Annexure 5
General Methodology

Development, testing and dissemination of the Flexible Interview for ICD-11 (FLII-11)

***Phase 1*:**
As discussed in the initial application, a FLII-11 international advisory group (FIAG) will be appointed to provide scientific, clinical, and experience-based consultation throughout the project. The FIAG will include diverse experts from LMIC, including PWLE recruited via various global networks and organizations. Phase 1 represents the primary activities under the proposed workplan and covers the initial testing of FLII-11's feasibility and clinical utility as well as its acceptability to PWLE. This section describes the general research methodology to be used for this part of the project, which supplements but does not include all of the specific steps and deliverables listed in the Timetable portion of the initial application. Phase 1 studies will be undertaken at sites in at least four countries: India, Mexico, South Africa and one additional country in Africa to be determined. Additional sites may be added depending on interest and available resources. The general protocol for Phase 1 testing may be adapted or modified in consultation with each partnering organization and of the applicable Institutional Review Board.

For sites that will be testing the FLII-11 in languages other than English, a methodology for translation/adaptation will be used following best practices in the area (Beaton et al., 2000; van Ommeren et al., 1999). This will include forward translation that represents as well as possible the range of application of the translated version. For example, for the Spanish translation, forward translators will be drawn from range of Spanish-speaking countries including at minimum Mexico, Spain, a South American country, and a Caribbean country. Following reconciliation of the different forward translations, a focus group of mental health professionals fluent in the target language but not involved in the translation will be convened to review and provide feedback about the translated document. The version resulting from modifications based on feedback from the clinician focus group will be the version used in Phase 1 field testing. (The translation/cultural adaptation protocol will be refined and modified based on the cumulative experience of the project and will be one of the products of this work stream.)

Participating sites are tertiary academic medical centres that encompass public mental health facilities (I.e., Groote Schur Hospital in Cape Town, South Africa; the National Institute of Psychiatry and the National Institute of Gerontology in Mexico City; the National Institute of Mental Health and Neurosciences in Bangaluru, India) that also have ties to community-based facilities. Specific recruitment processes will vary by centre, but participants will be drawn from clinical settings as well as participants not identified as mental health service recipients. Participants drawn from clinical settings will be individuals who are newly presenting for mental health services (Mexico, India) or recruited for a separate study of genetics and neurological disorders (South Africa). Participants not identified as mental health service recipients will include accompanying family members or friends and participants recruited directly from the community. The non-clinical sample is important because it will generally be more similar to likely participants in an epidemiological survey. For Phase 1, a minimum of 100 adult interviews per country will be conducted, including 50 individuals presenting for mental health services and 50 individuals who are not. This is similar to the international recruitment methodology we have previously used for field studies of the ICD-11 (Reed et al., 2018). Exclusion criteria will be limited to factors that would interfere substantially with being able to complete the interview (i.e., obvious intoxication or medical distress, delirium, moderate or severe dementia, immediate threats of suicide or violence, severe suspicion and lack of cooperation (e.g., based on persecutory ideas), severe mania leading to flight of ideas and motor acceleration), insufficient proficiency in the interview language). Interviewers will have some mental health experience but will not typically be mental health professionals qualified to make independent diagnosis (e.g., masters students in psychology). Interviewers will be trained during a 2-day session using the pilot training materials developed for this project.

The informed consent process for this study will emphasize that the interview is a new measure of mental health experience and that the purpose of the study is to evaluate the measure rather than to evaluate them individually. A key part of the study will be for participants to talk with us about their experience of the interview after they complete it. After informed consent is obtained, participants will provide basic demographic information and participate in the FLII-11. Participant will then be asked to complete the **Interview Experience Questionnaire** (IEQ). The IEQ is being developed in part based on our previous work related to PWLE experience of the process and content of diagnosis (Hackmann et al., 2019; Perkins et al., 2018). process. There are two parts of the IEQ: The Participant Portion

assesses the participant experience in detail, including any portions of the interview that they report having difficulty understanding or finding stigmatizing or upsetting. A series of additional questions will assess other aspects of their experience of the interview. (For example, I felt respected and valued during the interview; The interview took an appropriate amount of time; Some important aspects of my mental health experience were not asked about; I would participate in an interview like this again at some point in the future.) The interviewer will complete a different version of the IEQ for each interview, assessing the interviewer's observations of the participant during the interview, areas the participant appeared to find confusing or upsetting, and any difficulties that arose, as well as the interviewer's own experience in giving the interview. (For example, I found it difficult to get all the relevant information needed to make some of the ratings; I feel comfortable interviewing people with the FLII-11.) If the interviewer is professionally involved in the provision of mental health services, they will be asked another set of questions about their view of the FLII-11's clinical usefulness as it relates to that specific interview. The FLII-11 and the IEQ will be programmed in both Qualtrics and REDCap. We will work with the sites involved to determine which of these data platforms (or another one) is best suited to their needs.

At the end of the interview, participants who are identified as recipients of mental health services will be asked if they are willing to be contacted in several weeks to invite them to participate in a 90-minute group discussion of the experience of the interview. These groups will consist of between 5 to 10 people at each site. The focus group will be led by two trained facilitators. A discussion guide will be developed based on the IEQ and the preliminary results of the FLII-11. The discussion will focus on areas that have been identified as potentially difficult or problematic based. The goal will be to explore in more depth questions that participants most commonly identified as difficult to understand, unclear, stigmatizing, or upsetting, as well as questions identified by interviewers on the IEQ as being of concern. Group-administered cognitive interviewing techniques (Willis, 2015) will be used to gain a better understand of exactly how participants are understanding the questions. The focus group will also discuss possible alternative questions or formulations in areas of difficulty, as well as other aspects of their experience that they feel were not well reflected in the interview. Similar focus groups of adolescent participants will also be conducted in sites that are interviewing adolescents. PWLE focus groups will be thematically analysed based on a critical realist epistemological stance (Guba & Lincoln, 1994; Hackmann et al., 2019), recognising that participants have their own individual experience of reality, but analysing data at face value, using the perspectives of individuals as they represent themselves during the focus group discussions. This approach is intended to capture the nuance of individual experience and develop useful feedback for WHO.

Results from the PWLE focus groups as well as the IEQ and the data from the FLII-11 itself will be used to make changes to the FLII-11 for the next stage of the testing process. These changes will be made in consultation with the FIAG and further PWLE input as indicated by the data.

***Phase 2*:**

Phase 2 will include the development of additional protocols for testing the reliability and validity of the FLII-11 as revised at the end of Phase 1. Protocols for such testing will be developed in consultation with the FIAG and collaborating sites. Reliability protocols will include test-retest and interrater reliability. (Interrater reliability would be expected to be quite high given that the FLII-11 is fully structured, but it will be important to assess the effect of different interviewers.)

Two main types of validity protocols will be developed. The first will compare the results of the FLII-11 recorded clinical diagnoses. For these comparisons, it is important to keep in mind that clinical diagnoses may not be accurate or reliable, so differences between the result of the FLII-11 and recorded diagnoses do not necessarily reflect any deficiency in the FLII-11. However, these comparisons would still be of value in terms of their pattern and how they differ across settings and populations. Some differences are expected. For example, it is impossible for specified diagnostic algorithms to account for symptoms of one disorder being accounted by another disorder with partly overlapping symptoms. So, there will be a higher rate of co-occurring disorders based on the FLII-11 than in clinical settings. It will be important to evaluate different types of diagnostic disagreements. A disagreement that consists of two different diagnoses in the same grouping with equivalent treatments is not the same as one that consists of two unrelated diagnoses with incompatible treatment implications.

The second type of validity protocol will be developed to compare the results of the FLII-11 with other diagnostic instruments. The best comparison would be to a semi-structured clinician-administered interview based on the ICD-11 called the Structured Clinical Interview for ICD-11 (SCII-11, which we have developed in conjunction with our development of the FLII-11. Other available diagnostic interviews also be considered for possible use in validity studies, including as the Mini-International Neuropsychiatric Interview (MINI), the Structured Clinical Interview for DSM-5 (SCID-5), the Composite

International Diagnostic Interview (CIDI), and the Schedules for Clinical Assessment in Neuropsychiatry (SCAN).

It is important to note that funding for reliability and validity protocols is not included in the current grant application. We will of course be interested in the implementation of these protocols based on resources available to WHO or to government and institutional partners. Even if we do not obtain additional resources, we will make the protocols available for such use and will participate in their implementation and dissemination as actively as possible.

References

Beaton, D., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, *25*, 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Guba, E. G. & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln, Y.S (Eds.), *The landscape of qualitative research: Theories and issues* (pp. 105 – 117). Sage Publications.

Hackmann, C., Balhara, Y.P.S., Clayman, K., Nemec, P.B., Notley, C., Pike, K., Reed, G.M., Sharan, P., Rana, M.S., Silver, J., Swarbrick, M., Wilson, J., Zeilig, H., & Shakespeare, T. (2019). Perspectives on ICD-11 to understand and improve mental health diagnosis using expertise by experience (INCLUDE Study): An international qualitative study. *Lancet Psychiatry*, *6*, 778-785. https://doi.org/10.1016/S2215-0366(19)30093-8

Perkins, A., Ridler, J., Browes, D., Peryer, G., Notley, C., & Hackmann, C. (2018). Experiencing mental health diagnosis: A systematic review of service user, clinician, and carer perspectives across clinical settings. *Lancet Psychiatry*, *5*, 747–764. https://doi.org/10.1016/S2215-0366(18)30095-6

Reed, G.M., Sharan, P., Rebello, T.J., Keeley, J.W., Medina-Mora, M.E., Gureje, O., Ayuso-Mateos, J.L., Kanba, S., Khoury, B., Kogan, C.S., Krasnov, V.N., Maj, M., Mari, J. de J., Stein, D.S., Zhao, M., Akiyama, T., Andrews, H.F., Asevedo, E., Cheour, M., … & Pike, K.M. (2018). The ICD-11 developmental field study of reliability of diagnoses of high-burden mental disorders: Results among adult patients in mental health settings of 13 countries. *World Psychiatry*, *17*, 174-186. https://doi.org/10.1002/wps.20524

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*, 1999; 36: *285-301.* https://doi.org/10.1177/136346159903600304

Willis, G. B. (2015). The practice of cross-cultural cognitive interviewing. *Public Opinion Quarterly*, *79*(S1), 359–395. https://doi.org/10.1093/poq/nfu092