Modifying and validating the Composite International Diagnostic Interview (CIDI) for use in Nepal

DIRGHA J. GHIMIRE,[1] STEPHANIE CHARDOUL,[2] RONALD C. KESSLER,[3] WILLIAM G. AXINN[4] & BISHNU P. ADHIKARI[5]

1 Population Studies Center and Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, USA
2 The Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, USA
3 Department of Health Care Policy, Harvard Medical School, Boston, MA, USA
4 Population Studies Center, Survey Research Center, Institute for Social Research and the Department of Sociology, University of Michigan, Ann Arbor, MI, USA
5 Institute for Social and Environmental Research, Chitwan, Nepal

Key words
Composite International Diagnostic Interview (CIDI), validity, psychiatric diagnostic interview

Correspondence
Ronald C. Kessler, Department of Health Care Policy, Harvard Medical School, 180 Longwood Avenue, Boston, MA 02115, USA.
Telephone (+1) 617-432-3587
Fax (+1) 617-432-3588
Email: ncs@hcp.med.harvard.edu

Abstract

Background: Efforts to develop and validate fully-structured diagnostic interviews of mental disorders in non-Western countries have been largely unsuccessful. However, the principled methods of translation, harmonization, and calibration that have been developed by cross-national survey methodologists have never before been used to guide such development efforts. The current report presents the results of a rigorous program of research using these methods designed to modify and validate the Composite International Diagnostic Interview (CIDI) for an epidemiological survey in Nepal. *Methods*: A five-step process of translation, harmonization, and calibration was used to modify the instrument. A blinded clinical reappraisal design was used to validate the instrument. *Results*: Preliminary interviews with local mental health expert led to a focus on major depressive episode, mania/hypomania, panic disorder, post-traumatic stress disorder, generalized anxiety disorder, and intermittent explosive disorder. After an iterative process of multiple translations-revisions guided by the principles developed by cross-national survey methodologists, lifetime DSM-IV diagnoses based on the final Nepali CIDI had excellent concordance with diagnoses based on blinded Structured Clinical Interview for DSM-IV (SCID) clinical reappraisal interviews. *Conclusions*: Valid assessment of mental disorders can be achieved with fully-structured diagnostic interviews even in low-income non-Western settings with rigorous implementation of replicable developmental strategies. Copyright © 2013 John Wiley & Sons, Ltd.

Introduction

Although substantial challenges exist in constructing any multi-lingual survey (Harkness *et al.*, 2010), this is especially true when the survey deals with mental disorders. This is true both because the ways mental disorders are understood and talked about vary enormously across cultures and because mental disorders are broadly stigmatized (Prince,

2008). Many psychiatric epidemiological surveys have nonetheless been carried out throughout the world, but comparability of results can be called into question. The World Health Organization (WHO) World Mental Health (WMH) Surveys, for example, used standard WHO translation, back-translation, and harmonization procedures to administer a fully-structured diagnostic interview, the WHO Composite International Diagnostic Interview (CIDI; Kessler and Üstün, 2004), to community samples of respondents in 28 countries (Kessler and Üstün, 2008) but the prevalence estimates were implausibly low in some countries (Kessler *et al.*, 2007). Although clinical reappraisal studies show that diagnoses based on the CIDI have good concordance with blinded clinical diagnoses in developed European Diaspora countries (Haro *et al.*, 2006), methodological studies elsewhere have shown that diagnoses based on a standard translation of the CIDI are either poorly related to culturally competent diagnostic judgments (van Ommeren *et al.*, 1999) or fail to generate any diagnoses at all (Tausig *et al.*, 2011). A key reason for these failures is that the complexities involved in adopting diagnostic instruments to specific cultural contexts often require more than simple verbal translation to guarantee conceptual equivalence, criterion equivalence, and content equivalence (Schwarz, 2003).

This paper reports the results of an extensive program of developmental research to translate the CIDI into Nepali for purposes of a planned psychiatric epidemiological survey of Nepal. This program of research was specifically designed to address the complex challenges of culture-specific validation in a non-Western country while maintaining cross-cultural equivalence. The strategy used is highly structured and replicable, but requires each step in the process to be repeated until convergence is achieved on externally validated measures. Chitwan Valley in rural Nepal, the site of the study, is a challenging site because the population is culturally diverse and includes representatives of both South Asian Hindu culture and East Asian Buddhist culture (Bista, 1972). These groups, who together compose the largest fraction of the world's population, are quite different from the European Diaspora groups for whom the CIDI has been shown to be valid up to now. Because this site has also been the subject of intensive study over the past 15 years (Axinn and Pearce, 2006; Ghimire and Axinn, 2006), it provides a key opportunity for creation and validation of new survey measures.

Methods

Frameworks for translating survey measures

English-speaking Western cultures have been the main sources of measures for most community surveys. As a result, numerous English-language survey instruments have been translated into other languages (Harkness *et al.*, 2010). This work has generated several useful frameworks for considering the translation and adoption of survey instruments. The most general of these is that of Flaherty *et al.* (1988), who described five forms of equivalence between the original and translated survey instrument: (1) content equivalence: that is, whether the content of each item of the instrument is relevant to the phenomenon of each culture being studied; (2) semantic equivalence: that is, whether the meaning of each item is the same in each culture; (3) technical equivalence: that is, equivalence of such things as the mode of administration (e.g. face-to-face versus telephone survey) and the instructions used to introduce the survey to respondents; (4) criterion equivalence: that is, whether the interpretation of the measurement is the same across cultures; (5) conceptual equivalence: that is, whether the instrument is measuring the same theoretical construct in each culture. Most other conceptual models can be mapped into this one (Manson, 1997).

A number of approaches to the translation process have been proposed to achieve these different kinds of equivalence (Bravo *et al.*, 1991; Flaherty *et al.*, 1988; Manson, 1997; Mumford *et al.*, 1991; Steel *et al.*, 2009; van Ommeren *et al.*, 1999). The most highly-developed approach is that of Brislin (1986), who proposed a five-step process of translation and adoption that includes: (1) translation; (2) blind back-translation; (3) examination of original, translation, and back-translation for discrepancies that are then reconciled based on clarification of the basis of the discrepancies; (4) pilot study; (5) evaluation and revision based on pilot study results. The approach adopted by the WHO for its cross-national surveys expanded Brislin's (1986) first three steps to include establishment of a bilingual group of experts, examination of conceptual structure of the instruments by the experts, translation, examination of the translation by the experts, examination of the translation by a monolingual group representative of the population to be studied, and blind back-translation followed by examination of the blind back-translation by the experts and reconciliation of discrepancies (Sartorius and Janca, 1996). The WMH Survey Initiative additionally developed a series of guidelines for CIDI translation that emphasized cultural rather than literal equivalence. However, few of the CIDI translation efforts in the 28 countries that have so far participated in the WMH Survey Initiative have followed all these steps rigorously. The program of research described in this report did so.

Key challenges to success

Success in translating and adopting the CIDI to non-European Diaspora cultures, such as those in South Asia, has been

Ghimire *et al.*   Modifying and validating the CIDI for use in Nepal

limited by significant challenges to conceptual equivalence, criterion equivalence, and content equivalence. The first Nepali CIDI study, which was conducted in the late 1990s with Nepali-speaking Bhutanese refugees, concluded that in a cultural context such as Nepal, two basic assumptions of the CIDI core structure are violated: that respondents attribute their symptoms to mental, physical, or substance-related processes; and that doctors communicate diagnoses to their patients. The same team of researchers found that *comprehensibility*, *acceptability*, *relevance*, and *completeness* were all formidable challenges to achieving their goal of context-specific culturally appropriate translation in this population (van Ommeren *et al.*, 1999). Another study that used the CIDI in Nepal in 2002 encountered similar challenges in creating a culturally-appropriate translation (Tausig *et al.*, 2011).

The research setting

The research reported here was implemented as a pilot study for a planned Chitwan Health and Stress Study (CHSS) to be carried out by the Population Studies Center at the University of Michigan in collaboration with Tribhuvan University in Rampur, Chitwan, Nepal under the auspices of their jointly sponsored Population and Ecology Research Laboratory (PERL). PERL has carried out a number of large-scale community surveys in the Chitwan Valley region of rural Nepal since the early 1990s (http://perl.psc.isr.umich.edu) and maintains a professional interview staff to implement these surveys. The motivation for carrying out the CHSS can be traced to the fact that the decade-long armed conflict and recent resurgence of ethnic movements in Nepal have been continuous sources of psychological and emotional distress that the PERL team felt needed to be studied. In addition, the mismatch between the lack of economic opportunities and the rise of aspirations for modern living is fueling unrest, especially on the part of educated youths. Although the country has a history of social and economic inequalities, the decade-long insurgency and recent ethnic/minority movements have been continuous sources of violence, intimidation, and extortion, possibly leading to a rise in stress-related mental disorders. Yet data on the prevalence of mental disorders and service use for these disorders in Nepal are virtually non-existent.

Nepal's historical isolation, socio-economic and political background of extreme poverty, gender and ethnic/caste-based inequalities, and patriarchal social system have led to the Nepalese having a conceptualization of mental health that is quite different from that in the West. The high caste Hindus historically practiced the homeopathic treatment system Ayurveda, which interprets physical and psychological health problems as outcomes of an imbalance in human physiology and cosmological space (Atreya, 2006). In contrast, most indigenous groups believe in spiritual super-natural power and a Tantric treatment system, leading them to interpret mental disorders as punishments for human wrong doings (Justice, 1986), while some other indigenous groups believe that a person's well-being is dependent on astrology. This diversity of basic beliefs and understandings makes the CIDI translation process much more complicated than it might be otherwise, requiring a multi-method, multi-step approach to translating the English version of the CIDI that is described later.

The five-step translation process

Strategic design with mental health professionals

We began with a series of consultations with key mental health professionals (including both physicians trained in Western medicine and traditional healers) in Chitwan Valley as well as elsewhere in Nepal guided by a three-pronged strategy of contextualizing the CIDI questions to local context, using local expertise, and requiring rigor in every step. These meetings were invaluable in learning a broad range of local expressions and cultural idioms of mental illness that assisted us in translating English terms. These meetings also led us to focus our attention on six DSM-IV syndromes thought to be of importance in Nepal: major depressive episode, mania and hypomania, panic disorder, generalized anxiety disorder, post-traumatic stress disorder, and intermittent explosive disorder.

Training a setting-specific survey research staff

Next, before starting the actual translation, the Nepali research staff at PERL was trained on (1) the WMH translation protocol and (2) the WMH training program for administering the CIDI. In addition to the standard CIDI training provided to all WMH investigator teams, the staff was trained in methods of cognitive testing of the translated instrument (Willis, 2004) and survey interview techniques (Fowler and Mangione, 1989). In addition, in order to validate the diagnoses being generated by the translated instrument, a team of four bi-lingual Nepali psychiatrists in Chitwan Valley was trained in the use of the Structured Clinical Interview for DSM-IV (SCID; First *et al.*, 1994). This training was conducted by one of the developers of the SCID (Michael First). These psychiatrists subsequently worked to validate the initial and final CIDI translations against the SCID.

Iterative CIDI translation, testing, and revision

Although it is a formidable challenge to create a single questionnaire for a multi-ethnic, multi-lingual society such as Nepal, we accomplished this through an iterative process of translation, testing, and revision. This work was facilitated by the fact that it was carried out in conjunction with the experienced survey research team at PERL. This team, which includes representatives from all the ethnic groups in the study, has experience translating over a dozen questionnaires on diverse topics (Axinn and Pearce, 2006; Ghimire and Axinn, 2006; Thornton *et al.*, 2010). As noted earlier, before beginning the translation, the staff was trained in the WMH translation guidelines and studied published reports on previous unsuccessful attempts to translate the CIDI into Nepali.

The main goal of the initial translation was to produce a conceptually equivalent, easily understandable, and culturally acceptable translation of the CIDI with an emphasis on contextual meaning rather than literal translation. This translation was then independently reviewed by three of our four collaborating local psychiatrists to evaluate *semantic equivalence* (i.e. meaning of the translated item is the same as the original CIDI item) and *content equivalence* (i.e. the content of each item is relevant in the local context). This initial translation was also reviewed by local health workers, traditional healers, and representatives of the local population in a focus group interview format. Because these groups had little or no knowledge of the English language, the primary goal of these sessions was to confirm simplicity, clarity, cultural sensitivity, and appropriateness of the questions in Nepali. The results were very useful in helping us *harmonize* the CIDI with colloquial terms (see Appendix).

Cognitive interviewing

A second draft of the questionnaire was created based on feedback from these focus groups along with the psychiatrist review. Ten cognitive debriefing interviews of this second draft were then conducted by professional interviewers with a larger illiterate group of monolingual Nepali community residents. Cognitive interviews are qualitative interviews that ask survey respondents to report on the extent to which they understood the questions asked of them and to probe for areas of confusion that can be rectified with question revisions (Fowler and Mangione, 1989; Willis, 2004). The primary goal of the cognitive interviews was to study question clarity with a focus on conceptual understanding of the descriptions of psychiatric symptoms as well as of the response categories. Interviewers probed question wording that the respondent had any difficulty understanding

and collected alternative terminologies. Interviews were audio recorded for further review.

Results

Pre-testing the revised CIDI

Based on the insights gained from the cognitive interviews, the instrument was again revised. This third draft was then pre-tested in a small ($n = 15$) sample of psychiatric outpatients treated at the Bharatpur Medical College Hospital. The goals were to determine whether the translation was easily understood (*criterion equivalence*) and whether the question items in Nepali were measuring the concepts the original questions intended to measure (*conceptual equivalence*). The 15 respondents were then blindly re-interviewed by one of our psychiatrists using the SCID. A high degree of mismatch was found between diagnoses based on the CIDI and those based on the SCID. (Detailed results are not reported, but are available on request.) We conducted debriefing meetings between the research staff and the psychiatrists that focused separately on each interview in order to understand these differences. Two important insights emerged from these meetings. First, as noticed in the CIDI survey interview, the SCID interview also encountered wording problems about certain states of stress and feelings. The clinical interviewers felt, though, that these problems stemmed from the fact that they were using the original English version of the SCID and were trying to make instant translations. Nonetheless, because the SCID allowed the clinical interviewers to ask follow-up questions, responses were very useful in identifying synonyms. Second, we discovered that respondents had great difficulty with diagnostic duration requirements, as people do not think in terms of duration in Nepal. Although the SCID interviewers also faced these same challenges, they were less severe in the SCID than the CIDI because the SCID, unlike the CIDI, allowed clinical interviewers to ask follow-up questions about duration. This resulted in a higher number of disorders being diagnosed in the SCID than in the CIDI.

Based on the insights gained from this pre-test, the research staff and collaborating local psychiatrists developed a fourth draft of the CIDI. They also translated the SCID into Nepali. Table 1 presents some of the examples of the changes we made in version four of the questionnaire. The revised CIDI was then administered to two focus groups, one consisting of members of the local population and the other consisting of local health workers. Based on the feedback from these focus groups, we revised the translation of the CIDI yet again. This fifth draft was pre-tested on another 20 cases (10 community residents and 10 clinical outpatients

Table 1. Examples of improvement/revision in CIDI translation process – between version 3 and version 4 of the Nepal Pilot Study questionnaire

	In Version 3 of the CIDI Questionnaire	In Version 4 of the CIDI Questionnaire
1. SCID – Translation[1]		
SCID Questionnaire	English	Nepali
2. Changes in question wording[2]		
SC20.2- loss of control	Aafno Hos Gumayera	Aankhai Nadekhera
SC21- sad, empty or depressed	Man Dukhi Hune, Khinna Hune, Naramailo Lagne	*Man Naramailo Lagne, Chinta Parne, Niras Hune, Bachnai Napare Pani Hunthyo Jasto Lagne*
SC.25- irritable	*Jhaurine*	*Dikka Lagne/Jhijo Lagne*
SC.26a- nervous or anxious	*Aattinu/Chatpataunu*	*Aattenu/Hadbadaunu*
SC 33- during childhood or adolescence	*Sano chhada wa kishor abastha*	Bis barsha pugnu bhanda aghi
3. Addition of words/phrases[3]		
SC.20–Have you ever in your life had an attack of fear or panic when all of a sudden you felt very frightened, anxious, or uneasy?	*Tapainko Jiban Bharima Kahelai Ekkashi Bina Karan Aattiyera Wa Chatpatiyera Ekdamai Darayeko, Pir Wa Chinta Pareko Theyo*	*Tapainko Jiban Bharima Kahelai Ekkashi Bina Karan Aateyanti* **Man Daraune, Mutu Dhukdhuk Garne, Mareyela Ki Jasto Hune, Sas Badhne Ra Yesto Aabastha Ek Patakma Karib 15 Minute Samma Rahane Ra Barambar Dohorine Bhayeko**
SC.21-Have you ever in your life had a period lasting several days or longer when most of the day you felt sad, empty or depressed?	*Tapaiko Jinabma Kahelai Dherai Dinsamma Man Dukhi hune, Khinna Hune Wa Naramailo Lagirahane Bhayeko Theyo*	*Tapaiko Jinabma Kahelai Dherai Dinsamma Man Naramailo Lagne, Chinta Parne, Niras Hune,* **Bachnai Napare Pani Hunthyo Jasto Lagne, Tauko Dukhne, Sarir Jhamjham Garne, Yaun Kriyama Kami Aaune, Sano Kam Pani Garna Sakdina Jasto Lagne, Aattine, Kamma Dhyan Najane Ra Dikka Lagne Bhayoko Theyo**
5. Time duration cues[4]		
	15 days	Two weeks
	Four weeks	One month

[1]Clinical interviewers (psychiatric doctors) were using the English version of the SCID questionnaire in version 3 of our work. For version 4, we translated the SCID into Nepali. Even though the clinical interview is not fully structured and interviewer could ask follow-up questions as necessary, having the opening question translated in Nepali seems to improve the respondents' answers.

[2]In version 4 of the questionnaire we replaced technical words that respondents had difficulty understanding or have ambiguous meaning with more colloquial worlds that were more easily understood by respondents.

[3]In addition to wording changes we also added several synonymous words/phrases to be more inclusive of a broad range of cultural variation, experiences and literacy levels.

[4]In order to improve the understandability we changed the timing duration to aggregated units (days to weeks and weeks to months).

from Kathmandu) in order to assess the applicability of our translation to a wider range of the Nepali population outside of Chitwan Valley. Just as with the first pre-test, a professional survey interviewer interviewed each respondent and then a psychiatrist blindly re-interviewed the respondent using the new Nepali translation of the SCID. Much closer CIDI-SCID concordance was found in this phase of the pre-test than in the earlier phase (Detailed results are not reported, but are available on request.) Based on this result, we moved to piloting the instrument on a larger sample.

Clinical reappraisal study results

The sample

The final Nepali translation-revision of CIDI 3.0 was administered to a systematically selected general population sample of 403 individuals living in the western Chitwan Valley. The goal was to carry out a final clinical reappraisal study of the instrument prior to large-scale field implementation. Respondents were chosen using a two-stage sampling procedure. First, the study area was divided into three distinct strata with successive distance from the urban center. Second, a sample of four to six neighborhoods, each consisting of between 10 and 15 households, was selected within each stratum. In addition, based on local psychiatrists' recommendation, we purposively selected two additional neighborhoods from high poverty areas that were thought by the psychiatrists to have high prevalence of mental disorders. This process yielded 17 neighborhoods with a total of 198 households. Once a household was selected, all individuals aged 18+ residing in the household were interviewed. This sampling procedure generated a set of 403 individuals who were eligible for interview.

The selected respondents were interviewed by professional interviewers who had been specially trained in CIDI administration. Interviews were carried out face-to-face with a paper and pencil instrument. The field period lasted three weeks and the response rate was 91%. The sample consisted of a higher proportion of women (58%) than men (42%), primarily because of the male out-migration for employment and schooling. (Table 2) The mean age was 37, which reflects the fact that even with the huge out-migration of young males for employment, Nepal still has a very young population. The high proportion of respondents who were married (82%) reflects the near universality of marriage in this population. The low mean years of schooling (4.9 years) with 36% of the sample having no schooling reflects the low level of formal education among the adult population of this area. These socio-demographic distributions are quite comparable to those

found in a recently-completed large-scale community survey of this area, increasing our confidence in the representativeness of the sample.

In order to evaluate clinical validity of diagnoses generated by the revised CIDI, a subsample of respondents was blindly administered SCID clinical reappraisal interviews. This subset included 100% of the respondents who were diagnosed positive for at least one disorder in the CIDI and a probability subsample of 10% of all remaining respondents. This sampling procedure yielded 80 individuals: 47 CIDI cases and 33 non-cases. We assigned a weight of 10 to each non-case to adjust for the fact that only 10% of non-cases were re-interviewed before analyzing CIDI-SCID concordance.

Aggregate concordance

We investigated whether CIDI prevalence estimates are biased in comparison to SCID prevalence estimates using McNemar χ^2 tests. As with all significance tests reported in this paper, McNemar tests were carried out using the Taylor series design-based estimation method to adjust for the effects of weighting and clustering and oversampling of CIDI cases (Wolter, 1985). Because neither the CIDI nor the SCID diagnosed any cases of mania and the SCID is not designed to diagnose intermittent explosive disorder, we focused on only the remaining four disorders for which we collected data. The lifetime prevalence estimate for having any of these four disorders was 8.2% in the CIDI and 10.1% in the SCID (Table 3). The McNemar test of CIDI versus SCID prevalence differences was not significant for this comparison ($\chi_1^2 = 2.3$, $p = 0.13$), demonstrating that the final Nepali translation of the CIDI produces similar overall prevalence estimates as the SCID. Comparable McNemar tests of CIDI versus SCID prevalence differences were significant only for one of the four disorders (post-traumatic stress disorder [$\chi_1^2 = 10.0$, $p < 0.001$], with the SCID estimate substantially higher than the CIDI estimate [6.6% versus 4.0%]).

Individual-level concordance

Individual-level concordance was then evaluated using the two most commonly used descriptive measures for individual-level concordance the area under the receiver operator characteristic curve (AUC; Hanley and McNeil, 1982) and Cohen's κ (Cohen, 1960). Although the κ statistic is the most widely-used measure of concordance in validity studies of psychiatric disorders, it is dependent on prevalence and consequently is often low in situations where there appears to be high agreement between low-prevalence measures (Byrt *et al.*, 1993; Cook, 1998). The AUC resolves this problem. We also examined sensitivity (SN; the proportion of clinical cases detected by the CIDI) and specificity (SP; the proportion of clinical

Table 2. Comparison of the demographic composition of the clinical reappraisal sample with that of the 2008 Chitwan Valley Family Study (CVFS)[1]

| | Clinical reappraisal sample | | | | CVFS | |
| | With high poverty area Over-sample | | Without high poverty area Over-sample | | | |
	Estimated	(Standard error)	Estimated	(Standard error)	Estimated	(Standard error)
Gender						
Female (%)	58	(2.4)	61	(2.6)	57	(0.5)
Hindu High Caste (%)	52*	(2.4)	52	(2.6)	47	(0.5)
Ethnicity						
Hill Indigenous (%)	20	(1.9)	14	(1.8)	16	(0.4)
Dalit (%)	13	(1.6)	12	(1.7)	10	(0.3)
Tarai Indigenous (%)	19	(1.9)	21	(2.1)	20	(0.4)
Age						
In years (mean)	37.4	(0.7)	37.9	(0.8)	37.8	(0.2)
Marital status						
Currently Married (%)	82	(1.8)	82	(2.0)	81	(0.4)
Education						
In years (mean)	4.9*	(0.2)	5.0	(0.2)	5.5	(0.1)
Illiterate (%)	36	(0.2)	36	(2.5)	32	(0.5)
(*n*)	(403)		(352)		(7072)	

[1]The Chitwan Valley Family Study (CVFS) (http://perl.psc.isr.umich.edu/researchprojects.htm) is an ongoing series of community surveys originally (1994) designed to study the influence of changing social contexts on the timing of marriage, child-bearing, and contraceptive use but subsequently expanded to address a wide range of issues in the social demography of the population of rural Nepal. The most recent survey carried out as part of the CVFS was implemented in 2008. The socio-demographic data shown in the last column of the table come from that survey.
*Significant difference between the clinical reappraisal sample and the CVFS sample at the 0.05 level, two-sided test.

Table 3. Aggregate concordance of lifetime DSM-IV prevalence estimates based on the CIDI and SCID in the clinical reappraisal sample (*n* = 80)[1]

| | CIDI | | SCID | | McNemar |
	Percentage	(Standard error)	Percentage	(Standard error)	χ_1^2
Major depressive episode	5.0	(0.01)	5.3	(0.02)	0.3
Bipolar disorder	0.0	—	0.0	—	0.0
Panic disorder	0.3	(0.002)	0.3	(0.006)	0.0
Post-traumatic stress disorder	4.0	(0.01)	6.6	(0.03)	10.0*
Generalized anxiety disorder	0.8	(0.004)	0.8	(0.01)	0.0
Any disorders	8.2	(0.01)	10.1	(0.03)	2.3

[1]The first-stage clinical reappraisal sample consisted of a probability sample of 403 respondents who were administered the CIDI. The 80 second-stage respondents were administered the SCID. These included all 43 of the original 403 respondents who met DSM-IV criteria for one or more of the disorders in the CIDI plus a 10% probability subsample of the remaining 370 respondents (*n* = 37). The 37 CIDI non-cases were given a weight of 10 to adjust for the fact that they represent only 10% of the CIDI non-cases in the first-stage sample.
*Significant difference between the CIDI and SCID prevalence estimates at the .05 level, two-sided test.

non-cases classified as non-cases by the CIDI), the key components of AUC and κ, along with positive predictive value (PPV; the proportion of CIDI cases confirmed by the SCID) and negative predictive value (NPV; the proportion of CIDI non-cases confirmed by the SCID).

Using terms widely employed to describe concordance between measures (Landis and Koch, 1977), individual-level concordance between CIDI and SCID diagnoses can be described as almost perfect or substantial for all four disorders (Table 4). Individual-level concordance was also substantial for any diagnosis overall. Sensitivity estimates of 95% to 100% for each diagnosis other than post-traumatic stress disorder (where sensitivity is 60%) show that the vast majority of clinical cases were detected by the CIDI other than cases of post-traumatic stress disorder. In addition, the vast majority of SCID non-cases (97–100%) were classified by the CIDI as not meeting lifetime criteria for the disorder. The high levels of positive predictive value (87–100%) confirm that the vast majority of CIDI cases were also clinical cases. The Nepali AUC estimates are comparable to those obtained in Western countries (Haro *et al.*, 2006).

Discussion

A key limitation of the results reported here is the small number of disorders investigated. However, the success of the translation process across even this small set of disorders demonstrates clearly that it is possible to develop valid translations of difficult concepts across cultural and linguistic groups if significant effort is invested in doing so. (Detailed examples comparing initial and final translations of key concepts are available on request.) Our lower level of success in developing a valid CIDI assessment of post-traumatic stress disorder than the other disorders is the focus of ongoing study along with the expansion of the translation process to include a broader range of disorders. The approach we used to achieve these positive results is highly structured and replicable, suggesting that similar success in translation could be achieved in other settings, especially in Asian countries where, as in Nepal, previous CIDI translations have yielded implausibly low prevalence estimates (Shen *et al.*, 2006; Simon *et al.*, 2002). The key is that each step in the process was repeated until convergence on externally validated measures was achieved. The greater success of past CIDI translation in European Diaspora settings than in less developed countries likely reflects the fact that the closely connected cultures and linguistic groupings in the former countries made repetition of the translation steps less necessary than in Africa or Asia. Nepal includes South Asian and Tibeto-Burmese populations among whom previous translations of the CIDI have not worked well. Although the cultures and languages represented in Nepal are linked to European cultures and

Table 4. Individual-level concordance of lifetime DSM-IV diagnoses based on the CIDI and SCID in the clinical reappraisal sample (*n* = 80)[1]

	SN[2]	SP[2]	PPV[2]	NPV[2]	TCA[2]	κ[2]	AUC[2]	AUC from European and US WMH surveys[3]	AUC from US NCS-A survey[4]
Major depressive episode	86	99	78	99	98	0.91	0.92	0.75	0.87
Panic disorder	100	99	63	100	99	0.50	1.00	0.72	0.87
Post-traumatic stress disorder	60	95	46	97	93	0.48	0.77	0.69	0.79
Generalized anxiety disorder	100	100	100	100	100	1.00	1.00	1.0	0.80
Any disorder	74	95	62	97	93	0.63	0.85	0.76	0.87

[1]The first-stage clinical reappraisal sample consisted of a probability sample of 403 respondents who were administered the CIDI. The 80 second-stage respondents were administered the SCID. These included all 43 of the original 403 respondents who met DSM-IV criteria for one or more of the disorders in the CIDI plus a 10% probability subsample of the remaining 370 respondents (*n* = 37). The 37 CIDI non-cases were given a weight of 10 to adjust for the fact that they represent only 10% of the CIDI non-cases in the first-stage sample.
[2]SN = Sensitivity; SP = specificity; PPV = positive predictive value; NPV = negative predictive value; TCA = total classification accuracy; κ = Cohen κ; AUC = area under the receiver operating characteristics curve.
[3]Based on World Mental Health (WMH) survey clinical reappraisal interviews carried out in community samples in France, Italy, Spain and the United States comparing diagnoses based on the CIDI and SCID (Haro *et al.*, 2006).
[4]Based on clinical reappraisal interviews carried out in a community sample of adolescents in the United States comparing diagnoses based on the CIDI and the Schedule for Affective Disorders and Schizophrenia for School-Aged Children (K-SADS; Kaufman *et al.*, 1997) as part of the US National Comorbidity Survey Replication Adolescent Supplement (NCS-A; Kessler *et al.*, 2009).

languages, the overlap is low (Bista, 1972). As a result, previous more superficial efforts to translate the CIDI for Nepalese speaking populations have worked poorly (Tausig *et al*., 2011; van Ommeren *et al*., 1999). This setting consequently provides a good trial for the intensive strategies used here to improve cross-cultural translation of survey-based assessments of DSM-IV disorders.

Acknowledgments

Declaration of interest statement

Dr Kessler has been a consultant for AstraZeneca, Analysis Group, Bristol-Myers Squibb, Cerner-Galt Associates, Eli Lilly & Company, GlaxoSmithKline Inc., HealthCore Inc., Health Dialog, Integrated Benefits Institute, John Snow Inc., Kaiser Permanente, Matria Inc., Mensante, Merck & Co, Inc., Ortho-McNeil Janssen Scientific Affairs, Pfizer Inc., Primary Care Network, Research Triangle Institute, Sanofi-Aventis Groupe, Shire US Inc., SRA International, Inc., Takeda Global Research & Development, Transcept Pharmaceuticals Inc., and Wyeth-Ayerst; has served on advisory boards for Appliance Computing II, Eli Lilly & Company, Mindsite, Ortho-McNeil Janssen Scientific Affairs, Plus One Health Management and Wyeth-Ayerst; and has had research support for his epidemiological studies from Analysis Group Inc., Bristol-Myers Squibb, Eli Lilly & Company, EPI-Q, GlaxoSmithKline, Johnson & Johnson Pharmaceuticals, Ortho-McNeil Janssen Scientific Affairs., Pfizer Inc., Sanofi-Aventis Groupe, and Shire US, Inc. The remaining authors report nothing to disclose.

References

Atreya. (2006) Practical Ayurveda: Secrets for Physical, Sexual and Spiritual Health, Mumbai, JaicoPublishing House.

Axinn W.G., Pearce L.D. (2006) Mixed Method Data Collection Strategies, Cambridge, Cambridge University Press.

Bista D.B. (1972) People of Nepal, Kathmandu, Ratna Pustak Bhandar.

Bravo M., Canino G.J., Rubio-Stipec M., Woodbury-Farina M. (1991) A cross-cultural adaptation of a psychiatric epidemiologic instrument: the diagnostic interview schedule's adaptation in Puerto Rico. *Culture, Medicine and Psychiatry*, **15**(1), 1–18, DOI: 10.1007/BF00050825

Brislin R.W. (1986) The wording and translation of research instruments. In Lonner W.J., Berry J.W. (eds) Field Methods in Cross-Cultural Research, pp. 137–164, Newbury Park, CA, Sage Publications.

Byrt T., Bishop J., Carlin J.B. (1993) Bias, prevalence and kappa. *Journal of Clinical Epidemiology*, **46**(5), 423–429, DOI: 10.1016/0895-4356(93)90018-V

Cohen J. (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, **20**(1), 37–46, DOI: 10.1177/001316446002000104

Cook R.J. (1998) Kappa and its dependence on marginal rates. In Armitage P., Colton T.

(eds) The Encyclopedia of Biostatistics, pp. 2166–2168, New York, John Wiley & Sons.

First M.B., Spitzer R.L., Gibbon M., Williams J.B. W. (1994) Structured Clinical Interview for Axis I DSM-IV Disorders – Patient Edition (SCID-I/P, Version 2.0), New York, Biometrics Research Department, New York State Psychiatric Institute.

Flaherty J.A., Gaviria F.M., Pathak D., Mitchell T., Wintrob R., Richman J.A., Birz S. (1988) Developing instruments for cross-cultural psychiatric research. *The Journal of Nervous and Mental Disease*, **176**(5), 257–263.

Fowler Jr. F.J., Mangione T.W. (1989) Standardized Survey Interviewing: Minimizing Interviewer-Related Error, Newbury Park, CA, Sage Publications.

Ghimire D.J., Axinn W.G. (2006) Family change in Nepal: evidence from western Chitwan. *Contributions to Nepalese Studies*, **33**(2), 177–201.

Hanley J.A., McNeil B.J. (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, **143**(1), 29–36.

Harkness J., Mohler P., Pennell B.E., Johnson T., Lyberg L., Edwards B., Smith T.W., Braun M. (eds) (2010) Survey Methods in Multinational, Multiregional, and Multicultural Contexts, Hoboken, NJ, John Wiley & Sons.

Haro J.M., Arbabzadeh-Bouchez S., Brugha T.S., de Girolamo G., Guyer M.E., Jin R., Lepine J.P.,

Mazzi F., Reneses B., Vilagut G., Sampson N.A., Kessler R.C. (2006) Concordance of the Composite International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World Mental Health surveys. *International Journal of Methods in Psychiatric Research*, **15**(4), 167–180, DOI: 10.1002/mpr.196

Justice J. (1986) Policies, Plans, and People: Culture and Health Development in Nepal, Berkeley, CA, University of California Press.

Kaufman J., Birmaher B., Brent D., Rao U., Flynn C., Moreci P., Williamson D., Ryan, N. (1997) Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry*, **36**(7), 980–988, DOI: 10.1097/00004583-199707000-00021

Kessler R.C., Angermeyer M., Anthony J.C., de Graaf R., Demyttenaere K., Gasquet I., de Girolamo G., Gluzman S., Gureje O., Haro J.M., Kawakami N., Karam A., Levinson D., Medina Mora M.E., Oakley Browne M.A., Posada-Villa J., Stein D.J., Adley Tsang C.H., Aguilar-Gaxiola S., Alonso J., Lee S., Heeringa S., Pennell B.E., Berglund P., Gruber M.J., Petukhova M., Chatterji S., Ustun T.B. (2007) Lifetime

prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry*, **6**(3), 168–176.

Kessler R.C., Avenevoli S., Green J., Gruber M.J., Guyer M., He Y., Jin R., Kaufman J., Sampson N.A., Zaslavsky A.M. (2009) National Comorbidity Survey Replication Adolescent Supplement (NCS-A): III. Concordance of DSM-IV/CIDI diagnoses with clinical reassessments. *Journal of the American Academy of Child and Adolescent Psychiatry*, **48**(4), 386–399, DOI: 10.1097/CHI.0b013e31819a1cbc

Kessler R.C., Üstün T.B. (2004) The World Mental Health (WMH) Survey Initiative Version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research*, **13**(2), 93–121, DOI: 10.1002/mpr.168

Kessler R.C., Üstün T.B. (2008) The WHO World Mental Health Surveys: Global Perspectives on the Epidemiology of Mental Disorders, Cambridge, Cambridge University Press.

Landis J.R., Koch G.G. (1977) The measurement of observer agreement for categorical data. *Biometrics*, **33**(1), 159–174.

Manson S.M. (1997) Ethnographic methods, cultural context, and mental illness: bridging different ways of knowing and experience. *Ethos*, **25**(2), 249–258, DOI: 10.1525/eth.1997.25.2.249

Mumford D.B., Tareen I.A., Bajwa M.A., Bhatti M.R., Karim R. (1991) The translation and evaluation of an Urdu version of the Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, **83**(2), 81–85, DOI: 10.1111/j.1600-0447.1991.tb07370.x

Prince M. (2008) Measurement validity in cross-cultural comparative research. *Epidemiologia e Psichiatria Sociale*, **17**(3), 211–220.

Sartorius N., Janca A. (1996) Psychiatric assessment instruments developed by the World Health Organization. *Social Psychiatry and Psychiatric Epidemiology*, **31**(2), 55–69, DOI: 10.1007/BF00801901

Schwarz N. (2003) Culture-sensitive context effects: a challenge for cross-cultural surveys. In Harkness J.A., Van de Vijver F.J.R., Mohler P. P. (eds) Cross-cultural Survey Methods, pp 93–100, New York, John Wiley & Sons.

Shen Y.C., Zhang M.Y., Huang Y.Q., He Y.L., Liu Z.R., Cheng H., Tsang A., Lee S., Kessler R.C. (2006) Twelve-month prevalence, severity, and unmet need for treatment of mental disorders in metropolitan China. *Psychological Medicine*, **36**(2), 257–267, DOI: 10.1017/S0033291705006367

Simon G.E., Goldberg D.P., Von Korff M., Ustun T.B. (2002) Understanding cross-national differences in depression prevalence. *Psychological Medicine*, **32**(4), 585–594, DOI: 10.1017/S0033291702005457

Steel Z., Silove D., Giao N.M., Phan T.T., Chey T., Whelan A., Bauman A., Bryant R.A. (2009) International and indigenous diagnoses of mental disorder among Vietnamese living in Vietnam and Australia. *The British Journal of Psychiatry*, **194**(4), 326–333, DOI: 10.1192/bjp.bp.108.050906

Tausig M., Subedi J., Broughton C., Pokimica J., Huang Y., Santangelo S. (2011) The continued salience of methodological issues for measuring psychiatric disorders in international surveys. *International Journal of Mental Health and Addiction*, **9**(3), 229–239, DOI: 10.1007/s11469-010-9276-3

Thornton A., Achen A., Barber J.S., Binstock G., Garrison W.M., Ghimire D.J., Inglehart R., Jayakody R., Jiang Y., de Jong J., King K., Lesthaeghe R.J., Mehenna S., Mitchell C., Moaddel M., Ofstedal M.B., Schwarz N., Wang G., Xie Y., Yang L.S., Young-DeMarco L.C., Yount K. (2010) Creating questions and protocols for an international study of ideas about development and family life. In Harkness J.A., Braun M., Edwards B., Johnson T.P., Lyberg L., Mohler P.P., Pennell B. E., Smith T.W. (eds) Survey Methods in Multinational, Multiregional and Multicultural Contexts, pp. 59–74, New York, John Wiley & Sons.

van Ommeren M., Sharma B., Thapa S., Makaju R., Prasain D., Bhattarai R., de Jong J. (1999) Preparing instruments for transcultural research: use of the translation monitoring form with Nepali-speaking Bhutanese refugees. *Transcultural Psychiatry*, **36**, 285–301, DOI: 10.1177/136346159903600304

Willis G. (2004) Cognitive Interviewing: A Tool for Improving Questionnaire Design, Newbury Park, CA, Sage Publications.

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

Appendix

Table A1. Comparison of lifetime DSM-IV diagnoses based on the initial translation of the fully-structured Composite International Diagnostic Interview (CIDI) and independent diagnoses based on clinical reappraisal interviews with the Structured Clinical Interview for DSM-IV (SCID) among clinical outpatients in Bharatpur Medical College Teaching Hospital ($n = 15$)

	CIDI						SCID[2]				
	MDE[1]	M[1]	PD[1]	PTSD[1]	GAD[1]	IED[1]	MDE[1]	M[1]	PD[1]	PTSD[1]	GAD[1]
Case	2	0	1	1	0	0	9	4	2	0	4
Non-case	13	15	14	14	15	15	6	11	13	15	11
Total	15	15	15	15	15	15	15	15	15	15	15

[1]MDE = Major Depressive Episode; M = Mania or hypo-mania; PD = Panic disorder; PTSD = Post-traumatic stress disorder; GAD = Generalized anxiety disorder; IED = Intermittent explosive disorder.
[2]IED was not assessed in the SCID.

Table A2. Comparison of lifetime DSM-IV diagnoses based on the second draft translation of the fully-structured Composite International Diagnostic Interview (CIDI) and independent diagnoses based on clinical reappraisal interviews with the Structured Clinical Interview for DSM-IV (SCID) among psychiatric outpatients at Bharatpur Medical College Teaching Hospital and a Private Clinic in Kathmandu ($n = 20$)

	CIDI						SCID[2]				
	MDE[1]	M[1]	PD[1]	PTSD[1]	GAD[1]	IED[1]	MDE[1]	M[1]	PD[1]	PTSD[1]	GAD[1]
Case	7	0	1	1	0	0	9	4	2	0	4
Non-case	13	20	19	19	20	20	11	16	18	20	16
Total	20	20	20	20	20	20	20	20	20	20	20

[1]MDE = Major Depressive Episode; M = Mania or hypo-mania; PD = Panic disorder; PTSD = Post-traumatic stress disorder; GAD = Generalized anxiety disorder; IED = Intermittent explosive disorder.
[2]IED was not assessed in the SCID.