# Reliability and Validity of the Substance Abuse Outcomes Module

G. Richard Smith, M.D.
M. Audrey Burnam, Ph.D.
Cynthia L. Mosley, B.B.A.
Jan A. Hollenberg, M.S.
Mike Mancino, M.D.
Wen Grimes, M.A.

*Objective:* **The study sought to determine the validity and reliability of the Substance Abuse Outcomes Module (SAOM), a self-report tool designed to assess patient characteristics, process of care, and outcomes of care, using a minimum amount of information, in order to improve treatment.** *Methods:* **A longitudinal field test (baseline and three-month follow-up) compared the SAOM to seven other research instruments in the assessment of 100 substance-abusing patients who were entering a new treatment episode. Quota samples of patients were drawn from two private inpatient substance abuse treatment facilities and an outpatient methadone clinic. The study's primary outcome measures were diagnostic accuracy, internal and test-retest reliability of key constructs, concurrent and predictive validity, and sensitivity to change. Cronbach's alpha coefficients were calculated to examine internal consistency and reliability. Intraclass correlation coefficients and kappa coefficients were used to examine test-retest reliability. Concurrent validity of outcomes measures was examined with Pearson or Spearman correlation coefficients and chi square and kappa statistics. Changes between baseline and follow-up were examined as a function of case-mix measures with ordinary least-squares multiple regression. Sensitivity to change was examined by calculating effect size scores.** *Results:* **The SAOM had high internal consistency and a high level of agreement with research diagnoses at baseline and follow-up. The SAOM was found to be highly reliable, to have very strong validity, and to be sensitive to clinical change.** *Conclusions:* **The SAOM appears to be a reasonably reliable and valid self-report instrument when used to monitor substance abuse treatment among patients with a primary substance use diagnosis. (***Psychiatric Services*** 57:1452–1460, 2006)**

The estimated annual cost for substance abuse treatment in the United States ranges from $16 to $18 billion (1,2). In 2001 annual expenditures for alcohol treatment services and drug abuse treatment were estimated at $9.7 billion and $8.5 billion, respectively (1). As of March 31, 2003, approximately 1.1 million patients were in treatment (3). To better serve these patients and keep costs to a minimum, clinicians must be aware of treatments that work effectively and of outcomes expected from these treatments.

Observational studies indicate that, on average, substance abuse treatments are effective (4,5), and controlled studies support the efficacy of specific treatment interventions or approaches (6). However, a wide variety of substance abuse treatments exist, and costs are substantial. Additional information about types of patients who would benefit from specific treatments and the amount of treatment that is appropriate to achieve the best patient outcomes at an affordable cost could be very helpful.

Thus another assessment tool for routine care could provide knowledge about what works in substance abuse treatment settings, which, in turn, could help improve patient outcomes. Posttreatment follow-up of patients in substance abuse treatment is important because of the disorder's chronic, relapsing nature. Thus a self-report measure is particularly important. To address this concern, we evaluated the Substance Abuse Outcomes Module (SAOM), a self-report instrument developed several years ago that does not require a trained interviewer, is brief, assesses substance use and various social and functioning consequences, and collects data about substance use diagnosis, relapse, and remission.

This article presents field test results about SAOM's ability to assess the outcomes of substance abuse treatment as part of routine clinical care. A preliminary, prospective field

*Dr. Smith, Ms. Mosley, Ms. Hollenberg, Mr. Grimes, and Dr. Mancino are affiliated with the Department of Psychiatry, University of Arkansas for Medical Sciences, 4301 West Markham Street, SL 554, Little Rock, AR 72205 (e-mail: smithgrichard@ exchange.uams.edu). Dr. Mancino is also with the Central Arkansas Veterans Healthcare System. Dr. Burnham is with RAND Corporation, Santa Monica, California.*

test was used to examine the SAOM's utility, specifically the reliability and construct validity of its key patient characteristics and patient outcome measures and the SAOM's responsiveness to change over time among individuals treated for substance abuse and dependence.

The SAOM also has the potential to improve care on a broader scale and more indirectly—through systems of care or through measures of competencies or performance. Although routine assessment of outcomes in substance abuse treatment settings is not common, some provider organizations have started incorporating this information technology into their systems of care. Two examples are the Veterans Affairs clinical programs for substance abuse treatment (7) and the treatment outcomes and performance pilot studies supported by the Center for Substance Abuse Treatment (8), which require routine outcome assessment with the Addiction Severity Index (ASI) (9).

The Accreditation Council on Graduate Medical Education started requiring residency training programs, including psychiatry programs, to document resident competencies in practice-based learning and improvement (10). The American Board of Psychiatry and Neurology has plans to begin requiring documentation of competencies for psychiatrists who participate in the Maintenance of Certification Program (11). In addition, the Centers for Medicare and Medicaid Services of the U.S. Department of Health and Human Services is about to implement a pay-for-performance plan in which physicians will be reimbursed at a higher rate for meeting certain practice-based performance measures (12). Routine outcomes assessments, such as the SAOM, will likely be one way to measure outcomes for systems of care and to document competencies and performance measures.

## Methods

### Substance Abuse Outcomes Module

The SAOM is the latest in a series of six modules (13–18) developed to improve the understanding of relationships between patient characteristics, processes of care, and patient out-comes of care by using the minimum amount of information possible (19). The SAOM, like the other five outcomes modules, has both clinician and patient baseline assessment forms, which are completed upon diagnosis of a new episode of substance abuse. The patient follow-up assessment form is completed three and six months after baseline. A user's manual provides details about recommended protocols, scoring instructions, and instrument development. The SAOM is available for unlimited free use by permission: an electronic copy can be obtained at the authors' Web site (www.netoutcomes.net) and elsewhere (20).

The SAOM uses the "tracer condition" approach, in which a single disorder is closely scrutinized in a given treatment setting. This disease-specific approach is the standard throughout medicine (21). These characteristics make the SAOM most appropriate for use in substance treatment settings and also for use with patients who have primary substance problems in general psychiatric settings. The SAOM would not be an appropriate assessment tool for patients with a primary diagnosis other than substance abuse or dependence who have a comorbid or secondary substance use disorder.

A comprehensive review of the literature on key patient characteristics, processes of care, and outcomes of care for substance abuse treatment was the basis for the SAOM's development. The module includes constructs that have been shown in multiple studies to be related to treatment outcomes; the module emphasizes collection of the minimum amount of data needed to understand key relationships in order to improve care (19).

The SAOM was also designed to measure the types and outcomes of care received for substance abuse and characteristics that influence outcomes and types of care given; it includes criteria to identify diagnostically similar groups of patients. Table 1 lists the domains measured by each SAOM component as well as the instruments used to construct the module (22–28).

Preliminary feasibility testing included field testing of a prototype module in multiple substance abuse treatment settings. After revision based on field test results, the prototype module was used for the study reported here. The version evaluated in this study is a 113-item self-report assessment designed to be used in a paper-and-pencil format, although it can be administered orally in person or via telephone and also by computer. Written at a third-grade reading level and currently available only in English, the SAOM takes 20 minutes, on average, to complete, according to preliminary feasibility testing. The SAOM makes current diagnoses (substance abuse or dependence) according to the particular substance of abuse and includes a variety of response formats (Likert scales, yes-no formats, and list selections).

### Design of the field test

Data on the SAOM's reliability and validity were collected in 1999 and 2000 as part of a longitudinal study of 100 patients who were identified at the beginning of treatment for an episode of *DSM-IV* alcohol or drug abuse or dependence. At baseline patients completed the SAOM, and a trained research assistant administered accompanying validation instruments. Approximately three months later, patients were recontacted to complete the SAOM's self-administered follow-up portion and a research assistant again administered the accompanying validation instruments. The approval of the University of Arkansas for Medical Sciences Institutional Review Board was obtained before study initiation.

### Patient recruitment

Clinicians conducting intake evaluations of consecutive admissions at an outpatient methadone treatment program and at two private inpatient substance abuse treatment facilities in Little Rock, Arkansas, referred eligible patients to the study. To restrict the study to those beginning a new episode of treatment, the admission had to be the first within the past six months and patients could not have had any outpatient treatment in the previous six months. Quota samples were drawn at admission to test the module across conditions in five groups on the basis of type of drug problem identified—alcohol, heroin or other opiates, cocaine or

**Table 1**

Domains and content of the Substance Abuse Outcomes Module

| Domain and module component | Content | Instruments used or modified (reference)[a] |
|---|---|---|
| **Patient characteristic** | | |
| Diagnosis | | |
|   Patient baseline assessment | *DSM-IV* criteria for abuse and dependence | |
|   Clinician baseline assessment | *DSM-IV* criteria for abuse and dependence | |
| Prognosis and case mix: patient baseline assessment | Severity of alcohol problems | |
| | Severity of drug problems | |
| | Age of onset | Modified from the Typology Questionnaire (22) |
| | Parental substance abuse or dependence | |
| | Social support | |
| | Support for sobriety | |
| | Previous treatment | |
| | Depression symptoms | Modified from the DIS (23) |
| | Antisocial traits | Modified from the DIS (23) |
| | Substance-related physical problems | |
| | Parenting responsibilities | |
| **Patient outcome** | | |
| Outcomes: patient baseline and patient follow-up assessments | Alcohol quantity and frequency | Modified from AUDIT (24) |
| | Drug use frequency | Modified from the DIS (23) |
| | Most problematic substance | Modified from the DIS (23) |
| | Drinking in the past month | Modified from the DIS (23) |
| | Drug use in the past month | Modified from the DIS (23) |
| | Drinking in past 3 months | Modified from the DIS (23) |
| | Drinking or using drugs in the past month versus the past 3 months | Modified from the DIS (23) |
| | Consequences of drinking or using drugs | INDUC-2R (25) |
| | Physical consequences | INDUC-2R (25) |
| | Interpersonal consequences | INDUC-2R (25) |
| | Intrapersonal consequences | INDUC-2R (25) |
| | Impulse control | INDUC-2R (25) |
| | Social consequences | INDUC-2R (25) |
| | General health consequences | MOS SF-36 (26) |
| | Days spent in bed because of health problems | MOS (27) |
| | Reduced-activity days | MOS (27) |
| | Failed-expectation days | |
| | Lost work days | NHS (28) |
| | General health status | MOS SF-36 (26) |
| **Process of care** | | |
| Clinician baseline assessment | Type and extent of treatment by major provider or insurer | |
| Patient follow-up assessment | Type and extent of treatment by major providers, insurers, or others | |

[a] DIS= Diagnostic Interview Schedule; AUDIT=Alcohol Use Disorders Identification Test; INDUC-2R=Inventory of Drug Use Consequences; MOS-SF-36=Medical Outcomes Study 36-item Short-Form Health Survey; NHS=National Health Survey

crack, marijuana, and hallucinogens. Eligible patients were at least 18 years old, had a primary diagnosis of alcohol or drug dependence or abuse, were able to understand and speak English, were medically stable, had no active psychoses, and had sufficient cognitive function to report on lifetime drinking and drug use.

A research assistant met with each newly referred patient to fully explain the study's purpose and requirements and to obtain written informed consent to participate in the research. Patients were assured that all information collected from personal interviews and administered instruments would not be shared with the treatment team, that study results would be presented for groups of patients only, and that their course of treatment was in no way connected to information provided to the study researchers. Each participant agreed to be contacted for follow-up assessment regardless of treatment success. After meeting with the research assistant, six referred patients declined to participate in the study. The research assistant administered the Neurobehavioral Cognitive Status Examination (29) to verify that referred patients demonstrated sufficient cognitive function to report about substance use during the past three months. One person who scored in the range of mild to moderate impairment was excluded from the study.

*Validation instruments*
Widely used and highly regarded instruments that collect information on SAOM domains (substance use, diagnoses, and consequences) were used to evaluate the SAOM's content validity.

The substance abuse module of the Composite International Diagnostic Interview (CIDI-SAM) (30) is a highly structured, detailed interview designed to ascertain specific diagnoses of substance abuse and dependence. The Addiction Severity Index, 5th edition (ASI), is a comprehensive, semistructured, interviewer-administered assessment tool (31). The Alcohol Use Disorder Identification Test (AUDIT) is an alcohol screening procedure developed by the World Health Organization that quantifies alcohol consumption and harmful consequences (32,33). The Diagnostic Interview Schedule (DIS) is a highly reliable, structured, lay-administered interview developed to assess psychiatric conditions (34). The DIS sections on depressive disorders and antisocial personality traits and disorders were used.

The Timeline Follow-Back Assessment (TFBA) is a reliable, valid, interviewer-administered method to estimate alcohol consumption (35). The Inventory of Drug Use Consequences (INDUC-2R) is a self-report instrument designed to assess lifetime and recent consequences of drug or alcohol use, including physical consequences, intrapersonal consequences, interpersonal consequences, and impulse control (25). The Medical Outcomes Study Social Support Survey (MOS-SSS) is a self-report instrument developed as part of the Medical Outcomes Study to assess the extent of respondents' social support networks (36).

An experienced, trained research assistant administered the CIDI-SAM, ASI, DIS sections, and TFBA. Patients were asked to complete the AUDIT, INDUC-2R, and MOS-SSS.

### Data collection

At baseline patients were randomly assigned to complete the SAOM's patient baseline assessment either before or after the accompanying validation instruments to control for order effects; one-half (N=50) of the patients received the SAOM first. One-half of the sample, chosen by independent random selection from the groups with a drug problem, was asked to complete the module a second time on day 3 or 4 after the initial SAOM administration to examine test-retest re-

liability. Baseline assessments occurred in face-to-face interviews or were self-administered, depending on the administration protocol for each instrument. All patients completing the baseline assessment received a check for $30, and those completing the retest received an additional $30.

At the time of follow-up, the SAOM and the validation instruments were readministered. Patients were considered lost to follow-up if they did not respond after several attempts to make phone or mail contact. All follow-up assessments had a target date three months after the baseline interview. Follow-up was conducted within a six-week window (two weeks before and four weeks after the target date). Follow-up interviews were completed by telephone, personal interview, or by mail, as appropriate for each assessment. As in the baseline assessment, the SAOM was administered to one-half of the patients before the other validation instruments and to the other half after the other validation instruments.

### Data analysis

Cronbach's alpha coefficients were calculated to examine the internal consistency and reliability of the multi-item SAOM. Test-retest reliability statistics were calculated for SAOM outcome measures with intraclass correlation coefficients (ICCs), and kappa coefficients were calculated for continuous and dichotomous SAOM measures.

Concurrent validity of outcome measures was examined by analyzing the association between SAOM measures and the appropriate validation measures with Pearson or Spearman correlation coefficients for continuous measures and chi square and kappa statistics for dichotomous measures. To examine the extent of predictive utility for SAOM measures of case mix, changes in SAOM outcome measures between baseline and follow-up were examined as a function of case-mix measures with ordinary least-squares multiple regression. Sensitivity of SAOM outcome measures to change was examined by calculating effect size scores (37).

### Table 2

Number of patients with diagnoses of substance abuse and dependence among 100 patients on admission to two inpatient treatment facilities or a methadone program[a]

| Substance | Abuse | Dependence |
|---|---|---|
| Alcohol | 17 | 72 |
| Amphetamines | 6 | 23 |
| Cannabis | 12 | 26 |
| Cocaine | 4 | 36 |
| Hallucinogens | 5 | 4 |
| Inhalants | 0 | 1 |
| Opioids | 3 | 20 |
| Phencyclidine | 3 | 0 |
| Sedatives | 4 | 11 |
| Any drug (including alcohol) | 42 | 91 |

[a] As measured by the Composite International Diagnostic Interview. Some patients had multiple abuse or dependence diagnoses. No patient had both an abuse and a dependence diagnosis for the same substance.

### Results

#### Sample

Of the 100 patients recruited at baseline, 32 were women. Ages ranged from 18 to 75 years, with a mean age of 40.7±9.6. The sample included 23 African-American patients, one Hispanic patient, and 76 white patients. Thirty-two were married or cohabitating, 19 were single and never married, 28 were divorced, 18 were separated, and three were widowed. Forty-six reported an annual household income of less than $20,000, 42 had incomes between $20,000 and $80,000, 11 reported incomes higher than $80,000, and one refused to answer this question. Twenty-nine had not completed high school, 22 were high school graduates, 28 attended some college, and 21 were college graduates.

Of the 100 patients, who were all beginning treatment for a new episode of substance abuse or dependence, 64 were new admissions to inpatient treatment program A, 26 were new admissions to inpatient treatment program B, and ten were new admissions to the outpatient methadone treatment program. Many patients had multiple substance abuse and dependence diagnoses as is typical in substance abuse treatment settings. Ninety-one patients had a CIDI-SAM diagnosis of sub-

**Table 3**

Reliability of key constructs of the Substance Abuse Outcomes Module

| Domain and construct | Items | Internal reliability (alpha) | Test-retest reliability[a] |
|---|---|---|---|
| Patient characteristic | | | |
| Diagnosis | 17 | .89 | $\kappa$=.56 |
| Case mix | | | |
| Severity of dependence | 17 | .89 | ICC=.95 |
| Severity of abuse | 12 | .90 | ICC=.95 |
| Parental substance abuse or dependence | 1 | na | $\kappa$=.96) |
| Age of onset (years) | 1 | na | ICC=.92 |
| Social support | 6 | .89 | ICC=.92 |
| Co-occurring medical disorder | 20 | .69 | ICC=.92 |
| Previous treatment | 1 | na | $\kappa$=.71 |
| Support for sobriety | 3 | .58 | ICC=.79 |
| Antisocial traits | 19 | .85 | $\kappa$=.90 |
| Patient outcome | | | |
| Alcohol consumption | | | |
| Past month | | | |
| Drinking days | 1 | na | ICC=.97 |
| Drinks on drinking days | 1 | na | ICC=.84 |
| Heavy drinking days | 1 | na | ICC=.85 |
| Drinks on heavy drinking days | 1 | na | ICC=.86 |
| Past 3 months | 1 | na | $\kappa$=.91 |
| Quantity and frequency | | | |
| Total drinks in the past month | 1 | na | .63 |
| Total drinks on heavy drinking days in the past month | 1 | na | ICC=.59 |
| Drug consumption | | | |
| Past month | | | |
| Any drug | 1 | na | $\kappa$=.85 |
| Drug use days | | | |
| Marijuana | 1 | na | ICC=.97 |
| Cocaine or crack | 1 | na | ICC=.99 |
| Heroin | 1 | na | ICC=.96 |
| Hallucinogens | 1 | na | ICC=.99 |
| Inhalants | 1 | na | |
| Methadone or opiates | 1 | na | ICC=.81 |
| Sedatives or hypnotics | 1 | na | ICC=.99 |
| Amphetamines | 1 | na | ICC=.90 |
| Total drug use days | 1 | na | ICC=0.93 |
| Most problematic drug | 1 | na | ICC=.86 |
| Alcohol and drug use | | | |
| Polydrug use days in the past month | 1 | na | ICC=.47 |
| Comparability of substance use in the past month versus the past 3 months | 1 | na | ICC=.77 |
| Consequences of alcohol or drug use | | | |
| Physical | 3 | .83 | ICC=.77 |
| Intrapersonal | 3 | .76 | ICC=.59 |
| Interpersonal | 3 | .72 | ICC=.64 |
| Impulse control | 3 | .75 | ICC=.53 |
| Social | 3 | .81 | ICC=.76 |

[a] Intraclass correlation coefficient or kappa coefficient

stance dependence, and 42 had a diagnosis of substance abuse (see Table 2 for specific diagnoses). Four did not meet *DSM-IV* criteria for dependence or abuse according to the CIDI-SAM interview, even though they had a clinical diagnosis of substance abuse or dependence when admitted to the treatment program. One patient who entered the study died before the three-month follow-up assessment, three could not be located, and three refused follow-up assessment. The remaining 93 patients completed the follow-up assessment.

*Accuracy of self-reports of diagnostic criteria*

The accuracy of the SAOM's diagnostic measure was compared with the accuracy of diagnoses obtained with the research standard, the CIDI-SAM. At baseline, the SAOM's diagnostic component's 17 items had high internal consistency, with a Cronbach alpha coefficient of .89. Agreement between the SAOM diagnostic measure and the CIDI-SAM about the presence of a substance use diagnosis (abuse or dependence) was 93 percent. Nonagreement between the measures was evenly balanced between over- and underdiagnoses. The SAOM diagnostic measure was highly sensitive (96 percent) in detecting whether any substance abuse diagnosis was present. Because the probability of not having a diagnosis in this baseline clinical sample was very low (4 percent), calculation of specificity was inappropriate.

A similar analysis was conducted for the accuracy of the SAOM's diagnostic measure at follow-up, when some patients' substance use disorders had remitted. The percentage agreement with the CIDI-SAM was high, with 90 percent agreement on the presence of any diagnosis and 89 percent agreement on the presence of abuse or dependence. The kappa coefficient (.76) suggested substantial agreement on the presence of any substance use diagnosis, based on the scoring criteria of Cohen (38,39).

*Reliability of key constructs*

Table 3 shows results of assessments of internal reliability (alpha) and test-retest reliability (ICC or kappa coefficient) for key constructs in the SAOM domains of patient characteristics and patient outcomes of care.

*Patient characteristics.* In addition to a diagnostic measure, the SAOM includes case-mix variables that have been shown in prior studies to predict the outcomes of substance abuse treatment. These variables include severity of abuse and dependence, parental abuse or dependence, age of onset of abuse or dependence, social support,

*Table 4*

Concurrent validity of patient variables at baseline and follow-up as measured by the Substance Abuse Outcomes Module

| Variable | Research standard[a] | Baseline correlation or other statistic[b] | Follow-up correlation or other statistic |
|---|---|---|---|
| Case-mix variable | | | |
|   Severity of alcohol abuse or dependence | ASI severity | .70 | .80 |
| | AUDIT | .81 | .88 |
| | CIDI-SAM severity of abuse | .66 | .92 |
| | CIDI-SAM severity of dependence | .72 | .94 |
|   Severity of drug abuse or dependence | ASI severity | .38 | .48 |
| | CIDI-SAM severity of abuse | .39 | .88 |
| | CIDI-SAM severity of dependence | .33 | .65 |
|   Age of onset | CIDI-SAM age of onset | .56 | na |
|   Previous treatment | ASI | $\chi^2=11.3, \kappa=.31$ | na |
|   Parental substance abuse or dependence | ASI family history | $\chi^2= 69.1, \kappa=.46$ | na |
|   Social support | MOS-SSS | .54 | na |
|   Medical comorbidity | ASI medical status | .36 | na |
|   Antisocial personality traits | DIS | $\chi^2=19.5, \kappa=.48$ | na |
| Patient outcome | | | |
|   Alcohol or drug use | | | |
|     Alcohol use in the past month | TFBA | .85 | .73 |
| | ASI | .85 | .81 |
|     Drug use in the past month | ASI | $\chi^2=72.0, \kappa=.84$ | $\chi^2=66.9, \kappa=.84$ |
|     Alcohol use in the past 3 months | TFBA | $\chi^2=77.3, \kappa=.87$ | $\chi^2=88.9, \kappa=.98$ |
|   Alcohol quantity | TFBA | .56 | .62 |
| | AUDIT | .53 | .86 |
| | ASI | .47 | .72 |
|   Alcohol frequency | TFBA | .60 | .65 |
| | AUDIT | .56 | .86 |
| | ASI | .53 | .72 |
|   Total drug use days | ASI | .70 | .80 |
|   Most problematic substance | ASI | .97 | .99 |
| Consequences of alcohol or drug use | | | |
|   Physical | INDUC-2R | .82 | .93 |
|   Interpersonal | INDUC-2R | .77 | .95 |
|   Intrapersonal | INDUC-2R | .75 | .95 |
|   Impulse control | INDUC-2R | .66 | .99 |
|   Social | INDUC-2R | .69 | .91 |

[a] ASI=Addiction Severity Index; AUDIT=Alcohol Use Disorders Identification Test; CIDI-SAM=substance abuse module of the Composite International Diagnostic Interview; MOS-SSS=Medical Outcomes Study Social Support Survey; DIS=Diagnostic Interview Schedule; TFBA=Timeline Follow-Back Assessment; INDUC-2R=Inventory of Drug Use Consequences

[b] The correlation values are intraclass correlation coefficients, except for chi square statistics (df=1) or kappa coefficients as noted.

comorbid medical conditions, previous treatment for abuse or dependence, support for sobriety, and presence of antisocial personality traits. Generally, the internal reliability of all case-mix variables was good to excellent, with alpha values from .58 to .90 and all but two values reaching .85 or above. The ICCs or kappa coefficients, as appropriate, were also strong, ranging from .56 to .96, with the values for most case-mix variables above .90 except for diagnosis, previous treatment, and support for sobriety.

*Patient outcomes.* Patient outcome measures included assessments of substance use and consequences of substance use, as shown in Table 3. Substance use variables included multiple measures of quantity and frequency of substance use. Because use measures were single items, only test-retest reliability is shown in Table 3 for those measures. Measures of consequence were multi-item scales that assessed physical, intrapersonal, interpersonal, impulse control, and social consequences of substance use. All consequence scales demonstrated high internal consistency, with alpha values from .72 to .83. Both categories showed moderate to high test-retest reliability, with most ICCs or kappa coefficients above .80 (range= .47–.99). Items below .60 were total number of drinks on heavy drinking days, polydrug use days in the past month, intrapersonal consequences, and impulse consequences.

### Validity of key constructs
*Concurrent validity.* Table 4 presents concurrent validity data for key patient case-mix characteristics and patient outcome variables assessed in the SAOM. Correlations of the SAOM case-mix variables with the corresponding research standard ranged from .36 to .94. Correlations were .54 or higher, except for severity of drug abuse or dependence at baseline and comorbid medical conditions. All chi square values were highly significant, and all kappa coefficient values were in the moderate to high range (35,36).

For patient outcomes, correlations between the module variables and the research standards were high, ranging from .47 to .99. All values exceeded .60, except the values for alcohol quan-

**Table 5**

Validity of baseline patient characteristics as predictors of patient outcomes at follow-up, as measured by the Substance Abuse Outcomes Module

| Outcome and case-mix variable at baseline | p | Overall $R^2$ |
|---|---|---|
| N of alcohol drinking days: severity of abuse | | .18 |
| or dependence | ns | |
| N of heavy drinking days at follow-up | | .35 |
| Severity of abuse | .03 | |
| Severity of dependence | .05 | |
| Social support | .01 | |
| Alcohol quantity or frequency at follow-up: severity | | |
| of abuse or dependence | .03 | .16 |
| N of drug use days at follow-up: antisocial | | |
| personality traits | ns | .14 |
| N of drug use days at follow-up: antisocial | | |
| personality traits | ns | .14 |
| Abstinence at follow-up | | .15 |
| Parental substance abuse | ns | |
| Medical comorbidity | ns | |
| Antisocial personality traits | ns | |
| Change in consequences from baseline to follow-up | | |
| Physical | | .34 |
| Severity of abuse or dependence | <.001 | |
| Social support | ns | |
| Depression | ns | |
| Intrapersonal | | .29 |
| Severity of abuse or dependence | .04 | |
| Social support | .004 | |
| Depression | .005 | |
| Interpersonal | | .29 |
| Severity of abuse or depend | ns | |
| Social support | ns | |
| Depression | .003 | |
| Impulse control: severity of abuse or | | |
| dependence | <.001 | .33 |
| Social | | .32 |
| Severity of abuse or dependence | ns | |
| Age of onset | ns | |
| Parental substance abuse | ns | |
| Social support | ns | |
| Depression | .002 | |
| Antisocial personality traits | ns | |
| Bed days: medical comorbidity[a] | .002 | .20 |
| Reduced-activity days | | .23 |
| Severity of abuse or dependence[b] | .03 | |
| Parental substance abuse | ns | |
| Failed expectation days | | .33 |
| Severity of abuse or dependence[b] | .05 | |
| Age of onset | ns | |
| Previous treatment | .04 | |
| Medical comorbidity[b] | .004 | |
| Depression | .03 | |

[a] Number of days in the past month that the patient was kept in bed because of health problems
[b] Prediction in the opposite direction to that expected

tity and frequency at baseline. The chi square and kappa coefficient comparisons were also highly significant.

The SAOM's validity in determining remission of substance use disorders according to *DSM-IV* criteria is important, because follow-up remission rates are often used as an indication of quality of care. The SAOM's follow-up diagnostic assessment was compared with that of the CIDI-SAM in a 2×2 analysis of remission. Agreement (no-no and yes-yes) occurred in 69 of the 81 cases (85 percent). In one case (1 percent), the SAOM diagnostic assessment indicated remission, but the CIDI-SAM did not. In 11 cases (14 percent), the SAOM assessment underreported remission compared with the CIDI-SAM. The SAOM diagnostic assessment, compared with the CIDI-SAM, had 83 percent sensitivity to determine remission, 94 percent specificity, 98 percent positive predictive value, 39 percent negative predictive value, and a kappa coefficient of .64.

*Predictive validity.* The predictive validity of baseline patient case-mix characteristics in regard to patient outcomes at follow-up was tested. The SAOM's case-mix variables in the baseline assessment included severity of dependence or abuse, age of onset of dependence or abuse, parental substance abuse or dependence, previous substance abuse treatment, social support, support for sobriety, number of comorbid medical conditions, depression, and antisocial traits. Table 5 lists patient outcomes and notes which case-mix variables significantly predicted them. Statistically significant findings ($p \leq .01$) and trends ($p > .01$ and $p \leq .05$) are reported for each case-mix measure, along with the overall multiple correlation coefficient ($R^2$) value when the entire set of case-mix predictors was included. More stringent statistical criteria were used because of multiple comparisons. The $R^2$ value represents the total amount of variance in patient outcomes explained by all case-mix measures.

Each case-mix measure significantly predicted at least one patient outcome. The majority of case-mix measures predicted outcomes in hypothesized directions. Two predicted in a direction opposite of what was expected, but the values were not statistically significant (between .05 and .10). These two measures were parental substance abuse, which predicted fewer social consequences (rather than more), and severity of substance abuse, which predicted fewer reduced-activity days (rather than more).

### Sensitivity of key constructs to clinical change

Analyses examined the SAOM's sensitivity to clinically important change. Change scores between baseline and follow-up outcome measures were calculated, then effect-size scores for each outcome variable were calculat-

ed as the mean change score divided by the standard deviation of the baseline mean score (34). The effect size was interpreted as defined by Cohen (34): .20 or less is a small effect, .20 to .50 is a moderate effect, and .80 or greater is a large effect. A positive effect size represents improvement, and a negative effect size denotes worsening of the condition. All SAOM effect-size scores were positive, indicating that the SAOM variables changed in the clinically expected direction. Most scores were in the moderate or large ranges, indicating that the SAOM measures were sensitive to important clinical changes.

## Discussion

On the basis of this preliminary study, the SAOM appears to be a reasonably reliable and valid instrument for use in routine substance abuse treatment settings, particularly those that are similar to settings reported here for patients with substance abuse or dependence as a primary diagnosis. The SAOM provides self-report data regarding the presence of any diagnosis of substance abuse or dependence, key patient case-mix characteristics, and specific outcomes related to substance abuse or dependence, including substance use, consequences of use, relapse, and remission.

The SAOM was not designed as a clinical research instrument; it was specifically designed to systematically assess outcomes of care in routine clinical settings in order to monitor or improve patient outcomes (19). These types of clinical assessment tools need to be brief, preferably based on patient self-report, and to measure key case-mix characteristics as well as outcomes. An ideal outcome module would take only moments of the patient's time and no clinician time. The 20 minutes for patient completion of the SAOM at baseline and follow-up and the 1.5 minutes for clinician completion (clinician assessment) at baseline appear to be a reasonable trade-off for the potential improvement in care and assurance of good treatment outcomes. Future studies are needed to determine the SAOM's usefulness in monitoring and improving patient outcomes.

### Advantages of the SAOM

Although making diagnoses is the province of clinicians, estimates of diagnoses can be helpful in some quality improvement efforts. The SAOM's ability to estimate substance use diagnoses and remission according to *DSM-IV* criteria eliminates the expense of training and labor for outcomes assessment interviewers. Other advantages are the convenience to patients, who can select times to complete the SAOM; diagnostic confirmation, which can help in the reimbursement process; and determination of remission, which is becoming an increasingly important measure of successful treatment in other areas of medicine and may also become a helpful concept in substance abuse treatment.

Assessing patient case-mix characteristics is critical when treatment programs are compared in nonexperimental evaluations. Systematic assessment of these variables is needed to show that differences in programs are due to differences in treatment effectiveness rather than differences in disease severity of patient populations. Outcome assessment that fails to take case-mix characteristics into account is, therefore, of little use in efforts to improve care. Tools such as the SAOM have the potential to shed light on this important concept and to focus on processes of care that determine patient outcomes.

### Limitations

This study had several limitations. First, the study was primarily of inpatients in substance abuse treatment who were beginning a new episode of care (ten patients were from an outpatient methadone clinic), and most substance abuse treatment in the United States is outpatient treatment. Although this design concept was intentional to provide a relatively homogeneous group of patients with sufficient diversity to have good external validity (generalizability), the number of outpatients brings into question the SAOM's applicability or utility in outpatient care.

In addition, most assessments were of patients who had insurance for at least detoxification treatment and were initially treated as inpatients in substance abuse treatment units. It is unknown whether these results can be generalized to public-sector or uninsured outpatients and their treatment programs.

Further research is also needed to clearly understand the effects of variations in case mix on outcomes. Finally, the SAOM may have utility in general psychiatric settings, but it was not designed for patients with primary diagnoses other than substance use diagnoses. Any utility the SAOM might have among these patents remains to be determined.

### Areas for improvement and further study

Comparison with research standards indicate that several SAOM items or domains need improvement in subsequent revisions of the module. These include assessment of the severity of drug abuse, alcohol quantity and frequency items, social support assessment, and determination of comorbid medical conditions. In addition, other investigators will need to assess the SAOM's strengths and weaknesses, and the SAOM must be proven in the clinical settings for which it was designed. Additional research is also needed on diverse patient populations to extend the generalizability of findings reported in this study, to assess the responsiveness of SAOM outcome measures to clinically significant change, and to establish norms to better interpret magnitude of change. The SAOM's utility as a reliable measure of change for a single individual (in which case reliability should exceed .90) looks promising for some measures but is weak for others.

## Conclusions

The SAOM appears to be a reasonably reliable and valid self-report instrument for use in routine substance abuse treatment settings to collect information about diagnosis, key patient characteristics, and outcomes related specifically to substance abuse or dependence and thus to monitor or improve substance abuse treatment. Evaluating outcomes of clinical care is essential to improving care and developing new knowledge about care in routine settings. With tools such as the ASI and now the SAOM, clinical programs have fewer

barriers to prevent them from engaging in the process of ensuring that patients achieve the best possible outcomes from their substance abuse treatment.

### References

1. Mark TL, Coffey RM, McKusick DR, et al: National Estimates of Expenditures for Mental Health Services and Substance Abuse Treatment, 1991–2001. Rockville, Md, Substance Abuse and Mental Health Services Administration, 2005

2. The Economic Costs of Drug Abuse in the United States, 1992–2002. Washington, DC, Office of National Drug Control Policy, 2004

3. National Survey of Substance Abuse Treatment Services (N-SSATS), 2003. Rockville, Md, Substance Abuse and Mental Health Services Administration, 2003

4. Hubbard RL, Marsden ML, Rachal JV, et al: Drug Abuse Treatment: A National Study of Effectiveness. Chapel Hill, NC, University of North Carolina Press, 1989

5. Flynn PM, Kristiansen PL, Porto JV, et al: Costs and benefits of treatment for cocaine addiction in DATOS. Drug and Alcohol Dependence 57:167-174, 1999

6. Project MATCH Research Group: Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes. Journal of Studies on Alcohol 58:7–29, 1997

7. Moos RH, Finney J, Cannon D, et al: Outcomes Monitoring for Substance Abuse Patients: I. Patients' Characteristics and Treatment at Baseline. Palo Alto, Calif, Department of Veterans Affairs, Program Evaluation and Resource Center and Health Services Research and Development Center for Health Care Evaluation, 1998

8. Treatment Outcomes and Performance Pilot Studies (TOPPS). Rockville, Md, Substance Abuse and Mental Health Services Administration, 1999

9. McLellan AT, Luborsky L, Woody GE, et al: An improved diagnostic evaluation instrument for substance abuse patients: the Addiction Severity Index. Journal of Nervous and Mental Disease 168:26–33, 1980

10. Jarvis RM, O'Sullivan PS, McClain T, et al: Can one portfolio measure the six ACGME general competencies? Academic Psychiatry 28:190-196, 2004

11. Davis GC: General and psychiatry-specific patient care core competencies, in Core Competencies for Psychiatric Practice: What Clinicians Need to Know. Washington, DC, American Psychiatric Publishing, 2003

12. Peterson ED: Should we link payment to quality? American Heart Journal 148(suppl 5):S56-S58, 2004

13. Smith GR Jr, Ross RL, Rost KM: Psychiatric Outcomes Module: Depression (DOM), in Outcomes Assessment in Clinical Practice. Edited by Sederer LI, Dickey B. Baltimore, Williams and Wilkins, 1996

14. Smith GR Jr, Babor T, Burnam MA, et al: Alcohol Outcomes Module. Little Rock, Ark, University of Arkansas for Medical Sciences, Department of Psychiatry, 1994

15. Fischer EP: Mental Health Outcome Module Development and Testing: Development and initial validation of the Schizophrenia Outcomes Module: Final Report. Service Directed Research Grant 91.005. Fischer, EP1–55. Washington, DC, Department of Veterans Affairs, 1993

16. Rost K, Smith GR Jr, Katon W, et al: Panic Outcomes Module: User's Manual. Little Rock, Ark, University of Arkansas for Medical Sciences, Department of Psychiatry, 1994

17. Robbins JM, Taylor JL, Burns BJ, et al: Adolescent Treatment Outcomes Module (ATOM): User's Manual. Little Rock, Ark, University of Arkansas for Medical Sciences, Department of Psychiatry, 1997

18. Smith GR Jr, Babor T, Burnam MA, et al: Substance Abuse Outcomes Module: User's Manual. Little Rock, Ark, University of Arkansas for Medical Sciences, Department of Psychiatry, 1996

19. Smith GR, Rost K, Fischer E, et al: Assessing the effectiveness of mental health care in routine clinical practice, in Outcomes Measurement in the Human Services: Cross-Cutting Issues and Methods. Edited by Mullen EJ, Magnabosco JL. Frederick, Md, National Association of Social Workers Press, 1997

20. Smith GR Jr, Ross RL, Rost KM: Psychiatric Outcomes Module: Substance Abuse Outcomes Module (SAOM), in Outcomes Assessment in Clinical Practice. Edited by Sederer LI, Dickey B. Baltimore, Williams and Wilkins, 1996

21. Kessner DM, Kalk CE, Singer J: Assessing health quality: the case for tracers. New England Journal of Medicine 288:189–194, 1973

22. Typology Questionnaire. Hartford, Conn, University of Connecticut School of Medicine, Department of Psychiatry, 1987

23. Robins LN, Helzer JE, Croughan J, et al: National Institute of Mental Health Diagnostic Interview Schedule: its history, characteristics, and validity. Archives of General Psychiatry 38:381–389, 1981

24. Saunders JB, Aasland OG, Babor TF, et al: Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption: II. Addiction 88:791–804, 1993

25. Miller WR, Tonigan JS, Longabaugh R: The Drinker Inventory of Consequences (DrInC): An Instrument for Assessing Adverse Consequences of Alcohol Abuse. Test Manual. Project MATCH Monograph Series: Vol 4. Rockville, Md, National Institute on Alcohol Abuse and Alcoholism, 1994

26. Ware JE Jr, Sherbourne CD: The MOS 36-item Short-Form Health Survey (SF-36): I. conceptual framework and item selection. Medical Care 30:473–483, 1992

27. Tarlov AR, Ware JE Jr, Greenfield S, et al: The Medical Outcomes Study: an application of methods for monitoring the results of medical care. JAMA 262:925–930, 1989

28. Current Estimates From the National Health Interview Survey, 1990. Vital and Health Statistics, Series 10: Data From the National Health Survey. Hyattsville, Md, Department of Health and Human Services, 1991

29. Schwamm LH, Van Dyke C, Kiernan RJ, et al: The Neurobehavioral Cognitive Status Examination: comparison with the Cognitive Capacity Screening Examination and the Mini-Mental State Examination in a neurosurgical population. Annals of Internal Medicine 107:486–491, 1987

30. Cottler LB, Robins LN, Helzer JE: The reliability of the CIDI-SAM: a comprehensive substance abuse interview. British Journal of Addiction 84:801–814, 1989

31. McLellan AT, Kushner H, Metzger D, et al: The fifth edition of the Addiction Severity Index. Journal of Substance Abuse Treatment 9:199–213, 1992

32. Babor TF, de la Fuente JR, Saunders J, et al: AUDIT: The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Health Care. Geneva, World Health Organization, 1992

33. Reinert DF, Allen JP: The Alcohol Use Disorders Identification Test (AUDIT): a review of recent research. Alcoholism Clinical and Experimental Research 26:272–279, 2002

34. Robins L, Marcus L, Reich W, et al: NIMH Diagnostic Interview Schedule, Version IV (DIS-IV): Interview and Specifications. St Louis, Mo, Washington University, Department of Psychiatry, 1996

35. Sobell LC, Sobell MB: Timeline followback: a technique for assessing self-reported alcohol consumption, in Measuring Alcohol Consumption. Edited by Litten R, Allen J. Totowa, NJ, Humana, 1992

36. Sherbourne CD, Stewart AL: The MOS social support survey. Social Science and Medicine 32:705–714, 1991

37. Cohen J: Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ, Erlbaum, 1988

38. Cohen J: A coefficient for agreement of nominal scales. Educational and Psychological Measurement 20:37–46, 1960

39. Cohen J: Weighted kappa: nominal scale agreement with provision for scale disagreement or partial credit. Psychological Bulletin 70:213–220, 1960