

Feasibility, acceptability and clinical utility of the Cultural Formulation Interview: mixed-methods results from the DSM-5 international field trial

Roberto Lewis-Fernández, Neil Krishan Aggarwal, Peter C. Lam, Hanga Galfalvy, Mitchell G. Weiss, Laurence J. Kirmayer, Vasudeo Paralikar, Smita N. Deshpande, Esperanza Díaz, Anel V. Nicasio, Marit Boiler, Renato D. Alarcón, Hans Rohlf, Simon Groen, Rob C. J. van Dijk, Sushrut Jadhav, Sanjeev Sarmukaddam, David Ndeti, Monica Z. Scalco, Kavos Bassiri, Sergio Aguilar-Gaxiola, Hendry Ton, Joseph Westermeyer and Johann M. Vega-Dienstmaier

Background

There is a need for clinical tools to identify cultural issues in diagnostic assessment.

Aims

To assess the feasibility, acceptability and clinical utility of the DSM-5 Cultural Formulation Interview (CFI) in routine clinical practice.

Method

Mixed-methods evaluation of field trial data from six countries. The CFI was administered to diagnostically diverse psychiatric out-patients during a diagnostic interview. In post-evaluation sessions, patients and clinicians completed debriefing qualitative interviews and Likert-scale questionnaires. The duration of CFI administration and the full diagnostic session were monitored.

Results

Mixed-methods data from 318 patients and 75 clinicians found the CFI feasible, acceptable and useful. Clinician feasibility ratings were significantly lower than patient ratings and other clinician-assessed outcomes. After administering one CFI, however, clinician feasibility ratings improved significantly and subsequent interviews required less time.

Conclusions

The CFI was included in DSM-5 as a feasible, acceptable and useful cultural assessment tool.

Declaration of interest

R.L.-F., N.K.A. and L.J.K. receive royalties from an edited book on the Cultural Formulation Interview, published by American Psychiatric Publishing, Inc.

Copyright and usage

© The Royal College of Psychiatrists 2017.

Unexamined cultural differences in how patients and clinicians frame illness and care may distort diagnosis and assessments of severity, impose communication barriers, compromise engagement, adherence and response, and unnecessarily prolong patients' suffering.^{1,2} Patient–clinician differences in age, gender, sexual orientation, socioeconomic status, race/ethnicity, religion, language, and/or national origin can contribute to cultural differences in all clinical interactions.^{3,4} The DSM-IV Outline for Cultural Formulation (OCF) is a conceptual framework that helps clinicians identify the impact of culture on illness and care during a clinical evaluation.^{5,6} The OCF is widely used in clinical training and cultural competence initiatives.^{7–10} However, its implementation in routine care has proved challenging:¹¹ clinicians had to improvise questions to collect the information, received limited guidance on which patients would benefit most, and faced uncertainty about whether to implement the OCF as a separate assessment or embed it in a standard clinical evaluation.^{12–14} The lack of a structured instrument also impeded research on cultural assessment and inclusion of cultural information in clinical trials.^{15,16} In response, the American Psychiatric Association's DSM-5 Cross-Cultural Issues Subgroup (DCCIS) developed the Cultural Formulation Interview (CFI)¹⁷ to operationalise the OCF for routine use in the clinical assessment of any patient, based on a literature review and consensus-building discussions with designers of OCF-based interviews.¹⁸ The CFI instruments comprise an initial assessment interview (core CFI), an informant interview for collateral information and 12 supplementary modules that expand on these basic assessments. The core CFI consists of an introduction,

open-ended questions for patients and instructions to clinicians for each question. Acknowledging the need for global relevance and recognising international work on the OCF, sites in six countries participated in the field trial.

This report presents findings from the international field trial that tested the 14-item pilot version of the core CFI (online supplement DS1) in three service domains based on patient and clinician feedback. Together with other field trial data not reported here, this process resulted in the final 16-item version in DSM-5.¹⁹ We assessed several factors related to successful implementation of clinical innovations in service settings,²⁰ including patient and clinician perceptions of the CFI's feasibility ('Can it be done in clinical settings?'), acceptability ('Do patients and clinicians like it?'), and potential clinical utility ('Is it helpful?'). We also considered whether closed- and open-ended assessments yielded similar results, and whether outcomes showed a practice effect, improving with experience. Our study is the first to examine these service domains for a tool to enhance cultural competence in multiple international settings.

Method

Study design and settings

The CFI field trial was designed by the DSM-5 DCCIS via regular teleconferences.^{19,21} The study was conducted from November 2011 to September 2012; the New York site coordinated logistics for all sites. The study design purposively included samples of diverse patients, clinician disciplines and types of out-patient

services, because a goal of the DSM-5 trials was to test the feasibility, acceptability and utility of proposed diagnoses and assessments under varied clinical conditions to determine inclusion in DSM-5.^{22,23} Each site aimed to enrol at least 30 patients from affiliated psychiatric out-patient clinics in Canada (one site), India (two), Kenya (one), The Netherlands (one), Peru (one) and the USA (five). Sites were chosen based on involvement of a principal investigator in the DCCIS and aimed to include diverse cultural populations and types of out-patient services (general community, immigrant/refugee and ethnic-focus clinics).

An opportunity sample of new and existing patients at each site was enrolled using a standard recruitment script. Clinicians who had no prior contact with their study patient conducted the interviews ('study clinicians'). Clinicians did not interview their own patients because prior knowledge and a pre-existing relationship would confound study aims focusing on an initial assessment. Current patients were referred by treating clinicians to local study clinicians. Each study clinician was expected to interview 3–6 patients during the trial to assess practice effects. Each patient participated only once. Patients and clinicians could also invite companions (for example relatives) to participate in the interview and subsequent assessments.²⁴

All study clinicians participated in a 2 h CFI training session at their site consisting of (a) reviewing the core CFI's written guidelines; (b) a 24 min video demonstration; (c) interactive behavioural simulations with coaching and feedback from local principal investigators; and (d) a question-and-answer period.

The study clinician administered the CFI followed by a routine diagnostic assessment. Topics of the CFI comprise four cultural domains: (a) definition of the problem; (b) perceptions of cause, context and support; (c) factors affecting self-coping and past help-seeking; and (d) factors affecting current help-seeking. All sessions were audiotaped with patient consent. The study was approved by each site's institutional review/ethics board and followed local informed consent regulations. All patients completed their locally approved consent process.

Participants

Eligible patients were aged 16 or older and fluent in the language of the local clinicians. We required the language match to avoid using interpreters who might introduce cultural information not obtained through the CFI. Patients were excluded if they were acutely suicidal or homicidal, intoxicated or in substance withdrawal, or if their condition seriously limited the assessment (such as dementia). Eligible study clinicians had a clinical degree permitting them to see patients, consistent with each country's requirements.

Assessments

Pre-interview, patients and clinicians completed demographic surveys. Clinicians also indicated their professional training and cultural competence experiences. Local principal investigators identified demographic factors recognised by their governments as indicators of social differences, avoiding a USA-based characterisation.^{19,25} After every session, study clinicians provided patients' DSM-IV diagnoses and patients and clinicians completed follow-up questionnaires and semi-structured qualitative interviews. All assessments were translated into the local languages at each site and reviewed by a bilingual committee of mental health professionals for consensus.²⁶

Quantitative

Participants completed two brief questionnaires: the Debriefing Instrument for Patients (DIP) and the Debriefing Instrument

for Clinicians (DIC), which comprise self-administered, Likert-scale items assessing feasibility, acceptability and clinical utility (online supplement DS2) coded as 'Strongly disagree', 'Disagree', 'Agree' and 'Strongly Agree'. As with other DSM-5 trials,²² these instruments were created for use in the CFI field trial. Items were selected for measurement by the DCCIS with reference to three domains (feasibility, acceptability and clinical utility) likely to affect the implementation of assessments such as the CFI.^{20,22} The same content was included in each instrument, with wording adapted for each stakeholder group. As a measure of feasibility independent of self-report, we assessed the duration of the CFI and the total diagnostic interview (including the CFI), based on session audio files.

Qualitative

Separate semi-structured qualitative interviews (8–9 questions, previously reported¹⁹) with patients and clinicians conducted by research assistants at each site provided more detailed accounts of the impact of the CFI on the initial evaluation. These interviews assessed participants' perceptions of the most and least helpful aspects of the CFI, its impact on interview quality and outcomes, and its role in clinical practice, including diagnosis and treatment planning. Each site provided written English summaries of the interviews to the coordinating site.

Analysis

Quantitative

SAS version 9.4 (Cary, NC) was used for all analyses.

Descriptive information. Patient and clinician characteristics were compared cross-nationally using ANOVA for continuous variables and Chi-square (or Fisher's exact test) for categorical variables; the Kruskal–Wallis test was used for ordinal or continuous variables with skewed distributions.

DIC/DIP. Negative DIC/DIP responses were coded as -2 (strongly disagree) or -1 (disagree) and positive responses as $+1$ (agree) or $+2$ (strongly agree).^{24,27} Missing responses were imputed using the mean of the non-missing items within the assessment domain for the individual. Mean proportion of missing responses was 4.5% (s.d. = 1.4) for the DIP (range 2.8–7.6% for a single item) and 2.2% (s.d. = 1.0) for the DIC (range 0.9–4.1%). Cronbach's alpha was used to assess the internal consistency of the three DIC/DIP domains. For domains with $\alpha < 0.70$, inter-item correlation matrices, item correlation with total and changes to alpha by item were examined to detect problematic items; these items were excluded from subsequent analyses.

Mean DIC/DIP scores for feasibility, acceptability and utility were compared within patient and clinician cohorts, cross-nationally and overall. We also compared the overall patient and clinician mean scores for each assessment domain; remaining items in domains with excluded items were also compared individually. To account for site-specific effects, clinicians seeing several patients and the inclusion of new and existing patients to the clinic, we used generalised linear mixed-effects models (PROC GLIMMIX in SAS), with random intercepts for site and clinician and a fixed effect for new patient status. Tukey–Kramer *post hoc* tests that adjust for multiple comparisons were used to identify significant patient–clinician differences.²⁸

Duration. Durations of the CFI and the full diagnostic interview (including the CFI) were compared separately cross-nationally

using PROC GLIMMIX to adjust for new patient status and clinician effects. The proportion of total interview time devoted to the CFI was also calculated.

Practice effect. To determine whether clinicians' accumulated experience with the CFI affected their perceptions of the outcomes, we analysed changes in DIC scores over subsequent CFI interviews; we also analysed interview duration and the proportion of time devoted to the CFI in the full interview for each clinician. A mixed-effects model adjusted for clinician and site effects (but not patient newness, since patients were always new to study clinicians). Separate mixed-effects models and Tukey–Kramer *post hoc* tests contrasted DIC assessment domains between and within each administration, respectively.

Qualitative

Qualitative analyses were conducted by a three-person multi-disciplinary team (public health, sociology and psychiatry) using deductive content analysis and working independently of the quantitative analysis team. Deductive content analysis codes qualitative data using pre-established categories based on theoretical frameworks.^{29,30} Each debriefing interview was coded for feasibility, acceptability and utility according to a codebook (developed by N.K.A.): feasibility and acceptability were defined as per Proctor *et al*²⁰ and their definition for appropriateness was used to define utility, consistent with the terminology of the DSM-5 trials.²⁵ Coder training consisted of two 1 h sessions. Each coder labelled each interview phrase with one unique code for feasibility, acceptability or utility to minimise bias.³¹ Interrater reliability of 80% was achieved using a random 10% selection of transcripts. Iterative revision of the codebook was conducted over 5 weeks by reviewing concordance among codes and concepts, developing new subcodes, memoing, specifying code definitions with parameters (appropriate and inappropriate use), and reviewing data examples until new information produced no change to coding categories. All debriefing interviews were uploaded into NVivo (QSR International 2012) and randomly assigned for coding. NVivo reports were generated for codes, exploring patterns and drafting analytical memos by theme. Qualitative codes were counted by individual respondent and by number of mentions per text to analyse data by session and for the total sample.

Results

Patient characteristics

The field trial enrolled 321 patients; 3 were under 16 and were excluded, leaving 318 for analysis, of whom 189 were new and 129 existing patients. They had a mean age of 41.4 and 10.6 years of education; half were female (Table 1 and Table DS6). Most countries had an even distribution of employed, unemployed and participants who were out-of-the-labour-force (for example retired), except for the USA where nearly half were disabled. Marital status differed by country. Proportion of foreign-born individuals ranged widely, from 0% in Peru to 97% in Canada. Patients' primary language varied by site. Significant cross-national differences were observed for all sociodemographic variables (gender: $P < 0.05$; all others: $P < 0.001$). Clinically, 70% of patients received one DSM-IV Axis I diagnosis, 20% received two, 7% three or more, and 2% none (Table 1); this proportion varied significantly across countries ($P < 0.001$). Depressive disorders were diagnosed most frequently, followed by anxiety disorders.

Clinician characteristics

In total, 75 clinicians were enrolled, with an average age of 38.4; over 50% were female, except in The Netherlands and Peru (Table 2). Nearly 50% were psychiatrists or psychiatric trainees, 28% psychologists, and 15% social workers. Countries differed substantially on several indices. Kenyan clinicians had a mean of 3 years of practice, had seldom/never treated patients of different cultures, and all had < 10 h of cultural training. By contrast Dutch clinicians had 15.6 years of practice, 91% had daily cross-cultural contacts, and half had > 50 h of cultural training. The proportion of foreign-born clinicians ranged from 0% in India and Peru to 57% in Canada. All variables differed significantly across countries, except for age and gender.

Self-report outcome ratings

Cronbach's alphas for the DIC were high: 0.78 (feasibility), 0.80 (acceptability) and 0.89 (utility). DIP internal consistency was high for utility (0.82) but minimal for feasibility (0.18) and acceptability (0.17). Item-based analyses identified one problematic item under feasibility ('Took more time to share my perspective than I wanted') and acceptability ('Were too personal'); both items were negatively worded. Removing these items²⁷ increased Cronbach's alpha for feasibility (0.45) and acceptability (0.48) (online supplement DS2), these domains each now containing two items. Prior research on cross-cultural variation with negatively worded survey items supports this approach.³²

Patient and clinician ratings of feasibility, acceptability and clinical utility were positive, but varied significantly cross-nationally (online Table DS7). Once adjusted for site effects, mean overall results for all three outcomes (Table 3) were positive among patients – scoring 1.26–1.33 on a scale from -2 to $+2$ – but evaluations were less positive among clinicians, with scores of 0.93–0.98 on utility and acceptability and 0.75 on feasibility. Overall, feasibility was significantly lower than the other indices among clinicians, and significantly lower than patients' feasibility rating. Clinicians also rated acceptability and utility lower than patients, but not significantly. By contrast, patient scores across assessment domains were nearly identical.

After excluding the two problematic DIP items, comparison of remaining single-item ratings of feasibility (easy to understand, $t(10) = 5.27$, $P < 0.001$; improved flow, $t(10) = 2.32$, $P = 0.043$) and acceptability (encourage clinician use, $t(10) = 2.17$, $P = 0.055$; felt at ease, $t(10) = 21.3$, $P = 0.059$) across patient and clinician assessments revealed the same pattern as the analysis of means. DIC single-item results (online supplement DS2) identified clinician concerns about CFI comprehensibility and interview flow (feasibility) and about CFI impact on clarification of diagnosis, cultural background, severity, and patient–clinician differences (utility). DIP single-item results did not indicate specific concerns, although identification of barriers to care (utility) scored somewhat lower than other items.

Duration

Average CFI duration ranged from 18.8 min in The Netherlands to 29.2 in Kenya ($P < 0.001$) and total interview duration ranged from 37.6 min in Kenya to 88.2 in The Netherlands ($P < 0.001$). Average overall CFI duration was 23.4 min, within a 54.1 min intake. Cross-nationally, the proportion of the interview devoted to the CFI varied significantly (online Table DS7).

Table 1 Patient sample characteristics (n = 318)

Patients	Canada (n = 33)	India (n = 101)	Kenya (n = 29)	Netherlands (n = 30)	Peru (n = 34)	USA (n = 91)	Total (n = 318)	Test statistic ^a		P
								F (d.f.)	χ^2 (d.f.)	
Age, mean (s.d.)	51.12 (15.85)	35.42 (12.85) ^b	31.97 (10.77)	41.87 (15.33)	36.50 (10.47) ^b	49.25 (13.62) ^b	41.44 (14.95)	17.56 (5,306)		< 0.001***
Years of education, mean (s.d.)	7.53 (4.94) ^c	11.37 (4.21)	9.83 (3.37)	12.03 (4.97)	12.56 (2.83)	9.94 (4.78) ^d	10.64 (4.52)	6.92 (5,101)		< 0.001***
Female, n (%)	21 (63.64)	42 (41.58)	14 (48.28)	9 (30.00)	20 (58.82)	50 (54.95)	156 (49.06)			0.011*
Foreign-born, n (%)	32 (96.97)	1 (0.99)	1 (3.45)	17 (56.67)	0 (0)	61 (67.03)	112 (35.22)			< 0.001***
New to CFI clinic, n (%)	33 (100)	101 (100)	0 (0)	20 (66.67)	34 (100)	1 (1.10)	189 (59.43)			< 0.001***
Axis I diagnoses, n (%)										
0	2 (6.06)	3 (2.97)	0 (0)	1 (3.33)	0 (0)	1 (1.10)	7 (2.20)			< 0.001***
1	25 (75.76)	91 (90.10)	26 (89.66)	12 (40.00)	20 (58.82)	49 (53.85)	223 (70.13)			
2	5 (15.15)	6 (5.94)	3 (10.34)	9 (30.00)	10 (29.41)	32 (35.16)	65 (20.44)			
3 or more	1 (3.03)	1 (0.99)	0 (0)	8 (26.67)	4 (11.76)	9 (9.89)	23 (7.23)			

CFI, Cultural Formulation Interview.
 a. For the variable female: Fisher's exact test was used.
 b. Data unavailable for two participants.
 c. Data unavailable for one participant.
 d. Data unavailable for ten participants.
 *P<0.05; **P<0.01; ***P<0.001.

Table 2 Clinician sample characteristics (n = 75)

Clinicians	Canada (n = 7)	India (n = 21)	Kenya (n = 5)	Netherlands (n = 11)	Peru (n = 5)	USA (n = 26)	Total (n = 75)	Test statistic, ^a		P
								χ^2 (d.f.)	P	
Age, mean (s.d.)	37.57 (7.76)	34.67 (7.48)	33.40 (4.39)	43.64 (11.46)	39.60 (8.26)	40.08 (10.26) ^b	38.35 (9.12)	10.39 (5)		0.065
Years providing mental healthcare, mean (s.d.)	10.14 (5.24)	7.48 (7.09)	3.00 (1.22)	15.55 (12.64)	6.60 (4.34)	10.16 (8.61) ^b	9.47 (8.58)	13.12 (5)		0.022*
Female, n (%)	6 (85.71)	12 (57.14)	3 (60.00)	5 (45.45)	1 (20.00)	14 (53.85)	41 (54.67)			0.377
Professional discipline, n (%)										
Psychiatrist/psychiatry trainee	2 (28.57)	15 (71.43)	5 (100)	2 (18.18)	5 (100)	8 (30.77)	37 (49.33)			< 0.001***
Psychologist	1 (14.29)	2 (9.52)	0 (0)	6 (54.55)	0 (0)	12 (46.15)	21 (28.00)			
Social worker	1 (14.29)	4 (19.05)	0 (0)	3 (27.27)	0 (0)	3 (11.54)	11 (14.67)			
Other mental health clinician ^c	3 (42.86)	0 (0)	0 (0)	0 (0)	0 (0)	3 (11.54)	6 (8.00)			
Foreign-born, n (%)	4 (57.14)	0 (0)	1 (20.00)	2 (18.18)	0 (0)	11 (42.31)	18 (24.00)			< 0.001***
Frequency of contact with patients of different cultures, n (%)										
Daily	4 (57.14)	12 (57.14)	0 (0)	10 (90.91)	1 (20.00)	19 (73.08)	46 (61.33)			< 0.001***
Weekly or monthly	0 (0)	9 (42.86)	0 (0)	1 (9.09)	4 (80.00)	4 (15.38)	18 (24.00)			
Seldom or never	3 (42.86)	0 (0)	5 (100)	0 (0)	0 (0)	3 (11.54)	11 (14.67)			
Hours of cultural training										
< 10 h	2 (28.57)	6 (28.57)	5 (100)	3 (50.00) ^d	0 (0)	3 (11.54)	19 (27.14)			0.015*
10–50 h	1 (14.29)	11 (52.38)	0 (0)	0 (0) ^d	3 (60.00)	11 (42.31)	26 (37.14)			
> 50 h	4 (57.14)	4 (19.05)	0 (0)	3 (50.00) ^d	2 (40.00)	12 (46.15)	25 (35.71)			

a. For the variables female, professional discipline, foreign-born; frequency of contact with patients of different cultures: Fisher's exact test was used.
 b. Data unavailable for one participant.
 c. Other clinicians: licensed marriage and family therapist (n = 1), rehabilitation counselor (n = 1), psychology trainee (n = 2), unspecified clinician (n = 1).
 d. Data unavailable for five participants.
 *P<0.05; **P<0.01; ***P<0.001.

Table 3 Comparing feasibility, acceptability and clinical utility of the Cultural Formulation Interview (CFI) from Likert-scale debriefing questionnaires, by clinicians and patients ($n = 315$)^a

	Domain			Test statistic, <i>F</i> (d.f.)	<i>P</i>
	Feasibility	Acceptability	Clinical utility		
Patients, mean (s.d.)	1.33 (0.57)	1.27 (0.71)	1.26 (0.53)	1.41 (2,833)	0.246
Clinicians, mean (s.d.)	0.75 (0.90) ^{†‡}	0.98 (0.75) [†]	0.93 (0.70) [‡]	13.37 (2,864)	< 0.001***
Test statistic, <i>t</i> (d.f.)	3.53 (10)	1.65 (10)	2.14 (10)		
<i>P</i>	0.005**	0.131	0.058		

a. Mixed-effect models compared domain score differences within and between groups, controlling for clinicians seeing multiple patients, multiple clinicians within a site and whether the patient seen was new to the clinic. Data unavailable for the following parameters: patient acceptability ($n = 16$), patient feasibility ($n = 13$), patient utility ($n = 5$), and clinician acceptability ($n = 3$).
^{†‡}, Values with paired superscripts in the same row differ significantly ($P < 0.05$) after adjusting for multiple comparisons, Tukey–Kramer test.
^{*} $P < 0.05$; ^{**} $P < 0.01$; ^{***} $P < 0.001$.

Practice effects

Clinician (DIC) feasibility ratings improved significantly with practice, from an average of 0.59 at first use to 0.96 at the sixth or subsequent administration (Table 4). Acceptability and utility scores, by contrast, were stable and positive over time. Feasibility differed significantly from acceptability and utility ratings only for the first administration. Mean CFI duration decreased significantly, by over 4 min, consistent with clinicians' reports of increasing confidence in feasibility. This effect on CFI duration was evident by clinicians' second CFI administration, and remained stable at 22–23 min thereafter. Mean total diagnostic interview duration also decreased significantly but gradually, by over 12 min from first to last administration. CFI proportion of the total interview time increased slightly with practice.

Qualitative interviews

Qualitative coding of the post-CFI open-ended debriefing interviews identified a pattern similar to the closed-ended quantitative DIC/DIP analysis (online Table DS8). Clinicians had a more negative perception of CFI feasibility than patients: 107 of 318 clinician interviews included negative feasibility

comments about the CFI as a tool, and 39 negative feasibility comments concerning prospects for clinical implementation, compared with only 26 and 7 negative comments, respectively, among 318 patients. By contrast, patients made 81 positive feasibility comments about the CFI and 14 positive feasibility comments about its implementation prospects, whereas clinicians only made 30 and 9 positive comments, respectively. Clinicians' concerns focused on feasibility; acceptability and utility elicited more positive views. By contrast, patients' comments were largely positive across all assessment domains. These patterns were identical whether views were coded by participant or by total number of utterances.

Clinicians were concerned about the CFI's feasibility as a tool, faulting its organisation ('jumbled') and its placement early in the clinical interview. They also worried about implementation-related issues, such as time burden and whether the format was overly structured. Patients were more positive about feasibility, praising the CFI structure ('from basic questions to more complex . . . in the sense of how you feel') and clinicians' non-'pressured' administration. However, some patients found 'all the details' confusing; they also worried the CFI might be too time-consuming for busy clinicians.

Table 4 Practice effects on feasibility, acceptability, clinical utility, interview duration and proportion of total interview devoted to the Cultural Formulation Interview (CFI), in successive clinician interviews using the CFI ($n = 316$)^a

	CFI administration, mean (s.d.)						Beta (95% CI)	<i>P</i>
	First ($n = 74$)	Second ($n = 68$)	Third ($n = 67$)	Fourth ($n = 42$)	Fifth ($n = 26$)	≥ Sixth ^b ($n = 39$)		
Feasibility	0.59 (1.02) ^{†‡}	0.81 (0.95)	0.72 (0.92)	0.84 (0.66)	0.72 (0.94)	0.96 (0.67) ^c	0.053 (0.003 to 0.103)	0.039*
Acceptability	1.01 (0.72) [†]	0.98 (0.78)	0.97 (0.76) ^c	0.98 (0.79) ^d	0.87 (0.74)	0.98 (0.70) ^c	−0.011 (−0.051 to 0.029)	0.591
Clinical utility	0.96 (0.65) [‡]	0.92 (0.82)	0.84 (0.66)	0.91 (0.74)	0.98 (0.66)	1.06 (0.66) ^c	−0.013 (−0.046 to 0.021)	0.458
Duration of CFI, min	26.44 (10.40) ^e	22.23 (9.64) ^f	22.87 (9.38) ^d	22.16 (8.77) ^c	23.42 (9.57) ^c	22.28 (8.39)	−1.017 (−1.616 to −0.418)	0.001**
Duration of full diagnostic interview, min	62.70 (27.41) ^g	54.26 (25.95) ^h	53.67 (23.58) ⁱ	48.21 (21.49) ^j	47.92 (22.55)	50.43 (28.61) ^d	−1.609 (−2.708 to −0.510)	0.004**
CFI proportion (%) of total diagnostic interview	47.49 (21.95) ^g	47.62 (22.47) ^h	48.91 (22.72) ⁱ	51.67 (21.62) ^j	54.07 (17.69) ^c	51.94 (18.61) ^d	0.046 (−0.753 to 0.845)	0.910

a. Mixed-effect model comparisons control for clinicians seeing multiple patients and multiple clinicians within a site.
b. Combines the sixth administration or greater into one group. Sixth interview $n = 18$ individuals; seventh $n = 9$, eighth $n = 5$, ninth $n = 4$ and tenth $n = 3$.
c. Data unavailable for one participant.
d. Data unavailable for two participants.
e. Data unavailable for six participants.
f. Data unavailable for four participants.
g. Data unavailable for ten participants.
h. Data unavailable for five participants.
i. Data unavailable for three participants.
^{*} $P < 0.05$; ^{**} $P < 0.01$; ^{***} $P < 0.001$.
^{†‡}, Values with paired superscripts in the first-administration column differ significantly ($P < 0.05$) after adjusting for multiple comparisons, Tukey–Kramer test. No other values differed significantly within administrations.

Regarding acceptability, clinicians praised the CFI's ability to generate empathy but found some questions difficult to administer (for example on the clinician–patient relationship). Patients liked the flow and person-centeredness of the CFI questions ('I felt like I was talking to someone I knew'), although some became upset by the life content elicited. The views on CFI utility were the most positive. Generally, both groups of participants found the CFI useful with respect to diagnosis, treatment planning and understanding the patient's situation, including the role of culture in mental illness (for example 'will help me get better treatment;' 'will help me understand the patient's problem extensively on the basis of cultural, religious things').

Discussion

Main findings

The DSM-5 Cultural Formulation Interview field trial was the first international study to examine clinician and patient perceptions of the feasibility, acceptability and clinical utility of a cultural assessment interview designed for use in routine clinical practice in diverse cross-national settings. The international trial included 318 patients and 75 clinicians over 11 sites in six countries. Mixed-methods analyses showed that both patients and clinicians found the CFI to be feasible, acceptable and clinically useful and these findings supported its inclusion in DSM-5. The diversity of the samples and sites – and the fact that both closed-ended and open-ended assessments yielded similar results when analysed masked to one another – enhance the clarity, robustness and generalisability of our findings.

The strategy for our quantitative analysis was developed at one of the study sites in India and used here with minor modifications.²⁷ Site-specific analyses of the field trial data have also found positive perceptions of implementation-related outcomes.^{19,24,27} In the full sample, patients assessed the CFI more positively than clinicians, and the difference was significant for feasibility. Clinicians were more concerned about feasibility than about acceptability or utility. The qualitative data, based on post-CFI open-ended interviews, likewise showed greater clinician concern about feasibility, compared with patient views and other clinician-rated outcomes.

To be successfully implemented, a new assessment should address the concerns of all stakeholders;³³ our design enabled us to examine views of both clinicians and patients. Differing views of feasibility among stakeholders probably reflect practical concerns and limited time of busy clinicians,³⁴ relevant for effective allocation of health system resources that must balance clinical values and practical constraints.³⁵ Although stakeholders' perceived acceptability and utility of an assessment or intervention may conceivably differ,^{20,36} we found no significant differences in our field trial.

Our mixed-methods design identified barriers to implementation of the CFI field trial version. DIC single-item analysis and qualitative data largely converged. They also confirm a previously published subanalysis of New York-site qualitative data, which had identified lack of differentiation of the CFI from routine clinical assessments, question clarity and ordering, and the time required for the interview as main concerns.¹⁹ The consistency of these concerns in our cross-national analysis is striking, given the cultural and clinical diversity among study participants. Many of these issues were addressed in the revised version of the CFI published in DSM-5. Based on the field trial results, the revision clarified confusing wording, improved the flow of questions and distinguished the intent of the CFI from other aspects of clinical management. Four questions were condensed into two, and one question on cultural identity and three on the views of the

patient's social network were added. Future research should examine the impact of implementing the CFI on clinical practice and outcomes, and in cultural competence training.

The practice effect identified from self-report and interview-duration data has important implications for questions about feasibility. Findings suggest that 2 h of training followed by experience administering a few interviews may be sufficient to address clinicians' concerns about feasible use of the instrument, even in a diverse sample of provider disciplines and of cultural competence experience across sites.²⁵ Consideration of the practice effect may facilitate uptake of the CFI, mindful that implementing any new tool may initially evoke resistance,²⁷ which may lessen over time if its relative advantage becomes clear in routine practice.³⁷ Indeed, by the second CFI administration, clinician feasibility scores increased substantially and no longer differed significantly from clinician acceptability and utility scores. Duration of the CFI interview, an objective indicator of feasibility, showed a similar practice effect, decreasing by 4 min by the second administration and remaining stable thereafter.

Duration of the full diagnostic interview also decreased significantly albeit more gradually. By the last administration, the duration of the full intake assessment, including 22 min for the CFI, was 50 minutes. This is comparable to the time required for an initial assessment in many mental health settings. In the USA, for example, average duration of community-based psychiatric visits (initial and follow-up combined) was 32–38 min in 1989–2006;^{38–40} intakes are often 45–50 min. Our study found substantial international variation in intake duration. Some of this variation may derive not only from resource constraints – few clinicians for many patients – but also from clinic characteristics. The sites with the longest intakes (Canada and The Netherlands) included specialised programmes for immigrants and refugees, whereas most other sites operated in general community clinics. Sites also differed significantly in the proportion of total interview time devoted to the CFI, yet all were able to integrate the CFI into routine intake procedures. The proportion of the interview devoted to the CFI increased slightly with experience, suggesting clinicians continued to find it useful and that the information it yielded was relevant to other aspects of the diagnostic interview, inasmuch as less time was required for the overall interview as a practice effect.

Limitations

This study has several limitations. Participating clinics were recruited purposively and may pay higher-than-average attention to cultural issues; clinicians who were most interested may have done more interviews, potentially confounding the positive practice effect. However, clinicians' interest did not prevent them from stating their concerns candidly in the qualitative interviews. Second, we developed our own self-report measures of service outcomes because at the time of the field trial there were no psychometrically validated quantitative measures of implementation-related outcomes.⁴¹ The DIP feasibility and acceptability domains of assessment had psychometric limitations. One-time use of these assessments is consistent with the DSM-5 field trial goal of testing proposed diagnostic criteria (or tools such as the CFI) for inclusion or revision in the final manual.²² The congruence of the qualitative and quantitative results as a benefit of the mixed-methods design supports the robustness of the DIP data. Third, the study interview consisted of the CFI session followed by the routine diagnostic assessment. All clinicians were asked to inform patients when they transitioned from the CFI to the routine assessment. It is possible that some patients did not distinguish the CFI component of their

evaluation from the routine diagnostic component when responding to questions in their debriefing interviews.

Implications

Despite these limitations, the DSM-5 international field trial results support the feasibility, acceptability and clinical utility of the CFI. The positive valuation by patients and clinicians suggests that it is worth investing about 20 min of an initial evaluation on a cultural assessment that holds promise for enhancing clinical communication, diagnostic accuracy, effective treatment planning, patient satisfaction, engagement and clinical response.^{19,21} The promise of such benefits argues for further study of CFI implementation effects on clinical and service outcomes (such as cost and sustainability).²⁰ As a practical matter, the field trial suggests an attractive learning curve, with clear benefits after 2 h of training and a single interview. A 2014 *Lancet* commission on culture and health advocated use of the CFI in all medical subspecialties, not just psychiatry,⁴² highlighting its broad relevance. Although further studies of implementation outcomes are needed, our findings indicate good prospects for meeting these acknowledged needs.

Roberto Lewis-Fernández, MD, MTS, Department of Psychiatry, Columbia University and New York State Center of Excellence for Cultural Competence, Anxiety Disorders Clinic, and Hispanic Treatment Program, at the New York State Psychiatric Institute, New York, New York, USA; **Neil Krishan Aggarwal**, MD, MBA, MA, Department of Psychiatry, Columbia University and New York State Center of Excellence for Cultural Competence at the New York State Psychiatric Institute, New York, New York, USA; **Peter C. Lam**, MPH, New York State Center of Excellence for Cultural Competence at the New York State Psychiatric Institute, New York, New York, USA; **Hanga Galfalvy**, PhD, Departments of Psychiatry and Biostatistics, Columbia University and Division of Biostatistics, New York State Psychiatric Institute, New York, New York, USA; **Mitchell G. Weiss**, MD, PhD, Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute and University of Basel, Basel Switzerland; **Laurence J. Kirmayer**, MD, Division of Social & Transcultural Psychiatry, McGill University and Culture & Mental Health Research Unit, Institute of Community & Family Psychiatry, Jewish General Hospital, Montreal, Quebec, Canada; **Vasudeo Paralikar**, MBBS, MD, PhD, Psychiatry Unit, King Edward Memorial Hospital, Pune, India; **Smita N. Deshpande**, MD, DPM, Department of Psychiatry, Post-Graduate Institute of Medical Education and Research, Dr. Ram Manohar Lohia Hospital, New Delhi, India; **Esperanza Díaz**, MD, Department of Psychiatry, Yale University, New Haven, Connecticut, USA; **Andel V. Nicasio**, MSEd, New York State Center of Excellence for Cultural Competence at the New York State Psychiatric Institute, New York, New York, USA; **Marit Boiler**, MPH, New York State Center of Excellence for Cultural Competence at the New York State Psychiatric Institute, New York, New York, USA; **Renato D. Alarcón**, MD, MPH, Department of Psychiatry, Mayo Clinic College of Medicine, Rochester, Minnesota, USA and Universidad Peruana Cayetano Heredia, Lima, Peru; **Hans Rohloff**, MD, Centrum '45, Oegstgeest, the Netherlands; **Simon Groen**, MA, Department of Anthropology, University of Amsterdam, Amsterdam, The Netherlands and De Evenaar Centre for Transcultural Psychiatry, GGZ Drenthe Mental Health Care, Assen, The Netherlands; **Rob C. J. van Dijk**, MSc, Parnassia Psychiatric Institute, The Hague, The Netherlands; **Sushrut Jadhav**, MBBS, MD, MRCPsych, PhD, Division of Psychiatry, University College London, London, UK; **Sanjeev Sarmukaddam**, PhD, Department of Biostatistics, Maharashtra Institute of Mental Health, Sassoon Govt. Hospital Campus, Pune, India; **David Ndetee**, MBChB, DPM, MRCPsych, FRCPsych, MD, DSc, Department of Psychiatry, University of Nairobi and Africa Mental Health Foundation, Nairobi, Kenya; **Monica Z. Scalco**, MD, PhD, Department of Psychiatry, University of Toronto and Toronto Western Hospital, Toronto, Ontario, Canada; **Kavoos Bassiri**, MS, LMFT, LPCC, CGP, Richmond Area Multi-Services Inc. and Department of Psychiatry, University of California, San Francisco, San Francisco, California, USA; **Sergio Aguilar-Gaxiola**, MD, PhD, Department of Internal Medicine and Center for Reducing Health Disparities, University of California, Davis, Sacramento, California, USA; **Hendry Ton**, MD, MS, Department of Psychiatry and Behavioral Sciences and Center for Reducing Health Disparities, University of California, Davis and Asian Pacific Community Counseling-Transcultural Wellness Center, Sacramento, California, USA; **Joseph Westermeyer**, MD, PhD, Department of Psychiatry, University of Minnesota and Minneapolis VA Medical Center, Minneapolis, Minnesota, USA; **Johann M. Vega-Dienstmaier**, MD, Department of Psychiatry, Universidad Peruana Cayetano Heredia and Hospital Nacional Cayetano Heredia, Lima, Peru

Correspondence: Roberto Lewis-Fernández, MD, Department of Psychiatry, Columbia University and New York State Psychiatric Institute, 1051 Riverside Drive, Unit 69, New York, NY 10032, USA. Email: rlewis@nyspi.columbia.edu

First received 29 Apr 2016, final revision 27 Sep 2016, accepted 28 Sep 2016

Funding

This research was supported by the American Psychiatric Association, the New York State Office of Mental Health (N.K.A., P.C.L., H.G., A.V.N. and M.B.) and institutional funds from the New York State Psychiatric Institute (R.L.-F.). The Pune site received support from

the KEMH Research Centre and the Kenya site was partially funded by the Africa Mental Health Foundation. The funders did not have any input into the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review or approval of the manuscript.

Acknowledgements

The authors are grateful for the assistance of Oanh Meyer, Victoria Mutiso, Lincoln Khasakhala, Anne Mwayo, N. N. Mishra, Triptish Bhatia, Antonio Lozano, Luis Fiestas, Adelguisa Mormontoy, Martín Arévalo, Spencer Case, Seung-Hee Hong, Samantha Díaz, Ravi DeSilva, Venkat Bhat, Kwame McKenzie, Lauren Olsen, Ladson Hinton, Devon E. Hinton, Sophie Bäärnhielm, James Boehlein, Cécile Rousseau, Jaswant Guzder, Darrel A. Regier, David J. Kupfer, William Narrow, Diana Clarke, Jennifer Shupinka and Francis Lu.

References

- Adeponle AB, Thombs BD, Groleau D, Jarvis E, Kirmayer LJ. Using the cultural formulation to resolve uncertainty in diagnoses of psychosis among ethnoculturally diverse patients. *Psychiatric Serv* 2012; **63**: 147–53.
- Bhui K, Bhugra D. Explanatory models for mental distress: implications for clinical practice and research. *Br J Psychiatry* 2002; **181**: 6–7.
- Ayonrinde O. Importance of cultural sensitivity in therapeutic transactions: considerations for healthcare providers. *Dis Manage Health Outcomes* 2003; **11**: 233–48.
- Schouten B, Meeuwesen L. Cultural differences in medical communication: a review of the literature. *Patient Educ Couns* 2006; **64**: 21–34.
- Lewis-Fernández R. Cultural formulation of psychiatric diagnosis. *Cult Med Psychiatry* 1996; **20**: 133–44.
- Mezzich J. Cultural formulation and comprehensive diagnosis: clinical and research perspectives. *Psychiatr Clin North Am* 1995; **18**: 649–87.
- Aggarwal NK, Rohrbaugh RM. Teaching cultural competency through an experiential seminar on anthropology and psychiatry. *Acad Psychiatry* 2011; **35**: 331–2.
- Scarpinati Rosso M, Bäärnhielm S. The cultural formulation: a model to combine nosology and patients' life context in psychiatric diagnostic practice. *Transcult Psychiatry* 2009; **46**: 406–28.
- Hansen H, Dugan TM, Becker AE, Lewis-Fernández R, Lu FG, Oquendo MA, et al. Educating psychiatry residents about cultural aspects of care: a qualitative study of approaches used by US expert faculty. *Acad Psychiatry* 2013; **37**: 412–6.
- Rohloff H, Knipscheer J, Kleber R. Use of the cultural formulation with refugees. *Transcult Psychiatry* 2009; **46**: 487–505.
- Groen S, Richters A, Laban K, Deville W. Implementation of the Cultural Formulation through a newly developed Brief Cultural Interview: pilot data from the Netherlands. *Transcult Psychiatry* (in press).
- Lewis-Fernández R. The cultural formulation. *Transcult Psychiatry* 2009; **46**: 379–82.
- Caballero-Martínez L. DSM-IV-TR Cultural formulation of psychiatric cases: two proposals for clinicians. *Transcult Psychiatry* 2009; **46**: 506–23.
- Aggarwal NK. Adapting the cultural formulation for clinical assessments in forensic psychiatry. *J Am Acad Psychiatry Law* 2012; **40**: 113–8.
- Alarcón RD. Culture, cultural factors and psychiatric diagnosis: review and projections. *World Psychiatry* 2009; **8**: 131–9.
- Weiss M. Cultural epidemiology: an introduction and overview. *Anthropol Med* 2001; **8**: 5–29.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*, 4th edn. American Psychiatric Association, 1994.
- Lewis-Fernández R, Aggarwal N, Bäärnhielm S, Rohloff H, Kirmayer LJ, Weiss MG, et al. Culture and psychiatric evaluation: operationalizing cultural formulation for DSM-5. *Psychiatry* 2014; **77**: 130–54.
- Aggarwal NK, Nicasio AV, DeSilva R, Boiler M, Lewis-Fernández R. Barriers to implementing the DSM-5 cultural formulation interview: a qualitative study. *Cult Med Psychiatry* 2013; **37**: 505–33.
- Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health* 2011; **38**: 65–76.
- Lewis-Fernández R, Aggarwal N, Hinton L, Hinton D, Kirmayer. *The DSM-5 Handbook on the Cultural Formulation Interview*. American Psychiatric Publishing, 2016.
- Clarke DE, Narrow WE, Regier DA, Kuramoto SJ, Kupfer DJ, Kuhl EA, et al. DSM-5 field trials in the United States and Canada, Part I: Study design, sampling strategy, implementation, and analytic approaches. *Am J Psychiatry* 2013; **170**: 43–58.

- 23 Regier D, Narrow W, Clarke D, Kraemer HC, Kuramoto SJ, Kuhl EA, et al. DSM-5 field trials in the United States and Canada, Part II: test-retest reliability of the selected categorical diagnoses. *Am J Psychiatry* 2013; **170**: 59–70.
- 24 Hinton L, Aggarwal N, Losif A-M, Weiss M, Paralikar V, Deshpande S, et al. Perspectives of family members participating in cultural assessment of psychiatric disorders: findings from the DSM-5 international field trial. *Int Rev Psychiatry* 2015; **27**: 3–10.
- 25 Aggarwal NK, Lam P, Castillo EG, Weiss MG, Díaz E, Alarcón RD, et al. How do clinicians prefer cultural competence training? Findings from the DSM-5 Cultural Formulation Interview field trial. *Acad Psychiatry* 2016; **40**: 584–91.
- 26 Bravo M, Canino GJ, Rubio-Stipec M, Woodbury-Farina M. A cross-cultural adaptation of a psychiatric epidemiologic instrument: the diagnostic interview schedule's adaptation in Puerto Rico. *Cult Med Psychiatry* 1991; **15**: 1–18.
- 27 Paralikar V, Sarmukaddam S, Patil K, Nulkar A, Weiss M. Clinical value of the cultural formulation interview in Pune, India. *Indian J Psychiatry* 2015; **57**: 59–67.
- 28 Kramer CY. Extension of multiple range tests to group means with unequal numbers of replications. *Biometrics* 1956; **12**: 307–10.
- 29 Krippendorff K. *Content Analysis: An Introduction to its Methodology*. Sage Publications, 2013.
- 30 Elo S, Kyngas H. The qualitative content analysis process. *J Adv Nurs* 2008; **62**: 107–15.
- 31 Barbour RS. Checklists for improving rigour in qualitative research: a case of the tail wagging the dog? *BMJ* 2001; **322**: 1115–7.
- 32 Wu C. An examination of the wording effect in the Rosenberg Self-Esteem Scale among culturally Chinese people. *J Soc Psych* 2008; **148**: 535.
- 33 Proctor E, Landsverk J, Aarons G, Chambers D, Glisson C, Mittman B. Implementation research in mental health services: an emerging science with conceptual, methodological, and training challenges. *Adm Policy Ment Health* 2009; **36**: 24–34.
- 34 Aggarwal N. Clinical implementation of the Cultural Formulation Interview: planning and assessment. In *The DSM-5 Handbook on the Cultural Formulation Interview* (eds R Lewis-Fernández, NK Aggarwal, L Hinton, DE Hinton, LJ Kirmayer): 191–207. American Psychiatric Publishing, 2016.
- 35 Saxena S, Thornicroft G, Knapp M, Whiteford H. Resources for mental health, scarcity, inequity, and inefficiency. *Lancet* 2007; **370**: 878–89.
- 36 Fischer EP, Shumway M, Owen RR. Priorities of consumers, providers, and family members in the treatment of schizophrenia. *Psychiatr Serv* 2002; **53**: 724–9.
- 37 Greenhalgh T, Robert G, MacFarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q* 2004; **82**: 581–629.
- 38 Mechanic D, McAlpine D, Rosenthal M. Are patients' office visits with physicians getting shorter? *N Engl J Med* 2001; **344**: 198–204.
- 39 Olfson M, Cherry D, Lewis-Fernández R. Racial differences in visit duration of outpatient psychiatric visits. *Arch Gen Psychiatry* 2009; **66**: 214–21.
- 40 Olfson M, Marcus S, Pincus H. Trends in office-based psychiatric practice. *Am J Psychiatry* 1999; **156**: 451–7.
- 41 Bird VJ, Le Bouillier C, Leamy M, Williams J, Bradstreet S, Slade M. Evaluating the feasibility of complex interventions in mental health services: standardised measure and reporting guidelines. *Br J Psychiatry* 2014; **204**: 316–21.
- 42 Napier A, Ancarno C, Butler B, Calabrese J, Chater A, Chatterjee H, et al. Culture and health. *Lancet* 2014; **384**: 1607–39.



psychiatry in the movies

One Flew Over the Cuckoo's Nest: Dean Brooks

Aidan Collins

Hardly a Sunday goes by without Milos Forman's film of Ken Kesey's *One Flew Over the Cuckoo's Nest* getting a mention in the weekend supplements, most usually in the context of the application or misapplication of electroconvulsive therapy. Its influence is so great that if future reconstruction archaeologists were attempting to recreate a 20th-century psychiatric hospital, this film would likely be the keystone of the project.

For doctors, perhaps, the bits that grab us most (especially if we are middle-aged male psychiatrists) are the scenes where the film's protagonist, R. P. McMurphy, is interviewed by his psychiatrist, Dr John Spivey. Having already walked out a considerable distance on thin ice by allowing Forman to use his functioning state hospital in Oregon, the hospital medical director, Dr Dean Brooks, began stomping around testing the limits of that ice by agreeing to play the part of Spivey. Moreover, he agreed to the interviews with Jack Nicholson's McMurphy being ad-libbed. Aware that elements of the storyline were anachronistic (I'm looking at you, prefrontal leucotomy), Dr Brooks insisted that the credits would state that the film was not a true depiction of a modern (1970s) psychiatric hospital.

In late-life interviews Dean Brooks comes across as a fearlessly humorous man. In the film he was less the weak-willed Dr Spivey of Kesey's book and perhaps a little more like himself. This was a man who had island-hopped in the most terrifying way during World War II and who had advocated and effected considerable institutional reform in his own hospital. After all, what sort of weak-willed psychiatrist sits alone in his office with an apparent psychopath (who has the stare of stares) and accuses him of messing?

The British Journal of Psychiatry (2017)
210, 297. doi: 10.1192/bjp.bp.116.196634