Psychological Medicine (2013), **43**, 471–481. © Cambridge University Press 2012
doi:10.1017/S0033291712001511

REVIEW ARTICLE

Global variation in the prevalence and incidence of major depressive disorder: a systematic review of the epidemiological literature

A. J. Ferrari[1,2]*, A. J. Somerville[3], A. J. Baxter[1,2], R. Norman[1,4], S. B. Patten[5], T. Vos[1] and H. A. Whiteford[1,2]

[1] *The University of Queensland, School of Population Health, Herston, QLD, Australia*
[2] *Queensland Centre for Mental Health Research, The Park Centre for Mental Health, Wacol, QLD, Australia*
[3] *The University of Queensland, Institute of Social Science Research, St Lucia, QLD, Australia*
[4] *The University of Queensland, Queensland Children's Medical Research Institute, Herston, QLD, Australia*
[5] *Department of Community Health Sciences, University of Calgary, Calgary, AB, Canada*

Background. Summarizing the epidemiology of major depressive disorder (MDD) at a global level is complicated by significant heterogeneity in the data. The aim of this study is to present a global summary of the prevalence and incidence of MDD, accounting for sources of bias, and dealing with heterogeneity. Findings are informing MDD burden quantification in the Global Burden of Disease (GBD) 2010 Study.

Method. A systematic review of prevalence and incidence of MDD was undertaken. Electronic databases Medline, PsycINFO and EMBASE were searched. Community-representative studies adhering to suitable diagnostic nomenclature were included. A meta-regression was conducted to explore sources of heterogeneity in prevalence and guide the stratification of data in a meta-analysis.

Results. The literature search identified 116 prevalence and four incidence studies. Prevalence period, sex, year of study, depression subtype, survey instrument, age and region were significant determinants of prevalence, explaining 57.7% of the variability between studies. The global point prevalence of MDD, adjusting for methodological differences, was 4.7% (4.4–5.0%). The pooled annual incidence was 3.0% (2.4–3.8%), clearly at odds with the pooled prevalence estimates and the previously reported average duration of 30 weeks for an episode of MDD.

Conclusions. Our findings provide a comprehensive and up-to-date profile of the prevalence of MDD globally. Region and study methodology influenced the prevalence of MDD. This needs to be considered in the GBD 2010 study and in investigations into the ecological determinants of MDD. Good-quality estimates from low-/middle-income countries were sparse. More accurate data on incidence are also required.

Received 8 March 2012; Revised 5 June 2012; Accepted 7 June 2012; First published online 25 July 2012

Key words: Epidemiology, global burden of disease, incidence, major depressive disorder, meta-regression, prevalence.

Introduction

Depressive disorders as defined by the Diagnostic and Statistical Manual of Mental Disorders (DSM), fourth edition, text revision (DSM-IV-TR) typically present in adolescence or early adult life, can be chronic or episodic and are frequently recurrent and co-morbid with substance abuse or other mental and physical health conditions (APA, 2000). The burden attributed to depressive disorders worldwide is a major public health concern. The Global Burden of Disease (GBD), Injuries, and Risk Factors Study measures disease burden in terms of the disability-adjusted life year (DALY). This is the sum of years of life lost due to premature mortality (YLL) and years of life lived with a disability (YLD) (Murray & Lopez, 1996). The GBD 1990 study reported that depressive disorders were the fourth leading cause of burden, accounting for 3.7% of DALYs in 1990 (Murray & Lopez, 1996). By 2000 they were the third leading cause of burden, accounting for 4.3% of DALYs, and the leading cause of disability, accounting for 13.4% of YLDs in women and 8.3% in men (Üstün *et al*. 2004).

Our understanding of the public health importance of mental disorders, including their impact on other

* Address for correspondence: A. J. Ferrari, Policy and Evaluation Group, Queensland Centre for Mental Health Research, The Park Centre for Mental Health, Locked Bag 500, Sumner Park BC, Brisbane, QLD 4074, Australia.
(Email: alize_ferrari@qcmhr.uq.edu.au)

health conditions (Prince *et al*. 2007; Collins *et al*. 2011), will be enhanced by the GBD 2010 study currently underway (Murray *et al*. 2007). This will quantify burden for over 220 diseases (including 11 mental disorders) by country and 21 world regions for the years 1990, 2005 and 2010. It will use significantly improved methodology compared with the GBD 1990 study, including more representative disability weight estimations (see http://www.globalburden.org/). In the GBD 2010 study, a dimensional approach is being taken, with burden being calculated for major depressive disorder (MDD), i.e. one or more major depressive episodes (MDEs) using three severity levels, and dysthymia separately. This will enable burden estimates to encapsulate differences in morbidity within and between depressive disorder subtypes (Salomon *et al*. 2003; Global Burden of Disease, 2009; Salomon, 2010).

To quantify the morbidity (i.e. YLD) attributable to MDD in the GBD 2010 study, epidemiological data were required to describe the disease occurrence and course of illness (see Supplementary Appendix 1 for definitions of the data required). Although there is a wealth of literature on the different epidemiological parameters of MDD, this has yet to be systematically summarised at a global level. Aside from informing the GBD 2010 study, such an integrated summary has both clinical and public health applications (Kessler, 1999; Wittchen, 2000; Skapinakis & Lewis, 2001). However, this task is complicated by the presence of significant heterogeneity in the data. Heterogeneity refers to both the variability in epidemiological estimates resulting from true differences in the epidemiology of MDD and the variability produced by differences in the methodology used to capture data (Weich & Araya, 2004; Bromet *et al*. 2011).

Literature reviews on the prevalence and incidence of MDD have consistently raised the issue of heterogeneity. The GBD 2000 update estimated that the 12-month prevalence of a MDE was 1.6% in males and 2.5% in females. Predicted (from prevalence estimates) annual incidence was 3.2% in males and 4.9% in females (Üstün *et al*. 2004). An analytical review of the GBD 2000 update concluded that there was a lack of data on unipolar depression and significant heterogeneity across epidemiological estimates (Brhlikova *et al*. 2010).

Paykel and collaborators' review of MDD in Europe reporting a 12-month prevalence of 5% also revealed significant gaps in the literature. The incidence of MDD was unclear because of the lack of longitudinal follow-up of cases and there were limited data from Eastern Europe (Paykel *et al*. 2005). Another review investigating the prevalence of MDD in the USA, Canada, Puerto Rico, France, West Germany, Italy, Lebanon, Taiwan, Korea and New Zealand revealed significant regional variation in prevalence. The 12-month prevalence of MDD ranged from 0.8% in Taiwan to 5.8% in New Zealand (Weissman *et al*. 1996). Explaining this regional variation is difficult, given differences in the methodology used by the different studies (Weich & Araya, 2004).

In an attempt to control for this, the World Mental Health Survey (WMHS) Consortium conducted population surveys in 28 countries using a standard protocol for data collection and assessment. The World Health Organization's Composite International Diagnostic Interview (CIDI) was used to diagnose cases of MDE (Bromet *et al*. 2011). Results revealed regional variation in the 12-month prevalence of a MDE, ranging from 2.2% in Japan to 10.4% in Brazil with similar averages of 5.5% in developed and 5.9% in developing countries. The average duration of an episode estimated from WMHS ranged from 23.1 to 33.8 weeks (Kessler & Üstün, 2008; Seedat *et al*. 2009; Bromet *et al*. 2011). These differences may reflect true regional variability in the epidemiology of MDE. However, despite WMHS's standardised procedure, some limitations were reported. Notably, validation exercises involving the CIDI have been completed almost entirely in western countries; hence its cross-cultural reliability and validity remain unclear (Simon *et al*. 2002; Wang *et al*. 2007; Bromet *et al*. 2011).

The aim of this paper is to summarise the global prevalence and incidence of MDD, exploring the global distribution and sources of heterogeneity and, where feasible, adjusting for variability caused by differences in study methodology and design. This will help identify true differences in the global distribution of MDD that need to be considered in the integration of the epidemiological data. This work was undertaken by the GBD Mental Disorders Research Group (http://www.qcmhr.uq.edu. au/BODP).

Method

Literature review

The systematic review adhered to guidelines recommended by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Moher *et al*. 2009). Electronic databases Medline (via CSA), PsycINFO and EMBASE were used. Search strings were derived in consultation with a research librarian. See http://www.gbd.unsw.edu.au/gbdweb.nsf/resources/MD_Pt2_SearchStrings/$file/MDSearch+strings_2.pdf for more information on the search strings. Publications in languages other than English were included. Although the search was limited

to studies published between 1980 and 2007, we have continued to capture additional studies up to 2011 through continuing perusal of the literature and correspondence with experts in the field.

Inclusion and exclusion

Population-based surveys representative of the community, region or country were included. Non-representative subsets of the population such as samples based on in-patient admissions or treatment trials were excluded. Studies adhering to DSM or International Classification of Diseases (ICD) diagnostic classifications were preferred to allow for better consistency in the measurement of MDD; however, studies using symptom scales that only broadly map on to DSM/ICD diagnostic thresholds were included for comparison. Where the same data were reported across different papers the most informative paper was selected.

For prevalence, we included studies reporting point (i.e. current or past-month prevalence), 6- and/or 12-month prevalence. Lifetime estimates were excluded, as these are more prone to recall bias (Simon & VonKorff, 1995; Kruijshaar *et al*. 2005; Moffitt *et al*. 2010; Susser & Shrout, 2010). Estimates of MDD were the focus; however, if studies reported estimates of depression not otherwise specified (NOS) these were included for comparison. Studies reporting incidence hazards were included (i.e. those with person-years of follow-up in the denominator). Estimates of probability such as cumulative incidence based on longer risk intervals were excluded.

Extraction

Data extracted included information on the country and year of study, parameter value and type (point, 12-month prevalence), sample coverage (community, regional, national), sample urbanicity (rural, urban, mixed), sex (male, female, persons), age range, case ascertainment period (recorded as the midyear time point), response rate, depression subtype (MDD or MDD + depression NOS), diagnostic criteria (ICD, DSM) and survey instrument (diagnostic instruments, symptom scales). Countries were stratified into regions based on the GBD categorization of regions (see http://www.globalburden.com.au/project-description). Forced entry of key variables was required for quality assurance. A random sample of papers was reviewed by two researchers to check consistency of data extraction. Where differences occurred, reviewers discussed these with the study's primary investigator (H.A.W.) to arrive at a consensus.

Uncertainty [standard error (S.E.) or confidence interval] pertaining to each prevalence or incidence estimate was extracted if reported or otherwise calculated. If the denominator (sex- and/or age-specific sample size) was reported, S.E. was calculated using $\text{S.E.} = \sqrt{2.1 \times (P \times (1 - P)/N)}$, where P is the proportion of cases reported, 2.1 is the average design effect and N is the denominator. The average design effect accounted for any increase in uncertainty produced by a study's sampling methodology. It was calculated based on a sample of 110 design effects from the GBD Mental Disorders Research Group's affective disorders dataset. If the denominator was not reported by age and sex, the United Nation's country-, sex-, age- and year-specific population size was used to distribute the overall sample size across age and sex categories (United Nations, 2007).

Analysis

Prevalence

For prevalence, Stata II.2 software (StataCorp LP, USA) was used to conduct a meta-regression (Sterne *et al*. 2008) to help explain the variability between studies. We based our statistical methodology on previous applications of meta-regression to explore the effect of methodological and ecological variables on prevalence (Van Houwelingen *et al*. 2002; Sterne *et al*. 2008; Baker *et al*. 2009). Since the distribution of prevalence estimates was positively skewed, logarithmic transformation (natural log) was applied to meet the parametric assumption of normality (Tabachnick & Fidell, 2007).

Results from the meta-regression guided the stratification of pooled prevalence estimates in the subsequent meta-analysis (Sterne *et al*. 2008). As it is essential for GBD purposes to avoid any overestimation in burden estimates, point prevalence is considered as the 'gold standard' as it is less susceptible to recall bias compared with estimates of period prevalence (Global Burden of Disease, 2009). For this reason, it was also set as the primary summary measure here as well as in the upcoming GBD disease modelling of MDD for which this literature review was undertaken.

Statistical heterogeneity was quantified using the I^2 statistic, which indicates the total variation in the data attributable to heterogeneity (Higgins *et al*. 2003). A random-effects model was selected over the fixed-effects model to accommodate for heterogeneity (DerSimonian & Laird, 1986; DerSimonian & Kacker, 2007). We used the post-estimation 'predict' command to estimate overall prevalence by region, accommodating for methodological factors. This

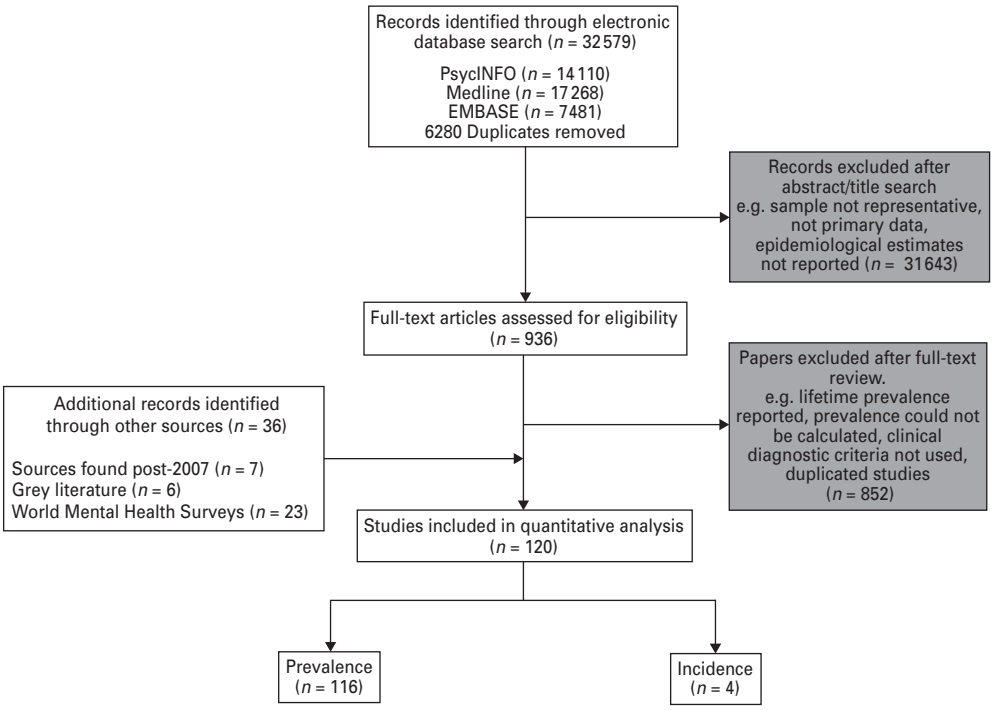

Fig. 1. Flowchart showing results of the systematic review for the prevalence and incidence of major depressive disorder.

command fitted a value to each reported prevalence estimate and the associated S.E. based on the coefficients from the meta-regression.

Incidence

Similar methods of pooling incidence were used as described for prevalence. However, the paucity of incidence estimates did not permit us to conduct a meta-regression.

Results

Out of 32579 data sources on the epidemiology of MDD, only 120 studies fitted the inclusion criteria for prevalence and incidence. The search and main reasons for exclusion are summarized in Fig. 1. A summary of the studies included for each parameter is given in Supplementary Appendix 2.

Prevalence

A total of 116 studies reporting the prevalence of MDD were included, the majority of which were from western countries, i.e. Europe and North America ($n = 74$). We identified 11 studies reporting both MDD and depression NOS and 22 studies using a symptom scale. There was considerable variability between estimates. Point prevalence ranged from 0.05% in males from Japan aged 65 years or older (Ihara *et al*. 1998) to 73% in females from Afghanistan aged 15 years or older (Lopes Cardozo *et al*. 2005).

To maximize inclusion, potential outliers in the dataset with no salient methodological limitation were retained and investigated further through the meta-regression. To ensure independence of observations, where person- as well as sex-specific estimates were reported by the same study, only the latter were included in analyses. Where age-specific and overall-age estimates were reported, only the latter were included. Where only age-specific estimates were reported, these were combined to calculate the overall-age prevalence, i.e. the summed number of cases across each age group was divided by the summed denominator across each age group. This reduced the final dataset from 783 to 274 prevalence estimates. To investigate the effect of age on prevalence, estimates were grouped into four broad categories: whole age range (e.g. 0–99 years), <18 years, 18–65 years and >65 years. Where reported age ranges could fit into more than one category they were allocated to the most representative one on a case-by-case basis.

In the meta-regression, study-level variables were inserted in the model first (model 1), and the country-level variable (region) second (model 2). Coverage and urbanicity variables were excluded from both models as they were signficantly correlated with the majority of other variables, particularly with the region variable, which is a major focus of this paper. Model 1 explained 43.1% of the between-study variance

(adjusted R^2 43.1%, $F_{11,265}=16.9$, $p<0.001$). Prevalence type, sex, midyear of case ascertainment, age range, depression subtype and survey instrument were statistically significant determinants of reported prevalence (Table 1). Diagnostic criteria were not significantly associated with prevalence.

Prevalence was statistically higher for past year prevalence compared with point, in females compared with males and persons, in studies capturing MDD and depression NOS compared with MDD alone, and in studies with more recent case ascertainments. Persons under 18 years had lower prevalence than persons across the overall age group and studies that used symptom scales reported a higher prevalence than studies where DSM or ICD criteria were used. The survey instrument variable was originally made up of eight categories, summarizing the most frequent instruments in the dataset. The CIDI was used as the reference as it was the most commonly reported. This variable was dichotomized in the final model because symptom scales were the only instruments to yield signicantly different prevalence estimates at $p<0.001$. Results were similar when clinical interview (interviews conducted by a clinician) were used instead of the CIDI as the reference, with only symptoms scales ($p<0.001$) and the Geriatric Mental State Schedule ($p<0.041$) yielding statistically significant results.

The inclusion of the region variable in model 2 explained an additional 14.6% of the between-study variance. The overall variance explained by study- and country-level variables combined was 57.7% (adjusted R^2 57.7%, $F_{19,257}=16.5$, $p<0.001$). Prevalence period, sex, midyear of case ascertainment, depression subtype, survey instrument and age range remained statistically significant determinants of reported prevalence in model 2. When Western Europe, the region with the most data, was set as the reference, a statistically significant effect of region emerged such that prevalence from South America, South Asia and Africa/Middle East was statistically higher than prevalence from Western Europe. Estimates from East/Southeast Asia and Asia Pacific (high income) were statistically lower. Similar results were obtained when North America was used as the reference, except that South America and North Africa/Middle East no longer yielded statistically higher prevalence.

A series of meta-analyses was used to further illustrate the effect of statistically significant determinants of prevalence. Reported prevalence was pooled according to each statistically significant study-level determinant (Table 2). For each determinant, reported prevalence was also pooled by sex, but there were insuffient data to simultaneously stratify prevalence by all other determinants. To address this limitation, prevalence was also predicted for each region while adjusting for the effects of study-level determinants in model 2 (Fig. 2). Although reduced, cross-national differences in prevalence persisted after adjusting for study-level determinants of prevalence. Notably, although adjusted downwards, estimates from South Asia and Africa/Middle East remained the highest in the dataset, and although adjusted upwards, estimates from East/Southeast Asia remained amongst the lowest. Estimates from Asia Pacific were no longer amongst the lowest. The adjusted and unadjusted prevalence by region is summarised in Supplementary Appendix 3.

Incidence

Only four studies (Eaton *et al*. 1989; Lewinsohn *et al*. 1993; Patten, 2001; Mogga *et al*. 2006) reporting annual incidence rates of MDD from USA, Canada and Ethiopia were identified. Estimates ranged from 1.6% in females aged 18 years or above to 7.1% in females aged 15 to 19 years, from the USA. Male and female estimates across overall age groups were pooled to calculate an overall estimate of annual incidence (Table 3). Total estimates were only included if sex-specific estimates were not reported. Although pooled female estimates were higher than pooled male estimates, there was no statistically significant effect of sex. Since there were only three estimates in each group, more data are required to make a definitive statement on whether a difference exists.

Discussion

The majority of the literature on prevalence was from Western Europe and North America, with much less from non-Western regions. Prevalence data were highly sensitive to elements of study design and methodology. Consistent with existing literature, prevalence in females was higher than in males (Culbertson, 1997; Angst *et al*. 2002; Kuehner, 2003) and 12-month prevalence was higher than point prevalence (Faravelli *et al*. 1990; Stefánsson *et al*. 1994; Bromet *et al*. 2005). Given these results, it would be reasonable to assume that when pooled by sex and prevalence type simultaneously, pooled 12-month prevalence would be higher than pooled point prevalence; however, this was not the case. For persons pooled 12-month prevalence was lower than pooled point prevalence, although this result was not statistically significant. Pursuing the reason for this finding was outside the scope of this study; however, similar unanticipated results have been reported in the literature pertaining to the prevalence of schizophrenia (Saha *et al*. 2005). It must also be noted that prevalence

Table 1. *Results of meta-regression models showing odds ratios for reported prevalence by study- and country-level determinants*[a]

	Unadjusted[b]		Adjusted (model 1)[c]		Adjusted (model 2)[d]	
Determinant	Odds ratio (95% uncertainty)	*p* value	Odds ratio (95% uncertainty)	*p* value	Odds ratio (95% uncertainty)	*p* value
Prevalence period	Reference: point					
12-month prevalence	1.02 (0.79–1.32)	0.853	1.2 (1.0–1.5)	0.048*	1.3 (1.1–1.6)	0.007*
Sex	Reference: female					
Male	0.6 (0.5–0.7)	<0.001*	0.6 (0.5–0.7)	<0.001*	0.6 (0.5–0.7)	<0.001*
Total	0.7 (0.5–0.9)	0.009*	0.7 (0.6–0.9)	0.012*	0.7 (0.6–0.9)	0.008*
Midyear of case ascertainment[e]	1.03 (1.01–1.05)	0.003*	1.02 (1.00–1.03)	0.036*	1.02 (1.01–1.04)	0.002*
Age, years	Reference: overall age groups (0–99 years)					
18–65	1.02 (0.78–1.34)	0.864	1.2 (0.9–1.5)	0.145	1.2 (0.9–1.5)	0.067
<18	0.7 (0.5–1.0)	0.025*	0.6 (0.4–0.7)	<0.001*	0.6 (0.4–0.7)	<0.001*
>65	0.5 (0.4–0.7)	<0.001*	0.98 (0.73–1.33)	0.913	1.1 (0.8–1.4)	0.568
Depression subtype	Reference: MDD only					
MDD + depression NOS	0.6 (0.4–0.8)	0.002*	1.96 (1.47–2.62)	<0.001*	2.1 (1.6–2.8)	<0.001*
Diagnostic criteria	Reference: DSM					
ICD	0.8 (0.6–1.2)	0.281	0.95 (0.72–1.25)	0.697	0.9 (0.7–1.2)	0.390
Not specified	3.3 (1.9–5.8)	<0.001*	1.5 (0.9–2.5)	0.145	1.5 (0.9–2.5)	0.106
Survey instrument	Reference: DSM/ICD tools					
Symptom scale	3.4 (2.6–4.6)	<0.001*	3.6 (2.6–5.0)	<0.001*	2.9 (2.1–4.0)	<0.001*
Region	Reference: Western Europe					
Central/Eastern Europe	1.5 (0.9–2.4)	0.132			1.3 (0.9–2.0)	0.143
North America	0.9 (0.6–1.3)	0.563			1.2 (0.9–1.6)	0.200
South America	1.3 (0.9–1.9)	0.148			1.6 (1.5–2.1)	0.004*
Australasia	0.8 (0.5–1.2)	0.281			0.9 (0.6–1.3)	0.580
East/Southeast Asia	0.4 (0.3–0.6)	<0.001*			0.4 (0.3–0.6)	<0.001*
South Asia	3.3 (1.8–6.2)	<0.001*			2.1 (1.2–3.6)	0.006*
Asia Pacific, high income	0.6 (0.3–1.0)	0.045*			0.4 (0.3–0.7)	<0.001*
Africa/Middle East	1.9 (1.4–2.7)	<0.001*			1.4 (1.1–1.8)	0.042*

MDD, Major depressive disorder; NOS, not otherwise specified; DSM, Diagnostic and Statistcal Manual of Mental Disorders (APA, 2000); ICD, International Classification of diseases (World Health Organization, 1992).

[a] The response rate variable was exlcuded from the model as it did not contribute to the overall variance explained.

[b] Unadjusted results represent meta-regression ran on indiviual variables without countrolling for the effect of others.

[c] Adjusted model 1 results represent meta-regression ran on study-level variables only.

[d] Adjusted model 2 results represent meta-regression ran on study- and country-level variables.

[e] Midyear of case ascertainment modelled as a continuous variable.

* Statistically significant at $p < 0.05$.

Table 2. *Pooled reported prevalence stratified by study-level determinants statistically associated with prevalence*[a]

Determinant	Prevalence % (95% uncertainty)		
	Female ($n=82$)	Male ($n=79$)	Person ($n=34$)
Prevalence period			
Point ($n=82$)	5.9 (4.7–7.5)	3.8 (3.2–4.5)	4.7 (3.7–5.5)
12 months ($n=47$)	7.2 (6.0–8.9)	3.9 (3.0–5.1)	3.7 (2.7–5.0)
Depression subtype			
MDD only ($n=105$)	5.8 (4.7–7.1)	3.5 (2.9–4.2)	4.2 (3.4–5.2)
MDD + depression NOS ($n=11$)	10.2 (8.9–12.1)	6.4 (5.2–7.9)	7.2 (4.3–11.8)
Survey instrument			
DSM/ICD tool ($n=94$)	5.4 (4.9–6.1)	3.1 (2.7–3.6)	3.8 (3.1–4.6)
Symptom scale ($n=22$)	17.03 (11.46–25.28)	11.6 (7.5–18.0)	12.1 (9.3–15.7)

n, Number of studies in each group; MDD, major depressive disorder; NOS, not otherwise specified; DSM, Diagnostic and Statistcal Manual of Mental Disorders (APA, 2000); ICD, International Classification of diseases (World Health Organization, 1992).

[a] For each determinant, prevalence was additionally pooled only by sex, and hence did not account for the effect of all other determinants of prevalence. All I^2 statistics were >90%.

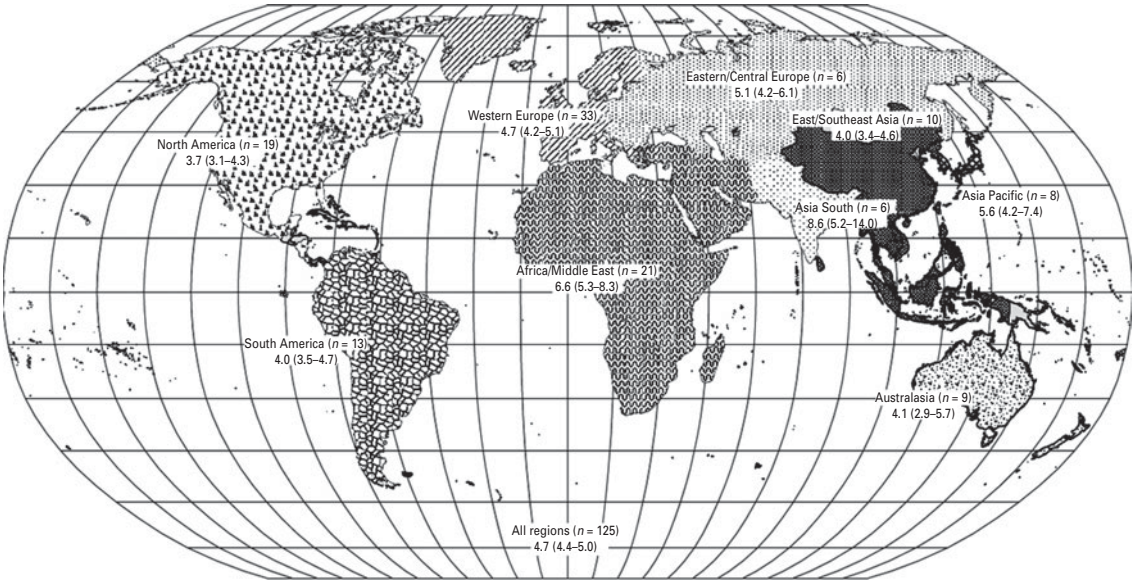

Fig. 2. Predicted point prevalence % (95% uncertainty) by region, adjusted for study-level determinants. Map derived using MapInfo Professional version 10.5.2. (2010) (Pitney Bowes Software, Inc., USA). Map projection: Robinson WGS84. *n*, Number of studies in each region.

was pooled by prevalence type and sex only due to the lack of data. Controlling for the other significant study-level determinants of prevalence, as was done in the meta-regression, may have yielded different results.

Symptom scales were the only survey instruments significantly associated with prevalence, suggesting adequate consistency between the other diagnostic tools in the MDD literature. There has been continuous debate as to whether symptom scales are better suited to measuring mental disorder symptoms or psychological distress (Hollander *et al*. 2007; Horwitz & Wakefield, 2007; Ventevogel *et al*. 2007). In this case, we found that symptom scales inflated the overall

Table 3. *Pooled estimates of annual incidence*[a]

Sex	Incidence, % (95% uncertainty)	Weight, %[b]
Female ($n=3$)	3.4 (1.9–6.3)	48.0
Male ($n=3$)	2.7 (2.0–3.7)	49.1
Total ($n=1$)	2.4 (0.7–8.4)	2.9
Overall ($n=7$)	3.0 (2.4–3.8)	100.00

n, Number of estimates in each group.
[a] All I^2 statistics were >86%.
[b] Random-effects weights.

prevalence of MDD. That said, they were often the only tools used to capture prevalence, particularly in conflict settings where epidemiological data are sparse (Hollifield *et al*. 2002). In order to maximize the global representativeness of our findings we chose to include prevalence estimates derived from symptoms scales. We adjusted for any inflation to pooled prevalence by specifying that prevalence derived from diagnostic tools was the 'gold standard'. This is consistent with the methodology used by other authors specifying that data from symptoms scales need to be 'recalibrated' relative to data from diagnostic tools (Ventevogel *et al*. 2007; Rodin & Van Ommeren, 2009; Steel *et al*. 2009). Despite some differences in the DSM and ICD definitions of MDD, diagnostic criteria did not have a statistically significant effect on prevalence.

We detected a time effect suggesting that the prevalence of MDD had increased over time. This was based on an ecological comparison of the midpoint of the case ascertainment period, as a continuous variable. It is possible that this finding represented a true increase in prevalence. Alternatively, it could be due to other methodological or ecological differences across time that we were unable to capture. More in-depth investigation is required to confirm this finding. The only age effect found was from the <18 years age group, which yielded lower estimates compared with estimates across the entire lifespan. This was probably due to estimates from very young children (e.g. 8–9 years) in the <18-year-old group. However, drawing conclusions from this variable is also problematic, given that we could only categorize age using four broad age categories and some age ranges could be allocated to multiple categories. We used broad as opposed to age-specific estimates, as the latter would violate the parametric assumption of independent observations required in a meta-regression. A more detailed comparison of prevalence across the lifespan is required for better conclusions.

We detected considerable regional differences in the prevalence of MDD, some of which were reduced when study-level sources of variability were controlled for. Although our finding of higher prevalence of MDD in developing regions (except for Asia East/Southeast) compared with developed regions corresponded to WMHS results, our overall adjusted point prevalence of 4.7% was higher than the WMHS finding of 1.8% point prevalence in developed countries and 2.6% in developing countries. Instead, it was closer to the 5.5% 12-month prevalence in developed countries (Bromet *et al*. 2011). The higher point prevalence obtained here may be due to the adjustments made for differences in study methodology. Despite WMHS efforts to enforce a standardized methodology, differences still occurred. Response rates and the amount of missing data varied substantially across countries, which may have reduced the representativeness of some samples (Simon *et al*. 2002; Wang *et al*. 2007; Bromet *et al*. 2011). The WMHS used data collected by the CIDI only. Our broader focus allowed us to include data from countries that were not part of the WMHS, using other diagnostic instruments. That said, our inclusion of prevalence derived from symptom scales must be treated with caution, given the possibility of inflating final results with presentations of MDD symptoms rather than diagnoses.

Estimates from East/Southeast Asia remained much lower than other regions even after adjusting for methodological differences. This may reflect a true difference in the global distribution of MDD. Alternatively, it may be due to unidentified sources of measurement bias that we were unable to control for. One possibility is that DSM/ICD diagnostic criteria are not sensitive to cross-cultural presentations of MDD (Simon *et al*. 2002; Wang *et al*. 2007; Bromet *et al*. 2011). Another is that MDD may be miscoded as depression NOS in less developed countries (Chang *et al*. 2008; Phillips *et al*. 2009), underestimating prevalence. We recommend that in upcoming GBD burden calculation this limitation be addressed through the inclusion of estimates of depression NOS as was done here.

Ecological factors may also contribute to regional differences in prevalence. Consistently high prevalence in Africa/Middle East and South Asia raises the possibility that conflict in countries such as Afghanistan, Iraq and Sudan increase the prevalence of MDD. This is consistent with literature suggesting that exposure to torture and other trauma in conflict settings increases the prevalence of depression and post-traumatic stress disorder (Steel *et al*. 2009). Based on this, we also recommend that the effect of conflict status be investigated further.

Whilst we found numerous naturalistic studies of the annual incidence of MDD, very few follow-up studies using representative community samples were available. The GBD 2000 update predicted an average incidence of 3.2% in males and 4.9% in females (Ayuso-Mateos, 2000; Üstün *et al*. 2004), which was higher than our results of 2.7% in males and 3.4% in females. If a duration of 30 weeks for an episode of MDD (Vos *et al*. 2004) is taken into account, there is a clear inconsistency between the few incidence estimates we obtained and our adjusted prevalence estimate, in that incidence was lower than prevalence instead of higher. This problem illustrates the importance of internal consistency between epidemiological parameters. While summaries of individual epidemiological parameters of MDD are useful, they may be inaccurate, particularly where data are limited. In this case, parameters need to be considered simultaneously for an internally consistent epidemiological profile of MDD. The upcoming GBD disease modelling of the epidemiology of MDD will help clarify this (Institute for Health Metrics and Evaluation, 2012). In the meantime, our incidence findings are indicative only. More cross-national data are required for stronger conclusions. Although more data were available for prevalence, there were few good-quality estimates from less developed parts of the world. This prevented us from conducting region-specific analyses of variance. We were also unable to control for all sources of variability in prevalence. Further investigation into other determinants of prevalence, for instance human development indicators outside the scope of this review, is required.

Our literature review addressed a range of issues central to the epidemiology of MDD. It identified the data sources required for burden estimation in the GBD 2010 study. It also provided an epidemiological summary of MDD, considering, where feasible, sources of heterogeneity in the data. We recommend that the statistically significant study-level determinants of prevalence identified be considered when generating other ecological models of MDD prevalence. We were also able to identify salient gaps in the literature that need further consideration. There were very few incidence data and very few studies from non-developed parts of the world across all parameters. We were also unable to comprehensively assess the effect of age on prevalence. Further investigation of these limitations is required for a clearer epidemiological profile of MDD.

Supplementary material

For supplementary material accompanying this paper visit http://dx.doi.org/10.1017/S0033291712001511.

Acknowledgements

Special thanks to Angela Higginson from the Institute for Social Science Research, University of Queensland for providing us with statistical advice. We are also grateful to the contribution of Holly E. Erskine, An Pham and Amanda Brown to data collection and to Roman Scheurer for assistance in generating maps. A.J.F., A.J.B. and H.A.W. are affiliated with the Queensland Centre of Mental Health Research (QCMHR), which receives its core funding from the Queensland Department of Health. A.J.S. was also affiliated with QCMHR during the data collection phase of this paper. T.V. and R.N. have received funding from the Global Burden of Disease project.

Declaration of Interest

None.

References

Angst J, Gamma A, Gastpar M, Lépine JP, Mendlewicz J, Tylee A (2002). Gender differences in depression. Epidemiological findings from the European DEPRES I and II studies. *European Archives of Psychiatry and Clinical Neuroscience* **252**, 201–209.

APA (2000). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn, text revision. APA: Washington, DC.

Ayuso-Mateos JL (2000). *Global Burden of Bipolar Disorder in the year 2000*, version 1.World Health Organization: Geneva.

Baker WL, White CM, Cappelleri JC, Kluger J, Coleman CI (2009). Understanding heterogeneity in meta-analysis: the role of meta-regression. *International Journal of Clinical Practice* **63**, 1426–1434.

Brhlikova P, Pollock AM, Manners R (2010). Global burden of disease estimates of depression – how reliable is the epidemiological evidence? *Journal of the Royal Society of Medicine* **104**, 25–35.

Bromet E, Andrade LH, Hwang I, Sampson NA, Alonso J, de Girolamo G, de Graaf R, Demyttenaere K, Hu C, Iwata N, Karam AN, Kaur J, Kostyuchenko S, Lepine JP, Levinson D, Matschinger H, Mora ME, Browne MO, Posada-Villa J, Viana MC, Williams DR, Kessler RC (2011). Cross-national epidemiology of DSM-IV major depressive episode. *BMC Medicine* **9**, 90.

Bromet EJ, Gluzman SF, Paniotto VI, Webb CPM, Tintle NL, Zakhozha V, Havenaar JM, Gutkovich Z, Kostyuchenko S, Schwartz JE (2005). Epidemiology of psychiatric and alcohol disorders in Ukraine: findings from the Ukraine World Mental Health Survey. *Social Psychiatry and Psychiatric Epidemiology* **40**, 681–690.

Chang SM, Hahm BJ, Lee JY, Shin MS, Jeon HJ, Hong JP, Lee HB, Lee DW, Cho MJ (2008). Cross-national difference in the prevalence of depression caused by the diagnostic threshold. *Journal of Affective Disorders* **106**, 159–167.

Collins PY, Patel V, Joestl SS, March D, Insel TR, Daar AS, Anderson W, Dhansay MA, Phillips A, Shurin S, Walport M, Ewart W, Savill SJ, Bordin IA, Costello EJ, Durkin M, Fairburn C, Glass RI, Hall W, Huang Y, Hyman SE, Jamison K, Kaaya S, Kapur S, Kleinman A, Ogunniyi A, Otero-Ojeda A, Poo MM, Ravindranath V, Sahakian BJ, Saxena S, Singer PA, Stein DJ (2011). Grand challenges in global mental health. *Nature* **475**, 27–30.

Culbertson FM (1997). Depression and gender. An international review. *American Psychologist* **52**, 25–31.

DerSimonian R, Kacker R (2007). Random-effects model for meta-analysis of clinical trials: an update. *Contemporary Clinical Trials* **28**, 105–114.

DerSimonian R, Laird N (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials* **7**, 177–188.

Eaton WW, Kramer M, Anthony JC, Dryman A, Shapiro S, Locke BZ (1989). The incidence of specific DIS/DSM-III mental disorders: data from the NIMH Epidemiologic Catchment Area Program. *Acta Psychiatrica Scandinavica* **79**, 163–178.

Faravelli C, Degl'Innocenti BG, Aiazzi L, Incerpi G, Pallanti S (1990). Epidemiology of mood disorders: a community survey in Florence. *Journal of Affective Disorders* **20**, 135–141.

Global Burden of Disease (2009). Global burden of diseases, injuries and risk factors study operations manual (http://www.globalburden.org/gbdops.html).

Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003). Measuring inconsistency in meta-analyses. *BMJ* **327**, 557–560.

Hollander AC, Ekblad S, Mukhamadiev D, Muminova R (2007). The validity of screening instruments for posttraumatic stress disorder, depression, and other anxiety symptoms in Tajikistan. *Journal of Nervous Mental Disease* **195**, 955–958.

Hollifield M, Warner TD, Lian N, Krakow B, Jenkins JH, Kesler J, Stevenson J, Westermeyer J (2002). Measuring trauma and health status in refugees: a critical review. *Journal of the American Medical Association* **288**, 611–621.

Horwitz AV, Wakefield JC (2007). *The Loss of Sadness: How Psychiatry Transformed Normal Sorrow into Depressive Disorder*. Oxford University Press: New York.

Ihara K, Muraoka Y, Oiji A, Nadaoka T (1998). Prevalence of mood disorders according to DSM-III-R criteria in the community elderly residents in Japan. *Environmental Health and Preventive Medicine* **3**, 44–49.

Institute for Health Metrics and Evaluation (2012). Burden of disease (http://www.healthmetricsandevaluation.org/research/team/burden-disease#/overview).

Kessler RC (1999). The World Health Organization International Consortium in Psychiatric Epidemiology (ICPE): initial work and future directions – the NAPE Lecture 1998. Nordic Association for Psychiatric Epidemiology. *Acta Psychiatrica Scandinavica* **99**, 2–9.

Kessler RC, Üstün TB (editors) (2008). *The WHO World Mental Health Surveys: Global Perspectives on the Epidemiology of Mental Disorders*. Cambridge University Press: New York.

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

Kuehner C (2003). Gender differences in unipolar depression: an update of epidemiological findings and possible explanations. *Acta Psychiatrica Scandinavica* **108**, 163–174.

Lewinsohn PM, Hops H, Roberts RE, Seeley JR, Andrews JA (1993). Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III-R disorders in high school students. *Journal of Abnormal Psychology* **102**, 133–144.

Lopes Cardozo B, Bilukha OO, Gotway CA, Wolfe MI, Gerber ML, Anderson M (2005). Mental health of women in postwar Afghanistan. *Journal of Women's Health* **14**, 285–293.

Moffitt TE, Caspi A, Taylor AJ, Kokaua J, Milne BJ, Polanczyk G, Poulton R (2010). How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective *versus* retrospective ascertainment. *Psychological Medicine* **40**, 899–909.

Mogga S, Prince M, Alem A, Kebede D, Stewart R, Glozier N, Hotopf M (2006). Outcome of major depression in Ethiopia: population-based study. *British Journal of Psychiatry* **189**, 241–246.

Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *PLoS Med* **6**, e1000097.

Murray CJL, Lopez AD (editors) (1996). *The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020*. Published by the Harvard School of Public Health on behalf of the World Health Organization and the World Bank; distributed by Harvard University Press: Cambridge, MA.

Murray CJL, Lopez AD, Black R, Mathers CD, Shibuya K, Ezzati M, Salomon JA, Michaud CM, Walker N, Vos T (2007). Global Burden of Disease 2005: call for collaborators. *Lancet* **370**, 109–110.

Patten SB (2001). The duration of major depressive episodes in the Canadian general population. *Chronic Diseases in Canada* **22**, 6–11.

Paykel ES, Brugha T, Fryers T (2005). Size and burden of depressive disorders in Europe. *European Neuropsychopharmacology* **15**, 411–423.

Phillips MR, Zhang J, Shi Q, Song Z, Ding Z, Pang S, Li X, Zhang Y, Wang Z (2009). Prevalence, treatment, and associated disability of mental disorders in four provinces in China during 2001–05: an epidemiological survey. *Lancet* **373**, 2041–2053.

Prince M, Patel V, Saxena S, Maj M, Maselko J, Phillips MR, Rahman A (2007). No health without mental health. *Lancet* **370**, 859–877.

Rodin D, Van Ommeren M (2009). Commentary: Explaining enormous variations in rates of disorder in trauma-focused psychiatric epidemiology after major emergencies. *International Journal of Epidemiology* **38**, 1045–1048.

Saha S, Chant D, Welham J, McGrath J (2005). A systematic review of the prevalence of schizophrenia. *PLoS Medicine* **2**, e141.

Salomon JA (2010). New disability weights for the global burden of disease. *Bulletin of the World Health Organization* **88**, 879.

Salomon JA, Murray CJ, Üstün B, Chatterji S (2003). Health state valuations in summary measures of population health. In *Health Systems Performance Assessment: Debates, Methods and Empiricism* (ed. C. J. Murray and D. B. Evans), pp. 408–436. World Health Organization: Geneva.

Seedat S, Scott KM, Angermeyer MC, Berglund P, Bromet EJ, Brugha TS, Demyttenaere K, de Girolamo G, Haro JM, Jin R, Karam EG, Kovess-Masfety V, Levinson D, Medina Mora ME, Ono Y, Ormel J, Pennell BE, Posada-Villa J, Sampson NA, Williams D, Kessler RC (2009). Cross-national associations between gender and mental disorders in the World Health Organization World Mental Health Surveys. *Archives of General Psychiatry* **66**, 785–795.

Simon GE, Goldberg DP, Von Korff M, Üstün TB (2002). Understanding cross-national differences in depression prevalence. *Psychological Medicine* **32**, 585–594.

Simon GE, VonKorff M (1995). Recall of psychiatric history in cross-sectional surveys: implications for epidemiologic research. *Epidemiologic Reviews* **17**, 221–227.

Skapinakis P, Lewis G (2001). Epidemiology in community psychiatric research: common uses and methodological issues. *Epidemiologica e Psichiatria Sociale* **10**, 18–26.

Steel Z, Chey T, Silove D, Marnane C, Bryant RA, van Ommeren M (2009). Association of torture and other potentially traumatic events with mental health outcomes among populations exposed to mass conflict and displacement: a systematic review and meta-analysis. *JAMA* **302**, 537–549.

Stefánsson GE, Líndal E, Björnsson JK, Gudmundsdóttir A (1994). Period prevalence rates of specific mental disorders in an Icelandic cohort. *Social Psychiatry and Psychiatric Epidemiology* **29**, 119–125.

Sterne JAC, Egger M, Sutton AJ (2008). Meta-analysis software. In *Systematic Reviews in Health Care: Meta-Analysis in Context*, 2nd edn (ed. M. Egger, G. D. Smith and D. G. Altman), pp. 336–346. BMJ Publishing Group: London.

Susser E, Shrout PE (2010). Two plus two equals three? Do we need to rethink lifetime prevalence? *Psychological Medicine* **40**, 895–897.

Tabachnick BG, Fidell LS (editors) (2007). *Using Multivariate Statistics*. Pearson/Allyn and Bacon: Boston.

United Nations (2007). *World Population Prospects: The 2006 Revision, Highlights*. Working Paper No. ESA/P/WP.202. United Nations Department of Economic and Social Affairs, Population Division: New York.

Üstün TB, Ayuso-Mateos JL, Chatterji S, Mathers C, Murray CJ (2004). Global burden of depressive disorders in the year 2000. *British Journal of Psychiatry* **184**, 386–392.

Van Houwelingen HC, Arends LR, Stijnen T (2002). Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statistics in Medicine* **21**, 589–624.

Ventevogel P, De Vries G, Scholte WF, Shinwari NR, Faiz H, Nassery R, van den Brink W, Olff M (2007). Properties of the Hopkins Symptom Checklist-25 (HSCL-25) and the Self-Reporting Questionnaire (SRQ-20) as screening instruments used in primary care in Afghanistan. *Social Psychiatry and Psychiatric Epidemiology* **42**, 328–335.

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

Wang PS, Aguilar-Gaxiola S, Alonso J, Angermeyer MC, Borges G, Bromet EJ, Bruffaerts R, de Girolamo G, de Graaf R, Gureje O, Haro JM, Karam EG, Kessler RC, Kovess V, Lane MC, Lee S, Levinson D, Ono Y, Petukhova M, Posada-Villa J, Seedat S, Wells JE (2007). Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. *Lancet* **370**, 841–850.

Weich S, Araya R (2004). International and regional variation in the prevalence of common mental disorders: do we need more surveys? *British Journal of Psychiatry* **184**, 289–290.

Weissman MM, Bland RC, Canino GJ, Faravelli C, Greenwald S, Hwu HG, Joyce PR, Karam EG, Lee CK, Lellouch J, Lepine JP, Newman SC, Rubio-Stipec M, Wells JE, Wickramaratne PJ, Wittchen H, Yeh EK (1996). Cross-national epidemiology of major depression and bipolar disorder. *JAMA* **276**, 293–299.

Wittchen HU (2000). Epidemiological research in mental disorders: lessons for the next decade of research – the NAPE Lecture 1999. Nordic Association for Psychiatric Epidemiology. *Acta Psychiatrica Scandinavica* **101**, 2–10.

World Health Organization (1992). *The ICD-10 Classification of Mental and Behavioural Disorders. Clinical Descriptions and Diagnostic Guidelines*. World Health Organization: Geneva.