

# Development of the SBIRT checklist for observation in real-time (SCORE)

Janice A. Vendetti, Bonnie G. McRee & Frances K. Del Boca

UConn Health, School of Medicine, Department of Community Medicine and Health Care, Farmington, CT, USA

## ABSTRACT

**Background and aims** Screening, Brief Intervention and Referral to Treatment (SBIRT) programs have been implemented widely in medical settings, with little attention focused on how well providers adhere to evidence-based service delivery in everyday practice. The purposes of this paper were to: (1) introduce a flexible, relatively simple methodology, the SBIRT Checklist for Observation in Real-time (SCORE), to assess adherence to evidence-based practice and provide preliminary evidence supporting its criterion validity; and (2) illustrate the feasibility and potential utility of the SCORE by analyzing observations of providers within four large-scale SBIRT programs in the United States. **Methods** Eighteen potential adherence judges were trained to recognize SBIRT service elements presented in realistic taped portrayals constructed to serve as criterion coding standards. Across the four SBIRT programs, 76 providers were observed performing 388 services in three types of medical settings; emergency departments ( $n = 10$ ), hospital out-patient/ambulatory clinics ( $n = 16$ ) and hospital in-patient settings ( $n = 5$ ). **Results** Across two exercises, trainees identified 81% of screening and 75% of brief intervention (BI) elements correctly; for the six FRAMES components (Feedback, Responsibility, Advice, Menu of options, Empathy, Self-efficacy), agreement ranged from 69% to 91%. Across programs, 56% of screening, 54% of brief intervention (BI) (81% of FRAMES) and 53% of referral to treatment elements were observed. Programs differed significantly in adherence [screening,  $P = 0.024$ ; BI,  $P < 0.001$ ; FRAMES,  $P < 0.001$ ; referral to treatment (RT),  $P < 0.001$ ]; medical setting differences were minimal. **Conclusions** The Screening, Brief Intervention and Referral to Treatment Checklist for Observation in Real-time provides a flexible method for assessing adherence to evidence-based Screening, Brief Intervention and Referral to Treatment service protocols. Preliminary evidence supports the criterion validity, feasibility and potential utility of the Screening, Brief Intervention and Referral to Treatment Checklist for Observation in Real-time protocol.

**Keywords** Adherence, brief intervention, observation, real-time, SBIRT, screening.

*Correspondence to:* Janice A. Vendetti, UConn Health, School of Medicine, Department of Community Medicine and Health Care, 263 Farmington Avenue, Farmington, CT 06030–6325, USA. E-mail: vendetti@uchc.edu

Submitted 15 April 2015; initial review completed 17 June 2015; final version accepted 7 December 2015

## INTRODUCTION

There has been substantial growth in research on Screening, Brief Intervention and Referral to Treatment (SBIRT) services applied to medical patients at risk for health and other problems because of psychoactive substance use [1,2]. Screening identifies individuals at varying levels of risk; brief intervention (BI) is a time-limited, patient-centered strategy focusing on reducing substance use by increasing insight and awareness [1,3]. Because screening in health-care settings will identify people with suspected substance use disorders who may warrant referral to specialized treatment, referral to treatment (RT) fills the gap between secondary prevention and intensive intervention

(for detailed descriptions of SBIRT services, see [4]). Numerous reviews have established the efficacy and effectiveness of SBI and SBIRT for alcohol misuse [1,2]; however, evidence regarding illicit drug use is more limited and conflicting [1,5–7].

SBIRT programs have been implemented widely in medical settings, and training curricula for providers (e.g. medical staff, behavioral health counselors) have also been developed [8]. However, little attention has focused upon how well SBIRT providers, once trained, adhere to evidence-based service delivery models in everyday (i.e. non-research, non-training) practice. Distractions and unexpected circumstances can require on-the-spot modifications to prescribed protocols, which may compromise

evidence-based practice and, in turn, patient outcomes. If SBIRT is to be used routinely in medical settings, a feasible and psychometrically sound methodology is needed to assess the degree to which providers can implement evidence-based services within the time and environmental constraints that characterize these venues.

Most large-scale SBIRT programs have adopted validated instruments and service protocols; however, these alone are insufficient to deploy and maintain evidence-based practice. Effective SBIRT implementation requires training, ongoing supervision, monitoring and feedback regarding protocol adherence to prevent drift (e.g. [4,9,10]). Although most SBIRT effectiveness studies report that providers were trained and monitored, specific details are usually not reported. Additionally, adherence assessment has focused almost exclusively upon BI, with little attention to screening or RT.

Until recently, there have been no empirically validated instruments for assessing SBIRT adherence [11], i.e. the extent to which specific, pre-defined activities are utilized as specified by evidence-based protocols. Most existing methods measure the fidelity of substance abuse treatments in the context of clinical trials. Fidelity is a multi-faceted construct that, in addition to adherence, subsumes domains such as patient responsiveness and program differentiation [12]. Fidelity protocols typically include complex, multi-dimensional scales reflecting the theoretical underpinnings of the specific treatments tested and require that provider–patient interactions be audio- or video-taped and reviewed by highly trained experts (e.g. Motivational Interviewing Treatment Integrity: MITI [13]; Motivational Interviewing Skills Code: MISC [14]; Motivational Interviewing Supervision and Training Scale: MISTS [15]). These procedures are an accepted standard for assuring the internal validity of clinical trials, but less appropriate for monitoring routine SBIRT service provision. The MD3 SBIRT Coding Scale provides an innovative scheme for evaluating training outcomes and fidelity using standardized patients; however, it is not an observational tool intended for use during everyday service delivery [16].

Finally, a methodology to assess adherence is needed to facilitate SBIRT implementation and dissemination research [11]. Adherence has been identified as an important implementation outcome for SBIRT program evaluation [4]. Reliable and valid measures are required to investigate whether adherence varies across different medical settings that differ in patient volume, flow or other characteristics. Similarly, studies are needed that compare SBIRT programs that may differ in terms of management activities, such as staff hiring, training and monitoring, identified as ‘implementation drivers’ [10] in prior research.

The aims of this paper are: (1) to introduce a flexible, relatively simple and easily implemented methodology, the SBIRT Checklist for Observation in Real-time (SCORE),

to assess adherence and provide preliminary evidence for its criterion validity; and (2) to examine the feasibility and potential utility of this protocol. Criterion validity is assessed by examining how accurately trained observers using the SCORE identify key elements of evidence-based service delivery in highly realistic taped provider–patient interactions, pre-coded by experts for validation purposes. The feasibility and utility of the SCORE are addressed by examining the results of observations of actual service delivery within the largest SBIRT program initiative implemented in the United States. SCORE procedures are considered feasible to the extent that they are acceptable to both patients and providers, and can be implemented easily and unobtrusively by observers during routine practice. Finally, the utility of the protocol is supported to the extent that it can demonstrate adherence to evidence-based practice in SBIRT research; facilitate staff supervision and monitoring; and generate measures that can be used in implementation studies to investigate the relationships between adherence and other program variables.

## METHOD

### SCORE construction and preliminary criterion validation

#### *Protocol development*

The SCORE was developed by the cross-site evaluation team to investigate adherence to evidence-based practice in the third cohort of SBIRT grant recipients funded by the US Substance Abuse and Mental Health Services Administration (SAMHSA) [2]. Checklists were constructed based on evidence-based protocols adopted by the four cohort III programs. Adherence judges mark whether particular service elements are observed (yes), not observed (no), unable to be determined/don’t know (DK) or not applicable (NA). Figure 1 illustrates the checklists for screening, BI and RT. Setting attributes (e.g. ED versus clinic; hospital room versus waiting area) and noteworthy patient conditions (e.g. drowsy, agitated) are recorded on a coversheet. A detailed protocol includes descriptions of the thresholds for meeting item criteria, including appropriate use of the ‘DK’ and ‘NA’ designations (see Supporting information, Appendix).

*Screening.* SBIRT screening begins typically with a small number of ‘pre-screening’ (PS) questions asked by medical staff at intake (e.g. AUDIT-C [17]; single drug question [18]). The PS is a filter for administration of a ‘full’ screening instrument that assesses gradations of risk. All SAMHSA cohort III programs used the Alcohol, Smoking and Substance Involvement Screening Test (ASSIST) [19] for full screening. The SCORE screening checklist was therefore based on this instrument, but it includes items relevant to most validated screening tools (e.g. provides a standardized introduction; asks questions as written).

Screening Components				
Yes	No	DK	NA	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	1. Utilizes full screen (if partial describe in Comments)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	2. Establishes rapport and introduces the session
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	3. Provides a rationale for asking the ASSIST questions
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	4. Addresses confidentiality
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	5. Provides a standardized introduction to assessment
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	6. Defines alcohol and/or drug use parameters
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	7. Defines time window of interest
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	8. Asks questions as written
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	9. Provides Response Card and Drug List to patient
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	10. Uses patient's "language" in defining substances
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	11. Accurately follows skip patterns
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	12. Accurately classifies drugs or standard drinks
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	13. Uses probing techniques to clarify ambiguities
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	14. Scores the assessment accurately
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	15. Accurately categorizes patient risk
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	16. Screen Positive? Score: _____
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Other components: (describe in Comments)
Comments				

Brief Intervention Items				
Yes	No	DK	NA	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	1. Establishes rapport and introduces the session
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	2. Asks to show the patient the screening scores
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	3. Shows screening scores to the patient
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	4. Helps patient identify target substance(s)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	5. Describes the risk levels associated with the scores, and their meaning
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	6. Describes the risks associated with substance use: health, legal, financial, social, etc.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	7. Describes lower-risk drinking guidelines
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	8. Promotes personal responsibility/choice
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	9. Provides advice related to limits of consumption: maintain, reduce, abstain
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	10. Provides a menu of change options
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	11. Reinforces low-risk drinking and/or drug use, if applicable
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	12. Utilizes importance/readiness/confidence rulers, decisional balance, pros/cons
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	13. Helps patient set goals/develop a plan of action
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	14. Provides take-home/resource materials
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	15. Expresses empathy
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	16. Reduces resistance
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	17. Develops discrepancy
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	18. Supports self-efficacy (2+)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	19. Utilizes open-ended questions (2+)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	20. Utilizes affirmations (2+)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	21. Utilizes reflective listening (2+)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	22. Utilizes summary reflections (2+)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	23. Generates change talk (2+)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	24. Avoids lecturing, warning, convincing - asks permission to educate, suggest or advise
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	25. Informs patient about additional BIs/BTs and makes appointment, if applicable
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	26. Closes with summary of session
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Other components: (describe in Comments)
Comments				

Referral to Treatment Components				
Yes	No	DK	NA	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	1. Uses MI techniques to determine patient's interest in additional treatment
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	2. Collaborates with patient to assess preferences for treatment options (e.g., modality, gender specific, schedule, location)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	3. Determines logistical barriers (e.g., insurance, transportation, child care, employment)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	4. Offers on-site SA assessment and treatment (e.g., co-located within medical or other SBIRT host setting)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	5. Facilitates warm handoff to therapist (e.g., sends therapist to patient's bedside, walks patient to therapist's office)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	6. Facilitates telephone conversation between patient and therapist (either in-house or at outside agency) for assessment and treatment
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	7. Facilitates telephone conversation between patient and outside agency for assessment and treatment (i.e., no specific therapist contact)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	8. Makes appointment for patient for further assessment and treatment at in-house department or outside agency
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	9. Provides patient agency contact information (where and when to go for assessment, but no set appointment)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	10. Provides patient with list of available treatment options and contact information in catchment area
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	11. Provides transportation to treatment agency directly from medical setting (e.g., sends patient to treatment agency in taxi or van)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	12. Conducts follow-up with patient, therapist, or agency to determine treatment initiation
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	13. Links patient to peer support to facilitate treatment engagement
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	14. Other components: (describe in Comments)
Comments				

Pre-screen Items from the Screening Checklist: 2, 4-8, 12-15

FRAMES Items from the Brief Intervention Checklist:

Provides personal Feedback: 2, 3, 5, 6

Promotes Responsibility: 8, 17, 23

Provides Advice: 7, 9, 11

Presents a Menu of options: 10, 14, 25

Empathetic counseling style: 15, 16, 19, 21

Supports Self-efficacy: 12, 13, 18, 20

**Figure 1** SCORe adherence checklists for Screening, Brief Intervention and Referral to Treatment

Screening elements are listed in the order in which they are likely to be observed.

*Brief intervention.* Cohort III programs proposed similar BI protocols based on a motivational interviewing (MI) model. Within this framework, attempts are made to explore substance use from the patient's perspective, increase motivation to change and address next steps and goals [20]. Core provider skills (e.g. asking open-ended questions, reflective listening) foster a patient-centered collaborative relationship to promote self-change [21].

All BI protocols incorporated the components summarized by the FRAMES acronym [22,23]: personalized Feedback on the patient's risk for negative consequences of substance use; emphasis on individual Responsibility for making decisions about substance use; clear and straightforward Advice on modifying substance use; a Menu of change options, fostering the patient's involvement in decision-making; an Empathetic provider, who is respectful and non-judgmental; and the nurturing of the individual's Self-efficacy by expressions of optimism that the patient can make positive changes [24]. The FRAMES model combines MI-informed stylistic elements (e.g. empathic counseling style, promoting self-efficacy) with specific BI content-related components (e.g. feedback, responsibility) that are regarded as essential ingredients for the intervention. The SCORE checklist specifically includes both service content elements, activities that are unambiguous and common to BI (e.g. reviews risks related to substance use) and aspects of MI spirit/style (e.g. expresses empathy).

*Referral to treatment.* Because efficacious substance abuse treatments are available, RT goals are to facilitate appropriate treatment program identification, engage the patient and accelerate access to care. Item selection for the SCORE RT checklist was informed by the substance abuse and medical treatment referral literatures that have identified procedures that enhance patient motivation for initiating treatment (e.g. make initial appointments prior to medical discharge) or address barriers to access (e.g. lack of transportation or insurance coverage). The RT checklist comprises a comprehensive list of procedures, although the availability of ideal options may be dependent upon other factors (e.g. level or location of treatment services). The checklist can be tailored to reflect program-specific procedures.

*Training of adherence judges.* The authors developed a 2-day training course. These senior members of the cross-site evaluation team had expertise conducting SBIRT clinical trials and implementation research, having participated in the World Health Organization (WHO) cross-national studies that produced and validated the ASSIST instrument and BI protocols [25] and co-authored MI-based BT manuals [26,27] and training protocols [28,29]. Training encompassed didactic sessions, video and audio demonstrations and role-playing. Standardized service delivery scripts

were designed to replicate real-world scenarios, while purposely including and omitting key elements for training purposes. The authors created pre-coded standards for each exercise.

Trainees completed two certification exercises that provided the data for this report. The first was based on an audiotape of an actual (not staged) BI conducted in a busy dental clinic waiting room. Only the BI was recorded, so trainees were not privy to information that might have been obtained during screening. The second scenario was a professionally produced SBI videotape. SCORE coding forms were completed by the authors to serve as criterion validity standards. After the training work-shop, the instruction protocol was refined, focusing primarily on use of the DK and NA codes.

Eighteen members of the evaluation team were trained. Most were knowledgeable regarding SBIRT; however, educational backgrounds varied (e.g. economics, behavioral health).

#### *Data analyses*

Using Microsoft Excel 2011 (version 14.4.1), percentage agreement to the criterion standard was calculated for each checklist item. Several items were not relevant to the specific taped scenarios (e.g. an item regarding drinking guidelines for a BI targeting marijuana use) and were omitted from the analyses. Three of the original 16 screening items [10,12,16] and four of the original 26 BI items (exercise 1: 1, 7, 26; exercise 2: 4, 7, 26) were excluded. Total agreement scores were based on the percentage of components identified correctly as present or absent across the two certification exercises, the percentage agreement to the BI content components [1–14,25] and percentage agreement to the MI style/spirit elements [15–24]. At least two provider occurrences of the behavior were required to meet the adherence threshold for MI spirit/style elements.

Scores representing each of the FRAMES elements were also constructed. For example, four items reflect personalized feedback [2,3,5,6]. Percentage agreement scores were calculated for each checklist item associated with feedback, as well as for the judges' ability to identify any one of the feedback items present in the criterion standard. The BI checklist items representing each FRAMES element are shown in Fig. 1.

#### **Real-time observations of SBIRT provider–patient service delivery**

##### *Programs, performance sites and participant providers*

The eight judges with the best certification performance visited the cohort III SAMHSA SBIRT programs: Georgia, Missouri, West Virginia and Tanana Chiefs Conference, in Alaska (for program descriptions, see [2]). For reporting

purposes, programs have been assigned numerical identifiers. Immediately prior to the visits, two additional exercises were conducted to ensure that coding refinements were utilized correctly, to minimize drift and to provide additional feedback. Each adherence judge visited multiple programs and different types of medical settings.

The cross-site evaluators worked with program staff to determine which performance sites and providers to observe and to ensure that visits would occur at times representative of typical patient flow and service delivery. In one program, all performance sites were visited. To obtain a representative sample of sites at two programs with geographically dispersed sites, observation locations were selected based on patient flow, type of medical setting, staffing arrangements and patient characteristics. In the remaining program, three sites with very low patient flow were excluded. Providers were shadowed during all hours of operation, including late night/early morning shifts.

Across programs, 76 SBIRT providers (program 1 = 13; program 2 = 14; program 3 = 21; program 4 = 28) were observed at 31 performance sites, representing three types of medical settings: 10 emergency department (ED), 16 hospital out-patient/ambulatory clinic (OP) and five hospital in-patient (IP). Observed providers had varying positions, and included general medical staff, who primarily conducted pre-screens during intake, as well as dedicated, bachelor's- and master's-level SBIRT specialists, many of whom had backgrounds in behavioral health or counseling.

#### Observation procedures

Participating providers were encouraged to go about their day as usual, deviating as little as possible from typical activities. SBIRT providers gave verbal consent to be observed, and no information about individual performance was shared with other program staff. Similarly, patients were asked to give verbal consent for the observation, and no protected health information was recorded. In no instance did SBIRT providers or patients object to being observed although, in very rare cases, service providers requested that observers not be present for particular patient interactions (e.g. immunocompromised patients). There was no evidence that observations interfered with clinical service delivery.

Across programs, 388 separate SBIRT provider-patient interactions were coded. To minimize error, data were entered immediately following observation shifts into a Microsoft Access form replicating the SCORE instrument. Missing data were minimal.

#### Data analyses

IBM SPSS statistics (version 21.0) was used for analyses. Mean count [standard deviation (SD)] and mean

percentage (SD) of PS, screening, BI and RT elements checked were calculated to produce total scores for each service component for each SBIRT provider. For BIs, the percentage of patient interactions where the provider utilized at least one of the items in each FRAMES element was calculated. BI items were also used to create three additional scores: mean count (SD) and mean percentage (SD) of content components, MI spirit/style items and FRAMES elements, utilized during the interaction.

Separate one-way analyses of variance of scores by program and by setting were computed. *Post-hoc* tests using Bonferroni adjusted alpha levels (0.05/n groups) were calculated to identify significant between-group differences. Tamhane's T2 test was used when between-group variance was unequal.

## RESULTS

### Adherence judge certification exercises

Table 1 shows how successfully trainees identified service components and style elements. For screening items, the median percentage agreement with the criterion across the two exercises was 81% (range = 54–100%). All trainees identified correctly: 'Establishes rapport and introduces the session', 'Provides a rationale for asking the ASSIST questions', 'Defines time window of interest' and 'Provides response card and drug list to patient'. Items with the lowest agreement were: 'Provides a standardized introduction to assessment' (56%), 'Defines alcohol and/or drug use parameters' (50%), 'Asks questions as written' (56%) and 'Accurately follows skip patterns' (56%).

For all BI items, the median percentage agreement with the criterion was 75% (range = 59–91%). Percentage agreement for the content components was 81%; high

**Table 1** Certification exercises ( $n = 18$  trainees): median percentage agreement with the screening items, all brief intervention (BI) items, BI content components and motivational interviewing (MI) spirit/style elements.

	Checklist items <sup>a</sup>	% Agreement to the criterion standard
Screening items ( $n = 13$ )	1–16 (omit 10, 12, 16)	81%
All BI items ( $n = 23$ )	1–26 (omit 1, 7, 26 exercise 1; omit 4, 7, 26 exercise 2)	75%
BI content components ( $n = 13$ )	1–14, 25 (omit 1 and 7 in exercise 1; omit 4 and 7 in exercise 2)	81%
MI spirit/style elements ( $n = 10$ )	15–24	70%

<sup>a</sup>Refer to Fig. 1 for items. SCORE = Screening, Brief Intervention and Referral to Treatment Checklist for Observation in Real-time.

agreement items included: 'Shows screening scores to the patient' (100%), 'Describes the risks associated with substance use' (94%), 'Promotes personal responsibility/choice' (94%) and 'Provides advice related to limits of consumption' (94%). Percentage agreement for MI spirit/style was lower at 70%.

Mean percentage agreement varied for the FRAMES elements. Across the two exercises, the percentage agreement with the criterion for any feedback item was highest at 91%. agreement with any responsibility item was 73%; advice, 71%; menu of options, 79%; empathy, 69% and self-efficacy, 69%.

### Observations of SBIRT service delivery of SAMHSA cohort III programs

Table 2 summarizes the 388 observations (161 pre-screens, 118 full-screens, 97 BIs, 12 RTs) that occurred during visits to the SAMHSA SBIRT programs. Most observations (51%) took place in EDs, with 37% in OP and 6% in IP; 6% of coded interactions were not live observations, but pre-recorded tapes used initially for on-site quality assurance.

As shown in Table 3, 50% of BI content and 61% of MI style/spirit elements were checked. The FRAMES composite measure was notably higher (81%). Additionally, 85% of the interactions included at least one feedback element and 76% at least one responsibility item. Advice was provided in 73% and a menu of options presented in 77% of interactions. The most widely utilized element was empathy (97% of sessions), while self-efficacy was observed in 80% of the interactions. Of the 97 BIs observed, in only 1% were fewer than two FRAMES elements present. All six elements were utilized in 40%, and an additional 32% contained five elements.

#### SBIRT program differences in adherence

Significant program differences were found on all seven adherence measures (Table 3). With the exception of PS, program 3 had the highest scores across indices, although differences between it and other programs were not always statistically significant in pairwise comparisons.

Conversely, program 1, with the highest PS score, had the lowest means for all other measures. Although program 1 did not differ significantly from the others in all comparisons, differences between programs 1 and 3 were statistically significant across the seven measures.

#### Medical setting differences in adherence

One-way analyses of variance (ANOVAs) indicate setting differences only for PS (Table 4). ED providers utilized an average of 65% of PS elements (SD = 26.5%), compared to 45% (SD = 26.0%) in OP and 50% (SD = 14.1%) in IP ( $F_{(2,158)} = 10.78, P < 0.001$ ). *Post-hoc* comparisons using Bonferroni-adjusted alpha levels of 0.0167 (0.05/3) indicate a significant difference between EDs and OPs. For RT, providers utilized 57% of elements (SD = 30.6%) on average. Small numbers of RT observations did not allow for meaningful comparisons.

## DISCUSSION

Trained adherence judges observed a large number of actual service delivery sessions at four programs within the most ambitious SBIRT program implemented in the United States. Although there are not yet established thresholds for evaluating the adequacy of evidence-based SBIRT practice, 50% of BI content and 61% of MI style/spirit, elements were delivered across programs. Further, FRAMES elements, an essential component in most SBIRT protocols, were highly prevalent. The FRAMES composite measure was 81%; all six FRAMES elements were utilized in 40% of observed BIs and an additional 32% contained five. These results suggest that SBIRT practice within the SAMHSA cohort III programs was indeed evidence-based.

The results also support the criterion validity, feasibility and utility of the SBIRT SCORE as a methodology for measuring adherence to evidence-based SBIRT protocols. Judges with differing educational and professional backgrounds identified correctly SBIRT service elements presented in highly realistic taped portrayals serving as objective criterion validity standards. Because this finding is limited to a particular set of judges and service scenarios,

**Table 2** Services observed by program.

	Observed provider/patient interactions n (%)	Pre-screen <sup>a</sup> n (%)	Full screen n (%)	Brief intervention n (%)	Referral to treatment n (%)
Program 1	115 (29.6)	65 (40.4)	31 (26.3)	17 (17.5)	2 (16.7)
Program 2	31 (8.0)	6 (3.7)	16 (13.6)	8 (8.2)	1 (8.3)
Program 3	149 (38.4)	28 (17.4)	54 (45.7)	60 (61.9)	7 (58.3)
Program 4	93 (24.0)	62 (38.5)	17 (14.4)	12 (12.4)	2 (16.7)
Total	388 (100)	161 (100)	118 (100)	97 (100)	12 (100)

<sup>a</sup>Pre-screens represent pre-screens only. If a full-screen was also observed, the observation was categorized as a full-screen.

Table 3 One-way analysis of variance of checklist item counts and percentages utilized by program.

	Total		Program 1		Program 2		Program 3		Program 4		Significance	
	n	Mean % (SD)	n	Mean % (SD)	n	Mean % (SD)	n	Mean % (SD)	n	Mean % (SD)	F	P
Pre-screen items utilized <sup>a</sup> [10]	161	56 % (27.8 %)	65	70 % (16.1 %)	6	50 % (21.0 %)	28	30 % (28.6 %)	62	53 % (28.8 %)	18.51	0.000
Full screen items utilized <sup>b</sup> [15]	118	56 % (21.3 %)	31	48 % (24.7 %)	16	52 % (21.6 %)	54	60 % (19.3 %)	17	64 % (15.8 %)	3.28	0.024
BI: all items utilized <sup>c</sup> [21]	97	54 % (22.7 %)	17	27 % (15.6 %)	8	52 % (16.2 %)	60	64 % (18.7 %)	12	47 % (21.0 %)	18.30	0.000
BI: content components utilized <sup>d</sup> [13]	97	50 % (23.6 %)	17	23 % (15.1 %)	8	51 % (16.4 %)	60	59 % (19.6 %)	12	42 % (25.1 %)	16.22	0.000
BI: MI spirit/style items utilized <sup>e</sup> [8]	97	61 % (27.3 %)	17	34 % (21.1 %)	8	53 % (24.8 %)	60	71 % (25.2 %)	12	56 % (18.8 %)	11.09	0.000
BI FRAMES elements utilized <sup>f</sup> [6]	97	81 % (21.5 %)	17	60 % (27.7 %)	8	88 % (11.8 %)	60	88 % (15.1 %)	12	74 % (23.0 %)	11.28	0.000
RT: All items utilized [13]	12	53 % (32.4 %)	2	38 % (NA)	1	100 % (NA)	7	56 % (36.0 %)	2	31 % (NA)		NA

<sup>a</sup>Pre-screen: differences occur between programs 1, 3 and 4. <sup>b</sup>Full screen: differences occur between programs 1 and 3. <sup>c</sup>Brief intervention (BI) all items: differences occur between program 1 and all others; programs 3 and 4. <sup>d</sup>BI: content components: differences occur between programs 1 and 3; programs 3 and 4; programs 1 and 2. <sup>e</sup>BI: motivational interviewing (MI) spirit/style items: differences occur between programs 1 and 3. <sup>f</sup>BI: feedback, responsibility, advice, menu of options, empathy, self-efficacy (FRAMES) elements: differences occur between programs 1 and 3; programs 1 and 2. SD = standard deviation; NA = not applicable.

more research is needed to document further the reliability and validity of the SCORE.

In terms of feasibility, adherence judges observed a large number of actual service delivery sessions at four SBIRT programs. When applied during observations of live, fast-paced, ongoing service delivery, the SCORE successfully captured content-specific elements for screening, BI and RT and MI style. Differences in the number of observations across types of settings reflect the actual variations among them in terms of patient volume and flow (e.g. more patients were served in EDs than IP units). Providers and patients accepted readily the presence of observers who coded SBIRT sessions in different medical settings under varied conditions.

The feasibility of the SCORE is enhanced by its flexibility. Although they contain elements regarded as universally applicable across SBIRT programs, the forms used for the present report were tailored somewhat to protocols of the observed SAMHSA programs. Modifications in item phrasing and/or the addition or removal of items will probably be needed for the SCORE to become an effective training, monitoring and program evaluation tool in other venues. Alterations are also expected as the dominant BI approaches (e.g. MI) evolve over time and the critical ingredients for positive patient outcomes are identified.

The potential utility of the SCORE is twofold. First, effective SBIRT implementation requires routine monitoring of providers and feedback to prevent drift (e.g. [9,10]). The observation procedures are similar to 'shadowing', often used to evaluate provider performance. The SCORE offers a structured method for this activity that provides immediate feedback and facilitates systematic comparisons across time and service providers. It is unlikely that supervisory staff, already familiar with SBIRT, would require the intensive training provided to cross-site evaluation staff.

Secondly, the SCORE offers considerable promise as a tool for program evaluation and implementation science. As noted, the SCORE was used successfully to examine adherence in the evaluation of SAMHSA's SBIRT programs. Additionally, the SCORE proved sensitive to differences in service component utilization among the SAMHSA programs. With the exception of PS, which is performed typically by medical rather than dedicated SBIRT staff, one program performed consistently better than the others. It is noteworthy that this program had also established the most rigorous staff training, supervision and monitoring procedures.

Despite variations in patient volume and flow, as well as the urgency of patients' presenting problems, medical setting differences in adherence were limited. The only significant difference was for PS, where ED providers utilized more evidence-based components than those in OP settings. Collectively, these results suggest that SBIRT providers can contend well with the often-hectic conditions

**Table 4** One-way analysis of variance of checklist item counts and percentages utilized by setting.

	<i>Total</i>		<i>ED</i>		<i>OP</i>		<i>IP</i>		<i>Significance</i>	
	<i>n</i>	<i>Mean % (SD)</i>	<i>n</i>	<i>Mean % (SD)</i>	<i>n</i>	<i>Mean % (SD)</i>	<i>n</i>	<i>Mean % (SD)</i>	<i>F</i>	<i>P</i>
Pre-screen items utilized <sup>a</sup> [10]	161	56 % (27.8 %)	89	65 % (26.5 %)	70	45 % (26.0 %)	2	50 % (14.1 %)	10.78	0.000
Full screen items utilized [15]	115	57 % (21.5 %)	57	58 % (20.0 %)	49	55 % (23.4 %)	9	57 % (20.8 %)	.309	0.735
BI: all items utilized [21]	76	50 % (23.0 %)	47	51 % (24.0 %)	21	46 % (21.6 %)	8	53 % (21.8 %)	.469	0.627
BI: content components utilized [13]	76	47 % (24.6 %)	47	49 % (25.0 %)	21	42 % (24.3 %)	8	49 % (24.6 %)	.496	0.611
BI: MI spirit/style items utilized [8]	76	55 % (25.3 %)	47	56 % (27.1 %)	21	52 % (22.8 %)	8	59 % (22.9 %)	.308	0.736
BI FRAMES elements utilized [6]	76	79 % (23.0 %)	47	80 % (23.7 %)	21	75 % (23.3 %)	8	81 % (18.8 %)	.348	0.707
RT: all items utilized [13]	11	57 % (30.6 %)	5	49 % (21.8 %)	3	59 % (35.5 %)	3	67 % (39.5 %)		NA

<sup>a</sup>Pre-screen: differences occur between emergency department (ED) and out-patients (OP). FRAMES = feedback, responsibility, advice, menu of options, empathy, self-efficacy; IP = in-patients. SD = standard deviation; NA = not applicable; RT = referral to treatment.

in EDs and exhibit levels of adherence comparable to those observed in other settings.

Finally, this study did not examine the relationship between adherence and patient outcomes. However, outcome data for SAMHSA's cohort III programs indicate that BI adherence (number of elements utilized, aggregated across providers and extrapolated to the program-level) was associated with significant reductions in alcohol use [30]. A similar result was obtained for another SAMHSA SBIRT program in which the SCORE was used: provider adherence to BI content components predicted declines in binge drinking significantly 6 months later [31]. These results provide evidence that supports the predictive validity of the SCORE.

There are several caveats regarding the SCORE methodology and the evaluation of SAMHSA's SBIRT programs. First, because the procedure was designed for ease of use, it assesses adherence, and does not necessarily measure competence or how skillfully particular elements are delivered. Although competence assumes adherence, it may affect outcomes separately [11]. Secondly, the brief treatment component of SAMHSA-funded SBIRT programs was not evaluated. Brief treatment is conducted typically in sessions resembling traditional counseling, for which other fidelity measures are more appropriate. Thirdly, it is possible that the observed interactions were not representative of usual practice, or that the more adherent providers participated in the evaluation. However, SBIRT performance sites were sampled systematically, and almost all the providers from selected venues participated. Fourthly, although observers coded a large number of SBIRT sessions, relatively few involved RT, limiting our ability to make meaningful comparisons for that service component, and the number of IP interactions was relatively small. Finally, providers may have been more attentive to aspects of the SBIRT protocols than was typical. However, observers were very

unobtrusive, emphasized the importance of performing all tasks as usual, and offered assurances that no observation data would be shared with other staff.

Despite these limitations, the findings indicate that the SCORE offers a promising methodology that can fill a significant void. In addition to the important substantive findings regarding evidence-based practice in SAMHSA's SBIRT programs, the present study provides evidence to support its criterion validity, feasibility and utility.

#### Declaration of interests

None.

#### Acknowledgements

This research was funded by US Department of Health and Human Services contract number HHSS283200700002I. The authors acknowledge the efforts and support from the Substance Abuse and Mental Health Services Administration (SAMHSA), Center for Substance Abuse Treatment (CSAT), as well as administrators and staff from the third cohort of SAMHSA-funded Screening, Brief Intervention and Referral to Treatment (SBIRT) programs. Additionally, the authors thank other members of the SBIRT cross-site evaluation team for their assistance with data collection, analysis and review of the manuscript.

#### References

1. Babor T. F., McRee B. G., Kassebaum P. A., Grimaldi P. L., Ahmed K., Bray J. Screening, Brief Intervention, and Referral to Treatment (SBIRT): toward a public health approach to the management of substance abuse. *Subst Abuse* 2007; 28: 7–30.
2. Bray J., Del Boca E., McRee B., Hayashi S., Babor T. Screening, brief intervention and referral to treatment (SBIRT): rationale,



- program overview and cross-site evaluation. *Addiction*; 112 (Suppl. 2): 3–11.
3. Substance Abuse and Mental Health Services Administration (SAMHSA). Screening, brief intervention and referral to treatment (SBIRT) in behavioral healthcare. 2011. Available at: [http://samhsa.gov/sites/default/files/sbirtwhitepaper\\_0.pdf](http://samhsa.gov/sites/default/files/sbirtwhitepaper_0.pdf) (accessed 11 April 2015) (Archived at <http://www.webcitation.org/6cnMZYG59>).
  4. Del Boca E., McRee B., Vendetti J., Damon D. The SBIRT program matrix: a conceptual framework for program implementation and evaluation. *Addiction* 2017; 112 (Suppl. 2): 12–22.
  5. Madras B. K., Compton W. M., Avula D., Stegbauer T., Stein J. B., Clark H. W. Screening, brief interventions, referral to treatment (SBIRT) for illicit drug and alcohol use at multiple healthcare sites: comparison at intake and 6 months later. *Drug Alcohol Depend* 2009; 99: 280–95.
  6. Roy-Byrne P., Bumgardner K., Krupski A., Dunn C., Ries R., Donovan D. *et al.* Brief intervention for problem drug use in safety-net primary care settings: a randomized clinical trial. *JAMA* 2014; 312: 492–501.
  7. Saitz R., Palfai T. P., Cheng D. M., Alford D. P., Bernstein J. A., Lloyd-Travaglini C. A. *et al.* Screening and brief intervention for drug use in primary care: the ASPIRE randomized clinical trial. *JAMA* 2014; 312: 502–13.
  8. Pringle J., Kowalchuk A., Meyers J., Seale P. Equipping residents to address alcohol and drug abuse: the national SBIRT residency training project. *J Grad Med Educ* 2012; 4: 58–62.
  9. Chaudoir S. R., Dugan A. G., Barr C. H. Measuring factors affecting implementation of health innovations: a systematic review of structural, organizational, provider, patient, and innovation level measures. *Implement Sci* 2013; 8: 22.
  10. Fixsen D., Naoom S., Blase K., Friedman R., Wallace E. *Implementation Research: A Synthesis of the Literature*. University of South Florida, Tampa, FL: Louis de la Parte Florida Mental Health Institute; 2005.
  11. Pantalon M. V., Martino S., Dziura J., Li F., Owens P. H., Fiellin D. A. *et al.* Development of a scale to measure practitioner adherence to a brief intervention in the emergency department. *J Subst Abuse Treat* 2012; 43: 382–8.
  12. Dusenbury L., Brannigan R., Falco M., Hansen W. B. A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Educ Res* 2003; 18: 237.
  13. Moyers T. B., Martin T., Manuel J. K., Hendrickson S. M., Miller W. R. Assessing competence in the use of motivational interviewing. *J Subst Abuse Treat* 2005; 28: 19–26.
  14. Moyers T., Martin T., Catley D., Harris K. J., Ahluwalia J. S. Assessing the integrity of motivational interviewing interventions: reliability of the motivational interviewing skills code. *Behav Cogn Psychother* 2003; 31: 177–84.
  15. Madson M. B., Campbell T. C., Barrett D. E., Brondino M. J., Melchert T. P. Development of the motivational interviewing supervision and training scale. *Psychol Addict Behav* 2005; 19: 303.
  16. DiClemente C. C., Crouch T. B., Norwood A. E. Q., Delahanty J., Welsh C. Evaluating training of screening, brief intervention, and referral to treatment (SBIRT) for substance use: Reliability of the MD3 SBIRT Coding Scale. *Psychol Addict Behav* 2015; 29: 218–24.
  17. Bush K., Kivlahan D. R., McDonnell M. B., Fihn S. D., Bradley K. A. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. *Arch Intern Med* 1998; 158: 1789–95.
  18. Smith P. C., Schmidt S. M., Allensworth-Davies D., Saitz R. A single-question screening test for drug use in primary care. *Arch Intern Med* 2010; 170: 1155–60.
  19. Humeniuk R., Henry-Edwards S., Ali R., Poznyak V., Monteiro M. *The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): Manual for Use in Primary Care*. Geneva: World Health Organization; 2010.
  20. Center for Substance Abuse Treatment (CSAT). Brief Interventions and Brief Therapies for Substance Abuse. Series 34 (DHHS Publication no. (SMA) 99–3353). Rockville, MD: CSAT; 1999.
  21. Miller W. R., Rollnick S., editors. *Motivational Interviewing: Helping People Change*, 3rd edn. New York: Guilford Press; 2013.
  22. Bien T. H., Miller W. R., Tonigan J. S. Brief interventions for alcohol problems: a review. *Addiction* 1993; 88: 315–35.
  23. Miller W. R., Sanchez V. C. Motivating young adults for treatment and lifestyle change. In: Howard G. S., Nathan P. E., editors. *Alcohol Use and Misuse by Young Adults*. Notre Dame, IN: University of Notre Dame Press; 1994, pp. 55–81.
  24. Hester R. K., Miller W. R. *Handbook of Alcoholism Treatment Approaches*, 2nd edn. Boston, MA: Allyn & Bacon, Incorporated; 1995.
  25. Humeniuk R., Ali R., Babor T., Souza-Formigoni M. L. O., de Lacerda R. B., Ling W. *et al.* A randomized controlled trial of a brief intervention for illicit drugs linked to the Alcohol, Smoking and Substance Involvement Screening Test (ASSIST) in clients recruited from primary health-care settings in four countries. *Addiction* 2012; 107: 957–66.
  26. Steinberg-Gallucci K., Damon D., McRee B., editors. Connecticut SBIRT Brief Treatment: Let's Play. Hartford, CT: CT Department of Mental Health and Addiction Services; 2012.
  27. Steinberg K. L., Roffman R. A., Carroll K. M., McRee B., Babor T. F., Miller M. *et al.* editors. *Brief Counseling for Marijuana Dependence: A Manual for Treating Adults*. Rockville, MD: Center for Substance Abuse Treatment, Substance Abuse and Mental Health Services Administration; 2005.
  28. McRee B., Granger J. Implementing Screening and Brief Intervention Services for Smokers and At-risk Drinkers, A Training Package for Primary Care Practices. 2016-11-14. Available at: <http://www.prescriptionforhealth.org/results/toolkit.html> (accessed 14 November 2016) (Archived at <http://www.webcitation.org/6m1JhEec8>).
  29. McRee B., Vendetti J., Damon D. Connecticut SBIRT Program Protocol. Grant no. 1U79T1023459–01. Hartford, CT: CT Department of Mental Health and Addiction Services; 2013.
  30. RTI International. *Screening, Brief Intervention and Referral to Treatment (SBIRT) cross-site evaluation*. Research Triangle Park, NC: RTI International; 2014.
  31. McRee B., Vendetti J. Application of the SCORe methodology to assess the effect of SBIRT program adherence on binge-drinking outcomes. [Abstract]. *Association for Medical Education and Research on Substance Abuse*, Book of Abstracts, 2, 2014.

### Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

**Appendix S1** SBIRT Checklist for Observation in Real-time (SCORE) Coversheet.

**Appendix S2** SBIRT Checklist for Observation in Real-time (SCORE) Protocol.

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.