Substance use outcomes of patients served by a large US implementation of Screening, Brief Intervention and Referral to Treatment (SBIRT)

By: Arnie Aldridge, Robyn Linford, and Jeremy W. Bray

This is the peer reviewed version of the following article:

Aldridge, A., Linford, R., Bray, J. (2017). Substance use outcomes of patients served by a large US implementation of Screening, Brief Intervention and Referral to Treatment (SBIRT). *Addiction*, 112: 43– 53. doi: 10.1111/add.13651.

which has been published in final form at https://doi.org/10.1111/add.13651.

Abstract:

Aims: To estimate changes in the substance use behaviors of patients who received services as part of the US Substance Abuse and Mental Health Services Administration's (SAMHSA) Screening, Brief Intervention and Referral to Treatment (SBIRT) grant program. **Methods:** We use a pre–post design and performance monitoring data collected by SBIRT organizations. For a sample of 17 575 patients, we compare pre-SBIRT substance use with substance use 6 months after receipt of SBIRT services. SBIRT's correlation with changes in substance use was estimated using generalized linear mixed models to account for the clustering of patients within health-care facility and US state. **Results:** From pre- to post-SBIRT we found large and statistically significant decreases for almost every measure of substance use. Model-adjusted means indicate that the prevalence of alcohol use was lower 6 months later by 35.6%, heavy drinking by 43.4% and illicit drug use by 75.8%. Greater intervention intensity was associated with larger decreases in substance use. The study design does not support causal conclusions and estimated decreases in reported substance use are due, at least in part, to a well-known set of confounders and natural substance use patterns that may be unrelated to any particular SBIRT intervention. **Conclusions:** Compared with previously published findings on the Screening, Brief Intervention and Referral to Treatment grant program, our estimates of substance use reduction were smaller, but still consistently large in absolute magnitude and within ranges of estimates from past trials of Screening, Brief Intervention and Referral to Treatment.

Keywords: brief intervention | brief therapy | brief treatment | illicit drugs | moderator | SBI | SBIRT

Article:

Introduction

In 2003, the Substance Abuse and Mental Health Services Administration (SAMHSA) established the Screening, Brief Intervention and Referral to Treatment (SBIRT) grant program to screen patients systematically in medical settings and, based on screening results, provide brief intervention (BI), brief treatment (BT) or referral to treatment (RT) as appropriate. SAMHSA's SBIRT program is the largest SBIRT dissemination effort ever undertaken in the United States. In the initial cohort of projects, six states and one tribal organization were awarded 5-year grants to promote the adoption and sustained implementation of SBIRT in a variety of medical settings. The goals of these grants were to expand the continuum of care for substance use disorders, including hazardous, harmful and dependent use, and to promote the integration of substance use disorder care into the traditional medical care community.

In the only published study of patient outcomes for SAMHSA's SBIRT initiative, Madras *et al*.[1] reported preliminary outcomes for six of the first seven SAMHSA SBIRT projects. In analyses performed separately for each project, Madras *et al*. found that past-month substance use outcomes were dramatically lower 6 months after exposure to SBIRT services than they were at the time of SBIRT screening. In addition to being the first published, multi-site assessment of SAMHSA's SBIRT program, Madras *et al*. received attention because it has provided novel evidence for areas in which SBIRT has been less well studied; namely, patients screened in emergency departments and in-patient settings as opposed to primary care settings, as well as patients screened for illicit substance use.

Despite having the largest sample of patients ever examined in an SBIRT study (459 599), the Madras *et al*. findings have several limitations noted by other SBIRT researchers[2-4]. The primary criticism is the study's lack of a rigorous design to eliminate alternative explanations of patient improvements. Such alternative mechanisms of change following SBIRT are an important and unresolved area of research. Screening and assessment alone may produce changes[5, 6]. Study elements can be a factor and were in place to some extent in all the SBIRT projects. Informed consent procedures, the need to provide contact information for follow-up contacts and the desire to meet outcome expectations or please a researcher might influence treatment and control patients[7-9]. Regression to the mean[10, 11] probably explains a large portion of 'improvement' seen in all patients who qualified for the study based on a relatively high screening score. In addition to these potential problems, Saitz[3] notes that the magnitude of improvement in outcomes reported by Madras *et al*. are unrealistic and are five times the best absolute outcomes reported for randomized controlled trials (RCTs) (e.g. 50 versus 12% for alcohol in a primary care setting). Other criticisms include possible selection bias due to inconsistent follow-up rates across projects and the lack of evidence of receipt of SBIRT services, which undermines the validity of an intent-to-treat framework.

In addition, we note several criticisms of the Madras *et al*. study that have not been described in the literature. First, the average substance use of the full program population at baseline was compared to the average substance use of a follow-up sample representing fewer than 10% of the baseline population, substantially increasing the potential for selection bias. Secondly, the statistics reported were unconditional means that did not control for potential confounding

factors available in the data. Thirdly, the estimates and their statistical significance did not take into account the variation and clustering across projects, which could lead to bias and Type I errors.

This paper re-visits the Madras *et al*. analysis for two reasons. First, an analysis of the SAMHSA SBIRT grant program using all seven cohort I projects and covering the entire period of program implementation is needed to provide a full account of the initiative's outcomes. Secondly, while Madras *et al*. cannot be used as positive evidence for SBIRT's effectiveness[12, 13], we attempt to address some of the criticisms of the Madras *et al*. analyses and to extend them in several key areas. Specifically, we use multi-level models on matched pre–post patients to account properly for the clustering of observations and to explore possible bias from sample selection and attrition. We reiterate that because of the pre–post design, no causal inference can be made about SBIRT from these analyses. We do, however, provide context for the pre–post changes in the 'SBIRT-treated' groups by comparing them to the pre–post results for treated groups in previous RCTs.

Methods

Data

Data were obtained from the Services and Accountability Improvement System (SAIS), which collects data required under the Government Performance and Results Act (GPRA) for all Center for Substance Abuse Treatment (CSAT) discretionary grantees. SBIRT cohort I projects were required to collect select GPRA data from all patients receiving SBIRT services at each of their sites. Patients entering a site were eligible for screening if they were awake, medically stable, not in severe distress and not intoxicated. All patients screened by the cohort I projects between January 2004 and March 2010 were considered for our analyses. The projects submitted data on 754 525 adult patients, 171 921 of whom screened positive for hazardous or harmful use and were recommended to BI (n = 118 323), BT (n = 24 071) or RT (n = 29 527). Recommendations to these service levels were based on patients' screening scores and on the clinical discretion of the providers administering the screens.

Each individual SBIRT program developed its own SBIRT model. Models varied due to time and space constraints of the medical settings and due to providers' willingness to provide and patients' acceptance of specific SBIRT services. Evidence-based screening instruments were primarily the Alcohol Use Disorders Identification Test (AUDIT) for alcohol and Drug Abuse Screening Test (DAST) for illicit drug use, which offered tested and validated service recommendations based on patient risk. Five of the programs adopted pre-screening questions. Although the majority of SBIRT services were provided by specialists (e.g. social workers, substance abuse counselors, health educators) who were associated with the SBIRT project, pre-screening and screenings were also conducted at some sites by a combination of medical generalists (e.g. nurses) and self-administration.

For BIs, all projects used a motivational interviewing (MI)[14, 15] style and all used elements of the FRAMES model[16] (Feedback regarding personal risk or impairment, emphasis on personal Responsibility for change, clear Advice to change, providing a Menu of change options,

Empathy as a counseling style and enhancement of Self-efficacy or optimism). All projects incorporated MI techniques and the transtheoretical model of change[17] in their BT service. Two projects used cognitive–behavioral therapy (CBT)[18, 19] and two projects used a community reinforcement approach (CRA)[20]. Projects offered both on-site and telephonic BT sessions. All projects enhanced their process of referral to specialty treatment with MI techniques, and explicit BI sessions were often provided at the time of the referral.

Within the BI, BT and RT service categories, at least a 10% sample was selected randomly for follow-up by the projects, with some projects choosing higher percentages throughout the course of their program. Randomization was based on the patient's Social Security Number (SSN) or birthdate. Follow-up interviews were conducted between 150 and 240 days after baseline screening and were administered either in person or by telephone. As noted in Madras *et al.*[1] most were by telephone because projects found it difficult to have patients return to the original site after 6 months.

Measures

The GPRA data collection tool (http://www.samhsa.gov/grants/CSAT-GPRA/services.aspx) recorded SBIRT screening results, demographic information, substance use, family and living conditions, education, employment and income, crime and criminal justice involvement, mental and physical health problems and discharge status, including SBIRT services planned. Assessment items in the instrument were based in part on the Addiction Severity Index-Lite Version[21]. Substance use questions were the days used during the past 30 days for alcohol, alcohol to intoxication ('heavy drinking days' or HDD), illicit substances and both alcohol and illicit substances on the same day. For dependent variables in our statistical analyses, we use indicator variables for at least 1 day of use of a substance in the past 30 days ('any use') and counts of days of use for each substance conditional on any use (i.e. from 1 to 30 days).

The complete GPRA data collection instrument was not administered to all patients. For patients screening negative, only the demographic section (gender, age, race/ethnicity) was completed and no follow-up interview was conducted. For those recommended to BI, only the demographic, substance use and planned services sections were completed at baseline and at follow-up (for those patients in the follow-up sample). Patients recommended to either BT or RT were administered all GPRA sections at baseline and at follow-up (for those patients in the follow-up sample). A patient and grantee identifier was also included in the GPRA data.

Design

The study design is a pre–post comparison of outcomes within an individual patient who was screened and recommended to one of the screen-positive (SP) categories. Assignment to service recommendation categories is based on screener scores and provider discretion rather than on actual service received. Therefore, the analysis is of the intent-to-treat influence of SBIRT rather than the exact services received by each patient. The number of services that they would ultimately receive is a function of patients' substance use, possible initial response to the screen or first intervention and other unobserved characteristics that might be correlated with their acceptance of services and retention. Thus, estimates of the dose–response of SBIRT services

would be biased without some model or control for these confounding factors, which is beyond the scope of this study.

Data analysis

Analyses began with a comparison of the full SP sample at baseline to the sample selected for follow-up and to the sample actually completing a follow-up interview using simple descriptive statistics. Madras *et al*.[1] noted that response rates were particularly low in two projects, one due to program interruption and the other to a significant shift in the program's follow-up model. We therefore calculated alternative response rates that were adjusted based on the timing of these events. Analyses were conducted for all SP patients and by SBIRT service recommendation (BI, BT or RT).

Next, we estimated the change in substance use outcomes over time associated with SBIRT using two specifications of generalized linear mixed models (GLMM). Measures of whether or not a substance (e.g. alcohol, heavy alcohol, illicit drug) was used are discrete outcomes, and so were modeled assuming a binomial distribution and a probit link function. Measures of the number of days each type of substance was used were analysed using negative binomial models. Reflecting the repeated-measures design, we included individual random-effects in both types of models to estimate properly the sampling variance in clustered data. Clustering within each project was modeled using nested random-effects. Hausman tests were used to determine whether random-effects failed to meet the assumption of independence with respect to other covariates. When this occurred, we used fixed-effects instead and noted such in the results.

The main independent variable of interest is an indicator of whether an observation is from the 6-month follow-up interview. We controlled for gender, race and ethnicity, age, baseline educational attainment and household income. Because the data for BI patients were limited, models that included the sample of patients recommended for BI only controlled for gender, race and ethnicity and age. Analyses were conducted for all SP patients and separately by SBIRT service recommendation (BI, BT or RT).

Comparison with results from other studies

SAMSHA's SBIRT project did not include a control group, and hence effectiveness cannot be tested using project data; in our re-analysis of the data, we compare the project data to data from other studies. We used a review of 32 controlled trials of BIs targeting alcohol problems[16]. We selected the studies because they involved a BI group and a control group, and each had a calculated within- and between-group effect. Although there are larger and more recent reviews of SBIRT trials, Bien *et al*.[16] was the only one that reported the within-patient changes for ‘treated’ patients, which are design-equivalent measures to our sample outcomes. In total, these trials examined more than 6000 problem drinkers from 14 different nations. The average time between baseline and follow-up observations was 22 months, and follow-ups occurred between 2 and 120 months after baseline, depending on the study. The between-group effects compared changes in alcohol use among patients who received BI with changes among a control group of patients. All but one of the between-group effects reported in Bien *et al*. were positive and 10 were significant.

Because many different alcohol use measures were employed in these studies, Bien *et al*. calculated standardized effect sizes (Cohen's *d*) of the changes in alcohol use measures after patients received BI. While our alcohol use measures differed from most reported in Bien *et al*., we constructed analogous standardized effect sizes using our alcohol use measures and compared these effects with those reported by Bien *et al*.

Results

Response rate analysis

This paper uses the 17 575 individuals with valid baseline and follow-up information. The overall response rate among the randomly selected SBIRT screens was 47.5%. Follow-up rates, however, varied considerably across the projects—87.6, 76.1, 73.3, 69.4, 62.2, 41.7 and 28.5%. Table 1 shows how these follow-up rates change when they are calculated only for SBIRT participants seen during periods when programs were fully functional; Madras *et al*.[1] suggest that interruptions in full functionality are a primary cause of low response rates. The response rate of the project with the second lowest response rate (41.7%) increases dramatically to 74.9% when service interruption periods are removed from the data. The lowest response rate, however, increases only marginally from 28.5 to 38% when we remove these periods from the data. Moreover, this project had almost half the total eligible patients and thus has a large influence on the overall follow-up rate. When removed from the calculation, the overall response rate is 75.1%.

Sample descriptions

Table 2 presents patient demographic and substance use information for patients used in the analyses. For each service category (all SP, BI, BT and RT), the first column presents descriptive information for all patients in that category and subsequent columns present information for the subsample of patients who have both a baseline and a follow-up observation (i.e. the matched sample). Patients screening positive for SBIRT were, on average, middle-aged (mean age = 37.5) and more likely to be male than female. SP patients were primarily white (53.4%), black (24.6%) or Hispanic (17.4%). Among those for whom educational attainment and income were recorded (BT and RT patients), approximately 60% were high school graduates. As a whole, BT patients reported $1648 in total monthly household income, and RT patients reported $1514. Males reported higher monthly income than females ($324, $P < 0.01$). Patients with both a baseline and a follow-up observation (hereafter the matched sample) reported lower income than the full baseline samples.

Table 1. Baseline eligibility and response rates by project.

	Screen positive					Brief intervention					Brief treatment					Referral to treatment				
	Eligible at BL		Completed FU			Eligible at BL		Completed FU			Eligible at BL		Completed FU			Eligible at BL		Completed FU		
Project	Raw	adj.	#	%		Raw	adj.	#	%		Raw	adj.	#	%		Raw	adj.	#	%	
5	2507		1907	76.1		1124		878	78.1		363		270	74.4		1020		759	74.4	
6	17 546	13 179	5007	28.5	38.0	12 035	8866	3630	30.2	40.9	2803	2084	596	21.3	28.6	2708	2229	781	28.8	35.0
7	2133		1868	87.6		1192		1044	87.6		343		293	85.4		598		531	88.8	
8	2465		1710	69.4		1744		1293	74.1		564		318	56.4		157		99	63.1	
9	238		148	62.2		136		91	66.9		40		27	67.5		62		30	48.4	
10	6091	3393	2542	41.7	74.9	3926	2116	1728	44.0	81.7	754	571	362	48.0	63.4	1411	706	452	32.0	64.0
11	5990		4393	73.3		3551		2651	74.7		979		730	74.6		1460		1012	69.3	
All Projects	36 970	29 905	17 575	47.5	58.8	23 708	18 729	11 315	47.7	60.4	5846	4944	2596	44.4	52.5	7416	6232	3664	49.4	58.8

BL = baseline; FU = follow-up.

Table 2. Characteristics and substance use of screen positive patients: full and matched samples.

		Screen positive			Brief intervention			Brief treatment			Referral to treatment		
		Matched sample			Matched sample			Matched sample			Matched sample		
		All	BL	FU	All	BL	FU	All	BL	FU	All	BL	FU
n		171 921	17 575		118 323	11 315		24 071	2596		29 527	3664	
Age (years)		38.1	37.5***		37.7	36.9		37.7	37.3		40.1	39.6**	
		(0.03)	(0.10)		(0.04)	(0.14)		(0.08)	(0.25)		(0.07)	(0.19)	
Female		0.360	0.383***		0.377	0.396***		0.332	0.376***		0.313	0.349*	
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.01)		(0.00)	(0.01)	
Black		0.246	0.200***		0.238	0.187**		0.267	0.205		0.261	0.237***	
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.01)		(0.00)	(0.01)	
Hispanic		0.174	0.209***		0.188	0.231***		0.171	0.200***		0.116	0.144***	
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.01)		(0.00)	(0.01)	
Other		0.029	0.028**		0.033	0.029		0.023	0.028**		0.018	0.023***	
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)	
High school degree		0.604	0.605					0.602	0.612		0.606	0.600	
		(0.00)	(0.01)					(0.00)	(0.01)		(0.00)	(0.01)	
Income		214.6	237.9					738.9	782.5		664.8	586.5	
		(8.3)	(9.8)					(47.8)	(48.9)		(28.5)	(29.8)	
Alcohol	Any	0.711	0.715***	0.516	0.708	0.714***	0.543	0.662	0.688	0.478	0.762	0.739*	0.462
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)

		Screen positive			Brief intervention			Brief treatment			Referral to treatment		
		Matched sample			Matched sample			Matched sample			Matched sample		
		All	BL	FU	All	BL	FU	All	BL	FU	All	BL	FU
	Days	11.6	10.9	8.3	9.7	9.0	7.6	12.5	11.4	8.5	18.5	16.3***	10.7
		(0.03)	(0.09)	(0.09)	(0.03)	(0.10)	(0.10)	(0.08)	(0.24)	(0.24)	(0.07)	(0.21)	(0.24)
Heavy drinking	Any	0.510	0.536***	0.352	0.475	0.502***	0.347	0.500	0.540	0.354	0.662	0.637	0.369
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)
	Days	10.4	9.6***	7.0	7.4	6.7***	5.9	11.5	10.3	7.7	18.5	16.3	9.9
		(0.04)	(0.10)	(0.10)	(0.04)	(0.11)	(0.11)	(0.10)	(0.26)	(0.26)	(0.08)	(0.24)	(0.27)
Illicit drugs	Any	0.400	0.389*	0.191	0.338	0.327***	0.168	0.576	0.518***	0.239	0.511	0.487	0.227
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)
	Days	12.9	12.8	11.9	10.8	10.9	11.3	14.7	13.9	12.2	16.8	16.0***	13.0
		(0.04)	(0.14)	(0.19)	(0.05)	(0.18)	(0.25)	(0.10)	(0.31)	(0.45)	(0.10)	(0.27)	(0.39)
Both on same day	Any	0.280	0.284***	0.140	0.233	0.231***	0.122	0.381	0.374	0.174	0.390	0.383***	0.170
		(0.00)	(0.00)**	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)
	Days	7.5	7.4	6.4	5.5	5.0	5.9	7.4	6.9	5.9	10.9	10.4***	7.5
		(0.05)	(0.15)	(0.18)	(0.05)	(0.17)	(0.23)	(0.10)	(0.32)	(0.40)	(0.11)	(0.30)	(0.38)

BL = baseline; FU = follow-up. Standard errors in parentheses. Income is monthly household income. 'Both' refers to the use of alcohol and drugs on the same day. *t*-Tests and χ^2 tests of baseline values were used to compare matched subsample to the remainder of the sample (*** $P < 0.01$; ** $P < 0.05$; * $P < 0.1$).

Table 3. Model adjusted changes in substance use outcomes by SBIRT category.

		Screen positive			Brief intervention			Brief treatment			Referral to treatment		
		BL	FU	Δ	BL	FU	Δ	BL	FU	Δ	BL	FU	Δ
Alcohol	Any	0.827	0.532	0.296	0.839	0.576	0.263	0.782	0.469	0.313	0.825	0.454	0.371
	Days	10.7	8.0	2.7	8.8	7.1	1.7	11.2	8.1	3.1	16.0	10.4	5.6
Heavy drinking	Any	0.560	0.315	0.245	0.519	0.312	0.207	0.567	0.301	0.266	0.686	0.336	0.350
	Days	9.7	7.1	2.6	6.9	5.9	1.0	10.4	7.7	2.7	16.0	9.9	6.1
Illicit drugs	Any	0.343	0.079	0.264	0.244	0.051	0.193	0.539	0.146	0.393	0.495	0.141	0.354
	Days	12.3	10.6	1.7	10.3	9.6	0.7	13.4	11.0	2.4	14.9	11.4	3.5
Alcohol and illicit drugs	Any	0.207	0.050	0.157	0.134	0.031	0.103	0.332	0.086	0.246	0.354	0.101	0.254
	Days	7.3	6.3	1.0	4.7	5.5	−0.8	6.7	5.6	1.2	10.1	7.1	3.0

BL = baseline; FU = follow-up. All estimates for measures of any recent use are from probit models with individual random-effects and site fixed-effects. Estimates for measures of days of use are from general linear mixed models with individual random-effects and site fixed-effects. Recent use includes any use within the past 30 days. Number of days of using a given substance is conditional on using that substance on at least 1 day. Heavy drinking is defined as becoming intoxicated from five or more alcoholic drinks. Conditional measures only include observations where outcome occurred at least once. All differences were significant at 0.01 level unless specified. ** $P < 0.05$; * $P < 0.1$; NS = not significant; SBIRT = Screening, Brief Intervention and Referral to Treatment.

Substance use measures are presented for all SP patients and separately for the three intervention categories in Table 2. The most common type of substance reported at baseline was alcohol (71.1%). Illicit drug use was reported less commonly than alcohol, but was still highly prevalent (40.0% of all positive screens). Fewer than one-third of all screens were dual substance users (alcohol and illicit drugs). Heavy drinking was moderately higher among RT patients (66.2%) than among BT (50.0%) or BI patients (47.5%). Illicit drug use was more common among BT (57.6%) than BI (33.8%) or RT patients (51.1%). Conditional days of use also varied by substance and SBIRT intervention. RT patients reported the most days of use for all outcomes. RT patients also used alcohol, had a heavy drinking day or used illicit drugs on more than half of the past 30 days (17.4, 17.6 and 15.7 days, respectively). BI patients reported the fewest conditional days of use for all substance categories.

Table 2 also indicates where baseline differences between the matched sample (i.e. those with both a baseline and a follow-up observation) and the remainder of patients (i.e. those who were either not selected for a follow-up interview or those that were selected but did not complete a follow-up interview) are statistically significant based on simple *t*- and χ^2 tests. This comparison demonstrates how representative is the analysis sample of matched patients is of the project as a whole. Many of the demographic characteristics of the matched sample are significantly different from those patients not in the matched sample, suggesting an obvious need to control for these characteristics in any analyses of outcomes. Of greater interest, however, are statistically significant baseline differences for more than half of the substance use behaviors, which could indicate selection bias. We see that in most cases the matched sample variables had similar values to the remainder sample (e.g. 71.1 versus 71.5 for the percentage with any alcohol use). This suggests that the *t*- and χ^2 tests may be overly sensitive due to the large sample sizes and their failure to correct for clustering of observations within clinical sites and projects.

Analysis of changes in substance use is reserved for the Main outcomes section (below). We note here that prevalence of use is lower at follow-up. Conditional days of use also decrease, with two exceptions: among the BI group, days of illicit drug use and days of combined illicit drug/alcohol use increase, conditional on use.

Main outcomes

Table 3 shows changes in the model-adjusted means for substance use measures. The difference between these is the estimated change in each outcome. Changes for most substance use measures were statistically significant at the $P < 0.01$ level, so we specify significance only when this is not the case. The probability of using any alcohol in the past 30 days decreased by 0.296 (35.6 %) for all SP patients, from 0.827 to 0.532. Among those patients who did drink alcohol, they drank 1.8 fewer days at follow-up. The proportion of BI patients who drank alcohol decreased the least (0.263) of the three SBIRT groups, despite starting at a similar prevalence. Similarly, BIs only decreased conditional days of drinking by 1.0, compared with 2.2 for BT patients and 4.7 for RT patients. The proportion of patients reporting heavy drinking dropped by 0.245 (43.4%) for all SP patients, with RT patients reporting the largest decrease (0.371). RT patients also had the largest decrease in conditional days of heavy drinking, dropping from 14.9 to 9.8 days on average.

The prevalence of illicit drug use dropped by half or more for each service modality and 75.8% for SP patients as a group. Days of illicit drug use dropped by 2.1 for BT and 2.9 for RT patients. However, BI patients who continued to use illicit drugs used them for the same number of days as those who used at baseline. The proportion of patients with at least 1 day of consuming both alcohol and drugs also fell by approximately half. However, the conditional number of days of consuming both increased slightly for BI patients and decreased 0.6 days for BT patients ($P < 0.1$).

Moderator analyses

Two sets of moderator analyses were examined. First, we estimated all outcome models separately for males and females. Secondly, for a subsample of the data for which we had information on clinical setting (e.g. ED versus in-patient), we interacted indicators for clinical setting with the post-SBIRT indicator to assess differential effects by setting.

When estimated for males and females separately, changes in substance use outcomes were similar in magnitude and statistical significance to most pooled estimates. Female patients recommended to BI exhibited the most substance use outcomes that were statistically significantly different from males. Decreases in conditional days of drinking and heavy drinking were smaller among females, with males seeing almost a half day more reduction ($P < 0.05$). Females who continued to use drugs increased their conditional days of use. Although not statistically significant, female BT and RT patients reported a greater decrease in any heavy drinking and any illicit drug use than men.

To analyse the role of health-care setting, we requested data from each project that contained setting identifiers. These data represent a subsample of the data described previously with the addition of a set of setting indicators. Six of the projects provided data on patients screened between October 2005 and February 2008 that included identifiers for clinical setting. The remaining project provided data covering October to December 2005 and June 2007 to February 2008. Altogether, the total number of patients screened during these periods was 336 832. Of those, 81 538 screened positive for hazardous or harmful use and were recommended to BI ($n = 52\,531$), BT ($n = 13\,149$) or RT ($n = 15\,858$).

Because models examining the role of clinical setting were estimated using a subset of the overall data, we first note that the magnitude and statistical significance of changes on substance use behaviors are, for the most part, similar to the results presented earlier. Overall, changes in prevalence of substance use differed little by the setting in which patients were screened, particularly between out-patient (OP) and emergency departments (ED) patients. For example, the largest statistically significant difference between OP and ED patient outcomes was the proportion of any heavy drinking among BI patients: ED patients decreased by 0.114 versus 0.159 for OP patients. The most striking differences across settings were for any heavy drinking and illicit drug use among IP patients. The decreases in any heavy drinking among OP and ED patients were not realized by IP patients, for whom there was either no decrease (BI) or a much smaller magnitude of decrease (BT and RT). In contrast, IP patients saw a greater decrease in any illicit drug use than OP or ED patients (−0.367 versus −0.221 and −0.241, respectively).

Table 4. Standardized effects sizes of brief intervention on changes in alcohol measures (Bien *et al.*[16]).

Study	Measure	Effect sizes		
		Within intervention	Within control	Between
Babor & Grant[15]	Typical daily alcohol consumption	0.17	0.06	0.32[a]
Heather *et al.*[16]	Last month alcohol consumption	0.29	0.16	0.35[a]
Miller *et al.*[17]	Weekly alcohol consumption	0.33	0.25	-0.03[a]
Kristenson *et al.*[18]	Decreases in GGT levels	0.44[a]	–	0.1
Daniels *et al.*[19]	Alcohol consumption/problem score	0.47	–	0.25[a]
Heather *et al.*[20]	Last week alcohol consumption	0.53	–	0.27[a]
Wallace *et al.*[21]	Weekly alcohol consumption	0.63[a]	0.31	0.35
Heather *et al.*[22]	Avg. alcohol consumption	0.65	–	0.34[a]
Miller *et al.*[23]	Weekly alcohol consumption	0.7	–	0.64
Chick *et al.*[24]	Last week alcohol consumption	0.87	0.75	0.07[a]
Maheswaran *et al.*[25]	Alcohol consumption	1.11	0.21	1.04[a]
Scott & Anderson[27]	Weekly alcohol consumption	1.16	–	0.03[a]
Harris & Miller[28]	Weekly alcohol consumption	1.18[a]	0.82	0.64
Anderson & Scott[26]	Weekly alcohol consumption	1.35	0.50	0.42[a]
SAMHSA's SBIRT	Past 30 days alcohol measure			
Brief intervention	Any alcohol use	0.72[a]		–
	No. of days of alcohol use	0.38[a]		–
	Any heavy drinking days	0.52[a]		–
	No. of heavy drinking days	0.25[a]		–
Brief treatment	Any alcohol use	0.68[a]		–
	No. of days of alcohol use	0.77[a]		–
	Any heavy drinking days	0.53[a]		–
	No. of heavy drinking days	0.61[a]		–
Referral to treatment	Any alcohol use	0.85[a]		–
	No. of days of alcohol use	0.97[a]		–
	Any heavy drinking days	0.73[a]		–
	No. of heavy drinking days	0.87[a]		–

All within-group effect sizes are computed as (intake mean – follow-up mean)/weighted pooled standard deviation. All between-group effect sizes are computed as (control follow-up mean – brief intervention follow-up mean)/weighted pooled standard deviation.

[a] Effect sizes represent significant differences at the 0.05 level. GGT = gamma glutamyl transpeptidase; BIRT = Screening, Brief Intervention and Referral to Treatment; SAMHSA = Substance Abuse and Mental Health Services Administration.

Comparisons with results from other SBI studies

To place our estimates in context, we compared standardized effect sizes from our analysis to those presented in the literature. Standardized effect sizes from 14 of the studies reviewed by Bien *et al.*[16] are given in Table 4[22-35]. The within-group effects are described by the baseline to follow-up changes in alcohol use among patients receiving BI, and are thus similar in design to the standardized effect sizes from our analysis also given in Table 4. The standardized effect sizes for all of our alcohol measures, across all SBIRT service groups, fall within the range of statistically significant effect sizes calculated by Bien *et al.*[16]. In particular, BI patients are the group that is most comparable with the studies in Bien *et al.*'s review. Our effect size of 0.72 for

any drinking is larger than those of five studies that found statistical significance when BI was compared to a control group: Babor & Grant[22], Heather *et al*.[23, 27, 29] and Daniels *et al*.[26]. Finally, effect sizes for the studies for which pre–post changes in control-group outcomes were available, were consistently lower than SAMHSA's effect sizes.

Discussion

SAMHSA's SBIRT grant program is the largest SBIRT dissemination effort ever undertaken in the United States, yet the only published multi-site study of patient outcomes for this initiative is Madras *et al*.[1], which has been cited widely as evidence of SBIRT's effectiveness, despite critical evaluations by several commentators[2-4] who have noted several limitations of the Madras *et al*.[1] findings that limit their utility. In this paper, we re-visit the Madras *et al*. in an attempt to address these limitations. Using data from all seven projects in the first cohort of SAMHSA SBIRT grants, we examined response rates and sample selection concerns carefully, used more appropriate statistical methodology and explored potential moderators of the correlation between SBIRT services and substance use outcomes.

The findings from our study provide evidence, albeit imperfect, in support of SAMHSA's SBIRT being associated with changes in patient outcomes consistent with earlier, rigorous SBIRT trials. Significant improvements in substance use outcomes were demonstrated by changes in simple means and model-adjusted means. We conclude that SBIRT was associated with improvement in patient outcomes, but with several caveats. First, no data on comparison groups were available for our analyses. In an effort to address this issue, we calculated standardized effect sizes for different alcohol use measures in our data and compared them to those from the controlled trials summarized by Bien *et al*.[16]. We found that our estimated changes were similar to the within-group effects of several clinical trials, suggesting that the changes observed among SBIRT services recipients are consistent with the clinical trial literature.

Next, we did not observe the delivery of SBIRT services to the patients in our sample. This is a concern, because some patients might not have actually received services so that the changes we observed might have been due to factors other than SBIRT. A related concern is that because we did not observe the screening score or other screening factors, we do not know the true severity of substance use risk among our sample of patients. More importantly, we note that we were unable to validate the accuracy of self-reported substance use. At baseline, the use of screening instruments with strong psychometric properties provides some confidence that baseline substance use was reported accurately. However, at follow-up, the interview only collected the raw substance use quantities, via a different mode of data collection. Social acceptability bias, a lack of rapport with the interviewer, differential recall bias under the different modes and a desire to complete the interview more quickly, among other factors, could all influence (probably suppressing) the reported use. These challenges again highlight the vital need for comparison groups.

Finally, our samples are not perfectly representative of the entire SAMHSA initiative. Because of disruptions in several projects, there is a group of patients who received services but who are excluded from our analysis because their outcome data were not collected. More problematic is the follow-up response rate that varies by project. Patients who do not respond at follow-up may

be systematically different from responders. Such differences may, in turn, be associated with a poor response to SBIRT services and with worse outcomes. However, we note that, after adjusting for project disruptions, one project in particular, number 6, is responsible for the low overall response rate. In analyses not reported here, we re-estimated our outcome models by project and compared the results of project 6 with those of the other six. We found that the estimated pre–post change for project 6 was never an outlier with respect to those of the other six projects. Specifically, for each outcome, the project 6 estimate was always lower than the project with the highest estimate and higher than the project with the lowest estimate.

Despite these important caveats, our findings provide value to ongoing assessment of SAMHSA's SBIRT initiative and to the broader SBIRT research community. Large public health programs such as SBIRT are constrained in their ability to conduct rigorous research by the practicalities of implementation, human subjects protections and resources. Evaluating the effectiveness of evidence-based SBIRT services is a complex and ongoing project that requires further randomized controlled trials. A variety of research must be completed and be maintained over time to conclude that SBIRT is having its potential impact. Two recent RCTs highlight the complexity of the science around SBIRT. Saitz *et al*.[36] tested the efficacy of a brief negotiated intervention for any illicit drug use or prescription drug misuse versus no BI, and found no effect. Roy-Byrne *et al*.[37] found no effect of a BI plus telephone booster for drug use. These two studies emphasize the need for deliberate, targeted use of SBIRT public health rather than as a panacea. Other research needs range from fidelity and quality studies of SBIRT to subgroup studies with a comparison group design to time–series assessment of population level outcomes. For such research to accrue, SBIRT would need to continue as a provided service. While we cannot claim that SBIRT caused improvements in patient outcomes, our results support proceeding with its service provision and further more rigorous study.

References

1. 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– 295.
2. Bernstein E., Bernstein J. A., Stein J. B., Saitz R. SBIRT in emergency care settings: are we ready to take it to scale? *Acad Emerg Med* 2009; **16**: 1072– 1077.
3. Saitz R. Candidate performance measures for screening for, assessing, and treating unhealthy substance use in hospitals: advocacy or evidence-based practice? *Ann Intern Med* 2010; **153**: 40– 43.
4. Saitz R., Alford D. P., Bernstein J., Cheng D. M., Samet J., Palfai T. Screening and brief intervention for unhealthy drug use in primary care settings: randomized clinical trials are needed. *J Addict Med* 2010; **4**: 123– 130.
5. Kypri K., Langley J. D., Saunders J. B., Cashell-Smith M. L. Assessment may conceal therapeutic benefit: findings from a randomized controlled trial for hazardous drinking. *Addiction* 2007; **102**: 62– 70.

6. Walters S. T., Vader A. M., Harris T. R., Field C. A., Jouriles E. N. Dismantling motivational interviewing and feedback for college drinkers: a randomized clinical trial. *J Consult Clin Psychol* 2009; **77**: 64– 73.

7. Rosenthal R., Rosnow R. L. *Essentials of Behavioral Research: Methods and Data Analysis*. New York, NY: McGraw-Hill Book Company; 1984.

8. Makäla K. Ethical control of social research. *Nordic Stud Alcohol Drugs* 2006; **25**: 5– 19.

9. Babor T. F. Research ethics: when the tail wags the dog. *Nord Stud Alcohol Drugs [English Suppl]* 2006; **23**: 23– 24.

10. Cook T. D., Campbell D. T. *Quasi-Experimentation. Design and Analysis Issues for Field Settings*. Boston: Houghton Mifflin Company; 1979.

11. Bertholet N., Daeppen J.-B., Wietlisbach V., Fleming M., Burnand B. Reduction of alcohol consumption by brief alcohol intervention in primary care: systematic review and meta-analysis. *Arch Intern Med* 2005; **165**: 986– 995.

12. McLellan A. T., Turner B. J. Chronic noncancer pain management and opioid overdose: time to change prescribing practices. *Ann Intern Med* 2010; **152**: 123– 124.

13. Marlatt G. A., Witkiewitz K. Update on harm-reduction policy and intervention research. *Annu Rev Clin Psychol* 2010; **6**: 591– 606.

14. Miller W. R. Motivational interviewing with problem drinkers. *Behav Psychother* 1983; **11**: 147– 172.

15. Miller W. R., Rollnick S. *Motivational Interviewing: Preparing People to Change Addictive Behavior*. New York, NY: Guilford Press; 1991.

16. Bien T. H., Miller W. R., Tonigan J. S. Brief interventions for alcohol problems: a review. *Addiction* 1993; **88**: 315– 335.

17. Prochaska J. O., DiClemente C. C. Toward a comprehensive model of change. In: W. R. Miller, N. Heather, editors. *Treating Addictive Behaviors: Processes of Change*. New York: Plenum Press; 1986, pp. 3– 27.

18. Carroll K. M. A Cognitive Behavioral Approach: Treating Cocaine Addiction. In: *National Institute on Drug Abuse Therapy Manuals for Drug Addiction, Manual 1*. Rockville, MD: National Institutes of Health; 1998. Publication No. 98-4308

19. Sampl S., Kadden R. *Motivational Enhancement Therapy and Cognitive Behavioral Therapy for Adolescent Cannabis Users: 5 Sessions, Cannabis Youth Treatment (CYT) Series*, Vol. **1**. Rockville, MD: Center for Substance Abuse Treatment (CSAT), Substance Abuse and Mental Health Services Administration (SAMHSA); 2001.

20. Smith J. E., Meyers R. J. The treatment. In: R. Meyers, W. Miller, editors. *A Community Reinforcement Approach to the Treatment of Addiction*. Cambridge, UK: University Press; 2001, pp. 28– 62.

21. McLellan A. T., Cacciola J. S., Zanis D. *The Addiction Severity Index-'Lite'(ASI-'Lite'): Center for the Studies of Addiction*. Philadelphia, PA: University of Pennsylvania/Philadelphia VA Medical Center; 1997.

22. T. F. Babor, M. Grant, editors. *Project on Identification and Management of Alcohol-Related Problems. Report on Phase II: A Randomized Clinical Trial of Brief Interventions in Primary Health Care*. Geneva, Switzerland: World Health Organization; 1992.

23. Heather N., Campion P. D., Neville R. G., Maccabe D. Evaluation of a controlled drinking minimal intervention for problem drinkers in general practice (the DRAMS scheme). *J R Coll Gen Pract* 1987; **37**: 358– 363.

24. Miller W. R., Sovereign R. G., Krege B. Motivational interviewing with problem drinkers: II. The Drinker's Check-up as a preventive intervention. *Behav Psychother* 1988; **16**: 251– 268.

25. Kristenson H., Ohlin H., Hulten-Nosslin M. B., Trell E., Hood B. Identification and intervention of heavy drinking in middle-aged men: results and follow-up of 24–60 months of long-term study with randomized controls. *Alcohol Clin Exp Res* 1983; **7**: 203– 209.

26. Daniels V., Somers M., Orford J., Kirby B. How can risk drinking amongst medical patients be modified? The effects of computer screening and advice and a self-help manual. *Behav Psychother* 1992; **20**: 47– 60.

27. Heather N., Whitton B., Robertson I. Evaluation of a self-help manual for media-recruited problem drinkers: six-month follow-up results. *Br J Clin Psychol* 1986; **25**: 19– 34.

28. Wallace P., Cutler S., Haines A. Randomised controlled trial of general practitioner intervention in patients with excessive alcohol consumption. *BMJ* 1988; **297**: 663– 668.

29. Heather N., Kissoon-Singh J., Fenton G. W. Assisted natural recovery from alcohol problems: effects of a self-help manual with and without supplementary telephone contact. *Br J Addict* 1990; **85**: 1177– 1185.

30. Miller W. R., Benefield R. G., Tonigan J. S. Enhancing motivation for change in problem drinking: a controlled comparison of two therapist styles. *J Consult Clin Psychol* 1993; **61**: 455– 461.

31. Chick J., Lloyd G., Crombie E. Counselling problem drinkers in medical wards: a controlled study. *BMJ (Clin Res Ed)* 1985; **290**: 965– 967.

32. Maheswaran R., Beevers M., Beevers D. G. Effectiveness of advice to reduce alcohol consumption in hypertensive patients. *Hypertension* 1992; **19**: 79– 84.

33. Anderson P., Scott E. The effect of general practitioners' advice to heavy drinking men. *Br J Addict* 1992; **87**: 891– 900.

34. Scott E., Anderson P. Randomized controlled trial of general practitioner intervention in women with excessive alcohol consumption. *Drug Alcohol Rev* 1990; **10**: 313– 321.

35. Harris K. B., Miller W. R. Behavioral self-control training for problem drinkers: components of efficacy. *Psychol Addict Behav* 1990; **4**: 82– 90.

36. 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– 513.

37. 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.