# **Supplemental Materials**

## The CVFS Context

As explained in the main text, the design of the Chitwan Valley Family Study (CVFS) took place over many years of work by an interdisciplinary team of Nepalese population experts working to apply state-of-the-art mixed methods and survey methodology. The CVFS then followed all members of a large, population-representative sample of families for more than 15 years before the work to design a Nepal-CIDI began. Over that time, interactions between CVFS research team members and the study population taught crucial details about data collection from this population that enhanced the success of the Nepal-CIDI. One of these details is that the CVFS always carefully protected the confidentiality and safety of all respondents, no matter what the topic, no matter what the context, including 8 years of armed conflict in the local setting. The second of these details is appropriate setting-specific approaches for measuring potentially sensitive personal behaviors such as details of contraceptive method use, individual variations in participation in the choice of a spouse, and details of the husband-wife relationship, including intimate partner violence (Axinn, Ghimire, & Smith-Greenaway, 2017; Ghimire, Axinn, & Smith-Greenaway, 2015; Ghimire, Axinn, Yabiku, & Thornton, 2006). The third is a consistent content focus on the health and wellbeing of respondents, including occasional long interviews to collect in-depth data. The fourth is managing participants’ desire to please researchers or obtain services by supplying responses that produce that outcome: across 25 years from early design work to the present CVFS has consistently told respondents that no program or services will follow from their answers, and all respondents have that lived experience. Together, these multiple features of the CVFS create an ideal context for thorough measurement of experiences with mental disorders.

## Methods to Create a Nepal-CIDI

 Several different efforts to translate and adapt the WMH-CIDI to measure psychiatric disorders in various regions of Asia had failed before the CVFS team began this work. Challenges were significant because both language and culture of specific populations within Asia were not only quite different from the European diaspora, they were also different than other populations within Asia. These challenges have the potential to undermine conceptual equivalence, criterion equivalence, and content equivalence (Ghimire, Axinn, Gatny, & Chardoul, 2017). The list of failures in Asia included an early effort to translate and adapt the WMH-CIDI for Nepal (van Ommeren, Sharma, Makaju, Thapa, & de Jong, 2000). That early study of the CIDI identified comprehensibility, acceptability, relevance, and completeness as key challenges to constructing s context-specific, culturally appropriate translation in this population(van Ommeren et al., 1999). A second effort to apply a culturally appropriate translation of the CIDI in Nepal in 2002 identified challenges that were quite similar (Tausig et al., 2011). The results of these early attempts—and their clear identification of key obstacles to success—were crucial guidance to the CVFS team of methodologists.

The strategy employed the survey methodology of cross-cultural comparison and known best practices in the field of survey translation and adaptation (Harkness et al., 2010; Harkness, van de Vijver, & Mohler, 2002). The strategy involved the following steps:

1. consultative meetings with local mental health professionals and researchers;
2. training the local field staff;
3. iterative translation and revision of specific measures;
4. cognitive interviewing with representatives of the general population;
5. repeated pretest and revision of the translation;
6. piloting the survey; and
7. validation of the measures using a gold standard clinical research diagnostic interview, the Structured Clinical Interview for DSM-IV (SCID) (First, Spitzer, Gibbon, & Williams, 1994), administered by culturally competent (local) psychiatrists.

The key ingredient to achieving success was ***multiple repetition of these steps across three years***. Repetitions included all the steps—all the way through to repeated efforts at clinical validation—exercising rigorous scientific standards in each iteration, as explained in Ghimire et al. (Ghimire, Chardoul, Kessler, Axinn, & Adhikari, 2013).

Finally, in this Chitwan population-specific study, to conduct repeated validation studies of the diagnoses generated from the translated and adapted instruments, 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) (Ghimire et al., 2017).

However, as is true in all settings, respondents struggle to recall all of the circumstances in their lives that help them recall specific symptoms that contribute to either DSM or ICD diagnoses of specific disorders. Nepalese language and culture make identification of precise timing difficult—these issues were discovered early in the CVFS design in the 1990s. In the 1990s, these issues led the CVFS team to design a specially constructed linked-pair of historical recall tools—a focus group interview based local Neighborhood History Calendar and an individual interview Life History Calendar that, in addition to recall tools used in others settings, add the time of local neighborhood events to the matrix of cues used to help individuals recall the timing of events in their own lives (Axinn & Pearce, 2006; Axinn, Pearce, & Ghimire, 1999). Building on this, the team designed a special Nepal-CIDI linked LHC (directly integrated into the Nepal-CIDI) to improve respondents’ recall of mental disorder symptoms, especially the timing and duration of those symptoms. Early tests of this addition to the design were so promising the CVFS team launched a large-scale experiment to investigate this change and clinically validate the results. The experiment randomly assigned more than 1000 respondents to receive either the LHC-CIDI or the CIDI with no LHC. Results demonstrated the LHC-CIDI significantly and substantially increases reporting of symptoms of disorders even when beginning with a culturally appropriate adaptation (the Nepal-CIDI) (Axinn et al., 2019). Large-scale clinical validation with a carefully trained team of psychiatric professionals also yields high concordance between the Nepal LHC-CIDI and the SCID, comparable to the best validation results from American and European CIDI validation studies (Axinn et al., 2019). Note that careful integration of a LHC into the Nepal-CIDI, among a study population familiar with the objective of the LHC (correct reports of timing of events), also potentially reduces respondent tendencies to “over-report” symptoms as within the past 12 months, when in fact those symptoms took place more than 12 months ago. It was this version of the final instrument that was used at full scale in the CVFS sample.

## Implementation of the Nepal-CIDI

 The administration of the Nepal-specific, LHC-enhanced WMH-CIDI took place with extensive quality control measures during the fieldwork itself. This was accomplished by using professional interviewers, computer-assisted interview (CAI) methods, and a state-of-the-art web-based management and quality control tool. Professional interviewers received more than a month of rigorous training and were required to pass a formal certification at the end of training before they could begin work. The CAI system was designed not only to simplify the complex CIDI instrument by taking interviewers to the correct next question following each response, it also provided dynamic quality control during the interview by detecting errors in data entry and inconsistent responses. During field work the study used a high level of field supervision (1 supervisor to 5 interviewers), following best-practices for all survey data collection, with training to the University of Michigan’s Survey Research Center’s (U-M SRC) standards, and harnessing supervisors each having more than a decade of prior experience (Groves et al., 2009). Finally, the CVFS team used a web-based electronic sample management system that allows centralized supervisors both in Nepal and at U-M SRC to monitor interviewer work as it happens. The electronic management system measures all actions taken by field interviewers, providing the means for analyses of these survey “paradata” to reveal poor quality interviewer performance on a daily basis—just as is done in the highest quality surveys in the most highly resourced settings (Groves et al., 2009).

**References**

Axinn, W. G., Chardoul, S., Gatny, H., Ghimire, D. J., Smoller, J. W., Zhang, Y., & Scott, K. M. (2019). Using life history calendars to improve measurement of lifetime experience with mental disorders. *Psychological Medicine*, 1–8. https://doi.org/10.1017/S0033291719000394

Axinn, W. G., Ghimire, D. J., & Smith-Greenaway, E. (2017). Emotional variation and fertility behavior. *Demography*, *54*(2), 437–458. https://doi.org/10.1007/s13524-017-0555-5

Axinn, W. G., & Pearce, L. D. (2006). *Mixed method data collection strategies*. New York, NY: Cambridge University Press.

Axinn, W. G., Pearce, L. D., & Ghimire, D. J. (1999). Innovations in life history calendar applications. *Social Science Research*, *28*(3), 243–264. https://doi.org/10.1006/ssre.1998.0641

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 Yok: Biometrics Research Dept.

Ghimire, D. J., Axinn, W. G., Gatny, H., & Chardoul, S. (2017). Preparing a culturally appropriate translation of a survey questionnaire. *SAGE Research Method Cases*. https://doi.org/10.4135/9781526409515

Ghimire, D. J., Axinn, W. G., & Smith-Greenaway, E. (2015). Impact of the spread of mass education on married women’s experience with domestic violence. *Social Science Research*, *54*, 319–331. https://doi.org/10.1016/j.ssresearch.2015.08.004

Ghimire, D. J., Axinn, W. G., Yabiku, S. T., & Thornton, A. (2006). Social change, premarital nonfamily experience, and spouse choice in an arranged marriage society. *American Journal of Sociology*, *111*(4), 1181–1218. https://doi.org/10.1086/498468

Ghimire, D. J., Chardoul, S., Kessler, R. C., Axinn, W. G., & Adhikari, B. P. (2013). Modifying and validating the Composite International Diagnostic Interview (CIDI) for use in Nepal. *International Journal of Methods in Psychiatric Research*, *22*(1), 71–81. https://doi.org/10.1002/mpr.1375

Groves, R. M., Fowler Jr., F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology*. Hoboken, NJ: John Wiley & Sons.

Harkness, J. A., Braun, M., Edwards, B., Johnson, T. P., Lyberg, L. E., Mohler, P. Ph., … Smith, T. W. (Eds.). (2010). *Survey methods in multinational, multiregional, and multicultural contexts* (1st ed.). Hoboken, NJ: Wiley.

Harkness, J. A., van de Vijver, F. J. R., & Mohler, P. Ph. (Eds.). (2002). *Cross-cultural survey methods* (1st ed.). Hoboken, NJ: Wiley.

Tausig, M., Subedi, J., Broughton, C., Pokimica, J., Huang, Y., & Santangelo, S. L. (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. https://doi.org/10.1007/s11469-010-9276-3

van Ommeren, M., Sharma, B., Makaju, R., Thapa, S., & de Jong, J. (2000). Limited cultural validity of the Composite International Diagnostic Interview’s Probe Flow Chart. *Transcultural Psychiatry*, *37*(1), 119–129. https://doi.org/10.1177/136346150003700107

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*(3), 285–301. https://doi.org/10.1177/136346159903600304

**Supplementary Table A. Educational Attainment by Birth Cohort**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Birth Cohort** | **Gender** | **N** | **1957-1971** | **1972-1981** | **1982-1991** | **1992-2001** | **ANOVA Test** |
| **Age (2016 July)** |  |  | 45-59 | 35-44 | 25-34 | 15-24 |  |
| **Year of Schooling, Mean (SD)** | Total | 10,714 | 3.77 (4.47) | 6.30 (4.50) | 8.81 (4.07) | 10.01 (2.68) | \*\*\* |
| Male | 4,923 | 5.97 (4.55) | 8.00 (4.04) | 9.29 (3.80) |  9.84 (2.71) | \*\*\* |
| Female | 5,791 | 1.57 (3.11) | 4.57 (4.28) | 8.47 (4.23) | 10.15 (2.66) | \*\*\* |
| **No Education, %** | Total | 10,714 | 45.70 | 21.63 |  5.69 |  0.46 | \*\*\* |
| Male | 4,923 | 19.05 |  7.22 |  2.04 |  0.34 | \*\*\* |
| Female | 5,791 | 72.22 | 36.42 |  8.37 |  0.55 | \*\*\* |
| **S.L.C or more, %** | Total | 10,714 | 12.00 | 20.63 | 41.44 | 56.42 | \*\*\* |
| Male | 4,923 | 20.71 | 29.06 | 44.86 | 54.60 | \*\*\* |
| Female | 5,791 |  3.31 | 11.98 | 38.91 | 57.87 | \*\*\* |

1. ANOVA test is among four birth cohorts: \*p < 0.05; \*\* p < 0.01; \*\*\* p < 0.001

2. The cohort 1957-1971 is at high risk of missing due to death. The cohort 1992-2001 ages 15-24 at the time of survey, truncated exposure to risk.

3. S.L.C. = School Leaving Certificate