

A National Depression Index for Australia

Andrew Mackinnon, Anthony F Jorm and Ian B Hickie

Depressive disorders are now the most disabling illnesses in Australia¹ and make major contributions to premature death by suicide, injury and cardiovascular disease or other health problems.^{1,2} Community knowledge about the key risk factors, protective strategies and effective self-help or medical treatments for these disorders is limited.^{3,4} About half of those affected do not seek medical care.⁵

An effective strategy for promoting community debate on major health, social or economic issues is to report simple but high-impact summary statistics, such as the road toll or the consumer price index. In mental health, the only national data regularly reported are those for the annual suicide rate. Recent development of a national wellbeing index⁶ has proved to be a useful method for focusing media attention on mental health and other aspects of quality of life. As community awareness of *beyondblue*⁷ and other awareness programs has increased, and the reporting of these programs focuses increasingly on prevention of depression, there is a need for monitoring changes in actual depression rates and levels of depressive symptoms. We therefore sought to develop a statistically valid National Depression Index that could be readily understood by the general public.

To be useful, an index must be strongly related to clinically defined depression, yet be based on information that is easy to acquire in a range of situations.

METHODS

The steps involved in developing the National Depression Index are outlined in Box 1.

Subjects

Data for our study came from two extant surveys: the National Survey of Mental Health and Wellbeing (NSMHWB 1997⁸) and the 2001 National Health Survey (NHS 2001⁹). These are large (respectively, $n = 10\,641$ and $n = 17\,918$) surveys of adult Australians carried out by the Australian Bureau of Statistics. The NSMHWB was conducted in 1997 and involved persons aged 18 and over. The NHS 2001 is part of a series undertaken by the Australian Bureau of Statistics on a regular basis. The latter included participants under 18 years, but in our study only responses of adults were used.

Candidate index items

Potential items for the index came from the 10-item Kessler Psychological Distress Scale (K10).¹⁰ These items were chosen

Mental Health Research Institute of Victoria and Department of Psychological Medicine, Monash University, Parkville, VIC.

Andrew Mackinnon, PhD, Deputy Director.

Centre for Mental Health Research, Australian National University, Canberra, ACT.

Anthony F Jorm, PhD, DSc, Director.

Brain and Mind Research Institute, University of Sydney, Rozelle, NSW.

Ian B Hickie, MD, FRANZCP, Executive Director; and Clinical Advisor, *beyondblue: the national depression initiative*.

Correspondence: Professor Andrew Mackinnon, Mental Health Research Institute of Victoria, Locked Bag 11, Parkville, VIC 3052.

ajm@mhri.edu.au

ABSTRACT

Objective: To develop a National Depression Index for measuring the depression status of the Australian population.

Design: Cross-sectional data were analysed from two random samples of the Australian adult population — the National Survey of Mental Health and Wellbeing (2000) and the National Health Survey (2001).

Participants: The National Survey of Mental Health and Wellbeing (2000) — 10 641 participants; and the National Health Survey (2001) — 17 918 participants.

Main outcome measures: Selected items from the Kessler Psychological Distress Scale (K10); and diagnoses of a major depressive episode according to DSM-IV criteria using a computerised interview.

Results: Six items from the K10 that were most closely related to the DSM-IV diagnosis of “major depressive episode” were identified. Scores on an index calculated from these items were rescaled to form an index reflecting relative risk of depression and having a value of 100 for the Australian adult population. Taking into account sex, employment status and income, index values were higher in younger people, females, unemployed people and those socioeconomically disadvantaged. This pattern provides additional support for the validity of the index, as well as establishing benchmark levels to which index values from future surveys and in other groups may be compared.

Conclusions: The proposed *National Depression Index* is a valid indicator of depression and level of depressive symptoms. It is suitable for monitoring depression at the population level. The scaling characteristics of the measure ensure that it can be interpreted by members of the general public.

MJA 2004; 181: S52–S56

because they cover core depressive symptoms, have been shown to have good ability to detect non-specific psychiatric distress (as indicated by high values for the areas under their receiver operating characteristic curves),¹⁰ and were utilised in both the NSMHWB and the NHS 2001.

The items shown in Box 2 were administered in the two surveys, with each question commencing “During the past 30 days, how much of the time did you feel . . .”. Responses by the individuals sampled were on a five-point Likert scale running from “all of the time”, “most of the time”, “some of the time”, “a little of the time” to “none of the time”. “Item 3” was administered contingent on the response to “Item 2”, so that respondents who indicated that they were “nervous” “none of the time” were not asked whether they were “so nervous cannot calm down”. In the analyses below, responses to the contingent item were imputed as “none of the time”. This pattern of contingency also applied to “Item 6”.

Criterion diagnoses

The NSMHWB dataset includes diagnoses of depression meeting DSM-IV and ICD-10 criteria prevalent over the preceding 1 month or 1 year. Diagnoses were made using the computerised version of the Composite International Diagnostic Interview (CIDI-A).¹¹

Development of the Index

Development of the National Depression Index involved four steps.

1. *Confirmatory factor analyses* (Box 3) were undertaken of all K10 items using data from the NSMHWB. A diagnosis of depression (case/non-case) was also included in these analyses. The loading of the depression diagnosis on a single common factor was fixed at unity, while the K10 items on this factor were allowed to vary (see Box 2). This structure effectively aligns the factor with clinically defined depression, so that the loadings of the K10 items on this factor are those on a depression factor. A number of different diagnoses of mood disorder were used in separate factor analyses. These included ICD-10 and DSM-IV criteria for depression covering 1-month and 1-year prevalences, and for dysthymia. Diagnoses that differentiated between levels of severity were coded as “mild”, “moderate” or “severe”. The statistical program Mplus 2.1¹² was used for this analysis. Weighted least-squares estimation, with sampling weights reflecting the structure of the Australian population, was used together with robust standard errors and mean- and variance-adjusted χ^2 test statistics.

2. *Ways of using the items that loaded highly on the depression factor were investigated.* There are two basic types of scores that can be derived for individuals from the above analysis. These are (i) the sum of responses to the items on their original Likert scales, and (ii) the factor scores. Either of these scores could potentially be transformed to make them more tractable or interpretable. An index based on the sum of item responses has obvious simplicity both in its calculation and interpretation. However, factor scores make greater use of the available information.

3. *A transformation of the score was developed to facilitate easy understanding.* We considered several kinds of linear and non-linear transformations, but eventually settled on a “relative risk index” as being the simplest to interpret.

4. *Data from the NHS 2001 were used to standardise the National Depression Index.* NHS data were used in preference to the NSMHWB because the NHS is a regular survey. This will permit the construction of a time series for the National Depression Index. Methodological differences between the two surveys mean that the NSMHWB data cannot readily be used to provide a base for later NHS data.

RESULTS

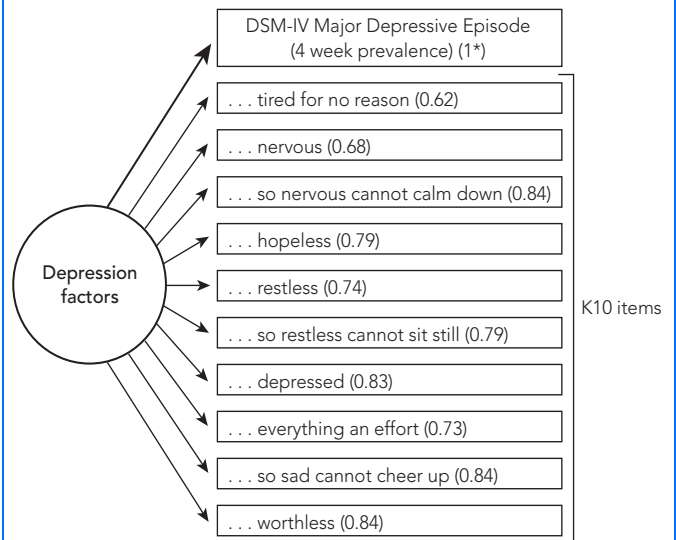
Factor analyses

Loadings on the factor model varied very little with the use of different diagnoses as the “anchor”. The loadings of all items were substantial. Those using 4-week prevalence of DSM-IV “major depressive episode” are shown in Box 2. All loadings were highly significant, reflecting the large sample size, and their standard errors were very small and consequently are not shown. A number of indices of model fit lay outside values indicative of acceptable fit. This partially reflected the large sample size, but also suggested that an adequate model incorporating all K10 items would require a more elaborate structure. Because the purpose of fitting this structure to the data was to identify items most strongly associated with depression, this was not a primary consideration. It was possible to identify a cluster of items having elevated loadings, the lowest of which was just below 0.80. Adopting these items as candidates for the index led to the exclusion of the stem items “nervous” and “restless” that are repeated in a more qualified format. Also excluded

1 Steps in constructing the National Depression Index

- Identify a database with the following characteristics:
 - contains a representative sample of the Australian population
 - includes diagnoses of depression according to recognised clinical criteria
 - contains survey items tapping or related to symptoms of depression
- Determine which survey items are most closely related to a depression construct defined using confirmatory factor analysis
- Calculate the location (factor scores) of individuals on the depression construct using the selected items and appropriate software
- Devise a method of expressing factor scores in a readily interpretable manner
- Calibrate the index to have a value of 100 in a second representative sample unrelated to the one used to construct the index
- Establish index values in subgroups of the population

2 Depression factor



Depression factor is defined by diagnosis of 4-week prevalence of DSM-IV “major depressive episode”, with loadings of the Kessler Psychological Distress Scale (K10) items shown in parentheses. Each item begins with “During the past 30 days, how much of the time did you feel ...”.

* The loading marked is fixed and defines the factor.

were two items concerning effort and being tired. This eliminates some redundancy in the K10. The fit of a single-factor model comprising just these six items was very good. The χ^2 value of the model was statistically significant ($\chi^2 = 401.20$, $df = 8$, $P < 0.0001$), reflecting the large sample size, but other indices had values indicating an excellent fit: comparative-fit index, 0.98; Tucker-Lewis index, 0.99; root mean square error of approximation, 0.07.

Scoring the scale

Box 4 shows the distribution of factor scores calculated from the index items. These values serve to locate each individual on the

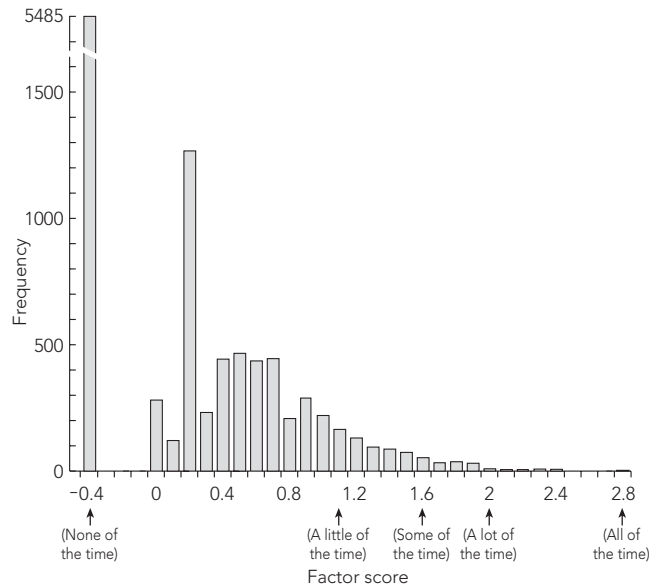
3 Confirmatory factor analysis

Factor analysis is a technique in which the existence of underlying dimensions (factors) is inferred by the statistical examination of patterns of correlations between observed variables. Early forms of factor analysis were exploratory tools. In the development of the National Depression Index, a more advanced form of factor analysis — confirmatory factor analysis — was used. In this technique the investigator specifies *a priori* a model involving one or more factors and stipulates which variables are related to each dimension. Data are fitted to the specified pattern to obtain estimates of factor loadings and statistics describing the fit of the model. The factor loading of a variable represents the correlation between that variable and the underlying dimension. Fit statistics are used to judge whether the pattern specified adequately represents the relationships in the data. If a model does not fit the data, it may be modified in a number of ways, such as allowing a variable to load on a factor on which it previously did not.

Specialised software is required to fit confirmatory factor models. The variables used in our study included binary (case/non-case) diagnosis status and responses to questions about symptoms, which were recorded on scales with only five points. Normal Pearson correlation coefficients are unsuitable for such data and so polychoric correlations were used. These estimate the correlation between the underlying continua on which these responses, with their limited resolution, were made.

Once a satisfactory model has been developed, it is possible to calculate factor scores. These scores estimate the location of each individual on each of the factors.

4 Distribution of factor scores of items comprising the National Depression Index



Factor scores corresponding to average endorsement levels of each point on the Kessler Psychological Distress Scale (K10) response scale are indicated with arrows.

depression factor. Consistent with responses to other inventories of depressive symptoms in epidemiological samples, the factor scores follow an exponential distribution: most individuals have a very low score, with decreasing numbers having scores indicating more severe depressive symptoms. We also examined the distribution of a scale score formed by summing responses to the six items. While this distribution showed a similar overall form, factor scores had greater ability to differentiate between individuals with differing levels of mild depressive symptoms. On this basis, subsequent development of the index proceeded using factor scores.

A depression index interpretable as relative risk of caseness

This index was developed by fitting a logistic regression model that predicted DSM-IV “major depressive episode” (4-week prevalence) from factor scores calculated from the index. As would be expected, given the manner in which the items for the index were selected, there was a very strong relationship between caseness and index factor scores (odds of caseness associated with a unit increase in factor score, 30.50; 95% CI, 23.63–39.37). Most importantly, the Hosmer–Lemeshow statistic indicated that the logistic model provided a good fit to the data ($\chi^2 = 2.69$; $df = 3$; $P = 0.44$).

Probability of caseness for each individual was calculated using parameter estimates from the model in the equation:

$$\text{Pr}(\text{case}) = \frac{1}{1 + e^{(5.728 - 3.418 \times \text{factor score})}}$$

The main effect of this procedure is to smooth the probability of caseness for given values of the factor score. Even though the total sample was very large, the number of cases of depression was

relatively modest ($n = 340$). Thus, the number of cases at each value of the factor score was quite small. In the NHS 2001 sample, the mean predicted probability of caseness estimated for the Australian population was 0.048. Predicted probabilities were divided by this value. This enables index values to be interpreted as relative likelihood of caseness. Index values were multiplied by 100 to eliminate the need to work with and report decimal values. Values greater than 100 indicate greater likelihood of depression than in the population as a whole, while lower values indicate lower likelihood. Scripts and files for calculating index values are available at <www.anu.edu.au/cmhr/NDI>.

Index values in subgroups

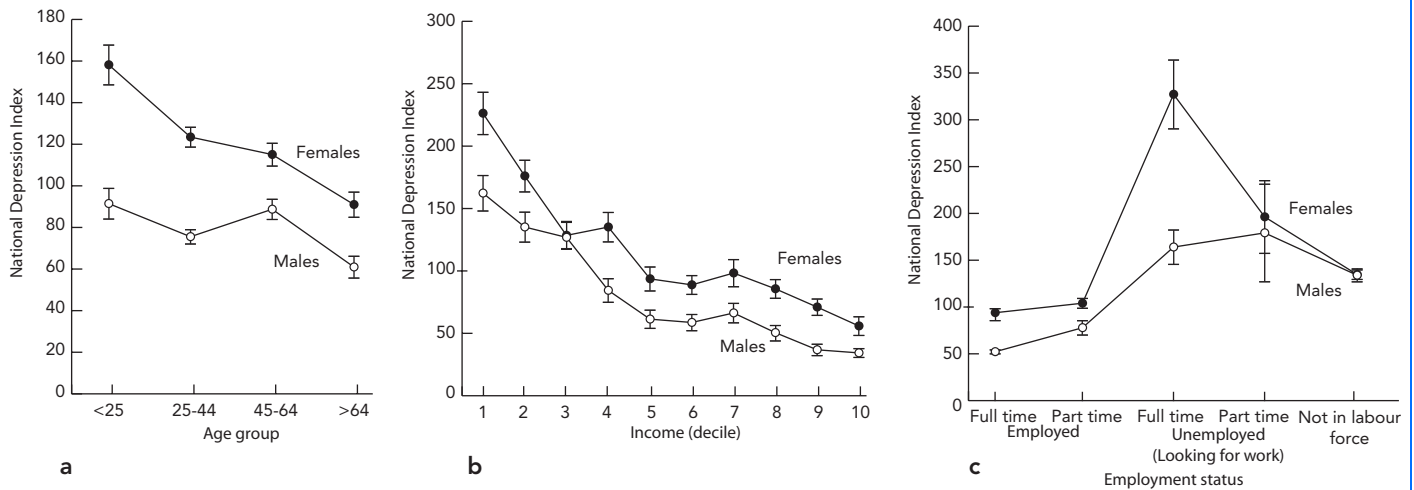
It is essential that any index is sensitive to the known characteristics of the disorder it taps.

In addition, the calculation of index values for subgroups of the Australian population in a large, representative sample creates benchmarks to which other groups may be compared and by which changes over time may be identified.

By definition, the mean index value for the Australian population estimated from the NSMHWB data was 100 (SD, 273.16). Box 5 shows mean values of the National Depression Index by sex and age group, income level, and employment status, respectively. Respondents who reported that they were taking prescribed antidepressant medication had higher values than those who did not (mean, 434.00; SD, 587.68; $n = 839$ versus mean, 83.58; SD, 235.74; $n = 17\,078$). No significant difference between persons living in major cities, inner regional areas and other, more remote locations was found ($F_{(2,17\,801)} = 1.25$, $P = 0.29$).

The cross-sectional nature of our study means that caution must be exercised in making causal inferences from these results.

5 National Depression Index (a) by sex and age, (b) by sex and income (higher deciles have greater income), and (c) by sex and employment status (error bars indicate ± 1 SE)



Error bars indicate ± 1 SE.

However, the pattern of index values across these subgroups is comparable to established patterns of response on other inventories and to the relative prevalence of depression in these groups.⁸ This serves to validate the index as a measure reflecting established associations between demographic and personal variables, and depression. It should be emphasised that the index is intended for characterising populations rather than individuals or small groups (eg, clinical samples). If the index is applied to small clinical samples, it may prove overly sensitive to the presence or absence of highly symptomatic individuals.

DISCUSSION

We have constructed a relative risk index for depression that has maximum possible validity as an indicator of DSM-IV-defined “major depressive episode” given its brevity and the available item pool. The use of factor scores rescaled to permit interpretation in terms of likelihood of depression relative to a large, well ascertained probability sample of the general community has achieved a high level of differentiation along the continuum tapped by the index items. From a community education perspective, it has the additional advantage of straightforward interpretation of index values. For the whole population, the level of the index is set at 100. The levels and dispersion of index values then vary with key subgroups of interest. Scores above 100 indicate populations or persons at greater risk (eg, females, the economically disadvantaged, and unemployed persons), while scores below 100 indicate those at lower risk.

Having developed the index, we will seek to utilise it in specific subpopulations that are perceived to be at risk or are the subject of specific depression prevention or early intervention programs (eg, unemployed persons, persons of low socioeconomic status). We will also seek to use it in groups that have not been subject to specific comparison with large normative populations (eg, persons living in smaller rural communities, women presenting in the antenatal period, and persons with comorbid substance misuse).

As yet we have not addressed another key feature, namely whether the index is responsive to genuine changes in population

prevalence. As both the existing national datasets involving the K10 were cross-sectional, we have no source for preliminary study. While the NSMHWB and the NHS 2001 were administered 3 years apart and returned different item endorsement rates, this is likely to have arisen from differences in the way the questions were administered rather than a substantial change in the prevalence of depression in the Australian community. These discrepant results also argue for our central proposition, namely that a reliable tool should be administered in the same manner on regular occasions to determine genuine variations in the population (or key subpopulations’) prevalence of depression. It is for this reason that we used the NHS 2001 to establish the baseline of the index. Our expectation is that, while the National Depression Index will demonstrate small variations over time, larger changes may occur within specific populations that are subject to specific depression prevention or early intervention programs. Through the funding mechanisms of *beyondblue* we will therefore seek to embed the index items within such longitudinal intervention projects. If, indeed, reduction in depression prevalence is evident, then this can be readily communicated to the general public and help to overcome one of the persistent myths about common mental disorders, namely that little can be done to reduce their burden within the Australian community.

ACKNOWLEDGEMENT

This work was supported in part by *beyondblue: the national depression initiative*.

COMPETING INTERESTS

None identified.

REFERENCES

- 1 Mathers C, Vos T, Stevenson C. The burden of disease and injury in Australia. Canberra: Australian Institute of Health and Welfare, 1999. (AIHW Catalogue No. PHE-17). Available at: <http://www.aihw.gov.au/publications/health/bdia/bdia.pdf> (accessed May 2004).

- 2 Bunker SJ, Colquhoun DM, Esler MD, et al. "Stress" and coronary heart disease: psychosocial risk factors. National Heart Foundation of Australia position statement update. *Med J Aust* 2003; 178: 272-276.
- 3 Highet NJ, Hickie IB, Davenport TA. Monitoring awareness of and attitudes to depression in Australia. *Med J Aust* 2002; 176 (5 Suppl): S63-S68.
- 4 Jorm AF. Mental health literacy: public knowledge and beliefs about mental disorders. *Br J Psychiatry* 2000; 177: 396-401.
- 5 Andrews G, Henderson S, Hall W. Prevalence, comorbidity, disability and service utilisation: overview of the Australian National Mental Health Survey. *Br J Psychiatry* 2001; 178: 145-153.
- 6 Cummins B. Introducing the Australian Unity Wellbeing Index. Available at: http://acqol.deakin.edu.au/index_wellbeing/index.htm (accessed May 2004).
- 7 *beyondblue: the national depression initiative*. <http://www.beyondblue.org.au>. Accessed May 2004.
- 8 Andrews G, Hall W, Teesson M, Henderson S. The mental health of Australians. Canberra: Department of Health and Aged Care, 2000.
- 9 Australian Bureau of Statistics. National Health Survey 2001: summary of results. Canberra: Australian Bureau of Statistics, 2002. (Catalogue no. 4364.0.)
- 10 Kessler RC, Andrews G, Colpe LJ, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychol Med* 2002; 32: 959-976.
- 11 World Health Organization. Composite International Diagnostic Interview — Version 2.1. Geneva: WHO, 1997.
- 12 Mplus [computer program], version 2.1. Los Angeles, Calif: Muthén & Muthén, 1998.

(Received 11 Nov 2003, accepted 1 Jun 2004)

