige patienten is het aan te raden om alternatieve toepassingen, zoals bijvoorbeeld via het Internet, te onderzoeken.

Hoofdstuk zeven is een bespreking van het hele ACE-Mental Health project waarbij de resultaten voor depressie naast die van andere psychiatrische aandoeningen gelegd worden. Voor schizofrenie kan een meer efficiente medicijnbehandelingsstrategie genoeg geld uitsparen om de introductie van gezinstherapie te financieren. Voor de aanbevolen behandelingen voor depressie en angststoornissen zijn aanzienlijke nieuwe fondsen nodig. De discussie waarschuwt voor een te letterlijke uitleg van de resultaten omdat niet alle behandelingsopties geëvalueerd zijn en omdat de analyses gebaseerd zijn op het ‘gemiddelde geval’ waarbij de diversiteit van patienten en reacties op behandeling niet optimaal ingebouwd zijn. Aan de andere kant zijn er overduidelijke tekenen dat de huidige geestelijke gezondheidszorg efficientie veel te weinig beloont.

Het meten van gezondheidswinst in DALYs wordt aangeduid als het belangrijkste methodologische probleem. Er bestaan geaccepteerde

methodes om de effectgrootte te meten in meta-analyses van trials. Het was echter noodzakelijk om nieuwe methodes te ontwikkelen om de effectgrootte om te zetten in DALY-eenheden , zodat de continue schaal van de effectgrootte vertaald kon worden naar een klein aantal categorische DALY waarden voor de ernst van ziekte. Daarom zijn de vergelijkingen voor dezelfde aandoening het meest valide omdat dezelfde (al dan niet adequate) omzetting gebruikt werd in alle analyses. Vergelijkingen tussen depressie en schizofrenie zijn minder geldig vanwege verschillen in de beschikbare epidemiologische data en als gevolg daarvan enigszins andere methodes om de effectgrootte om te zetten in DALYs. Vergelijkingen tussen de ACE-Mental Health resultaten, waarbij de gezondheidswinst met name morbiditeit betreft, en met die van een zusterproject, ACE-Heart Disease, zijn ondanks eenzelfde aanpak nog onzekerder omdat de gezondheidswinst voor hart- en vaatziektes met name door een vermindering in sterfte bepaald wordt.

De thesis besluit in hoofdstuk acht met een bespreking van de twee grote onderzoeksprojecten die als doel hadden om grondslagen te leveren voor de besluitvorming in de geestelijke gezondheidszorg in Australië. Er grote vooruitgang gemaakt in de ontwikkeling van geloofwaardige methodes om de ziektelast en kosteneffectiveit op het gebied van de geestelijke gezondheid te meten. Niettemin zijn verdere verbeteringen mogelijk en aanbevelingen voor toekomstig onderzoek worden gemaakt. Het hoofdstuk beëindigt met een korte analyse van de invloed van het hier gepresenteerde onderzoek op de besluitvorming in de geestelijke gezondheidszorg.

Word of Thanks

WORD OF THANKS

The origins of my career as a Burden of Disease researcher lie in a simple answer of 'yes' to two questions posed to me by Sharon Hutley and Betty Kirkwood of the London School of Hygiene and Tropical Medicine while I was a MSc student: (a) would I be interested in a job at the School on completion of my degree?; and (b) would I terribly mind taking on an assignment in Mauritius to help conduct a burden of disease study? Sharon and Betty, your support and encouragement over the three years that followed are much appreciated and have been influential in preparing me for this thesis. In particular, your meticulous guidance on scientific writing in a language other than my native tongue has made me a better researcher.

Next, I want to thank Mike Ackland who rang me with a few hours to spare to prepare a job application. Without your intervention Karla and I may still have been living in London with the mixed blessing of interesting work but poor quality of life. Our quality of life has certainly improved and I was given the rare opportunity as a bureaucrat to do meaningful research.

The Victorian Burden of Disease Study was an intensive effort of a small team of very dedicated researchers. Special thanks to Stephen Begg, Anne Magnus and Colin Mathers who are inspiring colleagues.

Rob Carter taught me how to become better at 'moonlighting' in health economics. Our complementary skills and knowledge helped shape the ACE projects. The ACE-Mental Health researchers (Cathy Mihalopoulos, Michelle Haby, Anne Magnus, Louise Heuzenroeder, Marie Donnelly and Maturot Chalamat) have been phenomenal. You have shown that serious research can be great fun and productive at the same time.

I am especially indebted Professor Gavin Andrews for his guidance on the use of the mental health survey to derive burden of disease estimates and many of the basic modelling concepts in the cost-effectiveness analyses such as the use of the effect size, the definition of evidence based treatment and the translation of effect size into a change in DALY weights using the 'conversion factor method'.

I give many thanks for the unconditional support throughout my life given by my parents, Hannie and Anthon, who have taught me compassion and the joys of learning and exploring.

WORD OF THANKS

Foremost, I want to thank my precious wife and mate Karla for always being there.

Curriculum

Vitae

CURRICULUM VITAE

After obtaining a medical degree in the Netherlands, Theo Vos was a 'bush' doctor in Lesotho and Zimbabwe for 9 years. Thereafter he did a MSc degree (Public Health in Developing Countries) at the London School of Hygiene and Tropical Medicine and stayed on as a lecturer for another three years. From 1997-2003 he was a Senior Epidemiologist at the Department of Human Services in the state of Victoria in Australia, with a part-time appointment at Monash University in Melbourne. Since Sept 2003 he is heading the Centre for Burden of Disease and Cost-Effectiveness at the School of Population Health of the University of Queensland. The Centre aims to provide health policy makers with best available evidence to guide the allocation of health resources.

Theo has carried out burden of disease studies in Mauritius and Victoria and has made major contributions to the national Australian study and studies in Zimbabwe, Thailand, South Africa and Malaysia. He is currently leading the update of the Australian Burden of Disease study including for the first time an Indigenous study. Burden of Disease studies the size and distribution of health problems in a population as well as the underlying causes. He has also led cost-effectiveness studies in the areas of cardiovascular disease and mental disorders and is currently driving the ACE-Prevention project which over the coming five years will comprehensively model the cost-effectiveness of prevention options for non-communicable disease in Australia. He is one of the principal investigators on a similar five-year grant, the Setting Priorities based on Information of Cost-Effectiveness (SPICE) project, to improve cause of death information and to model cost-effectiveness in Thailand. From 2006, Theo will lead the Global Burden of Disease activities as part of the newly created Ellison Institute

Publications

PUBLICATIONS IN PEER-REVIEWED JOURNALS

Teerawattananon Y, Vos T, Tangcharoensathien V, Mugford M (2005). Cost effectiveness of Models for Prevention of Vertical HIV Transmission: Voluntary Counseling and Testing and Choices of Drug Regimen. Cost Effectiveness and Resource Allocation, 3:5 (18 Jul 2005).

Haby MM, Donnelly M, Corry J, Vos T (2005). Cognitive behavioural therapy for depression, panic disorder and generalised anxiety disorder – a meta-regression of factors that may predict outcome. Australian and New Zealand Journal of Psychiatry (accepted for publication).

Nelson MR, Liew D, Vos T, Bertram M (2005). Epidemiological modelling of routine use of low-dose aspirin for the primary prevention of coronary heart disease and stroke in those aged 70 years and above. British Medical Journal, 330:1306–1311.

Chalamat M, Mihalopoulos C, Carter R, Vos T (2005). Assessing Cost-Effectiveness in Mental Health: Vocational rehabilitation for schizophrenia and related conditions. Australian and New Zealand Journal of Psychiatry, 39:693–700.

Vos T, Haby MM, Magnus A, Mihalopoulos C, Andrews G, Carter R (2005). Assessing Cost-Effectiveness (ACE)–Mental Health: Helping policy makers prioritise and plan health services. Australian and New Zealand Journal of Psychiatry, 39:701–712.

Vos T, Corry J, Haby MM, Carter R, Andrews G (2005). Cost-effectiveness of CBT and drug interventions for major depression. Australian and New Zealand Journal of Psychiatry, 39:683–692.

Magnus A, Carr V, Mihalopoulos C, Carter R, Vos T (2005). Assessing Cost-Effectiveness of Drug Interventions for Schizophrenia. Australian and New Zealand Journal of Psychiatry, 39:44–54.

Kruyshaar ME, Barendregt JJ, Vos T, de Graaf R, Spijker J, Andrews G (2005). Lifetime prevalence estimates of major depression: an indirect estimation method and a quantification of recall bias. European Journal of Epidemiology 20:103–111.

Danielsson N, Hoa DP, Thang NV, Vos T, Loughnan PM (2004). Intracranial haemorrhage due to late onset vitamin K deficiency bleeding in Hanoi province, Vietnam. Archives of Disease in Childhood. Fetal Neonatal Ed., 89: F546-F550.

Vos T, Haby MM, Barendregt JJ, Kruyshaar ME, Corry J, Andrews G (2004). The burden of major depression avoidable by longer-term treatment strategies. Archives of General Psychiatry, 61:1097–1103.

Haby M, Carter R, Mihalopoulos C, Magnus A, Andrews G, Vos T (2004). Assessing Cost-Effectiveness (ACE) – Mental Health: Introduction to the Study and Methods. Australian and New Zealand Journal of Psychiatry, 38:569–578.

Heuzenroeder L, Donnelly M, Haby M, Mihalopoulos C, Rossell R, Carter R, Andrews G, Vos T (2004). Cost-effectiveness of psychological and pharmacological interventions for generalised anxiety disorder and panic disorder. Australian and New Zealand Journal of Psychiatry, 38: 602–612.

Donnelly M, Haby MM, Carter R, Andrews G, Vos T (2004). Cost-effectiveness of dexamphetamine and methylphenidate for the treatment of childhood ADHD. Australian and New Zealand Journal of Psychiatry, 38: 592–601.

Haby MM, Tonge B, Littlefield L, Carter R, Vos T (2004) Cost-effectiveness of cognitive behavioural therapy (CBT) and selective serotonin reuptake inhibitors (SSRIs) for major depression in children and adolescents. Australian and New Zealand Journal of Psychiatry, 38: 579–591.

Mihalopoulos C, Magnus A, Carter R, Vos T (2004). Assessing Cost-Effectiveness (ACE) – Mental Health: Family Interventions for Schizophrenia and Related Conditions. Australian and New Zealand Journal of Psychiatry, 38:511–519.

Zhao Y, Guthridge S, Magnus A, Vos T (2004). Burden of disease and injury in Aboriginal and non-Aboriginal populations in the Northern Territory. Medical Journal of Australia, 180:498–502.

Stone CA, Carter R, Vos T, StJohn DJB (2004). Colorectal cancer screening in Australia: An economic evaluation of a potential biennial screening program using faecal occult blood tests. Australia New Zealand Journal of Public Health, 28:273–282.

Haas SJ, Vos T, Gilbert RE, Krum H (2003). Are β-blockers as efficacious in patients with diabetes mellitus as in patients without diabetes mellitus who have chronic heart failure? A meta-analysis of large-scale clinical trials. American Heart Journal, 146:848–53.

PUBLICATIONS

Barendregt JJ, van Oortmarssen GJ, Vos T, Murray CJL (2003) A generic model for the assessment of disease epidemiology: The computational basis of DisMod II. Population Health Metrics, 1:4.

Liew D,McNeil J, Peeters A, Lim SS, Vos T (2002) Epidemiological modelling (including economic modelling) and its role in preventive drug therapy. Medical Journal of Australia, 177:364–367.

Vos T, Mathers C, Herrman H, Harvey C, Gureje O, Bui D, Watson N, Begg S (2001). The burden of mental disorders in Victoria, 1996. Social Psychiatry and Psychiatric Epidemiology, 36:53–62.

Lim S, Vos T, Peeters A, Liew D, McNeil J (2001). Is prescribing statins according to the Pharmaceutical Benefits Scheme guidelines cost-effective? Medical Journal of Australia, 175:459–464.

McNeil JJ, Peeters A, Liew D, Lim S, Vos T (2001). A model for predicting the future incidence of coronary heart disease within percentiles of coronary heart disease risk. Journal of Cardiovascular Risk, 8:31–7.

Singh B, Hawthorne G, Vos T (2001). The role of economic evaluation in mental health care. Australian and New Zealand Journal of Psychiatry, 35:104-117.

Vos T (2001). Ranking reproductive health problems to define service priorities. African Journal of Reproductive Health, 5:24–32.

Mathers C, Vos T, Stevenson C, Begg S (2001). The burden of disease and injury in Australia. Bulletin of the World Health Organization, 79, 1076-1084.

Vos T, Begg S, Y Chen, Magnus A (2001). Socio-economic differentials in life expectancy and years of life lost in Victoria, 1992–1996. New South Wales Public Health Bulletin 12:126–130.

Begg S, Vos T, Stone C (2001). Mesothelioma as a marker for asbestos-related lung disease in Victoria. Victorian Chief Health Officer's Bulletin, 1:10–12.

Magnus A, Vos T, Begg S (2001). Improvement in the Life Expectancy of Victorians. Victorian Chief Health Officer's Bulletin, 1:16–20.

Mathers C, Vos T, Stevenson C, Begg S (2000). The Australian burden of disease study: measuring the loss of health from diseases, injuries and risk factors. Medical Journal of Australia, 172:592–6.

Mathers C, Vos T (2000). The burden of disease and injury among older Australians. Australasian Journal on Ageing, 19:54–5.

Vos T, Mathers C (2000). The burden of mental disorders: a comparison of methods between the Australian burden of disease studies and the Global Burden of Disease Study. Bulletin of the World Health Organization, 78:427–38.

Griekspoor A, Sondorp E, Vos T (1999). Cost-effectiveness analysis of humanitarian relief interventions: visceral leishmaniasis treatment in the Sudan. Health Policy and Planning, 14:70–76.

Deen JL, Vos T, Huttly S, Tulloch J (1999). Injuries and non-communicable diseases: Emerging health problems of children in developing countries. Bulletin of the World Health Organization, 77:518–524.

Vos T, Gareeboo H, Roussety F (1998). Ethnic differences in ischaemic heart disease and stroke mortality in Mauritius between 1989 and 1994. Ethnicity and Health, 3:45-54.

Lincetto O, Vos ET, Graça A, Cesar C, Tallarico M and Amparo A (1998). Impact of seasonality and importance of weight at discharge in Kangaroo Mother Care treated infants in Mozambique. Acta Paediatrica, 87:433-9.

Vos T, Mathers C (1998). Appropriate yardsticks for measuring population health. Letter to the Editor, Australia and New Zealand Journal of Public Health, 22:515-7.

Mabey D, Vos T (1997). Syndromic approaches to disease management. Lancet, 349 (suppl III): 26-28.

Sathananthan K, Vos T, Bango G (1996). Dental caries, fluoride levels and oral hygiene practices of school children in Matebeleland South, Zimbabwe. Community Dent Oral Epidemiol, 24:21-4.

Meursing K, Vos T, Coutinho O, Moyo M, Mpofu S, Oneko O, Mundy V, Dube S, Mahlangu T, Sibindi F (1995). Child Sexual Abuse in Matabeleland, Zimbabwe. Social Science and Medicine, 41:1693-1704

Wolf B, Ikeogu M, Vos ET (1995). Effect of nutritional and HIV status on bacteriaemia in Zimbabwean children who died at home. European Journal of Pediatrics, 154:299-303

Vos T (1994). Attitudes to Sex and Sexual Practices in Rural Matabeleland. AIDS CARE, 6:193-203

Wainer S, Vos ET (1991). A family with congenital arachnodactyly. Central African Medical Journal, 37:262-264

BOOK CHAPTERS

Rodgers A, Lawes C, Gaziano T, Vos T. The growing burden of risk from high blood pressure, cholesterol and body weight. In: Jamison D (Ed.) Disease Control Priorities in Developing Countries (2nd edition) (in press).

Vos T (2002). Shifting the goal post — normative survivorship goals in health gap measures. In Murray CJL, Salomon JA, Mathers CD, Lopez AD, eds. Summary measures of population health. Geneva: WHO.

Vos T (2002). The case against annual profiles for the valuation of disability weights. In Murray CJL, Salomon JA, Mathers CD, Lopez AD, eds. Summary measures of population health. Geneva: WHO.

Vos T (2001). Epidemiological estimates – General approach. In: Mathers C, Vos T, Lopez A, Salomon J, Lozano R, Ezzati M eds., National Burden of Disease Studies: a practical guide. Edition 2.0. Geneva: World Health Organization, 51–63.

Vos T (2001). Specific approaches to the calculation of YLD. In: Mathers C, Vos T, Lopez A, Salomon J, Lozano R, Ezzati M eds., National Burden of Disease Studies: a practical guide. Edition 2.0. Geneva: World Health Organization, 86–109.

Vos T (2001). Presentation and dissemination of results. In: Mathers C, Vos T, Lopez A, Salomon J, Lozano R, Ezzati M eds., National Burden of Disease Studies: a practical guide. Edition 2.0. Geneva: World Health Organization, 133–136.

OTHER PUBLICATIONS

Bundhamchareon K, Teerawatananon Y, Vos T, Begg S (2002). Burden of disease and injuries in Thailand. Nonthaburi, Thailand: Ministry of Public Health.

Vos T, Carter R, Andrews G, McNeil J (2002) The ACE studies: helping policy makers choose evidence-based health services. Proceedings of Australian Health Outcomes Conference 2002.

Vos T, Begg S, Magnus A (2002). Response to recent comments regarding burden of disease results in Victoria. Letter to the Editor, Australian and New Zealand journal of Public Health, 26:282–283.

Carter R, Stone C, Vos T, Hocking J, Mihalopoulos C, Peacock S, Crowley S (2000). Trial of Program Budgetting and Marginal Analysis (PBMA) to assist cancer control planning in Australia. Canberra: Commonwealth Department of Health and Aged Care.

Vos T, Begg S (1999). The Victorian Burden of Disease Study: Morbidity. Melbourne: Department of Human Services.

Vos T, Begg S (1999). The Victorian Burden of Disease Study: Mortality. Melbourne: Department of Human Services.

Mathers C, Vos T, Stevenson C (1999). The burden of disease and injury in Australia. Australian Institute of Health and Welfare, Canberra, AIHW.

Vos T (1998). Mauritius Health Sector Reform, Update of National Burden of Disease Study. Final report of consultancy to Ministry of Health, Mauritius.

Murray CJL, Mahapatra P, Ashley R, Michaud C, George A, Horborn P, Akhavan D, Shibuya A, Ssonyamontono K, Ngo D, Henscher M, Vos T (1997). The Health Sector in Mauritius: Resource use, interventions costs and options for efficiency enhancement. Consultancy report by Burden of Disease Unit, Harvard School of Public Health to the Government of Mauritius.

Vos T, Timæus I, Gareeboo H, Roussety F, Huttly S, Murray CJL (1995). Mauritius Health Sector Reform, National Burden of Disease Study. Final report of consultancy to Ministry of Health and Ministry of Economic Planning and Development, Mauritius.

Propositions

(Stellingen)

Stellingen

1. Research on priority setting in health is most likely to influence health policy decision-making if undertaken in close collaboration with government
2. Rational arguments for efficient use of health resources face formidable counterarguments by industry, politicians, bureaucrats, media, disease advocates and consumers trying to influence resource allocation.
3. Ignoring basic burden of disease information on prevalence, incidence and duration of disease when presenting marginal cost-effectiveness results to policy makers is an inadequate use of precious analytical resources.
4. Mental health epidemiologists ought to avoid measuring 'lifetime prevalence' at the expense of point or 12-month prevalence.
5. Frequent co-occurrence and similarity in treatment approach are two strong arguments to treat depression and anxiety disorders as a single mental disorder.
6. There is an inverse relationship between the frequency at which self-report health surveys are being conducted and their value to quantifying population health.
7. Epidemiologists should not assume that respondents have the same level of interest in the survey topic as they do.
8. The Australian 'no worries, mate' culture does not translate into lower rates of depression.
9. The public defense and the dissemination of the thesis' findings as a book make a PhD in the Netherlands much more rewarding than a PhD in Australia.
10. Research on burden of disease and cost-effectiveness is conducted in the spirit of Voltaire's 'Le mieux est l'ennemi du bien' — 'The best is the enemy of the good'.
11. The frequent accusations that the national broadcaster, the Australian Broadcasting Corporation, has a left-wing bias show the arrogance of the Liberal Party in a country where the media are dominated by the right-wing Packer and Murdoch empires,.