Psychological Medicine, 2006, **36**, 257–267. © 2005 Cambridge University Press
doi:10.1017/S0033291705006367 First published online 7 December 2005. Printed in the United Kingdom

Twelve-month prevalence, severity, and unmet need for treatment of mental disorders in metropolitan China

YU-CUN SHEN[1], MING-YUAN ZHANG[2], YUE-QIN HUANG[1], YAN-LING HE[2], ZHAO-RUI LIU[1], HUI CHENG[1], ADLEY TSANG[4], SING LEE[3]* AND RONALD C. KESSLER[5]

[1] *Peking University Institute of Mental Health, Peking*; [2] *Shanghai Mental Health Center, Shanghai*; [3] *Department of Psychiatry, The Chinese University of Hong Kong, HKSAR*; [4] *Hong Kong Mood Disorders Center, The Chinese University of Hong Kong, HKSAR, People's Republic of China*; [5] *Department of Health Care Policy, Harvard Medical School, Boston, MA, USA*

ABSTRACT

Background. Psychiatric epidemiological surveys in China have repeatedly found much lower prevalence estimates than in most other parts of the world.

Method. Face-to-face household interviews of 5201 subjects (2633 in Beijing and 2568 in Shanghai respectively) were conducted from November 2001 to February 2002 using a multistage household probability sampling method. A Chinese version of the World Health Organization Composite International Diagnostic Interview (CIDI) was used for assessment.

Results. Twelve-month prevalence of any DSM-IV mental disorder in metropolitan China is estimated to be 7·0%, with major depressive disorder (2·0%), specific phobia (1·9%), and intermittent explosive disorder (1·7%) the most common disorders. Of these, 13·9% are classified as serious, 32·6% moderate, and 53·5% mild. Only 3·4% of respondents with any disorder sought treatment within the previous 12 months.

Conclusions. Although the general pattern of disorders, risk factors, and unmet need for treatment are similar to those in other countries, a low prevalence of mental disorders is found in metropolitan China. Resolving methodological problems that cause downward bias in estimates, such as stigma-related under-reporting and diagnostic incongruity with a somatopsychic mode of symptom presentation may lead to more accurate and probably higher prevalence estimates in future epidemiological studies. As a low prevalence still translates into an enormous number of people in China, measures are urgently needed to address the huge unmet need for treatment of mental disorders.

INTRODUCTION

The prevalence of mental disorders in China has been a focus of controversy for several decades (Kleinman, 1986, 2004; Parker *et al.* 2001). In 1982 and 1993, leading psychiatrists in China conducted the twelve-region survey and the seven-region survey using the same Chinese version of the Present State Examination and nearly identical methodologies (Twelve-Region Psychiatric Epidemiological Study Work Group, 1986; Zhang *et al.* 1998). These surveys included large and nationally representative samples ($n = 38\,136$ and 19 223 respectively). The point prevalence of affective psychosis (comparable to bipolar disorder and severe depression

* Address for correspondence: Professor Sing Lee, Director Hong Kong Mood Disorders Center, 7A, Block E Staff Quarters, Prince of Wales Hospital, Shatin, NT, Hong Kong.
(Email: singlee@cuhk.edu.hk)

with psychotic/endogenous features combined) was estimated to be 0·076% and 0·052% in the two surveys, while the point prevalence of all neuroses [comparable to non-severe major depressive episodes and various anxiety disorders as conceptualized in the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (DSM-IV)] was estimated to be 2·21% and 1·51%.

Although these prevalence estimates are considerably lower than those found in recent psychiatric epidemiological surveys elsewhere in the world (Demyttenaere *et al.* 2004), direct comparison cannot be made with the results of other surveys because of numerous methodological differences. For example, many diagnostic categories used in the two surveys were either absent or different from those in international use. Instead of generalized or specific anxiety disorders, an all-inclusive category of 'neurosis' with a 3-month duration criterion was used (Lee, 1996). Unlike the DSM-IV definition of major depressive episode, affective psychosis in the surveys had a duration criterion of 4 weeks.

Many researchers have reservations about the Chinese findings because epidemiological studies of Chinese people in Hong Kong, Taiwan, and the USA do not indicate that mental disorders are as rare as the two surveys suggest (Cheng, 1989; Chen *et al.* 1993; Takeuchi *et al.* 1998). Two classes of explanation, albeit empirically little examined and themselves debated, have been offered for the low prevalence estimates in the surveys. The first emphasizes methodological factors that lead to downward bias in estimates, including errors in sampling, stigma-induced under-reporting, and culturally shaped symptom reporting such as somatization (Cheng, 1989; Shen *et al.* 1998; Parker *et al.* 2001; Kleinman, 2004). The second emphasizes substantive processes that cause prevalence to be genuinely low, such as a resilient family system and a cultural tradition of withstanding hardship that buffer Chinese people against mental disorders (Chen *et al.* 1993; Parker *et al.* 2001).

As a first step in a program of research aimed at improving the methodology of epidemiological research prior to carrying out a new national survey of mental disorders in China, we took part in the World Health Organization (WHO) World Mental Health (WMH) Survey Initiative (http://www.hcp.med.harvard.edu/wmh). The Chinese WMH survey was carried out in Beijing and Shanghai. Standardized methods of sampling, training, field procedures and measures were adopted. Our first aim was to re-examine the Chinese epidemiology of mental disorders in a cross-national context. The second aim was to generate a sample of respondents who could participate in further methodological research after the survey to refine the interview methods in preparation for a national survey.

Very brief summary prevalence estimates of four broad disorder classes were previously reported from this survey in the context of a cross-national comparison of WMH results (Demyttenaere *et al.* 2004). The current paper presents detailed data on prevalence, severity, correlates, predictors and treatment of individual disorders. It also discusses their implications for mental health policy in China and our future methodological studies.

METHOD

Measures

The survey used the WHO Composite International Diagnostic Interview (CIDI 3.0), a fully structured diagnostic interview, to assess disorders and treatment (Kessler & Üstün, 2004). Disorders include DSM-IV anxiety, mood, impulse-control, and substance use disorders. CIDI organic exclusion rules were applied in making all diagnoses. A severity gradient was created that defined 12-month disorders as serious, moderate, or mild. Details of this classification and its external validation can be found elsewhere (Demyttenaere *et al.* 2004). Briefly put, a serious disorder included such disorder-related conditions as bipolar disorder, substance dependence with a physiological dependence syndrome, and other disorders associated with a suicide plan, a suicide attempt, hospitalization, treatment with an antipsychotic medication, or severe role impairment. Respondents whose 12-month disorder did not meet criteria for serious were classified as either moderate or mild, based on their responses to the disorder-specific Sheehan Disability Scales (Leon *et al.* 1997).

The CIDI was translated into Chinese and back-translated using a standard WHO

protocol. An expert panel made up of three academic psychiatrists with epidemiological expertise (Y.-C.S., M.-Y.Z., S.L.) and a survey methodologist from the Research Center for Contemporary China (RCCC) in Beijing (fully trained in the CIDI at the University of Michigan and responsible for fieldwork) evaluated its content validity, tested it with Chinese patients, and revised it to ensure that the Chinese terms used were easily understood by lay people.

Subjects

Subjects were urban dwellers who met these criteria: adults aged 18–70 years; living in a family household; formally registered in a non-agricultural household; and residing within the urban districts of Beijing and Shanghai.

Sampling

A stratified two-stage systematic selection scheme was used to select the sample with equal probability for every eligible individual in the target population to be selected into the sample. Because this sampling procedure avoids the within-household clustering that is typical of community surveys, no within-household probability of selection weight was applied to the data, increasing the efficiency of the sample. Because the full interview is very long, the interview was divided into two parts in order to reduce respondent burden for respondents who did not have any mental health problems. Part I included the core diagnostic assessment. Part II included information about correlates and disorders of secondary interest. All respondents completed Part I. All Part I respondents who met criteria for any disorder plus a subsample of approximately 25% of those who did not were administered Part II.

'Neighborhood committees (NCs) are local community organizations in urban China that have been established since 1954 to promote communication and mutual help among residents on various issues of everyday living. Each NC consists typically of 100–700 households. Since all urban households belong to such committees, the latter were used as RCCCs preselected primary sampling units (PSUs). The sampling team first approached 50 NCs in each city to check their actual conditions against the obtained demographic data of the PSUs. Neighborhoods were selected using random sampling, which means that there is no selection bias at that stage. Of course, response bias could exist even though selection bias did not. That is, if the leaders of neighborhoods with poor mental health refused to participate in the survey, we would have response bias in the survey. However, because only a very small amount of refusal occurred at this level of the sample, bias of this sort is very unlikely.

A total of 47 of these NCs in Beijing and 44 in Shanghai were chosen as final PSUs for the survey. According to the pre-decided sampling interval and random starting points, the desired number of households was sampled from the household registration list in each NC. The sampler then checked the address and demographic information of all the eligible members of every selected household against the RCCC's standard Household Sampling Form (HSF). After editing the HSF, the field managers used the data randomly to select eligible persons as respondents. Of the 47 PSUs selected in Beijing there were 4024 eligible individuals who comprised the final master sample of respondents, while 3856 respondents out of the 44 selected PSUs were chosen in Shanghai. After a complete description of the study to the subjects, written informed consent was obtained and interviews were obtained with as many of these designated respondents as possible.

Sample size and weighting

The desired sample size was 2500 completed interviews in each city. The actual sample size was 2633 in Beijing and 2568 in Shanghai. The response rates were 74·8% and 74·6% respectively. Weights were applied to the data to adjust for discrepancies between the sample and population census data on the cross-classification of key sociodemographic variables. A second weight was applied to the Part II subsample to adjust for the oversampling of Part I respondents who met criteria for a core disorder.

Training and field procedure

RCCC's field managers conducted a one-day training session for field samplers in the two cities following their standard on-site training procedure. The trainers introduced the project, its sampling, principles, methods, and specific requirements for implementing the sampling

design. They also carried out briefings on the field conditions and potential difficulties at different sampling stages, demonstration, and exercise. RCCC's field managers then conducted 7 days of training for the interviewers in each city following their standard on-site training procedures. This included six main sessions (introduction; sampling design and procedures; review of the questionnaire; methods and techniques of field interviewing; briefing on field conditions; potential difficulties in the different neighborhoods) and in- and out-of-classroom exercises for interviewers.

Data analysis

Data are reported on prevalence, severity, association of severity with treatment, and sociodemographic predictors of prevalence, severity, disorder types and treatment. Simple cross-tabulations were used to calculate prevalence and severity. Associations of severity with days out of role and treatment were examined using analysis of variance. Standard errors were estimated using the Taylor series linearization method (Wolter, 1985) implemented in the SUDAAN software package (Research Triangle Institute, Durham, NC, USA) to adjust for the clustering and weighting of data. Logistic regression analysis (Hosmer & Lemeshow, 2000) was used to study the sociodemographic predictors of prevalence, severity, health-care treatment and disorder types. Multivariate tests were made using Wald χ^2 and F tests computed from design-adjusted coefficient variance/covariance matrices. Confidence intervals were estimated using the Taylor Series method with SUDAAN software to adjust for clustering and weighting. Statistical significance was based on two-sided tests evaluated at the 0·05 level of significance.

RESULTS

Demographic distribution

The demographic distribution of the weighted sample in Beijing is similar to that of Shanghai (results available on request). The demographic distribution of the combined and weighted sample is similar to the general population on post-stratification variables after being weighted (Table 1).

Table 1. *Demographic distribution of the sample compared with the population on post-stratification variables*

	Part I unweighted %	Part I weighted %	Census %
Age (yr)			
18–24	10·9	17·7	17·3
25–29	5·4	10·8	11·1
30–34	7·6	11·0	10·8
35–39	11·8	12·3	12·2
40–44	16·8	13·2	13·3
45–49	14·8	11·7	11·8
50–54	9·5	7·4	7·5
55–59	5·7	4·7	4·7
60–64	7·6	5·0	5·0
65–70	10·0	6·2	6·2
Sex			
Male	48·7	52·6	52·7
Female	51·3	47·4	47·3
	Part II unweighted %	**Part II weighted %**	**Census %**
Age (yr)[a]			
18–29	16·1	27·3	26·5
30–39	19·7	22·4	22·1
40–49	32·9	21·9	22·7
50–59	15·1	11·8	11·7
60–70	16·2	16·6	16·9
Sex			
Male	50·2	52·3	51·8
Female	49·8	47·7	48·2
Education[b]			
0–8	13·9	17·1	20·1
9–11	34·9	27·9	35·9
12	23·3	26·1	18·3
13–15	18·8	17·4	16·3
⩾16	9·2	11·5	9·4

[a] Age is less stratified in part II than in part I owing to smaller part II sample size.
[b] Education was only assessed in part II of the survey.

Prevalence and severity

The disorders with highest estimated 12-month prevalence are major depressive disorder (2·0%), specific phobia (1·9%), and intermittent explosive disorder (1·7%). Regarding class of disorder, impulse-control disorders are estimated to be the most prevalent (3·1%), followed by anxiety disorders (2·7%), mood disorders (2·2%), and substance use disorders (1·6%). The 12-month prevalence estimate of any disorder is 7·0% in the total sample.

Among all respondents with 12-month disorders, 13·9% are classified serious, 32·6% moderate, and 53·5% mild. The distribution

Table 2. *Twelve-month prevalence and severity of DSM-IV disorders*

	Any severity			Serious		Moderate		Mild	
Diagnosis	%	(S.E.)	χ^{2a}	%	(S.E.)	%	(S.E.)	%	(S.E.)
Anxiety disorder									
Panic disorder	0·2	(0·1)	—	10·0	(10·0)	64·0	(19·7)	26·0	(17·6)
Generalized anxiety disorder	0·8	(0·3)	0·0	3·7	(2·6)	59·9	(14·0)	36·4	(13·1)
Specific phobia	1·9	(0·5)	0·2	17·7	(9·5)	42·2	(14·0)	40·0	(10·1)
Social phobia	0·2	(0·1)	6·4*	24·0	(14·6)	44·4	(20·2)	31·7	(15·6)
Agoraphobia without panic	0·0	(0·0)	—	—	—	—	—	—	—
Post-traumatic stress disorder	0·2	(0·1)	1·1	0·0	(0·0)	90·5	(7·8)	9·5	(7·8)
Obsessive–compulsive disorder	—	—	—	—	—	—	—	—	—
Any anxiety disorder	2·7	(0·5)	0·3	14·1	(6·8)	44·6	(10·9)	41·3	(8·2)
Mood disorder									
Major depressive disorder	2·0	(0·3)	1·0	16·5	(5·4)	51·1	(7·9)	32·4	(7·7)
Dysthymia	0·1	(0·1)	0·0	46·4	(17·7)	53·6	(17·7)	0·0	(0·0)
Bipolar I–II disorders	0·1	(0·1)	0·3	100·0	(0·0)	0·0	(0·0)	0·0	(0·0)
Any mood disorder	2·2	(0·4)	0·7	21·4	(6·7)	48·6	(7·6)	29·9	(7·3)
Impulse disorder									
Oppositional-defiant disorder	—	—	—	—	—	—	—	—	—
Conduct disorder	0·1	(0·1)	—	0·0	(0·0)	0·0	(0·0)	100·0	(0·0)
Attention deficit hyperactivity disorder	—	—	—	—	—	—	—	—	—
Intermittent explosive disorder	1·7	(0·4)	13·4*	4·9	(2·7)	25·7	(15·8)	69·4	(16·5)
Any impulse-control disorder	3·1	(0·7)	8·3*	5·6	(3·1)	24·3	(17·1)	70·2	(18·0)
Substance disorder									
Alcohol abuse or dependence	1·6	(0·4)	25·8*	22·0	(8·0)	18·1	(6·2)	59·9	(9·4)
Alcohol dependence	0·6	(0·1)	5·2*	54·6	(14·4)	45·4	(14·4)	0·0	(0·0)
Drug abuse or dependence	0·1	(0·0)	1·4	10·6	(12·0)	15·9	(17·2)	73·4	(22·5)
Drug dependence	0·0	(0·0)	—	—	—	—	—	—	—
Any substance use disorder	1·6	(0·4)	27·5*	21·3	(7·7)	18·1	(6·0)	60·6	(9·0)
Any disorder									
1 or more disorder	7·0	(0·9)	8·3*	13·9	(2·7)	32·6	(4·0)	53·5	(4·5)
0 disorder	93·0	(0·9)	—	—	—	—	—	—	—
Exactly 1 disorder	5·4	(0·8)	—	9·3	(3·8)	27·7	(4·0)	62·9	(6·0)
Exactly 2 disorders	0·9	(0·3)	—	35·4	(8·3)	28·2	(8·8)	36·4	(8·6)
3 or more disorders	0·7	(0·2)	—	23·4	(12·3)	76·6	(12·3)	0·0	(0·0)
Total sample	—	—	—	1·0	(0·2)	2·3	(0·4)	3·7	(0·5)

[a] χ^2 tests between Beijing and Shanghai.
* Significant at the 0·05 level, two-sided test.

of severity across classes of disorder is different from the distribution of prevalence, with mood disorders having the highest percentage of cases classified serious (21·4%) and impulse-control disorders the lowest (5·6%). Individual disorders within each class with the highest percentage of seriousness are bipolar I–II disorders (100·0%) among the mood disorders, alcohol dependence (54·6%) among the substance disorders, social phobia (24%) among the anxiety disorders, and intermittent explosive disorder (4·9%) among the impulse-control disorders (Table 2).

Inter-city comparison showed that 12-month prevalence rates of most disorders were not significantly different except for social phobia (Beijing, 0%; Shanghai, 0·3%), intermittent explosive disorder (Beijing, 2·6%; Shanghai, 0·6%), alcohol abuse/dependence (Beijing, 2·5%; Shanghai, 0·4%) and alcohol dependence (Beijing, 0·8%; Shanghai, 0·3%) (detailed results available on request).

Association of severity with treatment types

The CIDI assessed 12-month treatment by asking respondents if they ever saw any of a long list of professionals either on an outpatient basis or as an in-patient for problems with their emotions, nerves, mental health, or use of substances. The present analysis grouped these professionals into four major groups: general medical (e.g. general practitioners), mental health (e.g. psychiatrists), healthcare (general practitioners and psychiatrists combined), and non-healthcare (e.g. spiritualists). Respondents with different levels of severity

Table 3. *Severity by treatment*

	Severe or moderate		Mild		Any		None	
Treatment	%	(S.E.)	%	(S.E.)	%	(S.E.)	%	(S.E.)
General medical	15·2	(7·9)	0·0	(0·0)	2·3	(0·5)	2·0	(0·5)
Mental health	4·6	(2·4)	1·7	(1·2)	0·6	(0·2)	0·4	(0·2)
Healthcare	18·6	(8·2)	1·7	(1·2)	2·9	(0·5)	2·4	(0·6)
Non-healthcare	9·1	(7·0)	0·0	(0·0)	1·0	(0·3)	0·7	(0·2)
Any treatment	19·8	(8·3)	1·7	(1·2)	3·4	(0·6)	2·9	(0·6)
No treatment	80·2	(8·3)	98·3	(1·2)	96·6	(0·6)	97·1	(0·6)
χ^{2a}	246·9*		1·2		0·2		—	

[a] *Versus* no-disorder group.
* Significant at the 0·05 level, two-sided test.

were compared on their treatment-seeking from the four major groups of professionals. As results were very similar in the two cites, only summary results for both cities combined are reported here.

Among all respondents with at least one 12-month disorder, only 3·4% sought treatment within 12 months of the interview. Cases classified severe or moderate had a higher proportion in treatment (19·8%) than those with no disorder (2·9%) and with a mild disorder (1·7%). The most common type of treatment for all cases as well as severe/moderate cases and non-cases was general medical treatment. Among the mild cases, general medical and mental health treatment were equally common (Table 3).

Sociodemographic predictors of disorders

Analysis documented that the odds ratio (OR) of any 12-month CIDI disorder is significantly related to age, sex, and family income. Young age is a risk factor and being female is a protective factor. The OR of severity is significantly related to sex and marital status. Being female and never married are risk factors of severity (Table 4). Regarding the four classes of disorders, low-average income and being separated/widowed/divorced are risk factors for mood disorders. Low income and no/primary school education are risk factors for anxiety disorders. Low-average income is a risk factor and being never married is a protective factor for impulse-control disorders. Young age is a risk factor, and being female, of average income and never married are protective factors for substance use disorders.

Some statistically significant interactions emerged in the analysis of age and education as risk factors for any disorders between the two cities. This also applied to income for severity of disorders, e.g. young age being the risk factor of any disorders was more significantly shown in Beijing than in Shanghai. However, inspection of within-city patterns in most of these cases showed that the interaction was due to an outlier of little significance (detailed results available on request).

Sociodemographic predictors of treatment

Respondents with severe and moderate disorders were significantly more likely to receive treatment than mild disorders in both cities. After controlling for disorder severity, analysis documented that the OR of receiving treatment is significantly related to age and family income, with lower income and higher age being associated with higher treatment rates (Table 5). Gender, education, and marital status were not significantly related to treatment after controlling for disorder severity.

Some statistically significant interactions emerged in analysis. Inspection of within-city patterns showed that the interaction was due to an outlier in most of these cases except the following: respondents with moderate to severe disorders were more likely to receive treatment in Shanghai than in Beijing; and older respondents were more likely to receive treatment in Beijing (detailed results available on request).

DISCUSSION

Methodological considerations

It is not commonly known that psychiatric epidemiological surveys in China routinely use psychiatric doctors as interviewers (Twelve-Region Psychiatric Epidemiological Study Work Group, 1986; Zhang *et al.* 1998). One reason is that the latter are relatively inexpensive to use. But there is another reason, namely, that Chinese epidemiological researchers are unfamiliar with, and at times skeptical about, the established practice where lay people are rigorously trained and able to make psychiatric diagnoses.

Although clinicians may make more valid diagnoses, there is a potential drawback of the above Chinese approach to community

Table 4. *Sociodemographic predictors of WMH-CIDI/DSM-IV 12-month prevalence, diagnosis and severity of diagnosis*

	Any disorder		Severity/disorder[a]		Mood		Anxiety		Impulse-control		Substance	
	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)
Sex												
Male	1·0	—	1·0	—	1·0	—	1·0	—	1·0	—	1·0	—
Female	0·5*	(0·3–0·9)	3·0*	(1·1–8·1)	0·8	(0·4–1·4)	1·3	(0·6–2·6)	0·2	(0·0–1·9)	0·1*	(0·0–0·2)
χ^2	6·7*		5·0*		0·8		0·4		2·2		21·2*	
Age (yr)												
18–34	2·9*	(1·5–5·8)	0·8	(0·3–2·3)	2·0	(0·8–4·8)	0·5	(0·2–1·6)	0·8	(0·2–2·8)	20·4*	(3·6–115·5)
35–49	1·7	(0·9–3·0)	1·8	(0·3–10·0)	0·9	(0·4–1·9)	0·7	(0·3–1·5)	1·0	—	6·7*	(1·7–26·8)
50–64	—	—	—	—	—	—	—	—	—	—	—	—
65+	1·0	—	1·0	—	1·0	—	1·0	—	—	—	1·0	—
χ^2	10·2*		1·1		3·5		1·7		0·1		13·1*	
Income												
Low	1·4	(0·6–3·6)	0·5	(0·1–2·6)	1·8	(0·7–4·6)	4·7*	(1·7–13·1)	1·7	(0·2–11·6)	0·6	(0·2–1·8)
Low-average	1·3	(0·6–2·6)	2·0	(0·6–6·4)	2·8*	(1·1–7·1)	3·1*	(1·6–5·9)	5·8	(1·0–33·9)	0·1*	(0·0–0·6)
High-average	0·8	(0·4–1·6)	1·4	(0·4–4·2)	1·0	(0·4–2·3)	2·5*	(1·1–5·4)	1·2	(0·2–10·1)	0·2*	(0·1–0·7)
High	1·0	—	1·0	—	1·0	—	1·0	—	1·0	—	1·0	—
χ^2	9·2*		2·6		8·0*		23·7*		4·6		18·7*	
Marital status												
Married/cohabiting	1·0	—	1·0	—	1·0	—	1·0	—	1·0	—	1·0	—
Separated/widowed/ divorced	1·2	(0·6–2·3)	2·5	(0·7–8·7)	2·8*	(1·0–7·6)	0·6	(0·2–1·8)	0·9	(0·1–9·0)	0·5	(0·1–3·3)
Never married	0·3*	(0·1–1·0)	6·1*	(1·2–29·7)	0·4	(0·1–1·2)	1·7	(0·5–5·3)	0·1*	(0·0–0·4)	0·1*	(0·0–0·4)
χ^2	4·8		6·5*		7·7*		2·0		9·5*		16·0*	
Education												
None and primary	2·1	(0·9–4·7)	1·1	(0·2–7·3)	1·5	(0·3–6·7)	2·6*	(1·1–6·2)	1·4	(0·0–85·4)	0·4	(0·1–4·0)
Junior High School	1·0	(0·5–2·1)	1·6	(0·5–5·0)	1·4	(0·5–3·4)	1·0	(0·3–3·1)	0·1*	(0·0–0·7)	1·4	(0·4–4·9)
Senior High School	1·3	(0·3–3·1)	0·7	(0·3–2·0)	1·4	(0·7–2·8)	1·3	(0·4–4·5)	1·1	(0·2–5·5)	0·9	(0·3–2·9)
College and beyond	1·0	—	1·0	—	1·0	—	1·0	—	1·0	—	1·0	—
χ^2	4·3		2·9		1·4		5·1*		7·4*		1·5	

OR, Odds ratio; CI, confidence interval.
[a] Serious or moderate (coded 1) *versus* mild (coded 0).
* Significant at the 0·05 level, two-sided test.

epidemiological surveys. This is that Chinese psychiatrists are clinically accustomed to dealing almost entirely with severe mental disorders. As such, they may be inclined to discount less severe depressive and anxiety disorders during diagnostic assessments (Lee, 1996). In fact, even mainland Chinese researchers have now acknowledged that interviews by psychiatrists can lead to underestimates of anxiety and depressive disorders (Shen *et al*. 1998).

The present study is the first Chinese community survey to use lay personnel to make fully structured psychiatric diagnoses. It shows that, when compared with similar Western surveys using basically the same methodology (Demyttenaere *et al*. 2004), quite low prevalence estimates are still obtained. This is despite our having used a state-of-the-art psychiatric epidemiological approach that includes rigorous sampling and interviewer training, a detailed fully structured diagnostic interview, careful field quality control procedures, computerized data entry and diagnostic algorithms, and meticulous statistical analysis.

Our findings are nonetheless instructive because they suggest that some of the methodological explanations previously raised to account for low prevalence in the 12-region and 7-region surveys are probably inappropriate. We obtain the same low prevalence estimates here despite the more sophisticated (albeit not necessarily more valid) methods. However, our results should not be uncritically taken as evidence of excellent mental health among Chinese people. We believe that previously unexamined methodological possibilities may explain the low prevalence estimates. These include stigma-induced conscious non-disclosure and conceptual and measurement problems leading to failure to detect disorders. We therefore plan to carry

Table 5. *Sociodemographic predictors of 12-month treatment*

	Treatment[a]	
	OR	(95% CI)
Sex		
Male	1·0	—
Female	1·3	(0·5–3·1)
χ^2	0·3	
Age (yr)		
18–34	0·1*	(0·0–0·3)
35–49	0·5	(0·2–1·2)
50–64	—	—
⩾65	1·0	—
χ^2	19·3	
Family income		
Low	9·3*	(2·2–38·7)
Low-average	2·7	(0·7–10·6)
High-average	1·6	(0·6–4·4)
High	1·0	—
χ^2	15·7	
Marital status		
Married/cohabiting	1·0	—
Separated/widowed/divorced	0·7	(0·1–3·1)
Never married	1·2	(0·4–3·6)
χ^2	0·6	
Education		
None and primary	0·8	(0·2–3·6)
Junior high school	0·6	(0·2–2·3)
Senior high school	0·9	(0·2–3·5)
College and beyond	1·0	—
χ^2	0·8	
Disorder severity		
None	1·0	—
Mild	0·6	(0·1–3·2)
Moderate	16·2*	(3·7–70·1)
Severe	4·8*	(1·0–22·4)
χ^2	20·6*	

OR, Odds ratio; CI, confidence interval.
[a] controlling for disorder severity.
* Significant at the 0·05 level, two-sided test.

out methodological follow-up studies in the WMH sample to investigate such possibilities.

As part of this methodological work, a preliminary study in which SCID (Structured Clinical Interview for DSM-IV Disorders) interviews (First *et al.* 2002) were conventionally administered by a trained psychiatrist to a small probability subsample of WMH respondents yielded similarly low prevalence estimates, arguing in favor of a substantive interpretation of results. However, our methodological studies will also explore the possibility that these SCID interviews missed clinically significant cases owing to the administrative incompatibility of the SCID with the subjective flow of psychopathological experience in Chinese people. In accordance with our clinical experience, anthropological research in China has long demonstrated that patients diagnosable with DSM-III major depression routinely initiated somatic distress that was readily linked to affective and cognitive psychopathology in skillfully conducted interviews (Kleinman, 1986).

It should be emphasized that this somatopsychic mode of reporting distress does not indicate that all Chinese people merely deny depression (Cheng, 1989). Nor does it tally with the hierarchical configuration of symptoms in the DSM schema (i.e. psychological distress is core whereas somatic distress is peripheral) upon which current psychiatric diagnostic schedules are rigidly based (Lee, 1996). If the similar interpersonal flow of psychopathological experience described above is typical of many respondents in the present study, then the fully structured WMH-CIDI, by focusing obligatorily on psychological core symptoms first, may generate an excessive number of false-negative cases. Our findings therefore motivate us to devise experience-near methods of eliciting diagnoses in community surveys without giving priority to psychological disturbance as an initial core symptom.

Patterns of findings

Despite the above methodological issues, the general pattern of our findings resembles those of other participating WMH countries in several ways (Demyttenaere *et al.* 2004). Because of small sample sizes in each of the two cities studied, we combined the two samples in order to generate enough statistical power to detect significant associations that could not be detected in either city separately. Despite some differences in the prevalence of certain disorders between Beijing and Shanghai, a similar pattern of findings is found in the two cities. These include the usual gradient of severity of disorders, higher prevalence of anxiety than mood disorders, sociodemographic risk factors, and much higher popularity of general medical over psychiatric help-seeking.

As in the previous Chinese surveys, we found a generally low prevalence of substance use disorders. Nonetheless, these disorders have elevated prevalence among younger people (especially male) and those with high income (Table 4). This may be one early sign of the

harmful effect of prosperity on people's mental health in urban China. Although the Chinese national surveys did not cover impulse disorder, the fact that intermittent explosive disorder has a comparatively high prevalence (1·7%) is of interest because uncontrollable irritability (*fanzao*) is reportedly a common symptom among Chinese patients with neuroses or neurasthenia (Lee, 1996).

Unmet treatment need

Previous Chinese surveys did not measure severity of disorder or unmet treatment need. Both of the national surveys used the ninth edition of the Present State Examination (PSE-9) psychosis and neurosis screening schedules that required significant functional impairment lasting over 1 month as a criterion of caseness (Shen *et al*. 1998). This would make psychiatric interviewers biased toward the identification of severe disorders. In contrast, we conducted the first survey in China that encompassed a spectrum of DSM-IV disorders classified by severity.

From a public health perspective, it is striking that unmet treatment need is almost routine, even in the two most developed cities in contemporary China. Given that China has a 1·3 billion population, the low prevalence estimates we demonstrated still translate into an enormous number of people with untreated mental disorders. The finding is indeed disturbing that 96·6% of people with any disorder and 80·2% of those with moderate and severe disorders received no 12-month treatment, even in Beijing and Shanghai where a higher than average concentration of health care resources including better-trained psychiatrists exist. For those people who did seek help from general medical practitioners (who usually have little mental health training in China), the outcome of treatment is unclear.

System barriers to care

Psychiatric stigma has routinely been blamed for the undertreatment of mental disorders, especially in non-Western societies (Lee, 2002). However, stigma cannot be the main reason for the high level of unmet treatment need in our study. This is because the great majority of people with disorders do not receive any form of treatment, including from general medical practitioners who do not stigmatize. Nor is low perceived need a convincing reason because it is most unlikely that some 80% of people with symptoms of moderate and severe mental disorders do not wish to seek care. Thus, stigma, low mental health literacy, and low perceived need cannot adequately explain the gross undertreatment.

Health economists have admonished that one of the most significant casualties during China's transition to a market economy has been the health care system. Once renowned for equity, it is buckling under the pressure to survive with waning Government subsidy. Medical insurance coverage is now non-existent in rural China and only available to about 40% of the urban population, some of whom have depleted their accounts. Consequently, out-of-pocket payment or co-payment for treatment fees is becoming increasingly common. Besides, hospitals have sought to increase revenue through charging patients for new diagnostic procedures, prolonged in-patient stay, or use of expensive new drugs (Yip & Hsiao, 1997).

The consequences of such an economic reality are wastage of already limited health-care resources and structural barriers to care. People suffering financial hardship may avoid seeking medical service or are declined such service. In contrast, those who are financially better off may receive treatment for subthreshold symptoms. Such a disparity may be reflected by the finding, unique among WMH participating countries, that China had a higher percentage of people with no disorder than mild disorder that received 12-month treatment (Table 3). It is also supported by the finding that high income is associated with high treatment rates (Table 5). Since untreated mental disorder causes substantial productivity loss, the unmet treatment need we demonstrated is ironic in the context of China's quest for continued economic growth. The need for multi-level and multi-sector efforts to address unmet treatment need can hardly be overemphasized.

Limitations and future research

First, owing to limited funding, we studied a metropolitan sample even though China is about 75% rural nowadays. While the study provides a foundation on which to build a large nationally representative survey, our results

should not be generalized indiscriminately. We believe that epidemiological studies are urgently needed in vast areas of rural China where suicide rates are high and health insurance is almost non-existent (Yip & Hsiao, 1997; Shen *et al*. 1998). Second, we have not conducted a meaningful number of clinical reappraisals to validate or calibrate the CIDI DSM-IV diagnoses and the Sheehan Disability Scale. As such, our prevalence and severity estimates should be considered preliminary. Third, because of the low prevalence and treatment rate, the small number of cases in specific diagnostic categories, even when data from the two cities are combined, limits the power of statistical analysis. In theory, we can overcome this limitation by studying a very large sample of individuals. However, the cost of doing so is prohibitive in China, where funding for mental health research is grossly inadequate. We therefore believe that it is important to devise methodological improvements that can lead to more accurate and presumably higher prevalence estimates. Regarding methodological issues, we suspect that community diagnostic interviews with Chinese people will be found to work better if they start with questions about physical health and those symptoms that bother them most, followed by asking the diagnostic stem questions on current psychological disturbance. This anthropologically informed 'bottom-up' or somato-psychic approach to symptom elicitation may 'warm up' Chinese respondents and facilitate the disclosure of mood and cognitive symptoms later on in the interview. This will be one key methodological enhancement strategy in our future studies. It may also help to resolve debates over the intricate issue of somatization in Chinese people (Cheng, 1989; Parker *et al*. 2001; Kleinman, 2004).

A further limitation of the present survey is that we excluded schizophrenia from the assessment because of the problems of non-response bias and overestimation of psychotic symptoms by lay people during structured interviews (Kessler *et al*. 2005). Finally, when several other non-Western countries (e.g. India, Indonesia, and Turkey) that have participated in the WMH Initiative complete their surveys, cross-national comparison of data may shed light on some of the issues raised in this paper. If problems such as those we have discussed likewise occur, the CIDI cannot be considered a globally satisfactory instrument until it is successfully adapted to such non-Western communities. The latter is the next task on our agenda.

ACKNOWLEDGMENTS

The Chinese World Mental Health Survey Initiative is supported by the Pfizer Foundation in conjunction with the World Health Organization World Mental Health (WMH) Survey Initiative. We thank the staff of the WMH Data Collection and Data Analysis Coordination Centres for assistance with instrumentation, fieldwork, and consultation on data analysis. These activities were supported by the John D. and Catherine T. MacArthur Foundation, the Pfizer Foundation, the USA Public Health Service (R13-MH066849, R01-MH069864, and R01 DA016558), the Pan-American Health Organization, Eli Lilly and Company, GlaxoSmithKline, and Ortho-McNeil Pharmaceutical. A complete list of WMH publications can be found at http://www.hcp.med.harvard.edu/wmhcidi.

DECLARATION OF INTEREST

None.

REFERENCES

Chen, C. N., Wong, J., Lee, N., Chan-Ho, M. W., Lau, J. T. F. & Fung, M. (1993). The Shatin community mental health survey in Hong Kong: II. Major findings. *Archives of General Psychiatry* **50**, 125–133.

Cheng, T. A. (1989). Symptomatology of minor psychiatric morbidity: a crosscultural comparison. *Psychological Medicine* **19**, 697–708.

Demyttenaere, K., Bruffaerts, R., Posada-Villa, J., Gasquet, I., Kovess, V., Lepine, J. P., Angermeyer, M. C., Bernert, S., Girolamo, G., Morosini, P., Polidori, G., Kikkawa, T., Kawakami, N., Ono, Y., Takeshima, T., Uda, H., Karam, E. G., Fayyad, J. A., Karam, A. N., Mneimneh, Z. N., Medina-Mora, M. E., Borges, G., Lara, G., Graaf, R., Ormel, J., Gureje, O., Shen, Y., Huang, Y., Zhang, M., Alonso, J., Haro, J. M., Vilagut, G., Bromet, E. J., Gluzman, S., Webb, C., Kessler, R. C., Merikangas, K. R., Anthony, J. C., Korff, M. R. V., Wang, P. S., Alonso, J., Brugha, T. S., Aguilar-Gaxiola, S., Lee, S., Heeringa, S., Pennell, B. E., Zaslavsky, A. M., Ustun, T. B. & Chatterji, S. (2004). Prevalence, severity and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys. *Journal of the American Medical Association* **291**, 2581–2590.

First, M. B., Gibbon, M., Spitzer, R. & Williams, J. B. W. (2002). *The Structured Clinical Interview for DSM-IV*. American Psychiatric Press: Washington, DC.

Hosmer, D. W. & Lemeshow, S. (2000). *Applied Logistic Regression Analysis*. Wiley: New York.

Kessler, R. C., Birnbaum, H., Demler, O., Falloon, I. R. H., Gagnon, E., Guyer, M., Howes, M. J., Kendler, K. S., Shi, L., Walters, E. & Wu, E. Q. (2005). The prevalence and correlates of non-affective psychosis in the National Comorbidity Survey Replication (NCS-R). *Biological Psychiatry* **58**, 668–676.

Kessler, R. C. & Üstün, T. B. (2004). The World Mental Health (WMH) Survey Initiative Version of the World Mental Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research* **13**, 93–121.

Kleinman, A. (1986). *Social Origins of Distress and Disease: Depression, Neurasthenia, and Pain in Modern China*. Yale University Press: New Haven.

Kleinman, A. (2004). Culture and depression. *New England Journal of Medicine* **351**, 951–953.

Lee, S. (1996). Cultures in psychiatric nosology: the CCMD-2-R and international classification of mental disorders. *Culture, Medicine and Psychiatry* **20**, 421–472.

Lee, S. (2002). The stigma of schizophrenia: a transcultural problem. *Current Opinion in Psychiatry* **15**, 37–41.

Leon, A. C., Olfson, M., Portera, L., Farber, L. & Sheehan, D. V. (1997). Assessing psychiatric impairment in primary care with the Sheehan Disability Scale. *International Journal of Psychiatry in Medicine* **27**, 93–105.

Parker, G., Gladstone, G. & Chee, K. T. (2001). Depression in the planet's largest ethnic group: the Chinese. *American Journal of Psychiatry* **158**, 857–864.

Shen, Q. J., Lee, S. & Gao, H. (1998). Rethinking the epidemiology of affective disorders in China [in Chinese]. *Chinese Journal of Psychiatry* **31**, 240–242.

Takeuchi, D. T., Chung, R. C. Y., Lin, K. M., Shen, H. K., Kurasaki, K., Chun, C. A. & Sue, S. (1998). Lifetime and twelve-month prevalence rates of major depressive episodes and dysthymia among Chinese Americans in Los Angeles. *American Journal of Psychiatry* **155**, 1407–1414.

Twelve-Region Psychiatric Epidemiological Study Work Group (1986). A national 12-region psychiatric epidemiological study – methodology and results [in Chinese]. *Chinese Journal of Nervous and Mental Diseases* **19**, 65–69.

Wolter, K. M. (1985). *Introduction to Variance Estimation*. Springer-Verlag: New York.

Yip, W. C. & Hsiao, W. C. (1997). Medical savings accounts: lessons from China. *Health Affairs* **16**, 244–251.

Zhang, W. X., Shen, Y. C., Li, S. R., Chen, C. W., Huang, Y. Q., Wang, J. R., Wang, D. P., Tu, J., Ning, Z. X., Fu, L. M., Ji, L. P., Liu, Z. G., Wu, H. M., Luo, K. L., Zhai, S. T., Yan, H. & Meng, G. R. (1998). Epidemiological investigation on mental disorders in 7 areas of China [in Chinese]. *Chinese Journal of Psychiatry* **31**, 69–71.