Draft: work in progress—do not cite or quote

Global Burden of Disease measures for depression - time for a rethink

Petra Brhlikova, Allyson Pollock, Rachel Manners
Centre for International Public Health Policy
University of Edinburgh

Abstract:

The Global Burden of Disease (GBD) Project has been developed in collaboration between the World Bank and WHO with the aim of providing unbiased and comparable summary measures of population health to inform priority setting in health policy interventions at national and international level. Since its first publication in 1993, the GBD project has become an important and widely used source of information on the burden of specific diseases; the WHO regularly publishes GBD estimates; and countries use national burden of disease estimates to adjust their national health policies and priorities to target specific disease interventions and treatments (National mental health programme in India, Tanzania Essential Health Interventions Project).

This paper reassesses the nature of the epidemiological evidence underpinning one of the GBD topics: the estimate for the global burden of depression. Specifically, we look at the quality and representativeness of data on prevalence and incidence of depression used for the 2000 GBD estimates, which stated that neuropsychiatric disorders contribute around 14% to the total disease burden and ranked depression as the fourth major cause of disease burden worldwide. These numbers are often cited to support further research and various initiatives in mental health - the 2001 World Health Report was devoted to mental health and in 2007 the Lancet published a series on Global Mental Health – and potentially divert resources from other health care priorities of developing countries.

1. Introduction

The Global Burden of Disease (GBD) Project has been developed in collaboration between the World Bank and WHO with the aim of providing summary measures of population health to inform priority setting in health policy interventions at national and international level. Since its inception in 1993, the GBD project has become an important and widely used source of information on the burden of specific diseases; the WHO regularly publishes GBD estimates; countries use national burden of disease estimates to adjust their national health policies and priorities to allocate resources, and to target specific disease interventions and treatments (National mental health programme in India, Tanzania Essential Health Interventions Project) (1).

Mental disorders were ranked as a leading cause of disability even in low and middle-income countries, with depression contributing 3.7% of all disability adjusted life years (DALYs), in the 1993 estimates (2). The 2000 GBD estimates stated that neuropsychiatric disorders contribute around 14% to the total disease burden (3) and ranked depression as the fourth major cause of disease burden worldwide. The 2001 World Health Report was devoted to mental health (4) and in 2007 the Lancet published a series on Global Mental Health, calling for increased attention to mental health globally, particularly in low and middle-income countries (5) citing Global Burden of Disease as an important reason for the focus on mental health. One year later, the WHO published its strategy for 'Scaling up care for mental, neurological, and substance use disorders' – the Mental Health Gap Action Programme (6) - and a website of an international scientific and social movement on global mental health was launched (7).

However the GBD is not without its critics. The value judgments used in weighing DALYs (8;9), complexity of the methodological approach that makes the interpretation of results difficult, and the absence of comprehensive, reliable and high quality data have all been highlighted. In particular, questions have been raised about the usefulness of the estimates for sub-Saharan Africa in health policy and the uncritical referencing of GBD in academic studies:

> "Can useful policy decisions for any region of the world be made without knowing any validity of the burden of disease measures? … The GBD study is rapidly being held up as the reference standard, despite the caveats in the accompanying multi-volume text. If these data are wrong the consequences are likely to be most damaging for the very populations under-represented in the fact-gathering process" (10).

In 2005 Mathers (11) conducted an uncertainty analysis for 16 leading causes of global burden of disease in 2000-2002 which showed that uncertainty arising from epidemiological estimates on incidence and prevalence of disease and severity distribution and from the extrapolation from sub-population to regional population ranged from 10 to 90 percent with a median of 41%. Mathers concluded that uncertainty was highest for low income countries and for diseases without appropriate health surveillance systems and there was a need for better epidemiological data and caution when interpreting global comparative epidemiological assessments. Nevertheless, Lopez et al. concluded that "[d]espite these uncertainties, the results of the 2001 GBD study suggest that further gains in health in poor countries could be achieved in relatively short time" (12).

In this paper we reassess the epidemiological evidence which underpins the estimate for the global burden of depression in 2000 published in the GBDep study (2) and used by the WHO.[1] Our three objectives are: to assess the representativeness of local studies in respect of regional estimates of burden of disease, to evaluate the studies against the four criteria specified in the GBDep study for

[1] New GBD estimates have been published in 2008 but no additional sources of epidemiological data on depression were added.

inclusion into GBD estimates, and to discuss the quality of data sources used in the studies.

2. Methods

The GBD estimate for depressive disorders in the year 2000 are derived from two key epidemiological sources: a WHO website publication described as the 'version one estimates' (13) and the GBDep study (2) which updates and excludes several of the data sources. Our study examines and evaluates the data sources listed in the second study.

Inclusion criteria used in the GBDep study: sample size; well defined study period; given age and sex distributions; clear random sampling methods.

Data Sources: Data sources for the GBD estimates were identified through a systematic literature search (2). The GBDep study drew on 24 published studies and 4 unpublished studies where information was collected via personal communication. We attempted to contact the authors of the unpublished studies but were only able to contact the author in Nepal. The authors of the studies conducted in Morocco, the Czech Republic and Russia did not reply and it is not known whether their data are unpublished studies or expert estimates.

The 24 published studies include 19 epidemiological country studies and 5 multi-country studies – WHO Multi-country Survey 2000-2001 (14), WHO International Consortium in Psychiatric Epidemiology 2000, Weissman et al. (1996), Copeland et al. (1999), and Ayuso-Mateos et al. (2001). These multi-country studies were used as sources of data on depression for 19 counties. With the exception of WHO Multi-country Survey 2000-2001, all multi-country studies drew on primary country-level studies published separately. We attempted to identify primary sources and complete information on sample in cases when secondary sources did not report all characteristics of interest.

We assigned each published and unpublished study to their country of origin and GBD region as adapted from the WHO. For the purpose of our analysis, country-level data from multi-country studies are treated as separate studies, i.e. the umbrella studies are ignored. For each study the population sample was extracted, if possible, and compared with the regional and country-level estimates of population derived from World Health Report 2000.

Data quality was analysed with respect to completeness and representativeness against the national population; study design and the four inclusion criteria used in the GBDep study; interview process – by medical practitioners or not; and instrument. In the present study, we summarised data on interview process; their analysis is subject to a separate study.

3. Analysis and results

3.1 Types of Study

Of the 42 country studies included in the GBDep study, two were incidence studies (Canada and Norway); the information could not be obtained for six studies. The remaining 34 studies were prevalence studies but in 18 country studies the duration of prevalence was unclear or unknown.

3.2 Representativeness of studies for regional estimates

The GBDep study required "a clearly specified method for sampling (a design that would yield a probabilistic national/regional representative sample) and implementation" (2). The 28 sources cited in the GBDep study give country data on 40 of the 191 countries in six regions (Table 1). No WHO region had complete country coverage. EURO has 15 out of 52 countries covered compared with only three out of 46 countries in the AFRO region, four in SEARO, three in EMRO, eight in AMRO, and seven in WPRO (see Table 2 for the list of country-level sources for each region).

Some countries did not have nationally representative data and local data based on a small village would be extrapolated to country and regional level. Of the three AFRO studies none was nationally representative. In EURO, the region with the best coverage, six out of 15 studies had a nationally representative sample. The number of individuals sampled in each region is very small ranging from 4.9 per 100,000 in EURO to 0.2 per 100,000 in AFRO.

3.3 Quality of data sources and comparison with GBDep criteria

Study objectives

All studies except the WHO multi-country survey, which provided data for 6 countries, aimed to determine prevalence or incidence of depression and other mental disorders in specified populations. However, the WHO multi-country survey (2000-2001) focused on health systems performance, measuring population health, health inequality, fairness in financial contribution, responsiveness and efficiency and data on prevalence of depression are not presented in the final publication.

Sampling frame

*A*ll 36 studies for which we have data stated their target population, but in sixteen the sampling frame was not clear. Of the 30 studies which presented their sampling method, 29 used random sampling and one addressed the whole targeted population.

Only twelve studies of the 36 (33%) used a nationally representative sample[2]; the other thirty studies were conducted in a region, town or village. Thirty studies sampled adults aged eighteen and over, one study was of children and three studies surveyed over-65s only. For eight studies the age group was not given, was unclear, or we were unable to contact authors of the study.

Sample size

For 17 of the 42 (41%) studies either the GBDep criteria were not met or the sample size was not known. Of those 17 studies, ten had a sample size less than 1000 (three AFRO studies, studies from Canada, Lebanon, Finland, Greece, Pakistan and two German studies). For the remaining seven we could not determine the sample size. These included Morocco, the Czech Republic, Russia and Nepal – where data had previously been sent by personal communication to authors of the GBDep study. We could not access the Peruvian study; and the data and sample for Japan and Cambodia are not described in sources referenced in the GBDep study.

Response rates

Various methods of calculating response rates were used and in several cases the reported response rate conflicted with the initial and final sample that was reported. Reported sample sizes, numbers of respondents and response rates are summarised in Table 2. Based on the reported numbers, response rate was at or over 90 percent for nine studies, between 70 and 90 percent for eighteen studies, less than 70 percent for four studies and unknown for eleven studies.

Sampling methods and interviewers

The epidemiological studies typically used multistage sampling method with face-to-face interviews with psychiatrist-lead interviews in the second stage and nurse or other trained interviewers involved in the first screening phase. The comparative cross-country study by Weissman et al. (16) that provided

[2] From the twelve, data sets from Mexico and Turkey might not be nationally representative. The WHO International Consortium in Psychiatric Epidemiology (15) reports that "the data sets in Canada, Germany, the Netherlands, and the USA were weighted to adjust for differences between the sociodemographic characteristics of the samples and the populations from which they were selected. These adjustments were not possible in other data sets owing to lack of population data" (p.415).

data sets for eight countries was a two-stage study (but without the screening phase) and the interviews were conducted by trained lay people. The WHO multi-country survey (14) that provided data sets for six countries used different sampling strategies as well as different modes of data collection (household individual interviews, household brief face-to-face interviews, computer assisted telephone interview, postal self-administered surveys; in Egypt, the survey was hand-couriered to the respondents and personally collected from them).

Epidemiological measures of depression: Measuring depression in the Community
The data sources included in the GBDep study used different depression scales and measures, which can be divided into two main groups: structured interviews can be undertaken by lay people, and do not allow room for clinical judgement and additional questioning in order to make them highly standardized (17), including the 'Diagnostic Interview Schedule' (DIS) and the 'Composite Diagnostic Interview Schedule' (CIDI); semi-structured are undertaken by psychiatrists with room for additional questioning and are not as standardized, such as the Present State Examination (18).

The most common measures used were CIDI (in 13), DIS (in 9) and the Present State Examination (in three). Many studies use different versions of the same measures and in the WHO International Consortium in Psychiatric Epidemiology study (15) which was the source of data for Mexico and Turkey, the use of different versions of CIDI was acknowledged as the source of wide variation in lifetime prevalence estimates across countries.

Multi-country studies
Although all multi-country studies attempted to collect comparable data, methods varied across countries. Sampling methods and results for individual countries were reported and published separately and the multi-country studies then referred to primary data sources (see Table in appendix) but reported 'processed' data. E.g., the Weissman et al. prevalence estimates from all study countries are adjusted to the US age and sex distribution rather than the country of origin (16).

4. Discussion

Comparability and generalisability of available data
A key difficulty in evaluating the data was the lack of consistency and proper referencing between the two main sources (2;13) and inaccessibility of data including 4 unpublished studies. There were only two incidence studies, one in Canada (19) and one in Norway (20). The GBDep study converts all prevalence data into point prevalence and then into incidence data. The methodology of the conversion was not fully explained in the GBDep study. The average episode duration of six months was derived from three studies, two from the USA (21;22), and one from the Netherlands (23) and none from other countries or regions. Thus the duration of depression was taken from three countries in the West and applied across the world to all regions.

Further, although the authors of the GBDep study state that "a systematic review of all available published and unpublished papers of meaningful population studies on depressive disorders" (2) was conducted to obtain up-to-date prevalence and incidence data, we could not trace all the studies. For data collected via personal communication we could contact only one author and we do not know whether the data from three unpublished studies were based on population-based studies or on 'informed estimates.' We could not locate data on Japan and the sources of data on Cambodia listed in the GBDep study did not contain any information on sample and prevalence or incidence of depression.

Only 40 out of 191 countries were included in the global burden of depression measures. The AFRO region had least coverage, with data on only 3 out of 46 countries relating to a total of 1,000 people. Likewise the South-East Asia region had data on only 6,000 people and 4 countries. Only 12 twelve

countries had nationally representative samples; all of the other country data are based either on unrepresentative samples or extrapolated from data from other countries.

For two regions (AFRO and SEARO), none of the studies met the requirement of 'a probabilistic national/regional representative sample' (2) and for 20 studies the sample was sometimes only a village, small area, or based on a GP register limiting the applicability of the data to the rest of the country. Twelve of the 42 studies fail to meet the sample criteria, with the three AFRO studies based on populations of 180, 356 and 501 participants respectively. The authors do not provide an explanation for their decision to include smaller studies nor were details given of studies which were excluded.

While 27 studies had a response rate of 70%, for 11 studies, the response rate was not given or unclear and one of the German studies had a response rate of 27% (24). These studies should not have been included.

The lack of country representation and proper samples must be a cause for concern when health interventions, resource allocation and health priorities are being set in response to the summary measures.

Depression Measures

The GBD studies are also problematic for their reliance on different depression measures many of which use different thresholds for reporting and different methods, including staff, for collecting the data. Other cross-national studies highlight this as a limitation of the applicability of the results (15;16) but the GBDep study does not mention it (2) even though it reduces the reliability of the estimate.

The importance of the socio-cultural context in which measures are used and the problems of using standardised tools as a substitute for clinical judgement has been highlighted previously (25). Sadness or depression may be associated with a range of factors including poverty and distress (26), out-with the limited model of economic growth. It is problematic to apply a standard weight to a disease like depression, which is experienced subjectively. In different cultures and economies the weight, significance and importance of different diseases may vary and so may the ability to diagnose. For example, Arnesen and Kapiriri showed that when a different weight is used for uni-polar depression, the burden of depression decreases significantly in comparison to developmental disability due to malnutrition (27). "If consideration of context is abandoned, all normal sadness responses can be seen as a sign of pathology; the very possibility of normal sadness is lost" (26).

While the Global Burden of Disease project as a whole acknowledges that there is a serious need for more reliable data (12), as Cooper et al. (10) point out, the results of the global burden of disease are largely treated as straightforward facts, without considering their limitations. A typical response to criticisms of methodology is:

> "Disciplines such as demography and economics often aim to make the best possible estimates using the available data … GBD 1990 has been criticised for using 'estimates' rather than 'actual data'. This is not a relevant discussion for comparative burden estimates because all epidemiological data relating to population are 'estimates of varying degrees of precision or uncertainty'" (2).

The question is why scientists and epidemiologists are prepared to depart from their normal standards and rely upon non-existent or unsatisfactory data. One explanation is that lack of evidence forces a make-do attitude. Another is that researchers themselves are coopted into the project by funding and academic requirements. Besides, studies publishing new GBD estimates have an impressive citation record. For example, according to Scholar Google the GBDep study (2) published in the British Journal of Psychiatry in 2004 had been cited in 198 studies by August 2008. Other articles in the same

issue of the journal have from 6 to 71 citations. The high number of citations is, however, not uncommon for other GBD estimates: e.g., Kearney et al. study published in the Lancet in 2004 (28) had 478 citations by August 2008.

The Disease Control Priorities Project (DCPP) initiated by the WB and backed by the WHO and the Bill and Melinda Gates Foundation among others builds on GBD estimates and analyses cost-effectiveness of available treatments for specific diseases. In 2006 the DCPP recommended the administration of tricyclic antidepressants in primary healthcare centres as the most cost-effective treatment for depressive and anxiety disorders (29;30). This approach can lead to over-medicalisation of disease, over treatment and a focus on pharmacological interventions rather than integration of culturally-specific and sensitive services for mental illness into existing health-care systems (3).

5. Conclusion

The current evidence to support the epidemiological calculations of global burden of disease estimates for depression are lacking and are based on inaccurate and often substandard data. The reductionist approach deriving single composite measures is highly problematic, as it conceals and hides uncertainty, biases and distortions in epidemiological data. On this basis the GBD measure for mental illness/depression is not a good basis for deciding on national or international priorities and resource allocation. The diagnosis, measurement approaches and response to mental illness should be part of national public health strategy and not global policy.

Table 1: Population coverage in the depression estimate, by region

WHO region	Countries in Region[a]	Number of countries included in the GBDep study	Total population of region[b]	Population in GBD depression estimate[c]: number of respondents	Regional coverage by population (in %)
AFRO	46	3	616,438,000	1,029	0.0002
AMRO	35	8	813,061,000	29,500	0.0036
EMRO	21	3	484,488,000	7,783	0.0016
EURO	52	15	872,622,000	42,987	0.0049
SEARO	10	4	1,508,241,000	6,016	0.0004
WPRO	27	7	1,673,575,000	40,084	0.0024
Total	191	40	5,968,425,000	127,399	

[a] Source: Mathers (2004)
[b] Source: World Health Report 2000 (31)
[c] Source: GBDep study (2); does not include data from Peru, Japan, Cambodia, Russia, the Czech Republic, Nepal or Morocco
Note: In the GBDep study (2), Pakistan and Singapore are included in the SEARO region instead of EMRO and WPRO respectively as per WHO regions. Therefore, the recalculated population numbers are: EMRO 332,157; SEARO 1,664,094; and WPRO 1,671,053. Recalculated numbers for the population of the region covered (in %) are: EMRO 0.0023; SEARO 0.0004 (unchanged); and WPRO 0.0024 (unchanged).

Table 2: Characteristics of included studies

Country	Study Design	Sample Size [No. of respondents] (age range)	Sample Population (sampling framework)	Random Sampling	Response rate (%)
AFRO					
Zimbabwe (32)	Prevalence (one month and one year)	181 [172] (18-65)	Women in a township - Harare city (dept. of work data- 5,000 households)	Yes	95
Lesotho (33)	Prevalence (point)	456 [356] (over 18)	One village	Whole population	78
Ethiopia (34)	Prevalence (lifetime and one month)	600 [501] (15-85)	Rural Butajira region ('Butajira rural health project')	Yes	85
AMRO					
USA (22)	Prevalence (lifetime and 12 month)	[8098] (15-54)	Nationally representative (not given)	Yes	82.6
USA (35)	Prevalence (3 month)	4500 [1338] (9, 11, 13)	Rural part of North Carolina (school records)	Yes	Not given
Canada (19)	Cumulative Incidence	489 [462] (40+)	One region (unclear)	Not stated	Unclear
Puerto Rico (36)	Prevalence (lifetime and six month)	2036 [1513] (18-64)	Nationally representative (unclear; 1701 household based)	Yes	91
Brazil (37)	Prevalence (lifetime)	6476 (14+)	Three urban areas (census data; housing units)	Yes	Not given
Mexico (15)	Prevalence (lifetime, 12 month and 30 days)	[1734] (18-54)	Mexico City (unclear)	Yes	60.4
Columbia (14)	Prevalence (type unclear)	6000 (18+)	Nationally representative (National Institute of Statistics)	Yes	82
Chile (38)	Prevalence (one week)	4300 [3870] (16-64)	Urban Santiago (unclear; Chilean Institute of National Statistics)	Yes	90
Peru (39)	Unable to access data				
EMRO					
Lebanon (16)	Prevalence (lifetime)	526 (18+)	Beirut (unclear)	Yes	77
Morocco (personal communication)	Unable to contact the author				
Egypt (14)	Prevalence (unclear)	8000 [7257] (18+)	Nationally representative (Central Agency for Public mobilisation and	Yes	99/92 (household

			statistics)		/ postal)
EURO					
UK (40)	Prevalence (one week)	[9792] (16-64)	Nationally representative (post-code address file)	Yes	79.4
Czech Republic (personal communication)	Unable to contact the author				
Russia (personal communication)	Unable to contact the author				
Ireland (24)	Prevalence (unclear)	1190 [936] (65+)	Elderly in Dublin (GP registers)	Unable to access data	85 (of those available)
The Netherlands (23)	Prevalence (lifetime, 12 month, 1 month)	[7076] (18-64)	Nationally representative	Yes	69.7
Germany (16)	Prevalence (lifetime and 1 year)	481 (26-64)	West Germany (not given)	Yes	76
Germany (24)	Prevalence (unclear)	516 in Berlin (70+) 358 in Munich (85+)	Berlin (The Registration office) Munich (electoral register)	Unable to access data	27 89
Finland (41)	Prevalence (unclear)	601 [339] (85+)	All old in one town (unclear)	n/a	92
Greece (42)	Prevalence (point)	618 [489] (18-74)	Two boroughs in Athens (census data) 70,000 households	Yes	83
France (16)	Prevalence (6 month and lifetime)	1746 (18+)	Small town near Paris (households determined from phone listings and field study)	Yes multistage	63 and 39%
Spain (43)	Prevalence (one month)	[1250/?] (17+)	Cantabria region (population register survey)	Yes	90/83 (phase 1/2)
Italy (16)	Prevalence (lifetime)	1000 (15+)	Florence (GP lists)	Yes	100
Norway (20)	Prevalence (2 week) and incidence (one year)	2727 [1879] (18+)	Nationally representative - Oslo and a rural area (Central bureau of statistics)	Yes	74/ 77 (phase 1/2)
Georgia (14)	Prevalence (unclear)	[9847] (18+)	Nationally representative -10 regions (unclear)	Yes	87
Slovakia (14)	Prevalence (unclear)	[1183] (18+)	Nationally representative - 8 regions (population register)	Yes	84

Turkey (15)	Prevalence (lifetime, 12 month, 1 month)	[6095] (18-54)	Nationally representative (unclear, household based)	Unable to access data	72.6
SEARO					
Singapore (44)	Prevalence (unclear)	1000 [612] (65+)	The elderly population in 3 constituencies (electoral roster)	Yes	61.2
India (14)	Prevalence (unclear)	5000 [5145] (18+)	Andra Pradesh, urban and rural (households, electoral roster)	Yes	98
Pakistan (45)	Prevalence (unclear)	259 (18+)	One semi-urban village (electoral roster)	Yes 2 stage study	98
Nepal (personal communication – Tausig)	Prevalence (unclear)		One village (unclear)		
WPRO					
Australia (46)	Prevalence (12 month, 1 month, 2 weeks)	10560 (adults)	Nationally representative (unclear but households)	Yes	78
New Zealand (47)	Prevalence (lifetime)	[1498] (18-64)	Christchurch (urban area) (unclear but households)	Yes	70
Japan (15)	Data not available in the referenced study				
China (14)	Prevalence (unclear)	10000 + 5000 [9442 + 2480] (18+)	Shandong, Henang and Gansu provinces (unclear but households)	Yes	99/50 (household / postal)
Taiwan (16)	Prevalence (lifetime and 12 month)	11000 [11004] (18+)	Nationally representative (household registration rosters)	Yes	90
Korea (16)	Prevalence (lifetime and 12 month)	[5100] (18+)	Seoul and rural areas (households via primary & secondary sampling units)	Yes	83
Cambodia (48)	No information on incidence or prevalence and the sample				

Table 3: Selected characteristics of multi-country studies

Multi-country study	Aim of the study	Countries included	Country data used in the GBDep study	Interview Measure and Score	Interviewer
Ayuso-Mateos et al. (2001)	Prevalence of depressive disorders in general population in five European countries	UK (Liverpool) Ireland (Dublin) Norway (Oslo) Finland (Turku) Spain (Santander)	Spain (Santander)	Beck Depression Inventory (Phase 1); Schedule for Clinical Assessment in Neuropsychiatry (Phase 2)	Psychiatrists, general practitioners, psychologists
Copeland et al. (1999)	Prevalence of depression in 65+ population in 9 European cities	Ireland (Dublin) Germany (Berlin, Munich) Netherlands (Amsterdam) Iceland UK (Liverpool, London) Italy (Verona) Spain (Zaragoza)	Ireland (Dublin) Germany (Berlin, Munich)	DSM-III-R version A3 in the three cities	Nurses and epidemiologists (Dublin); psychiatrists (Berlin and Munich)
Weissman et al. (1996)	'To estimate the rates and patterns of major depression and bipolar disorder based on cross-national epidemiological surveys'	Puerto Rico Lebanon Germany France Italy New Zealand Taiwan Korea United States Alberta (Edmonton)	Puerto Rico Lebanon Germany France Italy New Zealand Taiwan Korea	DIS (version III) and DSM-III; diagnosis generated by a computer algorithm	lay interviewers
WHO International Consortium in Psychiatric Epidemiology (2000)	'cross-national comparative studies of the prevalence and correlates of mental disorders'	Mexico Turkey Canada USA Brazil Germany Netherlands	Mexico Turkey	UM-CIDI with DSM-III-R (Mexico), WHO CIDI with DSM-III-R (Turkey)	Unclear
WHO Multi-country Survey, 2000-2001	Health systems performance assessment; the overall goal was '…to develop instruments that would allow the measurement of health, responsiveness, and other health-related parameters in comparable manner…' (p.763)	61 countries: 1 in AFRO, 11 in AMRO/PAHO, 8 in EMRO, 33 in EURO, 3 in SEARO, 5 in WPRO	Egypt Georgia Slovakia India China Colombia	CIDI used for mental health valuations (depression and alcohol use)	lay interviewers; self-administered survey

Reference List

(1) Lopez A, Mathers C, Ezzati M, Jamison C, Murray D. Measuring the global burden of disease and risk factors: 1990-2001. In: Lopez A, Mathers C, Ezzati M, Jamison C, Murray D, editors. Global Burden of Disease and Risk Factors. Oxford: The World Bank and Oxford University Press, 2006.

(2) Ustun TB, Ayuso-Mateos JL, Chatterji S, Mathers C, Murray CJL. Global burden of depressive disorders in the year 2000. The British Journal of Psychiatry 2004; 184(5):386-392.

(3) Prince M, Patel V, Saxena S, Maj M, Maselko J, Phillips MR et al. Global mental health 1 - No health without mental health. Lancet 2007; 370(9590):859-877.

(4) WHO. The World Health Report 2001 - Mental health, new understanding, new hope. Geneva: WHO, 2001.

(5) Lancet Global Mental Health Group. Scale up services for mental disorders: a call for action. The Lancet 2007; 370(9594):1241-1252.

(6) WHO. Mental Health Gap Action Programme: scaling up care for mental, neurological and substance use disorders. Geneva: World Health Organization, 2008.

(7) The Lancet. A Movement for Global Mental Health is launched. Lancet 2008; 372(9646):1274.

(8) Anand S, Hanson K. Disability-adjusted life years: a critical review. Journal of Health Economics 1997; 16(6):685-702.

(9) Paalman M, Bekedam H, Hawken L, Nyheim D. A critical review of priority setting in the health sector: the methodology of the 1993 World Development Report. Health Policy and Planning 1998; 13(1):13-31.

(10) Cooper RS, Osotimehin B, Kaufman JS, Forrester T. Disease burden in sub-Saharan Africa: what should we conclude in the absence of data? Lancet 1998; 351(9097):208-210.

(11) MATHERS C.D. Uncertainty and data availability for the global burden of disease estimates 2000-2002. 2005. World Health Organization. Evidence and Information for Policy Working Paper.

(12) Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJL. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. Lancet 2006; 367(9524):1747-1757.

(13) Ayuso-Mateos JL. Global burden of unipolar depressive disorders in the year 2000. 2000.

(14) Üstün T.B., Chatterji S., Villanueva M., Bendib L., Çelik C., Sadana R. et al. WHO Multi-country Survey Study on Health and Responsiveness 2000-2001. In: Murray C.J.L., Evans D.B., editors. Health Systems Performance Assessment: Debates, Methods and Empiricism. Geneva: World Health Organization, 2003: 761-796.

(15) WHO International Consortium in Psychiatric Epidemiology. Cross-national comparisons of the prevalences and correlates of mental disorders. Bulletin of the World Health Organization 2000; 78(4):413-426.

(16) Weissman MM, Bland RC, Canino GJ, Faravelli C, Greenwald S, Hwu HG et al. Cross-national epidemiology of major depression and bipolar disorder. JAMA 1996; 276(4):293-299.

(17) Lewis G, Pelosi AJ, Araya R, Dunn G. Measuring Psychiatric-Disorder in the Community - A Standardized Assessment for Use by Lay Interviewers. Psychological Medicine 1992; 22(2):465-486.

(18) Farmer AE, Katz R, Mcguffin P, Bebbington P. A Comparison Between the Present State Examination and the Composite International Diagnostic Interview. Arch Gen Psychiatry 1987; 44(12):1064-1068.

(19) Murphy JM, Monson RR, Laird NM, Sobol AM, Leighton AH. Studying the incidence of depression: an lsquointervalrsquo effect. International Journal of Methods in Psychiatric Research 2000; 9(4):184-193.

(20) Sandanger I, Nygard JF, Ingebrigtsen G, Sorensen T, Dalgard OS. Prevalence, incidence and age at onset of psychiatric disorders in Norway. Social Psychiatry and Psychiatric Epidemiology 1999; 34(11):570-579.

(21) Eaton WW, Kramer M, Anthony JC, Dryman A, Shapiro S, Locke BZ. The Incidence of Specific Dis Dsm-Iii Mental-Disorders - Data from the Nimh Epidemiologic Catchment-Area Program. Acta Psychiatrica Scandinavica 1989; 79(2):163-178.

(22) Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S et al. 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(1):8-19.

(23) Spijker J, de Graaf R, Bijl RV, Beekman ATF, Ormel J, Nolen WA. Duration of major depressive episodes in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). The British Journal of Psychiatry 2002; 181(3):208-213.

(24) Copeland JR, Beekman AT, Dewey ME, Hooijer C, Jordan A, Lawlor BA et al. Depression in Europe. Geographical distribution among older people. The British Journal of Psychiatry 1999; 174(4):312-321.

(25) Lapouse R. Problems in Studying Prevalence of Psychiatric Disorder. American Journal of Public Health and the Nations Health 1967; 57(6):947-&.

(26) Horowitz AW, Wakefield JC. The loss of sadness: How psychiatry transformed normal sorrow into depressive disorder. Oxford: Oxford University Press, 2007.

(27) Arnesen T, Kapiriri L. Can the value choices in DALYs influence global priority-setting? Health Policy 2004; 70(2):137-149.

(28) Kearney PM, Whelton M, Reynolds K, Muntner P, Whelton PK, He J. Global burden of hypertension: analysis of worldwide data. Lancet 2004; 365(9455):217-223.

(29) Chisholm D, Sanderson K, AYUSO-MATEOS JL, Saxena S. Reducing the global burden of depression: Population-level analysis of intervention cost-effectiveness in 14 world regions. The British Journal of Psychiatry 2004; 184(5):393-403.

(30) DCPP. Mental and neurological disorders. 2006.

(31) WHO. The world health report 2000 - Health systems: improving performance. Geneva: World Health Organization, 2000.

(32) Abas MA, Broadhead JC. Depression and anxiety among women in an urban setting in Zimbabwe. Psychological Medicine 1997; 27(1):59-71.

(33) Hollifield M, Katon W, Spain D, Pule L. Anxiety and depression in a village in Lesotho, Africa: a comparison with the United States. The British Journal of Psychiatry 1990; 156(3):343-350.

(34) Awas M, Kebede D, Alem A. Major mental disorders in Butajira, southern Ethiopia. Acta Psychiatrica Scandinavica 1999; 100:56-64.

(35) Costello EJ, Angold A, Burns BJ, Stangl DK, Tweed DL, Erkanli A et al. The Great Smoky Mountains Study of Youth: Goals, Design, Methods, and the Prevalence of DSM-III-R Disorders. Arch Gen Psychiatry 1996; 53(12):1129-1136.

(36) Canino GJ, Bird HR, Shrout PE, RubioStipec M, Bravo M, Martinez R et al. The Prevalence of Specific Psychiatric-Disorders in Puerto-Rico. Arch Gen Psychiatry 1987; 44(8):727-735.

(37) Almeida-Filho N, Mari Jd, Coutinho E, Franca JF, Fernandes J, Andreoli SB et al. Brazilian multicentric study of psychiatric morbidity. Methodological features and prevalence estimates. The British Journal of Psychiatry 1997; 171(6):524-529.

(38) Araya R, Rojas G, Fritsch R., Acuna J, Lewis G. Common mental disorders in Santiago, Chile: Prevalence and socio-demographic correlates. The British Journal of Psychiatry 2001; 178(3):228-233.

(39) Hayashi S, Perales A, Sogi C, et al. Prevalencia de vida de trastornosmentales en Independencia (Lima, Peru). Anales de Salud Mental 1985; 1:206-222.

(40) Bebbington PE, Dunn G, Jenkins R, Lewis G, Brugha T, Farrell M et al. The influence of age and sex on the prevalence of depressive conditions: report from the National Survey of Psychiatric Morbidity. Psychological Medicine 1998; 28(1):9-19.

(41) Paivarinta A, Verkkoniemi A, Niinisto L, Kivela SL, Sulkava R. The prevalence and associates of depressive disorders in the oldest-old Finns. Social Psychiatry and Psychiatric Epidemiology 1999; 34(7):352-359.

(42) Mavreas VG, Beis A, Mouyias A, Rigoni F, Lyketsos GC. Prevalence of Psychiatric-Disorders in Athens - A Community Study. Social Psychiatry 1986; 21(4):172-181.

(43) Ayuso-Mateos JL, Vazquez-Barquero JL, Dowrick C, Lehtinen V, Dalgard OS, Casey P et al. Depressive disorders in Europe: prevalence figures from the ODIN study. The British Journal of Psychiatry 2001; 179(4):308-316.

(44) Kua EH. Prevalence of mental disorders in the elderly in Singapore--a preliminary report. Annals of the Academy of Medicine, Singapore 1987; 16(2).

(45) Husain N, Creed F, Tomenson B. Depression and social stress in Pakistan. Psychological Medicine 2000; 30(2):395-402.

(46) 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(4):427-438.

(47) Wells JE, Bushnell JA, Hornblow AR, Joyce PR, OakleyBrowne MA. Christchurch Psychiatric Epidemiology Study .1. Methodology and Lifetime Prevalence for Specific Psychiatric-Disorders. Australian and New Zealand Journal of Psychiatry 1989; 23(3):315-326.

(48) Somasundaram DJ, van de Put WACM. Mental health care in Cambodia. Bulletin of the World Health Organization 1999; 77(3):275-277.

APPENDIX A: Sources of data

Country	Source of data referenced in the Ustun study	Primary source of data, if different
AFRO		
Zimbabwe	(Abas and Broadhead, 1997)	
Lesotho	(Hollifield et al., 1990)	
Ethiopia	(Awas et al., 1999)	
AMRO		
USA	(Kessler et al., 1994)	
USA	(Costello et al., 1996)	
Canada	(Murphy JM, 2000)	
Puerto Rico	Weissman et al 1996	(Canino et al., 1987)
Brazil	(Almeida-Filho et al., 1997)	
Mexico	WHO International Consortium in Psychiatric Epidemiology, 2000	Caraveo et al., 1998
Columbia	WHO Multi-country Survey, 2000-2001	
Chile	(Araya et al., 2001)	
Peru	(Hayashi S, 1985)	
EMRO		
Lebanon	(Weissman et al., 1996)	Karam (1992)
Morocco	(personal communication,)	
Egypt	WHO Multi-country Survey, 2000-2001	
EURO		
UK	(Bebbington et al., 1998)	
Czech Republic	(personal communication)	
Russia	(personal communication)	
Ireland	(Copeland et al., 1999)	Lawlor et al. (1994)
The Netherlands	Spijker et al., 2002	(Bijl et al., 1998)
Germany	Weissman et al., 1996	(Wittchen et al., 1992)
Germany	(Copeland et al., 1999)	Helmchen et al. (1996)
Finland	(Paivarinta et al., 1999)	
Greece	(Mavreas et al., 1986)	
France	Weissman et al., 1996	(Lepine JP, 1989)
Spain	Ayuso-Mateos et al., 2001	(Vazquez-Barquero et al., 1987)
Italy	Weissman et al., 1996	(Faravelli et al., 1990)
Norway	(Sandanger et al., 1999)	
Georgia	WHO Multi-country Survey, 2000-2001	
Slovakia	WHO Multi-country Survey, 2000-2001	
Turkey	WHO International Consortium in Psychiatric Epidemiology, 2000	Kylyc (1998)
SEARO		
Singapore	(Kua, 1987)	
India	WHO Multi-country Survey, 2000-2001	
Pakistan	(Husain et al., 2000)	
Nepal	(personal communication – Tausig)	
WPRO		
Australia	(Vos and Mathers, 2000)	
New Zealand	Weissman et al., 1996	(Wells et al., 1989)
Japan	WHO International Consortium in	Data not found; Japan not

	Psychiatric Epidemiology, 2000	included in the WHO 2000 study
China	WHO Multi-country Survey, 2000-2001	
Taiwan	Weissman et al., 1996	(Hwu et al., 1989)
Korea	Weissman et al,. 1996	(Lee et al., 1990)
Cambodia	(Somasundaram and van de Put, 1999)	

APPENDIX B: Data sources for India and Nepal

India

In the 'version one estimates' (13) there were two data sources for India listed: a household survey in Andhra Pradesh, which surveys 5000 people and a much smaller study set in West Bengal comparing depression prevalence in one tribe in two different settings - urban and rural (Nandi, 1992). The final version of GBD estimates for depression (GBDep study) includes only the survey from Andhra Pradesh.

The Andhra Pradesh survey was part of the large WHO study – the WHO Multi-Country Survey Study on Health and Responsiveness 2000-2001 - which was intended to provide comparable data on various health outcomes (including disease prevalence) in different countries, and to assess the responsiveness of health systems. The study was designed as a cross-sectional prevalence study. The DALY estimate, however, relies on incidence measures. It is unclear from the methodology of the GBDep study how this conversion was made, but it reduces the reliability of the results.

The Andhra Pradesh study used a multi-stage sampling procedure, stratified and random, to select clusters. Households were selected from electoral rosters and in each selected household, one adult (> 18 years) selected randomly was interviewed. The initial sample was 5,000 but the final sample reported was 5,145. And the reported response rate was 98% which does not correspond to the initial and final sample size. In contrast to the population ratio between male and female being 1.06, in the survey sample the ratio was 0.86 (46.7% males, 53.3% females). The survey sample was not nationally representative (14).

The mean duration of households interviews conducted in Andhra Pradesh was 87 minutes (14). The diagnosis of depression was made using the WHOs 'Composite International Diagnostic Interview (CIDI)', a standardized interview for assessing mental disorders, intended to be compatible with ICD and DSM criteria (Wittchen, 1994). This interview comprises 276 questions, is available in many languages, and can be used by non-clinicians. It is thought to be reliable as it is highly standardized. However, it has been criticised on the ground that is too long, lacks large-scale validation, and is inaccurate cross-culturally. A minor change in the order of questions can create a largely different prevalence (due to interviewees being able to avoid long lines of questioning by answering negatively to 'starter questions') (Wittchen, 1994). Like other measures of depression, a proportion of those diagnosed may not have a pathological condition but may be experiencing an appropriate response to their situation (Patten, 2003).

Nepal

Dr Tausig who provided data to authors of the GBDep study via personal communication replied to our query regarding the Nepali sample. In his response he made clear that the estimate for depression prevalence in Nepal was based on a small study, a single village, in rural Nepal and Dr Tausig emphasised that he "would not consider the estimate to have any validity beyond the village where the data were collected." For the study purposes a modified version of DIS was used.

SEARO region

The Andhra Pradesh study makes up a large portion of data used fore the SEARO estimate of depression: 5,145 out of 6,016 respondents, with an unknown sample size for Nepal. The other two studies included in the SEARO region were from Pakistan (612 respondents) and Singapore (sample size of 259). All of the information on the severity distribution of depression in SEARO comes from the study in Andra Pradesh (GBDep study), as the measure of depression used in that study groups depressive episodes into mild, moderate and severe categories according to ICD-10 criteria.