Addiction (1999) 94(9), 1361–1370

A validity study of the SSAGA—a comparison with the SCAN

MICHIE HESSELBROCK,[1] CAROLINE EASTON,[2] KATHLEEN K. BUCHOLZ,[3] MARC SCHUCKIT[5] & VICTOR HESSELBROCK[4]

[1]*University of Connecticut, School of Social Work,* [2]*Yale University, School of Medicine, Substance Abuse Treatment Unit,* [3]*Washington University, School of Medicine, Department of Psychiatry,* [4]*University of Connecticut, School of Medicine, Department of Psychiatry,* [5]*University of California at San Diego, Department of Psychiatry, USA*

Abstract

Objective. *This study examined the concurrent diagnostic validity of the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA) across alcohol and drug dependencies, major depression, anxiety disorders and ASPD. The Schedule for Clinical Assessment in Neuropsychiatry (SCAN) was selected as the comparison instrument because it arises from a different tradition and uses a different format for its administration. The SCAN has been shown to be valid and applicable cross-culturally.* **Method**. *Subjects included 38 men and 42 women volunteers from another study and from an outpatient psychiatry clinic. Selected sections of both the SSAGA and the SCAN interviews were administered to all subjects, approximately 1 week apart, in a randomized order. Because the SCAN does not assess Antisocial Personality Disorder (ASPD), the ASPD section of the Structured Clinical Interview for DSM-III-R (SCID) was substituted for this comparison.* **Results**. *The Kappa statistic was used to measure concordance between the two instruments. Kappa for alcohol dependence was in the acceptable range (0.63). Kappas were lower for sedative dependence (0.48) and for cannabis dependence (0.53), but higher for cocaine and stimulant dependence (0.85) and for opioid dependence (0.73). Kappa for major depression and the ASPD diagnoses were high (0.71 and 0.70), but slightly lower agreement was found for panic disorder (0.62). Kappa for social phobia was 0.47.* **Conclusion**. *These data, combined with results from two previous studies which examined reliability, indicate that the SSAGA is a highly reliable and valid instrument for use in studies of a variety of psychiatric disorders, including alcohol and drug dependence.*

Introduction

Accurate assessment of psychiatric disorders is well recognized as crucial for clinicians and researchers in order to predict clinical course and to provide the most appropriate treatment (Goodwin & Guze, 1996). While health-care providers may apply standardized diagnostic criteria to their patients in a less structured and informal manner, the need to obtain standardized data from a large number of subjects requires the use of valid and reliable instruments to assess psychiatric disorders. Structured interview schedules have the advantage of providing objectivity, validity and reliability for obtaining and interpreting relevant diagnostic information. Structured psychiatric interview schedules have been extensively used for both clinical and large scale epidemiological studies (Helzer, 1981), and

Correspondence to: Michie N. Hesselbrock PhD, School of Social Work, University of Connecticut 1798 Asylum Avenue, West Hartford, Connecticut 06117, USA.

Submitted 18th September 1998; initial review completed 27th November 1998; final version accepted 1st April 1999.

ISSN 0965–2140 print/ISSN 1360-0443 online/99/091361–10

Carfax Publishing, Taylor & Francis Limited

have demonstrated high agreement and consistency with other interview schedules (cf. Hesselbrock, 1982) as well as against clinical interviews (Helzer *et al.*, 1985; Wells *et al.*, 1995). While semi-structured psychiatric interview schedules approximate more closely conditions likely to be observed in clinical situations, the administration of the instrument typically requires considerable clinical skill and judgement. On the other hand, highly structured interview schedules offer the advantage of being cost-effective tools for use by trained non-clinical interviewers.

The need to assess alcohol problems accurately, including alcohol dependence, is a particularly important issue for genetic studies. Diagnostic imprecision and errors greatly reduce the power of genetic analyses (Maziade *et al.*, 1992) and may contribute to the lack of replication that has plagued psychiatric genetic studies (Ciaranello & Ciaranello, 1991). Further, different diagnostic definitions of alcoholism and related phenotypes appear to be differentially heritable (van den Bree *et al.*, 1998). With these issues in mind, family studies of psychiatric disorders, including alcohol dependence, should consider using a diagnostic instrument that can be economically administered to a large number of subjects, will would work equally well in assessing the psychiatric status of affected as well as unaffected family members, will provide a diagnosis in relation to several different formal diagnostic systems, will permit the phenotyping of the disorder of interest in relation to clinical features that are not necessarily included in diagnostic systems and will be useful in making a differential diagnosis. A survey of existing psychiatric interview schedules at the time the Collaborative Study on the Genetics of Alcoholism (COGA) was being developed failed to identify an instrument that had all these attributes. Consequently, the decision was taken to develop an new psychiatric interview schedule that would meet the needs of a large scale family study of the genetics of alcoholism.

The Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA) is a comprehensive psychiatric interview schedule designed by the Collaborative Study of the Genetics of Alcoholism (COGA) to assess the physical, psychosocial and psychiatric manifestations of alcohol abuse and dependence and related psychiatric disorders in adults. As described in more detail elsewhere (Bucholz *et al.*, 1994), the SSAGA is a highly structured diagnostic instrument, designed for use by well-trained lay people. It provides a detailed examination of alcohol and other drug use, a comprehensive assessment of their consequences and an assessment of comorbid psychiatric disorders. The base diagnostic system of the SSAGA is DSM-III-R. Since many diagnostic items in the SSAGA are based, in part, on other psychiatric assessments including the DIS (Robins *et al.*, 1981), HELPER (Coryell *et al.*, 1978), CIDI-SAM (Robins *et al.*, 1986; Cottler *et al.*, 1989), SADS (Endicott & Spitzer, 1978) and SCID (Spitzer *et al.*, 1992), it also provides poly-diagnostic data (e.g. ICD-10, DSM-IV) that can be compared with findings of other epidemiological studies using these instruments. To date, the SSAGA has been translated into seven different languages and has been used in more than 50 different studies of alcoholism.

The reliability of the SSAGA has been assessed in relation to both rater test–retest (Bucholz *et al.*, 1994) as well as a comparison of raters across COGA Centers (Bucholz *et al.*, 1995). Test–retest reliabilities for life-time DSM-III-R alcohol and other drug dependency diagnoses as well as major depressive disorder and the antisocial personality disorder were high, with agreement ranging from kappa = 0.70–0.90. The cross-center agreement was also acceptable for alcohol and other drug dependencies, with kappa ranging from 0.57 to 1.00, except for stimulant dependence (K = 0.44). Further, a test of the reliabilities of the individual criterion items for substance abuse dependencies in the SSAGA also revealed a high level of inter-rater consistency (Bucholz *et al.*, 1995).

While the reliability of the SSAGA has been established by those previous studies, the diagnostic ability of the SSAGA has not been compared to other diagnostic instruments. The ideal protocol for such a comparison would utilize administration of both the SSAGA and a comparison instrument by highly trained research interviewers to the same research subjects. The appropriate comparison instrument should be reliable, widely used in research and able to evaluate subjects utilizing a different format from the SSAGA. The Schedule for Clinical Assessment in Neuropsychiatry (SCAN) (Wing *et al.*, 1990) was selected as a comparison instrument as it meets many of these criteria. The first draft of the SCAN was completed in 1983 by a task force established jointly by the World Health

Organization (WHO) and the Alcohol, Drug Abuse and Mental Health Administration (ADAMHA). The instrument was designed to assess psychiatric disorders across many languages and cultures. The SCAN was intended for use only by mental health professionals trained to diagnose a range of psychiatric disorders. Compared to the SSAGA, the SCAN allows clinicians considerable flexibility for probing subject responses and requires clinical judgement when rating symptom severity according to the rating scales provided. Most SCAN items are rated on a standard three- or four-point scale, ranging from absent to mild to severe. Clinical judgements are made according to intensity (intrusiveness and extent of interference with mental functioning) and the frequency of the symptom. Many items in the SCAN were adopted from the PSE-9 and CIDI. Further, the SCAN permits the assessment of both current and lifetime psychiatric conditions according to the DSM-III-R, DSM-IV or ICD-10 diagnostic systems. A study of SCAN overall inter-rater reliabilities for life-time DSM-III-R alcohol and drug dependencies showed high agreement, with kappas ranging from 0.74 to 0.99. Kappa coefficients for life-time ICD-10 diagnoses of alcohol dependence were also high, ranging from 0.74 to 0.98 (Easton *et al.*, 1997). The diagnostic concordance between the SCAN and CIDI-SAM for alcohol and drug dependence also appears to be good, with kappas ranging from 0.49 for opiate dependence and 0.69 for alcohol dependence (Compton *et al.*, 1996). The selection of the SCAN (version 2.0) for the comparison was also based on the differences in the structure and administration of the two instruments. The SSAGA is highly structured and is administered by trained lay people using prescripted probes, while the SCAN is less structured and is to be administered only by trained mental health professionals. The format of the SCAN allows clinicians flexibility for probing and making clinical judgements. Further, all responses in the SSAGA are probed and coded in relation to scoring guidelines, while clinicians using the SCAN are required to use clinical judgement to determine symptom severity and whether to include a symptom for making a diagnosis. The use of clinical judgement in the SCAN is guided by specific probe questions, with a glossary that contains guidelines for coding item criteria. These differences in format and administration of the SCAN argued for its selection as the comparison instrument for examination of the diagnostic validity of the SSAGA.

Methods and procedures

Subjects for the present study were men and women who were 18 + years old with no obvious cognitive deficits. Volunteers were recruited from several different sources, including the COGA sample, inpatient and outpatient psychiatry services and people screened for, but not accepted into, the COGA project due to their family structure. The sampling scheme was designed to ensure that a sufficient number of subjects were examined within each diagnostic category rather than to provide a sample representative of any particular population. Quotas for each diagnostic category were established before data collection began. All volunteers were screened to provide a sample that included 10–12 subjects with no history of psychiatric disorders including alcoholism and at least 40 subjects with a history of varying degrees of involvement with alcohol. Alcoholic individuals were also selected on the basis of their co-existing psychiatric disorders including depressive disorder, anxiety disorders and ASPD. The individuals participating in this study were selected in order to provide a balance of symptom severity within different diagnostic categories for this validity test and, thus, do not reflect a typical clinical sample. All interviewers were blind to the screening information. Subjects were screened initially for admission to the study using the individual assessment module (IAM) of the Family History Assessment Module (FHAM; Rice *et al.*, 1995). Subjects were then selected for participation based upon their type and severity of symptoms. No attempt was made to diagnose subjects during screening. All potential subjects were screened by research technicians not responsible for conducting either the SCAN or SSAGA interviews.

The sample of 38 men and 42 women included 62.5% Caucasians, 27.5% African Americans, 5% Hispanic individuals with the remainder being either Asian or Native American (5%). The average age was 34.1 ± 10.7) (range from 18 to 62 years). Twenty-eight per cent of subjects were married; 48% were single and never married; the remaining 25% were divorced or separated. The majority of subjects

had a high school education or higher (70%). Slightly less than half of the subjects were working part time (9%) or full time (38%), while 54% of subjects reported being unemployed. Nearly half the subjects reported a current annual household gross income less than $20 000, 25% of subjects earned $20 000–50 000/year, and the remaining subjects reported earning more than $50 000/year.

Selected sections of both the SSAGA and the SCAN interviews were administered to all 80 subjects, approximately 1 week apart, in a randomized order as to which interview was given first. The sections included alcohol dependence and abuse, other substance dependence and abuse, major depression and anxiety disorders from both the SSAGA and the SCAN. Because the SCAN does not assess Antisocial Personality Disorder (ASPD), the ASPD section of the SCID was substituted for this comparison. All Axis I diagnoses include substance-induced symptoms as the SCAN interview does not screen for alcohol and drug-induced individual symptoms. Other Axis I diagnoses were not assessed due to their typically low prevalence in the general population.

Interviewers were blind to the subject's diagnoses and recruitment source. Further, the interviewers were trained to conduct the interviews according to prescribed procedures. The University of Connecticut is a WHO-sanctioned SCAN training site. The SCAN interviews were conducted by three people trained initially at the WHO SCAN training site at Washington University (St Louis) and had been certified as SCAN trainers. The SSAGA was administered by three experienced COGA interviewers who had completed the standardized training program developed by COGA. All interviews were audio-recorded to ensure that the interviewers followed standard procedures and for review if a scoring or transcription error was suspected. The previously developed DSM-III-R diagnostic scoring algorithms for each instrument were used to score the interviews after careful editing and checking of the interview data by the protocol monitor.

The SSAGA tended to identify higher lifetime rates of DSM-III-R alcohol and drug dependence, while similar rates of Axis I psychiatric disorders were identified by both instruments. The rate of alcohol dependence identified by the SSAGA was 46.2% compared to 32.5% by the SCAN, 25.0% vs. 18.8% for cannabis dependence, 32.5 vs. 27.5 for cocaine, 13.8% vs. 12.5% for opiates, 5.0% vs. 3.8% for stimulant dependence and 6.3% vs. 3.8% for sedative dependence. Affective and anxiety disorders were also examined. The rates of major depressive disorder was 50% by the SSAGA and 46% by the SCAN, panic disorder 15% vs. 11.3% and social phobia 5.0% by both instruments. Similarly, the rate of the ASPD identified by the SSAGA was 16.3% and 13.8% by the SCID.

The sensitivity, specificity and predictive values of the SSAGA compared to the SCAN are presented in Table 1. High sensitivity values were found for alcohol and substance dependence, ranging from 66.7% (sedative dependence) to 100% (stimulant dependence). Specificity ranged from 71.9% (alcohol dependence) to 98.7% (stimulant dependence). Positive predictive values ranged from 40.0% for sedative dependence to 84% for cocaine dependence. Negative predictive values were high, ranging from 93.3% (cannabis) to 100% for stimulants.

Similarly, high sensitivity, specificity, positive predictive values and negative predictive values were found for the diagnoses of major depressive disorder and antisocial personality disorder. Panic disorder had a low positive predictive value (58.3%), while the diagnosis of social phobia had both low sensitivity and low positive predictive values.

Since sensitivity, specificity, positive predictive values and negative predictive values tend to be influenced by the prevalence rates of a disorder in the population studied, 95% confidence intervals were computed. Kappa provides a measure of exact agreement (a special case of correlation) that is adjusted for chance occurrence; it was used to compare the concordance of diagnoses between the two interview schedules (Table 2). The first kappa statistic compared the SSAGA diagnoses and SCAN diagnoses only for symptoms identified by the SCAN interviewers that met clinical severity (i.e. scored as 2 or 3). The kappa statistics for alcohol dependence and all other drugs were found to range from acceptable to high, except for cannabis dependence and sedative dependence. Kappa for alcohol dependence was 0.63, while kappas for cocaine dependence, opioid dependence and stimulant dependence were somewhat higher (0.85, 0.73 and 0.85, respectively).

Table 1. *Sensitivity, specificity, positive predictive value, and negative predictive value of SSAGA compared to SCAN*

	SCAN	2 × 2 tables SSAGA		Sensitivity	Specificity	Positive predictive value	Negative predictive value
Alcohol & drug dependence		+ (n)	− (n)				
	+	24	2	%	%	%	%
Alcohol	−	13	41	92.3	71.9	64.9	95.3
	+	21	1				
Cocaine	−	4	54	95.4	93.1	84.0	98.2
	+	8	2				
Opiates	−	3	67	80.0	95.7	72.7	97.1
	+	11	4				
Cannabis	−	9	56	73.3	86.1	55.0	93.3
	+	3	0				
Stimulant	−	1	76	100.0	98.7	75.0	100
	+	2	1				
Sedative	−	3	74	66.7	96.1	40.0	98.7
Other Axis I disorders							
	+	30	4				
Major Depression	−	7	33	88.2	82.5	81.1	89.2
	+	7	2				
Panic Disorder	−	5	66	77.8	92.9	58.3	97.1
	+	2	2				
Social Phobia	−	2	74	50.0	97.4	50.0	97.4
Axis II disorder							
	+	9	2				
ASPD SCID	−	9	65	81.8	94.2	69.2	97.0

Table 2. *Diagnostic comparison of substance dependence between SSAGA and SCAN: kappa statistics*

	Only symptoms that met severity criteria on SCAN		All positive symptoms on SCAN
Diagnosis	Kappa	95% C.I.	Kappa
Alcohol dependence	0.63	0.46–0.80	0.49
Cocaine dependence	0.85	0.72–0.98	0.94
Opioid dependence	0.73	0.50–0.95	0.71
Cannabis dependence	0.53	0.30–0.75	0.71
Stimulant dependence	0.85	0.56–1.0	0.85
Sedative dependence	0.48	0.04–0.92	0.51

The second kappa statistic in Table 2 shows the diagnostic concordance between the SSAGA and the SCAN when the diagnosis was based upon all positively endorsed SCAN symptoms, including those that did not meet clinical severity criteria according to the SCAN rater's clinical judgement. When the kappa statistics were computed between the SSAGA and the SCAN diagnoses with inclusion of all positive symptoms, regardless of the symptom severity, the concordance rates were relatively unchanged or were higher, except for alcohol dependence. When all positively endorsed symptoms for alcohol dependence, regardless of severity rating, were included for the SCAN diagnosis of alcohol dependence, 100% ($n = 37$) of the positive diagnoses on the SSAGA were also diagnosed as alcohol dependent by the SCAN. However, 21 subjects who received positive diagnoses on the SCAN under this scoring scheme were not so diagnosed by the SSAGA.

The SSAGA/SCAN diagnostic concordance for the other Axis I disorders were found to be in the acceptable range, except for social phobia (Table 3) which ranged from $K = 0.62$ to 0.70. Similarly, a high kappa based upon the SSAGA

Table 3. *Diagnostic comparison of Axis I disorders and ASPD between SSAGA and SCAN: kappa statistics*

	SCAN	
Diagnosis	Kappa	95% C.I.
Depression	0.70	0.44–0.80
Social phobia	0.47	0.03–0.88
Panic disorder	0.62	0.36–0.88
Antisocial personality disorder (SSAGA vs. SCID)	0.71	0.49–0.93

and SCID comparison (0.71) was found for ASPD.

Examination of discrepancies in diagnoses

In order to examine the source of diagnostic discrepancies between the SSAGA and SCAN diagnoses, cases with discordant diagnoses were examined individually (see Table 4). Discrepancies that resulted from diagnostic discordance due to the presence or absence of symptoms without severity threshold considerations were not examined here. However, the majority of SSAGA/SCAN discrepancies across diagnoses were due to the differences in the rating of severity of symptoms. When a diagnosis was positive on the SSAGA, but negative on the SCAN, the discrepancy was due mainly to the differential assessment of symptom severity. In these discrepant cases, positively endorsed symptoms reported in the SSAGA interview which met the highest severity criteria were often rated as more mild (i.e. not meeting the severity criteria) by the SCAN interviewers. A small number of discrepancies were also due to subjects' inconsistency in reporting responses on either the SSAGA or SCAN interviews. The order of conducting the two interviews was also investigated in relation to the discrepancies, but no systematic bias was detected (data not shown).

The diagnosis of alcohol dependence deserves particular comment. When standard SCAN and standard SSAGA procedures were employed, there were 15 discrepant cases of alcohol dependence (see Table 1). Of the 13 positive cases identified by the SSAGA but not by the SCAN, 10 cases were positive for 'alcohol abuse' on the SCAN and the remaining three reported symptoms too mild for a diagnosis. On the other hand, the SCAN identified two alcohol-dependent cases not so identified by SSAGA; these two individuals denied having alcohol dependence symptoms (see Table 4). However, we were concerned that the difference in identifying alcohol dependence between the two interviews may be in relation to symptom severity rating. To examine this, we again compared SSAGA diagnoses of alcohol dependence in relation to a SCAN diagnosis based upon any and all positively endorsed alcohol symptoms, regardless of severity rating. In the usual application of the SCAN this type of procedure would not be done, since many of the symptoms now counted toward a 'diagnosis' were very mild and did not produce impairment. When all positively endorsed symptoms were counted, the number of discordant cases increased as expected, because the SSAGA cases only met clinical severity criteria while the SCAN cases were based upon any positive symptoms (including very mild and infrequent symptoms). This comparison produced a reduced kappa = 0.49 (see Table 2) since many of the 'discrepancies' in this comparison were not clinical cases.

With regard to the diagnosis of cannabis dependence, 13 subjects had a discrepant diagnosis. Seven subjects with cannabis dependence, according to the SSAGA, reported symptoms that were assessed as present, but not clinically significant, by the SCAN interviewers. Two subjects diagnosed as cannabis-dependent on the SSAGA interview reported too few or no symptoms related to cannabis use during the SCAN interview. Conversely, four subjects identified as cannabis-dependent by the SCAN interviews reported too few symptoms to meet cannabis dependence on the SSAGA interview. Similarly, the discrepancies in the diagnosis of depression were caused by the differential

Table 4. *Number of discrepancies between SCAN and SSAGA and explanation*

		Disagreement			
		Positive diagnosis on SSAGA, not SCAN		Positive on SCAN, not SSAGA	
Diagnoses	Total *N*	*N*	Explanation on SCAN interview	*N*	Explanation on SSAGA interview
Alcohol dependence	15	13	Symptoms mild (3) Met abuse criteria (10)	2	Denied symptoms
Cannabis dependence	13	9	Symptoms mild (7) Too few or no symptom (2)	4	Too few symptoms (4)
Cocaine dependence	5	4	Symptoms mild (4)	1	Experimental use without symptoms
Opiate dependence	5	3	Used $<5+$ times (1) Symptoms mild (2)	1	Denied symptoms
Sedative dependence	4	3	Symptoms mild (3) Denied use (1)	1	Denied any use
Stimulant dependence	1	1	Symptoms mild		
Depression	11	7	Depressive mood 1 week or less (4) No clustering of symptoms (2) Denied symptoms	4	No depressive symptoms (2) Depressive mood 1 week or less (2)
Panic disorder	7	5	SCAN symptoms mild (5)	2	No symptoms
Phobia	4	2	SCAN symptoms mild (2)	2	No symptoms
ASPD	6	4	Too few conducts Sx (3) Adult conduct only (1)	2	Too few conducts Sx (1) Adult conduct only (1)

reporting of the duration of the depressed/dysphoric mood, denial of symptoms by respondents, and by the failure for a sufficient number of symptoms to cluster.

Discussion

This study examined the validity of the SSAGA in relation to the SCAN across a variety of DSM-III-R life-time diagnoses. Both interviews are known to be highly reliable, even though they are quite different in relation to their format of administration. A comparison of the SSAGA against the SCAN revealed generally high concordance rates all across diagnoses examined. The results indicate that the SSAGA is a valid instrument to diagnose alcohol and other drug dependencies, major depressive disorder, anxiety disorders and antisocial personality disorder. While the kappa statistics for DSM-III-R alcohol dependence were not extremely high, the kappa statistics found for the SSAGA alcohol dependence compared favorably to reports which examined alcohol dependence using other diagnostic schedules. Cottler *et al.* (1997) compared the CIDI (WHO, 1993), SCAN and AUDADIS (Grant *et al.*, 1995) for the DSM-IV diagnoses of alcohol and other drug dependence. The pair wise SCAN–AUDADIS interview comparisons for alcohol dependence yielded kappas ranging between 0.62 and 0.67, and $K = 0.67$ for CIDI–AUDADIS. These values are similar to the kappa of 0.63 reported for SSAGA–SCAN in the present study. The CIDI and SCAN comparison for substance dependence diagnoses yielded kappas ranging from 0.37 for cannabis dependence to 0.67 for opiate dependence. The kappas for the SCAN and AUDADIS comparisons of substance dependence ranged from 0.35 to 0.62, while CIDI and AUDADIS resulted in kappas ranging from 0.38 to 0.67. The concordance rate between the SSAGA and SCAN on other Axis I diagnoses were much higher than those reported for the SCAN and CIDI for depression and panic disorder. Andrews *et al.* (1995) found the kappas for life-time diagnoses of the Axis I disorders ranged from 0.34 for depression to 0.61 for OCD among subjects who were accepted for treatment at an anxiety disorder clinic.

Diagnostic discrepancies found in the present study between the SSAGA and SCAN were

examined, and differences appeared to be caused by several factors. First, the SCAN is administered by trained clinicians. While scoring of subject responses when using the SCAN is guided by specific probe questions with a glossary that contains guidelines for coding item criteria, the clinical judgement of the interviewer actually determines whether a symptom meets the severity criterion. On the other hand, the SSAGA was designed to be administered by trained lay interviewers. Judgements made regarding symptom severity on the SSAGA are in relation to specific criteria (i.e. frequency of occurrence; seen a professional; interference with life tasks; or taken medications) rather than based upon clinical experience or skill.

Another factor which may have affected the diagnostic discrepancy rate is that the SSAGA interview, due to its poly-diagnostic features, examines a wider variety of symptoms than the SCAN or SCID. Thus, the SSAGA typically gathers more information related to a specific syndrome than other interview schedules because it includes a greater number of items to assess each diagnostic criterion. Similarly, the SSAGA has multiple items related to each diagnostic criterion item for other drug dependencies and for antisocial personality disorder. Since some discrepancies in diagnoses between SSAGA and SCAN tended to result from more people being diagnosed positive by the SSAGA than by the SCAN or SCID (ASPD), the difference may be due to the availability of more information provided by the SSAGA.

Discrepancies between the SSAGA and the SCAN may also be a function of the length of time to complete an interview. The SSAGA interview can be long and tedious due to the detail required for some diagnoses. Thus, fatigue may have led some subjects to deny certain symptoms, resulting in discrepant responses between two interview schedules or to report diminished symptom severity. However, the order of administering the interviews did not seem to provide a systematic bias in favor of one interview over the other.

No formal discrepancy protocol was employed to assess reasons for a discrepancy and we were unable to identify clearly which response was correct. However, judging from our examination of the cases with discrepant responses, the assumption of variation regarding symptom severity (i.e. clinical judgement vs. use of specific scoring criteria) seems reasonable.

Because the SSAGA was developed before the DSM-IV and the ICD-10 were officially published, the SSAGA is not able to make all Axis I diagnoses using these systems. Thus, we were unable to use either as the base criteria for comparison across all diagnoses. While the most appropriate comparison would have been to use the DSM-IV as the diagnostic criteria, recent empirical evidence indicates that the DSM-III-R, DSM-IV and ICD-10 diagnoses for alcohol and substance dependence are quite comparable. Schuckit *et al.* (1994) compared the ICD-10, DSM-III-R and DSM-IV diagnoses of alcohol and drug dependence for 1922 subjects from the COGA study using the SSAGA and found high agreement among all three criteria sets on the diagnoses of alcohol and drug dependence, while low agreement was found for both abuse and harmful use. More recently, Hasin *et al.* (1997) also found high agreement among the ICD-10, DSM-III-R, and DSM-IV alcohol and drug dependence diagnoses in a study using subjects from several countries using several different diagnostic instruments and criteria sets. Since the original SSAGA is capable of making DSM-IV and ICD-10 diagnoses only for alcohol and other drug dependencies, the SSAGA has recently been revised (SSAGA-II) to make DSM-IV and ICD-10 diagnoses for affective and anxiety disorders and ASPD by adding the requisite items.

The kappa for alcohol dependence is lower than for the other diagnoses. Is this possibly due to lower agreement on a particular criterion? In another study of the reliability of the SSAGA (Bucholz *et al.*, 1995), we examined the reliability of individual diagnostic criterion items for substance dependence in detail. We found that the inter-rater reliability for the nine DSM-III-R alcohol dependence criterion items was very good. The lowest was for 'use of alcohol to relieve withdrawal symptoms' with a kappa of 0.67; the eight other criterion items had kappas ranging from 0.75 to 0.92, Yule's statistics were in the same range. The present study indicates that the lower kappa is not due to the inability of the SSAGA to identify alcohol dependence, but is due to the lower rate of alcohol dependence identified by the SCAN (see Table 1). Of the 37 cases of alcohol dependence identified by SSAGA, 13 were 'missed' by SCAN. The difference in case detection rates appears to be due to the SCAN rating of a variety of symptoms as insufficiently severe for a diagnosis. However,

without a 'gold standard', we cannot tell whether the SSAGA is over-inclusive or if the SCAN is failing to detect clinical cases.

The results of the current study, along with the previous studies of its test–retest and between-study site reliabilities, support the SSAGA's utility for assessing a variety of Axis I psychiatric disorders, including alcohol and drug dependence. The poly-diagnostic feature of the SSAGA and its use by lay interviewers makes it an ideal instrument for use in a variety of applications, including epidemiological, genetic and family studies.

Acknowledgements

This work was supported by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) by USPHS grants NIAAA U10AA08401, U10AA08402, U10AA08403, 2 P05AA03510, and T32 AA07290. The Collaborative Study on the Genetics of Alcoholism (COGA) (Principal Investigator, H. Begleiter, SUNY HSCB and Co-Principal Investigator, T. Reich, Washington University), the National Collaborative Study includes six different centers where data collection takes place. The six sites and Principal Investigators and Co-Investigators are:

University of California at San Diego and Scripps Institute: M. Schuckit, F. Bloom.
University of Connecticut: V. Hesselbrock.
Indiana University: J. Nurnberger, Jr., T.K. Li, P.M. Conneally, H. Edenberg.
University of Iowa: R. Crowe, S. Kuperman.
State University of New York, Health Sciences Center at Brooklyn: H. Begleiter, B. Porjesz.
Washington University: T. Reich, C. R. Cloninger, J. Rice.

The authors would also like to thank the members of the COGA Assessment Committee who were responsible for the development of the SSAGA: Kathleen Bucholz, Remi Cadoret, C. Robert Cloninger, Stephen Dinwiddie, Victor Hesselbrock (chair), Louise Laster, John Nurenberger, Jr, Theodore Reich and Marc Schuckit.

References

Andrews, G., Peters, L., Guzman, A. & Bird, K. (1995) A comparison of two structured diagnostic interviews: CIDI and SCAN, *Australian and New Zealand Journal of Psychiatry*, 29, 124–132.

Bucholz, K.K., Cadoret, R., Cloninger, C.R. *et al.* (1994) A new, semi-structured psychiatric interview for use in genetic linkage studies: a report on the reliability of the SSAGA, *Journal of Studies on Alcohol*, 55, 149–158.

Bucholz, K.K., Hesselbrock, V., Shayka, J. *et al.* (1995) Reliability of individual diagnostic criterion items for psychoactive substance dependence and the impact on diagnosis, *Journal of Studies on Alcohol*, 56, 500–505.

Compton, W. M., Cottler, L. B., Dorsey, K. B., Spitznagel, E. L. & Mager, D. E. (1996) Comparing assessments of DSM-IV substance dependence disorders using CIDI-SAM and SCAN, *Drug and Alcohol Dependence*, 41, 179–187.

Ciranello, R.D. & Ciranello, A.L. (1991) Genetics of major psychiatric disorders, *Annual Review of Medicine*, 42, 151–158.

Coryell, W., Cloninger, C. R. & Reich, T. (1978) Clinical assessment: use of nonphysician interviewers, *Journal of Nervous and Mental Disease*, 166, 599–606.

Cottler, L., Grant, B., Blaine, J. *et al.* (1997) Concordance of DSM-IV alcohol and drug use disorder criteria and diagnoses as measured by AUDADIS-ADR, CIDID, and SCAN, *Drug and Alcohol Dependence*, 47, 195–205.

Cottler, L., Robins, L. & Helzer, J. (1989) The reliability of the CIDI-SAM: a comprehensive substance abuse interview, *British Journal of Addiction*, 84, 801–814.

Easton, C., Meza, E., Mager, D. *et al.* (1997) Test–retest reliability of the alcohol and drug use disorder sections of the schedules for clinical assessment in neuropsychiatry (SCAN), *Drug and Alcohol Dependence*, 47, 187–194.

Endicott, J. & Spitzer, R. (1978) A diagnostic interview, *Archives of General Psychiatry*, 35, 837–844.

Grant, B. Harford, T., Dawson, D., Chou, P. & Pickering, R. (1995) The Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): reliability of alcohol and drug modules in a general population sample, *Drug and Alcohol Dependence*, 39, 37–44.

Hasin, D., Grant, B., Cottler, L. *et al.* (1997) Nosological comparisons of alcohol and drug diagnoses: a multisite, multi-instrument international study, *Drug and Alcohol Dependence*, 47, 217–226.

Helzer, J.E. (1981) The use of a structured diagnostic interview for routine psychiatric evaluations, *Journal of Nervous and Mental Disease*, 169, 45–49.

Helzer, J.E., Robins, L.N., McEvoy, L.T. *et al.* (1985) A comparison of clinical and diagnostic interview schedule diagnoses: physician reexamination of lay-interviewed cases in the general population, *Archives of General Psychiatry*, 42, 657–666.

Hesselbrock, V., Stabenau, J., Hesselbrock, M., Mirkin, P. & Meyer, R. (1982) A comparison of two interview schedules: the Schedule for Affective Disorders and Schizophrenia-Lifetime and the National Institute of Mental Health Diagnostic Interview Schedule, *Archives of General Psychiatry*, 39, 674–677.

Maziade, M., Roy, M.A., Fournier, J.P. *et al.* (1992) Reliability of best-estimate diagnosis in genetic link

age studies of major psychoses, *American Journal of Psychiatry*, 149, 1674–1686.

RICE, J.P., REICH, T.R., BUCHOLZ, K.K. *et al.* (1995) Comparison of direct interview and family history diagnoses of alcohol dependence, *Alcoholism: Clinical and Experimental Research*, 19, 1018–1023.

ROBINS, L., HELZER, J., CROUGHAN, J. & RATCLIFF, K. (1981) The NIMH Diagnostic Interview Schedule: its history, characteristics, and validity, *Archives of General Psychiatry*, 28, 381–389.

ROBINS, L. N., COTTLER, L. B. & BABOR, T. (1986) *The WHO/ADAMHA CIDI-SAM Interview*, revised (St Louis, Washington University).

SCHUCKIT, M.A., HESSELBROCK, V., TIPP, J., ANTHENELLI, R., BUCHOLZ, K. & RADZIMINSKI, S. (1994) A comparison of DSM-III-R, DSM-IV and ICD-10 substance use disorders diagnoses in 1922 men and women subjects in the COGA study, *Addiction*, 89, 1629–1638.

SPITZER, R., WILLIAMS, J., GIBBON, M. & FIRST, M. (1992) The structured clinical interview for DSM-III-R (SCID) I: history, rationale, and description, *Archives of General Psychiatry*, 49, 624–629.

VAN DEN BREE, M.B.M., SVIKIS, D.S. & PICKENS, R.W. (1998) Genetic influences in antisocial personality and drug use disorders, *Drug and Alcohol Dependence*, 49, 177–181.

WING, J.K., BABOR, T., BRUGHA, T. *et al.* (1990) SCAN, Schedules for clinical assessment in neuropsychiatry, *Archives of General Psychiatry*, 47, 589–593.

WELLS, J.E., BUSHNELL, J.A. & OAKLEY-BROWNE, M.A. (1995) Using latent class analysis to investigate alcohol diagnosis, *International Journal of Methods in Psychiatric Research*, 5, 205–212.

WORLD HEALTH ORGANIZATION (1993) *Composite International Diagnostic Interview Schedule* (v2.0) (Geneva, Switzerland, World Health Organization).