The Burden of Major Depression Avoidable by Longer-term Treatment Strategies

Theo Vos, MD, MSc; Michelle M. Haby, MAppSc, PhD; Jan J. Barendregt, MA, PhD; Michelle Kruijshaar, MSc, PhD; Justine Corry, MSocSci; Gavin Andrews, MD

Background: Major depression is the largest single cause of nonfatal disease burden in Australia. Effective drug and psychological treatments exist, yet are underused.

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: Modeling 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 modeled through “what if” scenarios of no treatment, current treatment, and optimal treatment strategies with cognitive behavioral therapy or antidepressant 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 behavioral therapy could avert 28% (95% uncertainty interval, 19%-39%) of this disease burden, and with drugs 24% (95% uncertainty interval, 19%-30%) could be averted. During the 5 years after an episode of major depression, current episodic treatment patterns would avert 13% (95% uncertainty interval, 10%-17%) of Disability-Adjusted Life Years, whereas maintenance drug treatment could avert 50% (95% uncertainty interval, 40%-60%) and maintenance cognitive behavioral therapy could avert 52% (95% uncertainty interval, 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.

Arch Gen Psychiatry. 2004;61:1097-1103
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 modeled 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 during 5 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 use 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 vs 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 NSMHWB and apply these to 2000 population figures. The main outcome of the survey was the 1-year prevalence, ie, people qualifying for a diagnosis of major depression in the 12 months prior to the survey. An additional question on the recency of symptoms allows identification of respondents with current prevalence, ie, having had symptoms in the last 2 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, whereas 35.1% fulfilled our criteria for potentially having received evidence-based treatment: consulting a health professional at least 3 times and having had medication and/or CBT (“learning how to change thoughts, behaviors, 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) into normal (≥45), mild (35-44.9), moderate (25-34.9), and severe (<25). Disability weights for mild (0.14), moderate (0.35), and severe (0.76) depression, which were used in the Australian Burden of Disease Study and derived from a Dutch study,6 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 the time to the next episode. While there are many naturalistic studies of the duration of major depressive disorder episodes in clinical samples, there are few follow-up studies of major depressive disorder in community samples.9-13 The 4 US studies show a similar pattern of recovery over time after the start of an episode. The median time to recovery in the 4 studies ranged from 8 to 12 weeks, and at 1 year between 3% and 11% of cases had not yet recovered. The figures from Kendler et al10 have been adjusted for the 7% of excluded cases with an onset of more than 1 year prior to study. The fifth study from the Netherlands11 reports a considerably longer duration of episodes. Exclusion of subsyndromal depression and dysthymia in life chart histories is a possible explanation for this higher estimate. From the data reported in the US studies9,10,12,13 we fit a lognormal distribution17 that has the lowest sum of squared differences between modeled and observed data points (Figure 1).

Major depression is a chronic episodic disorder, and hence for our modeling purposes it is important to describe the pattern of time to a next episode after a previous episode. During a few decades of follow-up, major depression is reported as a recurrent disorder in 80% of cases.9,13 We assume that during a lifetime at least 90% of affected individuals experience a recurrence. Six naturalistic follow-up studies16-21 report on the risk of relapse during periods varying between 6 months and 2 years after cessation of drug treatment for an acute episode of major depression. We fit lognormal and Weibull distributions22 that give the best fit as determined by the lowest sum of squared differences between modeled and observed data points (Figure 2). We decide to use the lognormal distribution be-
cause 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 during a 6-month and a 5-year period after an episode.

IMPACT OF INTERVENTIONS

We separately evaluate drug treatment for episodes of major depression plus a continuation phase after remission of symp-toms and maintenance treatment of 5 years after remission of an episode, CBT treatment of major depressive episodes, and a maintenance variant of CBT with booster sessions during a period of 5 years.

A meta-analysis reporting on 48 trials estimated an effect size (ES) of 0.35 (95% confidence interval, 0.40-0.70) for selective serotonin reuptake inhibitors over placebo.24 No differences were found between 4 different selective serotonin reuptake inhibitors. Meta-analyses examining the efficacy of selective serotonin reuptake inhibitors and tricyclic antidepressants consistently show no significant differences between the 2 drug classes.25-26 Therefore, we assume the same efficacy for all antidepressants.

From the figures presented in a recent meta-analysis27 of the odds ratios of relapse in 26 maintenance drug studies and 7 continuation drug studies, we derive a pooled relative risk of 0.416 (95% confidence interval, 0.312-0.553) for relapse with continuation AD drug treatment and 0.437 (95% confidence interval, 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 (95% confidence interval, 0.44-1.10), close to the Gloaguen point estimate but with wider confidence intervals. We use these figures in our main analyses. Excluding 2 outlier studies (by the same author) with particularly high ES estimates reduces the Q statistic for het-erogeneity from 50.8 (df=16; P<.001) to 22.3 (df=13; P=.051). In a separate sensitivity analysis, we recalculate the model using the ES (0.54; 95% confidence interval, 0.29-0.79) calculated after excluding these 2 outliers.

Although 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 1.3 to 4 years) after randomized controlled trials,20-26 that were set up to compare CBT with AD drugs in the acute phase, we calculate a lower risk of relapse after CBT (relative risk, 0.64; 95% CI, 0.51-0.79).

Maintenance CBT is described in 2 trials. The first compares CBT maintenance with AD drug maintenance and during 1 year of follow-up found no difference in relapse.21 The other reports on a trial of maintenance CBT after acute CBT.21 At a 2-year follow-up, the groups who had maintenance CBT had 29% relapse (3/20) compared with 80% (16/20) in the group receiving case management only after CBT (AD drugs were tapered off and discontinued in both groups). The scanty evidence from these 2 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% to 6% lower rates for selective serotonin reuptake inhibitors in comparison with tricyclic antidepressants.23,24,26,31 However, because most trials are of short duration, representing what is possible with motivated patients and physicians, adherence rates may be lower than reported in the controlled trial literature. Adherence in 4 studies of primary care ranges from 30% 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 30% adherence in community settings.

We have found one community study of the attrition rate of CBT for depression in which volunteers were recruited via the local media for a 12-week course of CBT. The total drop-out rate was 47%, with almost half of those dropping out in the first 3 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 DALY

The health benefit of interventions is measured in DALY, which is the sum of a nonfatal component determined by the severity-weighted time lived with depression and a fatal component, years of life lost, calculated as the stream of life lost because of suicide.

As described elsewhere,40 we use 2 methods to translate ESs from trial literature into a reduction in the disability weights. Briefly, the first method relies on an estimate of disability weight change for each SD change in severity of depression, which we call the conversion factor.41 Because the ES quantifies the impact of an intervention in SD units, health gain in DALY units can be calculated as the product of the ES, the conversion factor, and the duration spent in the health state. The second (survey severity) method applies the ES to the mental component score of the SF-12 across eligible respondents in the mental health survey, after which the difference in average disability weight 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 disability weight are only applied to the time from the commencement of the intervention, ie, taking into account that there is a lag to treatment-seeking after the onset of symptoms. A UK study found a median 10-week interval between onset and care-seeking for patients with an affective disorder.42 We cannot assume a similar lag because 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 the 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 disability weight is modeled.

From the point prevalence of depression in the NSMHWB, a UK estimate of the relative risk 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 that a relative risk of 1.8 from Swedish routine data collection systems44 applies to time lived with depression while not effectively taking AD drugs. In the absence of long-term studies, we assume that suicide rates are similar in patients receiving CBT as in those taking AD drugs.

From these estimates, we derive suicide rates in those currently receiving effective treatment and those ineffectively treated. As in the Australian Burden of Disease Study,2 the years of life lost associated with a death are calculated as the cohort life expectancy for each age and sex category. We then divide the sum of years of life lost 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 receiving these treatments, taking into account the estimated lag to treatment and level of adherence.

UNCERTAINTY

We use simulation-modeling 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 Corp, 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 intervals 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 modeling. 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.

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 and fitted lognormal distribution of duration of episodes 2.4 and the mean proportion of time spent in major depression is 20.8%. The mean proportion of time spent with depression during 6 months after an episode is 19.5%.

The mean disability weights for mental health survey respondents on evidence-based treatment (0.429), those consulting health professionals but not receiving evidence-based treatment (0.364), and those not consulting physicians (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 disability weight from 0.490 (the hypothetical level of severity without treatment) to 0.429 for current treatment strategies.

<table>
<thead>
<tr>
<th>Table 1. Model Input Parameter Values and Sources of Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td>Relative risk for suicide in prevalent depression</td>
</tr>
<tr>
<td>Relative risk for suicide while receiving treatment vs while not receiving treatment Effect size</td>
</tr>
<tr>
<td>Antidepressant drugs</td>
</tr>
<tr>
<td>Cognitive behavioral therapy</td>
</tr>
<tr>
<td>Relative risk for relapse during 6 mo continuation of antidepressant drug treatment</td>
</tr>
<tr>
<td>Relative risk for relapse maintenance treatment</td>
</tr>
<tr>
<td>Relative risk for relapse during 18 mo after cognitive behavioral therapy</td>
</tr>
<tr>
<td>Average duration of episodes</td>
</tr>
<tr>
<td>Time depressed</td>
</tr>
<tr>
<td>6 mo after episode</td>
</tr>
<tr>
<td>5 y after episode</td>
</tr>
<tr>
<td>Average number of episodes during 5 y after episode</td>
</tr>
<tr>
<td>Disability weight conversion factor</td>
</tr>
<tr>
<td>Adherence with Antidepressant drugs</td>
</tr>
<tr>
<td>Cognitive behavioral therapy</td>
</tr>
<tr>
<td>Lag to treatment</td>
</tr>
<tr>
<td>Proportion of affected individuals seeking care</td>
</tr>
<tr>
<td>Proportion of affected individuals receiving evidence-based treatments</td>
</tr>
</tbody>
</table>

*Triangular distributions indicate the lower limit of the 95% confidence interval, the point estimate, and the upper limit of the 95% confidence interval.
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 men and 0.3% in women. For both sexes combined, the risk of suicide in those taking medication or receiving CBT is 0.26%, and in those not treated it is 0.47%. This translates on average across all ages into an annual loss of 0.093 years of life lost if treated and 0.167 years of life lost without treatment, ie, a net health gain of 0.074 years of life lost that we attribute to treatment per year lived with depression.

During the first year after the 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).

If all those seeking care for an episode are offered 5 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 with 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 bring 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 ES. Altering the ES for AD drugs by 25% alters results by 3 percentage points for episodic treatment and less than 1 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 disability weight, and, to a lesser extent, the ESs.

Prevention of suicide contributes to almost a third of the amount of health gain in DALY for each of the 4 intervention scenarios in comparison with no treatment. In the episodic treatment scenarios, reduction in severity is the main impact of treatments, whereas in maintenance treatment the impact on preventing relapse contributes more to overall health gain than reduction of severity while depressed (Table 2).

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 during 5 years after an episode of major depression can be averted. The main reasons for this favorable outcome are that maintenance treatment prevents relapses and that relapses that do occur are being treated from the start rather than after a lag time to seeking appropriate care. Because 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 for treating all depression as a chronic disorder and not just those with recurrent or more severe episodes as recommended in current treatment guidelines.3,4

We have made a conscious choice to simplify our modeling by using averages, eg, for the severity of episodes and the duration of the index episode, and by modeling all ages and both sexes together. Some of these decisions do not do justice to the great complexity and variation in the manifestation 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, sexes, 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 scrutinize and apply to 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 2 different methods of determining health benefit. The difference in burden averted between current practice and alternative treatment options is less affected because the same imperfect method is used for each treatment scenario. More accurate measurements of change in health status that can be attributed to interventions require 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 NSMHWB 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. Because 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 during 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 endeavored to follow up people over time to examine if our modeled 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 during 5 years of follow-up is higher than that from a clinical study in the United States, which found that 15% of time was spent with depressive symptoms at the level of major depression during 9 years of follow-up.45 Our results are rather insensitive to this finding because the treatment impact measures applied to a 15% or 20% amount of depression during follow-up give similar estimates of the proportion of depression burden averted.

We have limited our analyses to major depression, ignoring that during follow-up time many people will spend time with subsyndromal symptoms or dysthymia.45 If we assume that 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 these would apply equally to all people with depression as we have modeled. However, our results are not very sensitive to the estimates of ES for AD drugs or CBT, and thus our conclusions would not alter even if the effectiveness of treatment in primary care cases is estimated to be as much as 23% higher or lower. The information we used from 2 European studies43,44 to determine the risk of suicide is not so strong. However, the inclusion of years of life lost 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 with 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 the World Health Organization, available at http://www3.who.int/whosis/mort) and that we estimated the risk of suicide only during depression and not for all follow-up time, our estimates are only marginally higher (by 12% in men and 20% in women) than the US estimates.

Despite the limitations associated with a 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 it is 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 US47,48 and Australia1 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 United States.

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 and inform priority setting.

Submitted for Publication: December 8, 2003; final revision received March 25, 2004; accepted April 22, 2004.

Correspondence: Theo Vos, MD, MSc, School of Population Health, University of Queensland, Herston Rd, Herston QLD 4006, Australia (t.vos@sph.uq.edu.au).

Funding/Support: This study was conducted as part of the Assessing Cost Effectiveness (ACE)—Mental Health project funded by the Australian Department of Health and 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, Canberra, Australia, and Gavin Andrews, MD, provided a revised scoring algorithm to determine ICD-10 and DSM-IV diagnoses.

REFERENCES