Article

Prevalence of Major Depressive Episode in Two American Indian Reservation Populations: Unexpected Findings With a Structured Interview

Janette Beals, Ph.D.

Spero M. Manson, Ph.D.

Nancy R. Whitesell, Ph.D.

Christina M. Mitchell, Ph.D.

Douglas K. Novins, M.D.

Sylvia Simpson, M.D.

Paul Spicer, Ph.D.

American Indian Service Utilization, Psychiatric Epidemiology, Risk and Protective Factors Project Team

Objective: American Indian populations have often been considered to be at greater risk for major depressive episode than are other groups in the United States. The American Indian Service Utilization, Psychiatric Epidemiology, Risk and Protective Factors Project (AI-SUPERPFP), completed between 1997 and 1999, was designed to allow comparisons with the baseline National Comorbidity Survey (NCS), conducted in 1990–1992. The prevalence of lifetime and 12-month DSM-III-R major depressive episode was compared between the AI-SUPERPFP and NCS samples.

Method: A total of 3,084 tribal members (1,446 in a Southwest tribe [73.7% of eligible participants] and 1,638 in a Northern Plains tribe [76.8% of eligible participants]) age 15–54 years living on or near their home reservations were interviewed. An adaptation of the University of Michigan Composite International Diagnostic Interview and the NCS algorithm for diagnosis were used to estimate the prevalence of lifetime and 12-month major depressive episode in these groups.

Results: The prevalence estimates for lifetime and 12-month major depressive episode were substantially lower in the American Indian samples, compared to the NCS sample. Detailed analyses indicated differential endorsement of lifetime symptoms between the American Indian groups and the NCS participants. Furthermore, American Indians were substantially less likely than NCS participants to indicate that depressive symptoms had co-occurred during an episode lasting at least 2 weeks. The lifetime prevalence estimates based on the NCS algorithm ranged from 3.8% to 7.9% for men and women in the two tribes. The analogous rates based on an adapted AI-SUPERPFP algorithm ranged from 7.2% and 14.3%. Few tribe, age, and gender differences were found.

Conclusions: The findings underscore the need for careful examination of diagnostic instruments cross-culturally. Adaptation of the NCS algorithm for diagnosis appears necessary for estimation of the prevalence of major depressive episode in the American Indian populations included in this study. In striving to better reflect the clinical diagnostic process in epidemiological and services research, careful consideration of the resulting complexity becomes increasingly critical.

(Am J Psychiatry 2005; 162:1713–1722)

Major depressive episode is a common, chronic, and debilitating disorder. Recent estimates indicate that the lifetime prevalence of major depressive episode among adults in the United States is 16.2% and 6.6% in any given 12-month period (1). Yet, major depressive episode is not equally distributed: African Americans report lower rates of major depressive episode than do whites, especially if sociodemographic variation is taken into consideration (1, 2). The prevalence of major depressive episode among Hispanics is often reported to be comparable to that among whites but has been shown to vary by degree of acculturation (3). Furthermore, the rate of depression among Asian Americans may be as low as 3.4% (4).

To date, few studies have included sufficient numbers of American Indians to reach conclusions about the prevalence of major depressive episode in these populations. (We note that in 1977, the National Congress of American Indians and the National Tribal Chairmen's Association issued a joint resolution that the preferred term, in the absence of specific tribal designation, was American Indian rather than Native American when referring to the indigenous population of the lower 48 states.) Two efforts, both conducted with nonrandom samples of separate tribes with semistructured clinical interviews, reported lifetime rates of depressive disorder in excess of 25% (5, 6). The American Indian Service Utilization, Psychiatric Epidemiology, Risk and Protective Factors Project (AI-SUPERPFP) was designed to allow joint analyses with the baseline National Comorbidity Survey (NCS) by using DSM-III-R diagnoses generated by the University of Michigan Composite International Diagnostic Interview (CIDI) (7).

As we began AI-SUPERPFP, we hypothesized that the prevalence of major depressive episode among American Indians would be comparable to, if not higher than, that

TABLE 1. Lifetime Major Depressive Episode Symptoms in the American Indian Service Utilization, Psychiatric Epidemiology, Risk and Protective Factors Project (AI-SUPERPFP) and National Comorbidity Survey (NCS) Participants

	Male Participants								
	AI-SUPERPFP Samples								
	Southwest Tribe (N=617)			Northern Plains Tribe (N=792)			NCS Sample (N=3,847)		
Symptom	%	99% CI	Differing Groups[a]	%	99% CI	Differing Groups[a]	%	99% CI	Differing Groups[a]
Depressed mood	40.1	35.1–45.4	NM, SF, UF	24.8	20.8–29.2	SM, UM, SF, UF	43.8	41.0–46.7	NM, NF, UF
Anhedonia	43.5	38.3–48.8	NM, NF	27.5	23.3–32.1	SM, UM, SF, UF	37.2	34.4–40.0	NM, SF
Significant weight loss or weight gain	26.3	21.8–31.3	UF	17.8	14.4–21.8	UM, SF, UF	29.6	27.0–32.3	NM, UF
Insomnia or hypersomnia	28.5	23.9–33.6	NM, UF	19.2	15.6–23.3	SM, UM, SF, UF	32.7	30.1–35.5	NM, UF
Psychomotor agitation or retardation	14.9	11.5–19.2	UF	10.6	7.9–14.0	UM, UF	18.8	16.7–21.2	NM, NF
Fatigue or loss of energy	17.4	13.7–21.9	UM, SF, UF	11.3	8.5–14.8	UM, SF, NF, UF	25.8	23.4–28.4	SM, NM, NF, UF
Feelings of worthlessness or guilt	15.4	12.0–19.7	UF	10.1	7.5–13.6	UM, UF	17.4	15.3–19.7	NM, UF
Concentration problems	20.7	16.7–25.4	UF	13.4	10.4–17.1	UM, SF, UF	25.3	22.8–27.9	NM, NF, UF
Recurrent thoughts of death or suicide	22.5	18.3–27.3	UM, SF, UF	21.3	17.5–25.6	UM, SF, UF	30.9	28.3–33.6	SM, NM, UF

[a] Groups with nonoverlapping 99% confidence intervals ($p<0.01$), compared to the group listed in the column heading (SM=Southwest tribe male participants, NM=Northern Plains tribe male participants, UM=NCS [U.S.] male participants, SF=Southwest tribe female participants, NF=Northern Plains tribe female participants, UF=NCS [U.S.] female participants). Confidence intervals were constructed in Stata by using logit transformation to ensure that endpoints of the intervals were in the valid range for proportions (0 to 1).

reported in the NCS. Preliminary analyses using the NCS algorithms for diagnosis indicated that lifetime rates of major depressive episode among American Indians were only about 30% of those reported in the NCS. This startling finding led us to carefully review the component parts of the CIDI's operationalization of major depressive episode. We hypothesized that cultural factors would be at least partly responsible for the patterns found. Another likely factor in these differences was whether the major depressive episode symptoms co-occurred with one another. Our ethnographic reviews of the CIDI before fielding suggested that the co-occurrence questions were particularly problematic in these American Indian communities.

Method

AI-SUPERPFP Samples

The AI-SUPERPFP populations of inference were 15–54-year-old enrolled members of two closely related Northern Plains tribes and a Southwest tribe. To protect the confidentiality of the participating communities (8), we refer to them by these general descriptors rather than specific tribal names. The communities in question belong to different linguistic families, have different histories of migration, subscribe to different principles for reckoning kinship and residence, and historically pursued different forms of subsistence. Yet both tribes share experiences common to many American Indian groups. They have similar histories of colonization, including dramatic military resistance, externally imposed forms of governance, forced dietary changes, mandatory boarding school education, and active missionary movements. Thus, selection of these two tribes provided an opportunity to simultaneously examine both similarities and differences across tribes within the American Indian population, which is relatively small yet extremely diverse.

The AI-SUPERPFP methods are described in greater detail elsewhere (9). Briefly, tribal rolls formed the sampling universe; these records list all individuals meeting the legal requirements for recognition as tribal members. Stratified random sampling procedures were used, and strata were defined by cultural group, gender, and age (15–24-, 25–34-, 35–44-, and 45–54-year age groups, which also were used in NCS). Records were selected randomly for inclusion into replicates, which were then released as needed to reach our goal of approximately 1,500 interviews per tribe. The samples were restricted to those currently living on or near (within 20 miles) of their home reservations; in the Southwest and Northern Plains, 46.6% and 39.2%, respectively, of those listed in the tribal rolls met this residence requirement. Of those located and found eligible, 73.7% in the Southwest tribe (N=1,446) and 76.8% in the Northern Plains tribe (N=1,638) agreed to participate, with response rates slightly lower for male tribal members and younger tribal members. Sample weights were used in all inferential analyses presented here to account for differential selection probabilities across all strata and for patterns of nonresponse.

Tribal approvals were obtained before project initiation. Informed consent was obtained from all adult respondents; for minors, parental/guardian consent was obtained before we requested the adolescent's assent. Interviews were computer-assisted and were administered by tribal members with intensive training in research and interviewing methods. Extensive quality control procedures were used to verify that all interviews were conducted in a standardized, reliable manner. These procedures included confirming the location of 10% of those deemed ineligible because they lived away from the reservations, verification of the status of 10% of the refusals, and reviews of more than 10% of the audiotaped interviews to ensure that questions were read verbatim with appropriate tempo and that the interviewer established suitable rapport with the participant. Both the interview instrument and the training manual are available on our web site (http://www.uchsc.edu/ai/ncaianmhr/research/superpfp.htm).

NCS Sample

Our comparison to the general U.S. population utilized the baseline NCS, described in detail elsewhere (10). The baseline NCS was conducted in a stratified, multistage area probability sample of 8,098 U.S. residents age 15–54 years in 1990–1992.

Assessment of Major Depressive Episode

Mental health disorders were assessed in AI-SUPERPFP with items drawn from the CIDI and adapted for use in American Indian communities (9). As explained further later in this article, we focused chiefly on criterion A for the major depressive episode diagnosis. Mirroring DSM-III-R, this criterion requires endorsement of either depressed mood or anhedonia *and* the experience of at least five of nine depressive symptoms (which include depressed mood and anhedonia). Further, five criterion A symptoms must have co-occurred with one another for at least 2 weeks within the context of at least one episode of depression.

Female Participants								
AI-SUPERPFP Samples								
Southwest Tribe (N=829)			Northern Plains Tribe (N=848)			NCS Sample (N=4,251)		
%	99% CI	Differing Groups[a]	%	99% CI	Differing Groups[a]	%	99% CI	Differing Groups[a]
50.4	45.8–55.0	SM, NM, NF	33.5	29.2–38.1	UM, SF, UF	55.8	53.1–58.5	SM, NM, UM, NF
45.8	41.3–50.4	NM, UM, NF	32.6	28.3–37.2	SM, SF, UF	41.3	38.6–44.0	NM, NF
33.3	29.1–37.8	NM, UF	24.9	21.0–29.2	UF	42.5	39.8–45.3	SM, NM, UM, SF, NF
30.5	26.4–35.0	NM, UF	25.7	21.7–30.1	UF	39.6	37.0–42.3	SM, NM, UM, SF, NF
16.0	12.9–19.7	UF	10.8	8.2–14.1	UM, UF	23.2	20.9–25.6	SM, NM, SF, NF
26.2	22.3–30.4	SM, NM, UF	18.6	15.2–22.6	NM, UM, UF	37.2	34.6–39.9	SM, NM, UM, SF, NF
16.0	12.9–19.7	UF	12.8	9.9–16.4	UF	23.6	21.3–26.0	SM, NM, UM, SF, NF
22.1	18.6–26.2	NM, UF	15.7	12.5–19.6	UM, UF	32.6	30.0–35.2	SM, NM, UM, SF, NF
32.5	28.3–37.0	SM, NM	27.0	23.0–31.4	UF	39.4	36.7–42.1	SM, NM, UM, NF

Major depressive episode criteria B through D are exclusionary. Specifically, criterion B states that the symptoms are not initiated or maintained by physiological factors or medical conditions (operationalized in the CIDI as not being caused entirely by medication, drugs, alcohol, physical injury, or illness), nor are they attributable to a normal reaction to the death of a loved one. Criterion C excludes those with persistent hallucinations or delusions in the absence of mood symptoms, and criterion D excludes those with schizophrenia or related psychotic disorders. Criteria C and D were not assessed in AI-SUPERPFP and are not part of the major depressive episode rates reported here.

Generally, the AI-SUPERPFP CIDI's major depressive episode module was identical to that of the original CIDI. However, as explained elsewhere (9), we carefully reviewed the CIDI for cultural validity in the context of a previous study. Several concerns that were raised were general to the overall instrument; others were specific to the major depressive episode module. Our strategy for modification of the CIDI was conservative, and, wherever possible, we retained the original wording. In some cases, we provided a definition of problematic words (e.g., "period" as a time that has a definite beginning and end). In other cases, we altered wording; for instance, rather than asking, "Has there ever been a period of 2 weeks or more when you thought a lot about death—either your own, someone else's, or death in general?" we deconstructed this question into three separate questions and considered a response of "yes" to any of the questions to be equivalent to an affirmative response to the CIDI version of the question. This modification was especially critical in interviewing tribal members in the Southwest, where thinking about one's own or another's death is proscribed but where we still wanted people to have the opportunity to indicate they may have been thinking about death in general.

Focus group members indicated that the question about co-occurrence of depressive symptoms was especially problematic, with members noting the complexity of the wording and demands of the task of recalling whether symptoms had occurred together within a 2-week period. In particular, focus group members suggested that community members would find it difficult to answer the question that followed the endorsement of individual major depressive episode symptoms: "Now take your time to carefully review the list of problems (endorsed), and tell me the numbers of all the problems you had during the period(s) of [key phrase for depression]." A companion question that was used to ask about physiological factors raised similar concerns. Because of the importance of co-occurrence of symptoms to the definition of major depressive episode (and our interest in preserving comparability with the NCS), we chose not to change the wording of either of these questions. However, we remained aware of their possible difficulties for study participants.

AI-SUPERPFP Clinical Reappraisal

More than 10% (N=335) of the AI-SUPERPFP participants were reinterviewed by psychiatrists or clinical psychologists using the Structured Clinical Interview for DSM-III-R—Non-Patient Version (SCID-NP) (11). Approximately three-quarters of the clinical reappraisal group was chosen on the basis of a positive CIDI diagnosis of major depressive episode, posttraumatic stress disorder, or alcohol abuse/dependence. The remaining 25% did not qualify for any diagnosis assessed in AI-SUPERPFP, but most endorsed significant levels of depressed, anxious, or irritable symptoms on a scale independent of the CIDI. The eight clinician interviewers had extensive clinical experience (more than 15 years on average), at least some in American Indian communities. Each demonstrated a high level of interrater reliability (kappa ≥0.80) with a series of videotapes coded by an expert panel and also performed supervised interviews before entering the field. Furthermore, all clinical reappraisals were audiotaped and reviewed by master clinicians for quality assurance purposes. The response rate for the clinical reappraisal substudy was 72.3%; an average of 120 days elapsed between lay and clinical interviews. The clinicians were blind to the participants' CIDI diagnostic status.

Analyses

All inferential analyses were conducted with Stata's "svy" procedures (12) with sample and nonresponse weights (13). Table 1 and Table 2 present combined analyses of the AI-SUPERPFP and NCS data sets with the appropriate weights used for each data set. In Table 3 the unweighted concordance of two variations of the AI-SUPERPFP CIDI major depressive episode diagnosis with the SCID-NP is shown. The first variation replicates the NCS operationalization of the diagnosis; the second is the AI-SUPERPFP operationalization (explained in detail later). Table 4 presents lifetime and 12-month prevalence estimates for major depressive episode in the AI-SUPERPFP samples according to both the NCS and AI-SUPERPFP operationalizations. Nonoverlapping 99% confidence intervals highlight specific tribe and gender differences.

TABLE 2. Participants Who Met Specific Criteria for Major Depressive Episode in the American Indian Service Utilization, Psychiatric Epidemiology, Risk and Protective Factors Project (AI-SUPERPFP) and the National Comorbidity Survey (NCS)

	Male Participants								
	AI-SUPERPFP Samples								
	Southwest Tribe (N=617)			Northern Plains Tribe (N=792)			NCS Sample (N=3,847)		
Criterion	%	99% CI	Differing Groups[a]	%	99% CI	Differing Groups[a]	%	99% CI	Differing Groups[a]
Criterion A items									
Depressed mood or anhedonia (% of sample)	53.2	47.9–58.4	NM, NF, UF	34.3	29.8–39.1	SM, UM, SF, UF	52.6	49.7–55.5	NM, NF, UF
Five or more symptoms, of which at least one was depressed mood or anhedonia (% of sample)	23.5	19.2–28.3	UF	15.4	12.3–19.3	UM, SF, UF	28.6	26.1–31.3	NM, NF, UF
Criterion A: Five or more symptoms during same episode, of which at least one was depressed mood or anhedonia (% of sample)	6.5	4.3–9.7	UM, UF	5.4	3.6–7.9	UM, UF	17.3	15.3–19.6	SM, NM, SF, NF, UF
Criterion B items[b]									
Disturbance was not due to physiological factors or medical conditions (% of those meeting criterion A)	85.5	61.2–95.7		89.9	66.7–97.5		92.3	88.4–94.9	
Disturbance was not uncomplicated bereavement (% of those meeting criterion A)	90.7	70.0–97.6		88.4	70.6–96.0		95.5	92.1–97.5	
Criterion B: % meeting criterion A and not excluded because disturbance was due to physiological factors or was uncomplicated bereavement	76.2	52.4–90.3		78.2	56.3–90.9		87.8	83.1–91.3	

[a] Groups with nonoverlapping 99% confidence intervals ($p<0.01$), compared to the group listed in the column heading (SM=Southwest tribe male participants, NM=Northern Plains tribe male participants, UM=NCS [U.S.] male participants, SF=Southwest tribe female participants, NF=Northern Plains tribe female participants, UF=NCS [U.S.] female participants). Confidence intervals were constructed in Stata by using logit transformation to ensure that endpoints of the intervals were in the valid range for proportions (0 to 1).
[b] Criterion B was relevant for 37 Southwest tribe male participants, 45 Northern Plains male participants, 708 NCS male participants, 75 Southwest tribe female participants, 62 Northern Plains tribe female participants, and 1,143 NCS female participants.

TABLE 3. Concordance Between Two Variations of the Lifetime Major Depressive Episode Diagnostic Algorithm in the American Indian Service Utilization, Psychiatric Epidemiology, Risk and Protective Factors Project (AI-SUPERPFP) Composite International Diagnostic Interview (CIDI) and the Structured Clinical Interview for DSM-III-R—Non-Patient Version (SCID-NP) in 335 AI-SUPERPFP Participants[a]

Variation of CIDI	Number of True Positives (Both SCID-NP and CIDI Positive)	Number of False Negatives (SCID-NP Positive, CIDI Negative)	Number of False Positives (SCID-NP Negative, CIDI Positive)	Number of True Negatives (Both SCID-NP and CIDI Positive)	Kappa[b]	Sensitivity
NCS algorithm[f]	46	90	17	180	0.27	34%
AI-SUPERPFP algorithm[g]	61	75	29	168	0.32	45%

[a] A subgroup of AI-SUPERPFP participants (N=335) previously interviewed by lay interviewers using the CIDI were reinterviewed by psychiatrists or clinical psychologists using the SCID-NP. Statistical tests between algorithms were inappropriate because the two samples were not independent. Similar conclusions were reached whether or not the data were weighted back to the sample interviewed by lay interviewers; for ease of presentation, the unweighted results are reported.
[b] Cohen's kappa.
[c] Calculated using the Bayes's rule (using the SCID-NP prevalence weighted back to the full population).
[d] McNemar's chi-square test.

Results

Symptom Endorsement

Table 1 provides a detailed account of the endorsements of the criterion A lifetime symptoms for major depressive episode among the Southwest, Northern Plains, and NCS samples. Both Northern Plains men and women were less likely to endorse each of the major depressive episode symptoms than were their counterparts of the same gender in the NCS. Southwest men were less likely than men in the NCS to endorse two of the nine lifetime major depressive episode symptoms, while Southwest women were less likely than women in the NCS to endorse six of the nine symptoms. It is also noteworthy that the Southwest partic-

Female Participants								
AI-SUPERPFP Samples								
Southwest Tribe (N=829)			Northern Plains Tribe (N=848)			NCS Sample (N=4,251)		
%	99% CI	Differing Groups[a]	%	99% CI	Differing Groups[a]	%	99% CI	Differing Groups[a]
59.7	55.1–64.1	NM, NF	42.0	37.4–46.7	SM, UM, SF, UF	61.2	58.5–63.8	SM, NM, UM, NF
28.8	24.7–33.1	NM, UF	21.6	17.9–25.7	UM, UF	38.3	35.6–41.0	SM, NM, UM, SF, NF
9.5	7.1–12.6	UM, UF	7.6	5.4–10.5	UM, UF	25.8	23.5–28.3	SM, NM, UM, SF, NF
94.8	81.6–98.7		92.7	77.2–98.0		95.4	92.9–97.0	
91.8	79.6–97.0		86.4	70.3–94.5		96.1	93.9–97.6	
86.5	72.3–94.1		81.7	65.0–91.5		91.8	88.7–94.0	

Specificity	Positive Predictive Value[c]	Negative Predictive Value[c]	Bias[d]	Mean Global Assessment Scale Score[e]
91%	92%	32%	74.6	65.0
85%	90%	34%	56.0	64.4

[e] For participants classified as having major depressive episode according to the given algorithm.
[f] DSM-III-R definition of major depressive episode, including the requirement for co-occurrence of depressive symptoms during a 2-week period.
[g] DSM-III-R definition of major depressive episode, but no requirement for co-occurrence of depressive symptoms during a 2-week period.

ipants were more likely to endorse depressed mood or anhedonia than were their Northern Plains counterparts.

Meeting the Diagnostic Criteria

Criterion A comprises three components: at least one of the symptoms endorsed is depressed mood or anhedonia, at least five of the nine symptoms are endorsed, and five of these symptoms must co-occur within the same 2-week period. As the first row of Table 2 shows, although more than 50% of the Southwest and NCS samples endorsed depressed mood or anhedonia, fewer Northern Plains tribal members endorsed either of these symptoms. Respondents who did not endorse at least one of these symptoms were not asked the remaining questions about symptoms of major depressive episode; the patterns for the Northern Plains participants shown in Table 1 reflect, in part, that many in this group were not asked about the remaining symptoms.

The second row of Table 2 presents the second component of criterion A: the percentage of respondents in each sample who endorsed at least five of the nine possible symptoms (of which at least one is depressed mood or anhedonia). Women in the NCS were most likely to report this number of symptoms; Northern Plains men, Northern Plains women, and Southwest women were less likely than their NCS counterparts to do so. As shown in the third row, the application of the third component of criterion A—the co-occurrence of five symptoms—led to further variation across samples and yielded rates for criterion A for the

TABLE 4. Lifetime and 12-Month Prevalence of DSM-III-R Major Depressive Episode According to the National Comorbidity Survey (NCS) and American Indian Service Utilization, Psychiatric Epidemiology, Risk and Protective Factors Project (AI-SUPERPFP) Algorithms Among AI-SUPERPFP Participants, by Age and Gender

	Participants With Diagnosis According to NCS Algorithm[a]							
	Male Participants				Female Participants			
	Southwest Tribe (N=606)		Northern Plains Tribe (N=778)		Southwest Tribe (N=817)		Northern Plains Tribe (N=840)	
Diagnosis and Age Groups	%	99% CI	%	99% CI	%	99% CI	%	99% CI
Lifetime major depressive episode								
Total	4.5	2.7–7.4	3.8	2.4–6.0	7.9	5.8–10.9	5.6	3.8–8.3
Age groups (years)								
15–24	4.0	1.4–10.7	1.7	0.5–6.4	9.0	5.0–15.9	3.2	1.3–8.0
25–34	4.5	1.6–12.1	4.3	1.9–9.5	7.4	3.7–14.5	6.2	3.0–12.2
35–44	5.1	1.9–12.8	3.9	1.6–9.4	8.3	4.5–14.9	6.2	2.8–13.3
≥45	4.5	1.7–11.4	5.1	2.0–12.1	6.8	3.6–12.6	7.4	3.6–14.6
12-month major depressive episode								
Total	2.8	1.5–5.2	2.1	1.1–3.8	4.9	3.3–7.5	4.0	2.5–6.4
Age groups (years)								
15–24	1.3	0.2–7.8	1.7	0.5–6.4	5.3	2.4–11.4	2.4	0.8–6.9
25–34	3.0	0.8–10.2	1.6	0.5–5.0	5.3	2.3–11.9	4.4	1.8–10.5
35–44	4.4	1.5–11.8	2.5	0.8–7.3	5.1	2.3–11.1	4.6	1.8–11.4
≥45	2.4	0.6–8.5	2.5	0.7–9.0	3.8	1.6–8.8	4.9	2.0–11.7

[a] DSM-III-R definition of major depressive episode, including the requirement for co-occurrence of depressive symptoms during a 2-week period.
[b] DSM-III-R definition of major depressive episode, but no requirement for co-occurrence of depressive symptoms during a 2-week period.
[c] Groups with nonoverlapping 99% confidence intervals ($p<0.01$), compared to the group listed in the column heading (NM=Northern Plains tribe male participants, SF=Southwest tribe female participants). Differing groups were found for only the Northern Plains tribe male participants and the Southwest tribe female participants with a diagnosis according to the AI-SUPERPFP algorithm. Confidence intervals were constructed in Stata by using logit transformation to ensure that endpoints of the intervals were in the valid range for proportions (0 to 1).

American Indian samples that were approximately one-third of those for the NCS samples.

The final three rows in Table 2 show the effect of the criterion B exclusions for individuals who met criterion A. However, the diminished AI-SUPERPFP sample sizes caused by the low rate of endorsement of criterion A decreased power to identify significant differences here and rendered the effect of criterion B, including physiological factors, relatively minor.

Figure 1 graphically displays the differential effect of the components of the criterion A operationalization on major depressive episode diagnosis. For both the Southwest and NCS samples, many participants endorsed either depressed mood or anhedonia; the comparable rate for the Northern Plains was substantially lower. Approximately 50% of those endorsing depressed mood or anhedonia in the Southwest and Northern Plains also met the requirement for five or more major depressive episode symptoms, compared with almost 60% in the NCS sample. The co-occurrence requirement is where the largest discrepancies were noted. Less than one-third of the American Indians who endorsed five or more symptoms indicated that at least five co-occurred; the parallel percentage in the NCS sample was 64%.

As mentioned earlier, the focus groups that reviewed the CIDI identified the co-occurrence question as highly problematic, and the results described here bear out these qualitative findings. This convergence of ethnographic and epidemiologic findings led to consideration of an alternative algorithmic definition for major depressive episode for the AI-SUPERPFP—one that did not include the problematic formats involved in the co-occurrence question. The AI-SUPERPFP algorithm for the diagnosis simply required at least five major depressive episode symptoms—one of which was either depressed mood or anhedonia. We tested how this alternative definition fared, compared to the clinical reappraisal diagnoses. We focused on testing the following hypothesis: The NCS algorithm, which was more stringent and followed the DSM definition of major depressive episode more closely, would show superior agreement with the clinical reappraisal diagnosis (made with the SCID-NP), compared with the less stringent AI-SUPERPFP algorithm.

Validity Analyses

As Table 3 shows, the kappa statistic of chance-corrected agreement with the SCID-NP diagnosis was lower, not higher, for the more stringent NCS operationalization than for the AI-SUPERPFP operationalization. Furthermore, the increase in sensitivity for the AI-SUPERPFP version was greater than was the decrease in specificity, compared to the NCS operationalization. Bias, as measured by McNemar's test (14), was lower with the AI-SUPERPFP version of the major depressive episode algorithm. Also, Global Assessment Scale (15) scores were comparable across the major depressive episode algorithms; thus, participants classified as having major depressive episode under the less stringent AI-SUPERPFP algorithm were rated by the clinicians as having levels of impairment similar to those reported by participants classified as having major depressive episode with the NCS algorithm. In summary,

Participants With Diagnosis According to AI-SUPERPFP Algorithm[b]									
Male Participants					Female Participants				
Southwest Tribe (N=607)		Northern Plains Tribe (N=777)			Southwest Tribe (N=817)			Northern Plains Tribe (N=840)	
%	99% CI	%	99% CI	Differing Groups[c]	%	99% CI	Differing Groups[c]	%	99% CI
9.8	7.0–13.5	7.2	5.1–10.1	SF	14.3	11.4–17.9	NM	10.3	7.8–13.6
10.3	5.5–18.5	3.9	1.6–9.0	SF	15.0	9.5–22.7	NM	5.9	3.0–11.4
11.8	6.4–20.9	7.3	3.8–13.7		14.1	8.6–22.3		11.7	7.0–19.1
9.5	4.7–18.3	7.9	4.0–14.8		14.2	8.9–21.8		11.6	6.4–19.9
7.1	3.4–14.4	9.8	5.2–17.7		14.1	9.0–21.3		12.7	7.5–20.6
6.6	4.4–9.9	3.9	2.5–6.0	SF	8.8	6.5–11.9	NM	7.6	5.4–10.6
7.1	3.3–14.6	3.5	1.4–8.6		9.6	5.4–16.6		3.8	1.6–8.6
8.1	3.8–16.5	3.0	1.1–7.5		9.6	5.2–17.0		9.6	5.3–16.9
7.3	3.2–15.5	3.6	1.4–9.1		8.1	4.3–14.8		8.8	4.5–16.7
3.5	1.2–9.8	6.2	2.8–13.1		7.9	4.3–13.8		8.0	4.0–15.3

the hypothesis that the NCS version would prove superior to the AI-SUPERPFP algorithm was not supported.

Although the results are not reported in detail, we explored several additional hypotheses. Joint analyses of data for dysthymic disorder and minor depression indicated that the American Indian participants were not more likely to qualify for these depressive disorder diagnoses than were the NCS participants. Lifetime rates of dysthymic disorder were 7.0%, 3.7%, and 6.4% for the Southwest, Northern Plains, and NCS samples, respectively; the respective lifetime rates of minor depression were 7.8%, 5.3%, and 13.0%. Furthermore, tribe, age, gender, poverty status, marital status, ethnic identity, and stigma of mental illness did not differentiate individuals who qualified for a diagnosis with the AI-SUPERPFP algorithm from those who qualified with the NCS algorithm. Thus, for the AI-SUPERPFP samples, relying on simple symptom counts for defining major depressive episode appeared preferable to requiring co-occurrence of symptoms—especially as elicited by CIDI questions.

Prevalence of Major Depressive Episode in AI-SUPERPFP

Table 4 presents the prevalence rates of major depressive episode in the AI-SUPERPFP samples calculated with both the NCS algorithm and the AI-SUPERPFP adaptation that foregoes the co-occurrence requirement. Generally, the rates using the NCS algorithm were approximately one-half those using the AI-SUPERPFP algorithm. Few tribe, gender, or age differences were found.

A question arising from these analyses is whether the NCS prevalence figures would be similarly inflated if they were based on the AI-SUPERPFP major depressive episode algorithm. Figure 2 demonstrates that the relative increases were greater in the AI-SUPERPFP samples than in the NCS sample, further supporting the argument that the changes made in the AI-SUPERPFP major depressive episode algorithm had a differential effect across the two studies.

Discussion

The current study informs ongoing debates about cultural influences on experiences of disorder and the instruments that are increasingly used to assess such categories cross-culturally. Before embarking on AI-SUPERPFP, we predicted that the rates of major depressive episode in the American Indian populations in this study would be comparable to, if not higher than, those for the general U.S. population. At the same time, however, we were mindful of evidence of ethnic differences in major depressive episode rates (1–4) and of extensive ethnographic work suggesting that the social construction of depressive experiences might be different in at least some American Indian communities than in the general population (16).

Using almost identical case-ascertainment methods and the same diagnostic algorithms that were used in the NCS, we found that rates of major depressive episode for the American Indian groups were about 30% of the national estimates. The detailed analyses presented here suggest that an interface of cultural and methodological

FIGURE 1. Percentage of Southwest Tribe, Northern Plains Tribe, and National Comorbidity Survey (NCS) Samples Meeting Components of Criterion A for a DSM-III-R Major Depressive Episode Diagnosis

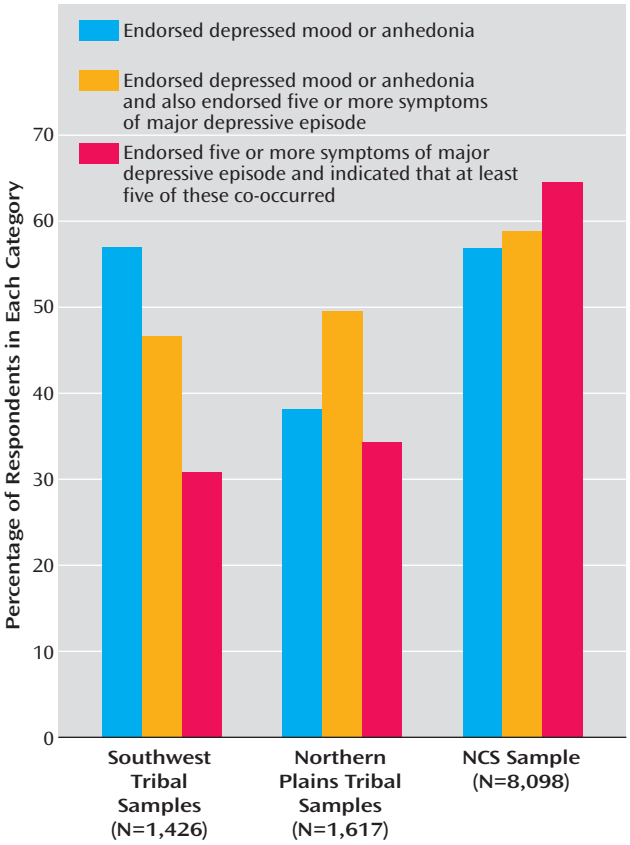
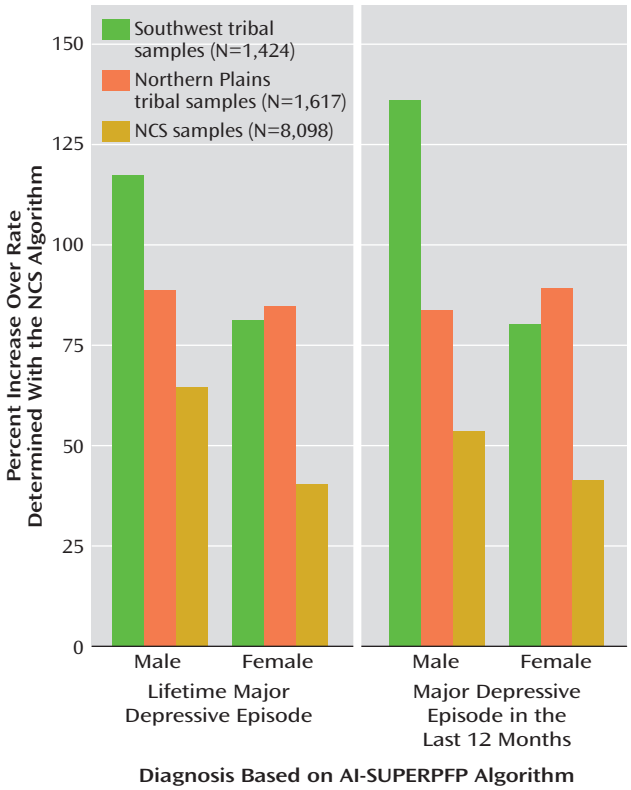


FIGURE 2. Relative Increase in Prevalence of Lifetime and 12-Month DSM-III-R Major Depressive Episode Based on the American Indian Service Utilization, Psychiatric Epidemiology, Risk and Protective Factors Project (AI-SUPERPFP) Algorithm, Compared With Rates Based on the National Comorbidity Survey (NCS) Algorithm, in Southwest Tribe, Northern Plains Tribe, and NCS Samples[a]



[a] The NCS algorithm used the DSM-III-R definition of major depressive episode, including requirement for co-occurrence of depressive symptoms during a 2-week period. The AI-SUPERPFP algorithm used the DSM-III-R definition of major depressive episode, without the requirement for co-occurrence of depressive symptoms during a 2-week period.

factors is responsible for these unexpected findings. As Figure 1 shows, members of the Northern Plains tribe were less likely to endorse the core major depressive episode symptoms of depressed mood or anhedonia. Historically, in this culture, admission of such symptoms may be relatively proscribed as signs of weakness (17). Given the structure of the major depressive episode diagnostic module in the CIDI, however, failure to endorse these symptoms precludes diagnosis. In the Southwest tribe, depressed mood or anhedonia was endorsed at levels similar to those in the NCS sample, but the rates of endorsement of the other major depressive episode symptoms were lower, suggesting a differential social construction of major depressive episode or a disparate patterning of depressive symptoms in this culture, compared with others in the United States or in the Northern Plains tribe.

Although tribal differences may account for the dissimilar symptom endorsements, both American Indian samples struggled with the interview question about co-occurrence of symptoms. Only about one-third of American Indian participants who endorsed five or more symptoms indicated at least five co-occurring symptoms, compared to more than 60% in the NCS sample. The mean number of symptoms endorsed differed by less than 1 across samples and is thus unlikely to account fully for this difference. In the ethnographic review of the CIDI, extensive discussion was devoted to understanding why participants found these questions difficult. In particular, the necessary conceptualization of time frames was problematic—both to comprehend and to fit with individuals' experiences. In fact, as explained earlier, because of the foreign nature of the task, we were asked to define the word "period" in the interview. During training and data collection, the co-occurrence question arguably received the greatest attention of the more than 3,000 items in the full interview. We were repeatedly advised that both interviewers and participants had difficulty understanding the required task. Indeed, our concern about administrative difficulties was so great that we completed a reinterview of a large subsample of both tribes with the major depressive episode module to assure ourselves that these questions were being administered appropriately and that rates did not change dramatically between administrations.

Memory recall, event boundaries, and perceived affinities among stimuli are key topics in cognitive psychology. The difficulty of AI-SUPERPFP participants in employing time as a frame of reference for judgments about the co-occurrence of depressive symptoms may illustrate differences in perceiving, remembering, and communicating structure in events (18), a process likely rooted in cultural traditions of learning (19). In particular, we were told that what we were labeling depression was not necessarily experienced as episodic; thus, questions that asked specifically whether each depressive symptom was present at a given time were very difficult to address. Responding to such questions is clearly a complex cognitive task that may be made more difficult because the diagnostic interview is structured with the assumption that depressive episodes and temporality are similarly constructed cross-culturally. We note that the more recent NCS replication (1) simplified the co-occurrence question and did not ask for a symptom-by-symptom accounting. At the same time, other aspects of the new NCS interview are clearly more complicated than previous versions; for instance, detailed judgments about the duration of some symptoms are required. We argue that at a minimum, the complex wording of diagnostic interview questions may present significant difficulties cross-culturally. More critically, however, these findings suggest that the cross-cultural validity of the measurement of major depressive episode merits further investigation.

These data and analyses have limitations, many of which have been discussed at length elsewhere (9). Briefly, the AI-SUPERPFP samples, while well defined and well justified, were limited. The restricted age range mirrored that of NCS. The limited tribal representation and the focus on only those tribal members currently resident on or near their home reservations were decisions driven by the need to balance cultural variability with feasibility (9). Joint analyses of data sets such as AI-SUPERPFP and NCS are limited methodologically. The data collection periods necessarily varied, as did the methods—at least to a certain extent. The NCS data were collected in 1990–1992; the AI-SUPERPFP data, between 1997 and 1999. Furthermore, the response rates for AI-SUPERPFP were somewhat lower than those of NCS: 75% combined across tribes, compared to 83% for NCS (10). However, the AI-SUPERPFP rates are similar to those of the Epidemiologic Catchment Area study—between 68% and 79% by site (20)—and the 73% response rate reported for the recent NCS replication (1).

Another limitation is the low concordance rates between the AI-SUPERPFP CIDI and the SCID-NP, although it should be noted that the kappa values in our study approximate those in other studies that have used this generation of instruments (21). More important here, however, was that our results did not support our hypothesis that the clinicians' diagnoses, which were less dependent on the use of specific question wording, would demonstrate greater concordance with the more stringent NCS operationalization.

In summary, the pattern of these findings is reasonable given existing knowledge about the cultural factors that influenced this study. We note that even with the more liberal AI-SUPERPFP operationalization, the rates of major depressive episode for the American Indian samples remained significantly lower than those reported for the NCS sample (22). Our results underscore the importance of integrating qualitative and quantitative methods for understanding depressive and other disorders within the purview of the DSM. As new versions of the CIDI (1) are modified to more closely reflect the clinical diagnostic process, the instrument is likely to become increasingly complicated and to rely more heavily on intricate time frames and language. Such changes are likely to increase cultural variability in endorsement patterns and in the subsequent estimates of mental illness.

Received Dec. 17, 2003; revision received Aug. 4, 2004; accepted Sept. 13, 2004. From the American Indian and Alaska Native Programs, University of Colorado at Denver and Health Sciences Center, Aurora; and the Department of Psychiatry, University of Colorado at Denver and Health Sciences Center, Denver. Address correspondence and reprint requests to Dr. Beals, American Indian and Alaska Native Programs, University of Colorado at Denver and Health Sciences Center, MS F800, P.O. Box 6508, Aurora, CO 80045-0508; jan.beals@uchsc.edu (e-mail).

Supported by NIH grants R01 MH-48174 (Dr. Manson and Dr. Beals, principal investigators), P01 MH-42473 (Dr. Manson, principal investigator), R01 DA-14817 (Dr. Beals, principal investigator), and R01 AA-13420 (Dr. Beals, principal investigator). National Comorbidity Survey (NCS) data were made available by the International Consortium of Political and Social Research; NCS was supported by the following grants with R. Kessler as principal investigator: NIH grants R01 MH/DA-46376 and R01 MH-49098 and grant 90135190 from the William T. Grant Foundation.

In addition to the authors named at the beginning of the article, the American Indian Service Utilization, Psychiatric Epidemiology, Risk and Protective Factors Project (AI-SUPERPFP) team includes Cecelia K. Big Crow, Dedra Buchwald, Buck Chambers, Michelle L. Christensen, Denise A. Dillard, Karen DuBray, Paula A. Espinoza, Candace M. Fleming, Ann Wilson Frederick, Joseph Gone, Diana Gurley, Lori L. Jervis, Shirlene M. Jim, Carol E. Kaufman, Ellen M. Keane, Suzell A. Klein, Denise Lee, Monica C. McNulty, Denise L. Middlebrook, Laurie A. Moore, Tilda D. Nez, Ilena M. Norton, Theresa O'Nell, Heather D. Orton, Carlette J. Randall, Angela Sam, James H. Shore, and Lorette L. Yazzie.

The authors thank interviewers and computer/data management and administrative staff members Anna E. Barón, Antonita Begay, Amelia T. Begay, Cathy A.E. Bell, Phyllis Brewer, Nelson Chee, Mary Cook, Helen J. Curley, Mary C. Davenport, Rhonda Wiegman Dick, Marvine D. Douville, Pearl Dull Knife, Geneva Emhoolah, Fay Flame, Roslyn Green, Billie K. Greene, Jack Herman, Tamara Holmes, Shelly Hubing, Cameron R. Joe, Louise F. Joe, Cheryl L. Martin, Jeff Miller, Robert H. Moran Jr., Natalie K. Murphy, Melissa Nixon, Ralph L. Roanhorse, Margo Schwab, Jennifer Settlemire, Donna M. Shangreaux, Matilda J. Shorty, Selena S.S. Simmons, Wileen Smith, Tina Standing Soldier, Jennifer Truel, Lori Trullinger, Arnold Tsinajinnie, Jennifer M. Warren, Intriga Wounded Head, Theresa (Dawn) Wright, Jenny J. Yazzie, and Sheila A. Young; Methods Advisory Group members Margarita Alegria, Evelyn J. Bromet, Dedra Buchwald, Peter Guarnaccia, Steven G. Heeringa, Ronald Kessler, R. Jay Turner, and William A. Vega; and the tribal members who so generously answered all the questions asked of them.

References

1. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, Rush AJ, Walters EE, Wang PS: The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA 2003; 289:3095–3105
2. US Department of Health and Human Services: Mental Health: Culture, Race, and Ethnicity, Supplement to Mental Health: A Report of the Surgeon General. Rockville, Md, Public Health Service, 2001
3. Vega WA, Kolody B, Aguilar-Gaxiola S, Alderete E, Catalano R, Caraveo-Anduaga J: Lifetime prevalence of DSM-III-R psychiatric disorders among urban and rural Mexican Americans in California. Arch Gen Psychiatry 1998; 55:771–788
4. Zheng YP, Lin KM, Takeuchi D, Kurasaki KS, Wang Y, Cheung F: An epidemiological study of neurasthenia in Chinese-Americans in Los Angeles. Compr Psychiatry 1997; 38:249–259
5. Robin RW, Chester B, Rasmussen JK, Jaranson JM, Goldman D: Prevalence, characteristics, and impact of childhood sexual abuse in a Southwestern American Indian tribe. Child Abuse Negl 1997; 21:769–787
6. Kinzie JD, Leung PK, Boehnlein J, Matsunaga D, Johnston R, Manson SM, Shore JH, Heinz J, Williams M: Psychiatric epidemiology of an Indian village: a 19-year replication study. J Nerv Ment Dis 1992; 180:33–39
7. Kessler RC, Wittchen H-U, Abelson JM, McGonagle KA, Schwarz N, Kendler KS, Knauper B, Zhao S: Methodological studies of the Composite International Diagnostic Interview (CIDI) in the US National Comorbidity Survey. Int J Methods Psychiatr Res 1998; 7:33–35
8. Norton IM, Manson SM: Research in American Indian and Alaska Native communities: navigating the cultural universe of values and process. J Consult Clin Psychol 1996; 64:856–860
9. Beals J, Manson SM, Mitchell CM, Spicer P, the AI-SUPERPFP Team: Cultural specificity and comparison in psychiatric epidemiology: walking the tightrope in American Indian research. Cult Med Psychiatry 2003; 27:259–289
10. Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, Wittchen H-U, Kendler KS: Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: results from the National Comorbidity Survey. Arch Gen Psychiatry 1994; 51:8–19
11. Spitzer RL, Williams JBW, Gibbon M, First MB: Structured Clinical Interview for DSM-III-R—Non-Patient Version (SCID-NP). New York, New York State Psychiatric Institute, Biometrics Research, 1987
12. Stata Reference Manual: Release 8.0. College Station, Tex, Stata Corp, 2003
13. Cochran WG: Sampling Techniques, 3rd ed. New York, John Wiley & Sons, 1977
14. Helzer JE, Robins LN, McEvoy LT, Spitznagel EL, Stolztman RK, Farmer A, Brockington IF: A comparison of clinical and Diagnostic Interview Schedule diagnoses: physician reexamination of lay-interviewed cases in the general population. Arch Gen Psychiatry 1985; 42:657–666
15. Endicott J, Spitzer RL, Fleiss JL, Cohen J: The Global Assessment Scale: a procedure for measuring overall severity of psychiatric disturbance. Arch Gen Psychiatry 1976; 33:766–771
16. O'Nell TD: Disciplined Hearts: History, Identity, and Depression in an American Indian Community. Berkeley, University of California Press, 1996
17. O'Nell TD: Coming home among Northern Plains Vietnam veterans: psychological transformations in pragmatic perspective. Ethos 2000; 27:441–465
18. Zacks JM, Tversky B, Iyer G: Perceiving, remembering, and communicating structure in events. J Exp Psychol Gen 2001; 130:29–58
19. Saxe GB: Cognition, development, and cultural practices. New Dir Child Adolesc Dev 1999; 83:19–36
20. Robins LN, Regier DA (eds): Psychiatric Disorders in America: The Epidemiologic Catchment Area Study. New York, Free Press, 1991
21. Brugha TS, Jenkins R, Taub N, Meltzer H, Bebbington PE: A general population comparison of the Composite International Diagnostic Interview (CIDI) and the Schedules for Clinical Assessment in Neuropsychiatry (SCAN). Psychol Med 2001; 31:1001–1013
22. Beals J, Novins DK, Whitesell NR, Spicer P, Mitchell CM, Manson SM, the AI-SUPERPFP Team: Prevalence of mental disorders and utilization of mental health services in two American Indian reservation populations: mental health disparities in a national context. Am J Psychiatry 2005; 162:1723–1732