

# Quality Concerns with Routine Alcohol Screening in VA Clinical Settings

Katharine A. Bradley, MD, MPH<sup>1,2,3,4,5</sup>, Gwen T. Lapham, MPH, MSW<sup>1,5</sup>, Eric J. Hawkins, PhD<sup>1,6</sup>, Carol E. Achtmeyer, MN, ARNP<sup>1</sup>, Emily C. Williams, PhD, MPH<sup>1,5</sup>, Rachel M. Thomas, MPH<sup>1</sup>, and Daniel R. Kivlahan, PhD<sup>1,2,6</sup>

<sup>1</sup>Health Services Research & Development (HSR&D), Veterans Affairs (VA) Puget Sound Health Care System, Seattle, WA, USA; <sup>2</sup>Center of Excellence in Substance Abuse Treatment and Education (CESATE), Seattle, WA, USA; <sup>3</sup>General Medicine Service, Veterans Affairs (VA) Puget Sound Health Care System, Seattle, WA, USA; <sup>4</sup>Department of Medicine, University of Washington, Seattle, WA, USA; <sup>5</sup>Department of Health Services, University of Washington, Seattle, WA, USA; <sup>6</sup>Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, USA.

**BACKGROUND:** Alcohol screening questionnaires have typically been validated when self- or researcher-administered. Little is known about the performance of alcohol screening questionnaires administered in clinical settings.

**OBJECTIVE:** The purpose of this study was to compare the results of alcohol screening conducted as part of routine outpatient clinical care in the Veterans Affairs (VA) Health Care System to the results on the same alcohol screening questionnaire completed on a mailed survey within 90 days and identify factors associated with discordant screening results.

**DESIGN:** Cross sectional.

**PARTICIPANTS:** A national sample of 6,861 VA outpatients (fiscal years 2007–2008) who completed the AUDIT-C alcohol screening questionnaire on mailed surveys (survey screen) within 90 days of having clinical AUDIT-C screening documented in their medical records (clinical screen).

**MAIN MEASURES:** Alcohol screening results were considered discordant if patients screened positive (AUDIT-C  $\geq 5$ ) on either the clinical or survey screen but not both. Multivariable logistic regression was used to estimate the prevalence of discordance in different patient subgroups based on demographic and clinical characteristics, VA network and temporal factors (e.g. the order of screens).

**KEY RESULTS:** Whereas 11.1% (95% CI 10.4–11.9%) of patients screened positive for unhealthy alcohol use on the survey screen, 5.7% (5.1–6.2%) screened positive on the clinical screen. Of 765 patients who screened positive on the survey screen, 61.2% (57.7–64.6%) had discordant results on the clinical screen, contrasted with 1.5% (1.2–1.8%) of 6096 patients who screened negative on the survey screen. In multivariable analyses, discordance was significantly increased among Black patients compared with White, and among patients who had a positive survey AUDIT-C screen or who received care at 4 of 21 VA networks.

**CONCLUSION:** Use of a validated alcohol screening questionnaire does not—by itself—ensure the quality of alcohol screening. This study suggests that the quality of clinical alcohol screening should be monitored, even when well-validated screening questionnaires are used.

**KEY WORDS:** alcohol screening; brief alcohol counseling; validation.  
J Gen Intern Med 26(3):299–306  
DOI: 10.1007/s11606-010-1509-4  
© Society of General Internal Medicine 2010

## INTRODUCTION

Alcohol misuse is common<sup>1</sup> and causes extensive morbidity and mortality.<sup>2</sup> Brief alcohol counseling decreases drinking,<sup>3,4</sup> and a National Commission on Prevention Priorities designated alcohol screening and brief alcohol interventions the third highest US adult prevention priority.<sup>5</sup> Routine alcohol screening is required to identify patients who might benefit from brief alcohol counseling.<sup>6</sup>

The Veterans Affairs (VA) Health Care System implemented routine screening for alcohol misuse in 2004,<sup>6</sup> and since 2006 has required that the Alcohol Use Disorders Identification Test—Consumption Questions (AUDIT-C) be used for screening. Each VA is expected to meet performance targets, but the approach used to implement alcohol screening is left up to individual facilities or networks (e.g. at triage, by primary care providers, paper questionnaire, etc.). Over 90% of VA outpatients nationwide are screened with the AUDIT-C and rates of documented brief alcohol counseling in VA are increasing.<sup>7</sup>

The AUDIT-C has been validated when interviewer-administered, and when completed on mailed questionnaires with results shared with primary care providers,<sup>8–13</sup> but little is known about its performance when implemented as part of routine clinical care. A prior small study raised concerns about the quality of clinical alcohol screening in VA despite the use of validated screening questionnaires.<sup>14</sup> In January 2008, an electronic clinical reminder was disseminated that prompted VA providers to ask AUDIT-C questions verbatim, in a private setting, and in a nonjudgmental manner.

The purpose of this study was to evaluate the quality of clinical alcohol screening in the VA from 2007–2008, by

Received January 13, 2010

Revised July 26, 2010

Accepted August 30, 2010

Published online September 22, 2010

comparing the results of the AUDIT-C documented during routine clinical care to the results of the AUDIT-C completed on a confidential mailed survey within 90 days of the clinical screen. A second aim was to evaluate factors associated with discordance between the results of clinical and survey alcohol screens, which was not possible in the previous smaller study.

## METHODS

The study sample included VA outpatients who were in two independent quality improvement programs conducted for the VA Office of Quality and Performance: the External Peer Review Program (EPRP)<sup>15</sup> and the Survey of Healthcare Experiences of Patients (SHEP).<sup>16</sup> Patients were eligible if they had AUDIT-C results available from both programs between October 1, 2006 and September 30, 2008. EPRP assesses VA care, including AUDIT-C alcohol screening, by reviewing VA medical records. SHEP is the VA's patient satisfaction survey (response rate of 54.5% during study), which also includes the AUDIT-C. The EPRP sampled a random sample of outpatients nationwide, but oversampled certain subgroups (e.g. women 52–69 years old and patients with specific chronic diseases). Both the EPRP and SHEP sampled patients with recent outpatient visits, but the samples were independent during this study, so the study sample resulted from chance overlap in patients included in the EPRP and SHEP (Fig. 1).

The study was approved by the VA Office of Quality and Performance and the University of Washington and VA Puget

Sound Institutional Review Boards, with waivers of written informed consent and HIPAA authorization.

## Measures

**Alcohol Screening.** AUDIT-C questions ask about typical quantity and frequency of drinking and the frequency of binge drinking and have established 3-month test-retest reliability.<sup>17,18</sup> AUDIT-C scores range from 0 to 12 points. Scores of 5 or more points were considered a positive alcohol screen in this study, consistent with the VA performance measure for brief alcohol counseling implemented in October 2007.<sup>7,19</sup> This threshold minimizes the burden of counseling patients with false-positive AUDIT-C screens.<sup>9,10</sup>

**Clinical Screen.** Clinical AUDIT-C screening documented in the medical record and abstracted by EPRP medical record reviewers could be conducted by telephone, at intake for appointments by medical assistants or nurses, by medical providers, or by self-report (e.g. on paper) with results later entered into VAs electronic medical record. Approaches to screening vary across VA clinics and facilities, but anecdotal reports suggest that most used a clinical reminder in the electronic medical record to score the AUDIT-C and document screening.

**Survey Screen.** AUDIT-C results from mailed SHEP surveys received within 90 days of the clinical screen were used as a comparison standard. The survey introduction stated “all information is strictly anonymous. It will not be shared with

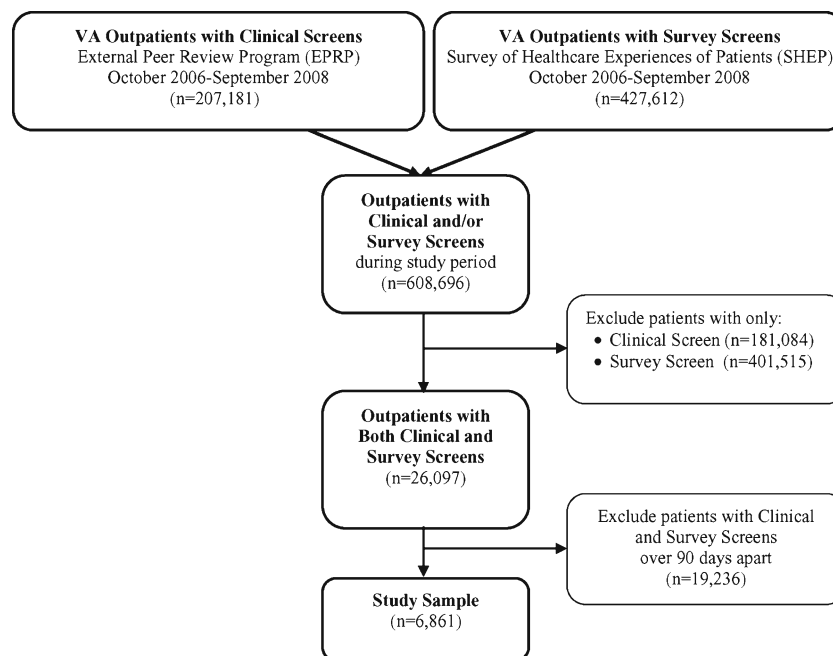


Figure 1. Study sample. The prevalence of positive survey screens among all of 427,612 survey respondents who completed the AUDIT-C was 13.4% (95% CI 13.3-13.5) and the prevalence of positive clinical screens in all 207,181 patients who had AUDIT-C results abstracted from the medical records was 6.4% (95% CI 6.3-6.5). These were higher than in the study sample, as expected, likely due to greater survey non-response in younger patients<sup>16</sup> and heavier drinkers.<sup>30</sup>

your doctor". Survey screens were classified as follows: no past year alcohol use (AUDIT-C score 0 points); low-risk drinking (1-3 points men; 1-2 women); possible alcohol misuse (4 points men; 3-4 women); alcohol misuse (5-7 points); and severe alcohol misuse (8-12 points).<sup>20-23</sup>

**Discordant Screening Results.** Screening results were considered discordant if one screen was positive (AUDIT-C  $\geq$  5) and the other negative.

## Other Measures

**Patient Characteristics.** Race, education, marital status, and income were self-reported on surveys. Gender, age, and tobacco use in the past year were obtained from medical record reviews. A measure of any mental health or alcohol or other substance use disorder diagnosis documented in the year prior to medical record reviews was based on ICD-9 codes from administrative data obtained by EPRP.

**Regional VA Networks.** Patients were assigned to one of 21 VA networks based on where they were screened clinically. Network directors oversee operations at all hospitals and freestanding community-based outpatient clinics in their region, and their annual contracts include financial incentives for meeting performance targets. If discordance varied across networks it could reflect different clinical cultures and alcohol screening strategies. Although greater variation was expected at the facility level ( $n=128$ ), the study sample included too few patients for precise facility-level estimates of discordance.

**Temporal Factors Potentially Associated with Discordance.** Three temporal measures were evaluated: the number of days between the two screens ( $\leq 14$ ; 15-30; 31-60; and 61-90 days); the order of the survey and clinical screens; and the date of the clinical screen relative to implementation efforts. A brief alcohol counseling performance measure initiated October 2007<sup>7</sup> was hypothesized to act as a disincentive to identify alcohol misuse and thereby increase discordance, whereas an alcohol screening clinical reminder disseminated 3 months later (January 2008), was hypothesized to decrease discordance because it prompted providers to ask screening questions verbatim, in a private setting and non-judgmentally.

## Analyses

The dichotomous results of survey and clinical screens were cross-tabulated to identify discordance in the total sample and across demographic, clinical, and temporal subgroups. Multivariable logistic regression was used to identify factors independently associated with discordance. Two models are presented—one with survey screen results and one without. The smaller model included demographic characteristics, tobacco use, prior year mental health or substance use disorders diagnosis, VA network where clinical screening occurred, and the three temporal measures. To evaluate

whether findings were confounded by severity of self-reported alcohol misuse, the second model added results of survey screens (5 AUDIT-C categories). The adjusted prevalence of discordance is presented to reflect the magnitude as well as the statistical significance of differences in prevalence rates of discordance across subgroups. All analyses were conducted using STATA, version 11.1.<sup>24</sup>

## RESULTS

The study included 6,861 patients with clinical and survey AUDIT-C screens within 90 days of each other (Fig. 1). Participants were predominantly older, white men, and 17% were women due to oversampling (Table 1). About twice as many patients screened positive for alcohol misuse on survey screens as on clinical screens: 11.1% (95%CI 10.4-11.9%) and 5.7% (5.1- 6.2%), respectively. On average, the time between clinical and survey screens was 54 days (SD =22), and clinical screens preceded survey screens for 71.7% of the sample, as expected because the clinical screen would precede the survey screen when both medical record reviews and patient satisfaction surveys were triggered by the same outpatient visit.

### Discordance Between Survey and Clinical Screening

Overall, 561 (8.2%; 7.5-8.8%) of 6,861 patients had discordant clinical and survey screens, with patients who screened positive on the survey much more likely to have discordant screening results (Table 2). Among the 561 patients with discordant screening results, 468 (83%) screened positive on the survey screens, and 93 (17%) screened positive on clinical screens. Whereas 468 (61.2%; 57.7-64.6%) of 765 patients with positive survey screens had discordant screens, only 93 (1.5%; 1.2-1.8%) of 6096 patients with negative survey screens had discordant screens (Table 2). In contrast, 93 (23.8%; 19.6-28.1%) of 390 patients who screened positive on clinical screens had discordant screens, and 468 (7.2%; 6.6-7.9%) of 6,471 who screened negative on clinical screens had discordant screens (Table 2). Among patients whose clinical screens indicated no alcohol use in the past year (AUDIT-C score 0), 21.9% reported drinking on survey screens. In comparison, among those whose survey screens indicated no past-year alcohol use, only 8.7% had clinical screens indicating past-year drinking.

**Association with Patient Characteristics.** Younger, male, and unmarried patients, those who smoked or had prior mental health or substance use disorder diagnoses, and patients who reported alcohol use on the survey AUDIT-C, especially those with positive survey screens, were more likely to have discordant alcohol screening results (Table 3). The pattern of discordance across AUDIT-C scores from clinical screens showed the expected normal distribution around the screening threshold (Fig. 2 Panel a). In contrast, there was a non-normal pattern of discordance across AUDIT-C scores from survey screens (Fig. 2; Panel b). For example, among those with the highest survey screens (AUDIT-C scores 8-12) 106/228 (46.5%) had discordant clinical screens (Table 3),

Table 1. Study Sample: VA Outpatients who had Both Clinical and Survey Alcohol Screens within 90 Days (N 6,861)

	Total Sample	
<b>Gender</b>		
Men	5,705	(83)
Women	1,156	(17)
<b>Age (years)</b>		
< 50	518	(8)
50-59	1,639	(24)
60-69	2,158	(31)
≥ 70	2,546	(37)
<b>Race (self reported)</b>		
White	5,891	(86)
Black or African American	472	(7)
Other	328	(5)
Missing	170	(2)
<b>Education (self-reported)</b>		
High school or less	3,250	(47)
At least some college	3,519	(51)
Missing	92	(1)
<b>Married (self-reported)</b>	4,083	(60)
<b>Income</b>		
< \$30,000/year	4,410	(64)
\$30,000-60,000/year	1,575	(23)
> \$60,000/year	465	(7)
Missing	411	(6)
<b>Tobacco use in the past year (medical record review)</b>	1,714	(25)
<b>Prior mental health diagnosis<sup>a</sup></b>	1,769	(26)
<b>Time between survey and clinical alcohol screen</b>		
0-14	589	(9)
15-30	628	(9)
31-60	2,486	(36)
61-90	3,158	(46)
<b>Order of screening</b>		
Survey screen first	1,943	(28)
Clinical screen first	4,918	(72)
<b>Timing of clinical screen vis a vis implementation efforts</b>		
Year before brief alcohol counseling performance measure (PM)	3,001	(44)
3 months after the PM <sup>b</sup> , before alcohol screening reminder	914	(13)
Period after dissemination of alcohol screening reminder	2,946	(43)
<b>Survey alcohol screen (AUDIT-C scores)</b>		
Nondrinker (0)	3,359	(49)
Low-risk drinker (1-3 men; 1-2 women)	2,133	(31)
Possible alcohol misuse (4 men; 3-4 women)	604	(9)
Alcohol misuse (5-7)	537	(8)
Severe alcohol misuse (8-12)	228	(3)

AUDIT-C=Alcohol Use Disorder Identification Test–Consumption;

<sup>a</sup>Includes mental health and substance use disorders within past year: VA administrative documentation of ICD 295-298.xx, 300-304.xx, 305.0, 305.2, 305.3, 305.4, 305.5, 305.6, 305.7, 305.8, 306.xx, 307.1, 308-309, 311.xx and 312.3x)

<sup>b</sup>PM=performance measure

reflecting at least 4-8 point differences between AUDIT-C scores from survey and clinical screens.

**Differences across VA Networks.** Rates of discordance differed across networks ranging from 4% to 13% ( $p=0.002$ ). The network prevalence of discordance among patients with positive survey screens ranged from 43 to 100% ( $p=0.002$ ), whereas the prevalence of discordance among patients with negative survey screens did not differ across networks (0.5–4%;  $p=0.236$ ). Compared to a large network with the lowest

Table 2. Discordant Clinical and Survey Alcohol Screening Results (N=6,861)

	Survey Screen Results <sup>a</sup>	
	Negative (n 6,096)	Positive (n 765)
<b>Clinical Screen Results<sup>a</sup></b>	N (%)	N (%)
Negative (N 6,471)	6,003 (98)	<b>468 (61)</b>
Positive (N 390)	<b>93 (2)</b>	297 (39)

<sup>a</sup>Bold numbers represent discordance. All percentages are column percents. “Negative”=AUDIT-C score<5 points; “positive”=AUDIT-C score≥5 points (chi-square  $p$ -value <0.001).

screening discordance (Network A), 11 of 21 networks had significantly higher discordance in bivariate logistic regression analyses.

**Timing and Order of the Two Screens.** There was no association between discordance and the time between, or order of, survey and clinical screens (Table 3). There was also no association between discordance and the timing of clinical screens with regards to implementation efforts, including a performance measure for brief alcohol counseling or dissemination of a screening clinical reminder that included recommendations aimed at improving the validity of screening (Table 3).

**Multivariable Analyses.** In multivariable analyses including all variables in Table 3 except survey AUDIT-C scores, the prevalence of discordance was increased in many patient subgroups (Table 3). However, many patient characteristics associated with discordance are known to be associated with AUDIT-C scores<sup>25</sup> and after adding survey AUDIT-C scores (5 categories) to the model only two factors aside from survey AUDIT-C scores were associated with discordance: self-reported Black/African American race and VA network (Table 3), suggesting that many associations were confounded by alcohol misuse. There were no significant interactions between race and VA network, or between race and AUDIT-C groups ( $p$ 's 0.99 and 0.95 respectively).

## DISCUSSION

This study of alcohol screening in the VA found that 61% of patients who screened positive for alcohol misuse on a mailed survey screened negative when screened clinically despite use of the same validated screening questionnaire, suggesting that many patients who could benefit from brief alcohol counseling are being missed by clinical screening in VA. Black race and the VA Network where clinical screening was conducted were the only factors other than survey AUDIT-C scores that were associated with discordance between clinical and survey AUDIT-C screens in fully-adjusted analyses.

Some discordance between clinical and survey screens is expected. Patients might be more motivated to report drinking honestly in clinical settings if they feel that their provider needs the information and/or it might be relevant to their health. In addition, some discordance is expected due to

Table 3. Prevalence of Discordant Results between Clinical and Survey Alcohol Screens (N=6,861)

	Discordance						
	Unadjusted Prevalence		p value	Adjusted Prevalence without Survey Screen <sup>f</sup>		Adjusted Prevalence with Survey Screen <sup>f</sup>	
	N	(%)		%	(95% CI)	%	(95% CI)
<b>Gender</b>			0.001				
Men <sup>e</sup>	495	(8.7)		8.3	(7.6–9.1)	1.2	(0.8–1.6)
Women	66	(5.7)		<b>3.6</b>	<b>(2.6–4.6)</b>	1.1	(0.5–1.6)
<b>Age</b>			<0.0005				
< 50	60	(11.6)		<b>11.4</b>	<b>(8.4–14.5)</b>	1.5	(0.7–2.4)
50–59	154	(9.4)		<b>8.5</b>	<b>(7.0–9.9)</b>	1.2	(0.7–1.7)
60–69	195	(9.0)		<b>8.0</b>	<b>(6.9–9.2)</b>	1.1	(0.7–1.5)
≥ 70 <sup>e</sup>	152	(6.0)		5.4	(4.5–6.4)	1.1	(0.7–1.5)
<b>Race</b>			0.082				
White <sup>e</sup>	463	(7.9)		7.0	(6.4–7.7)	1.1	(0.7–1.4)
African American	50	(10.6)		9.0	(6.4–11.7)	<b>2.1</b>	<b>(1.0–3.3)</b>
Other	29	(8.8)		7.1	(4.4–9.7)	1.1	(0.4–1.9)
Missing	19	(11.2)		11.1	(5.6–16.5)	2.2	(0.2–4.2)
<b>Education</b>			0.50				
High school or less <sup>e</sup>	255	(7.9)		7.1	(6.2–8.0)	1.2	(0.8–1.6)
At least some college	300	(8.5)		7.4	(6.5–8.3)	1.1	(0.7–1.5)
Missing	6	(6.5)		5.2	(0.0–10.8)	1.3	(0.0–3.3)
<b>Married</b>			<0.0005				
Yes <sup>e</sup>	282	(6.9)		6.2	(5.5–7.0)	1.1	(0.7–1.5)
No	269	(10.2)		<b>9.1</b>	<b>(8.0–10.3)</b>	1.2	(0.8–1.7)
Missing	10	(7.4)		6.6	(1.5–11.7)	1.1	(0.0–2.4)
<b>Income</b>			0.29				
< \$30,000/year <sup>e</sup>	360	(8.2)		7.0	(6.2–7.8)	1.2	(0.8–1.7)
\$30,000–60,000/year	127	(8.1)		7.4	(6.1–8.7)	1.0	(0.6–1.4)
> \$60,000/year	47	(10.1)		<b>10.1</b>	<b>(7.2–13.0)</b>	0.9	(0.4–1.4)
Missing	27	(6.6)		6.5	(4.1–8.9)	1.1	(0.4–1.8)
<b>Tobacco use, past year</b>			<0.0005				
No <sup>e</sup>	361	(7.0)		6.6	(5.9–7.3)	1.2	(0.8–1.6)
Yes	200	(11.7)		<b>9.7</b>	<b>(8.3–11.2)</b>	1.0	(0.6–1.4)
<b>Prior mental health dx<sup>a</sup></b>			0.008				
No <sup>e</sup>	390	(7.7)		7.3	(6.6–8.0)	1.2	(0.8–1.6)
Yes	171	(9.7)		7.1	(5.9–8.3)	0.9	(0.6–1.3)
<b>Time between Screens</b>			0.55				
≤ 14 days <sup>e</sup>	52	(8.8)		7.4	(5.4–9.4)	1.2	(0.6–1.9)
15–30	45	(7.2)		6.3	(4.4–8.2)	1.0	(0.5–1.5)
31–60	214	(8.6)		7.7	(6.7–8.8)	1.2	(0.8–1.7)
61–90	250	(7.9)		7.1	(6.2–8.0)	1.1	(0.7–1.5)
<b>Order of Screens</b>			0.386				
Survey screen first <sup>e</sup>	150	(7.7)		6.7	(5.6–7.8)	1.2	(0.7–1.7)
Clinical screen first	411	(8.4)		7.5	(6.7–8.2)	1.1	(0.8–1.5)
<b>Period of clinical screen</b>			0.18				
Year before PM <sup>b</sup> <sup>e</sup>	261	(8.7)		7.8	(6.8–8.8)	1.2	(0.8–1.6)
3 months interim	80	(8.8)		7.7	(6.0–9.4)	1.0	(0.6–1.5)
Period after reminder	220	(7.5)		6.6	(5.7–7.5)	1.1	(0.7–1.6)
<b>VA Network<sup>c</sup></b>			0.0019				
A (n=449) <sup>e</sup>	20	(4.5)		4.3	(2.4–6.1)	0.7	(0.2–1.1)
B (n=164)	8	(4.9)		4.2	(1.3–7.1)	1.3	(0.0–2.7)
C (n=318)	20	(6.3)		5.6	(3.2–8.1)	1.3	(0.4–2.1)
D (n=457)	29	(6.4)		6.3	(4.1–8.6)	0.5	(0.2–0.9)
E (n=339)	22	(6.5)		5.7	(3.3–8.1)	0.8	(0.3–1.3)
F (n=279)	19	(6.8)		6.3	(3.5–9.1)	1.0	(0.3–1.8)
G (n=219)	15	(6.9)		6.1	(3.0–9.1)	1.4	(0.3–2.6)
H (n=290)	20	(6.9)		6.4	(3.6–9.1)	1.3	(0.4–2.2)
I (n=195)	14	(7.2)		6.2	(3.0–9.5)	1.1	(0.2–2.0)
J (n=504)	38	(7.5)		7.0	(4.8–9.2)	1.1	(0.5–1.7)
K (n=468)	37	(7.9)		<b>7.4</b>	<b>(5.1–9.7)</b>	1.0	(0.5–1.6)
L (n=274)	22	(8.0)		6.7	(3.9–9.6)	1.2	(0.3–2.1)
M (n=612)	50	(8.2)		<b>8.0</b>	<b>(5.9–10.2)</b>	1.3	(0.6–1.9)
N (n=126)	11	(8.7)		7.6	(3.1–12.0)	0.8	(0.1–1.6)
O (n=324)	31	(9.6)		<b>8.2</b>	<b>(5.3–11.1)</b>	1.0	(0.4–1.6)
P (n=321)	31	(9.7)		<b>8.4</b>	<b>(5.5–11.3)</b>	<b>1.6</b>	<b>(0.6–2.5)</b>
Q (n=386)	40	(10.4)		<b>9.7</b>	<b>(6.7–12.6)</b>	1.3	(0.6–2.1)
R (n=373)	40	(10.7)		<b>9.5</b>	<b>(6.6–12.4)</b>	1.3	(0.6–2.1)
S (n=177)	19	(10.7)		<b>9.2</b>	<b>(5.1–13.2)</b>	<b>2.3</b>	<b>(0.5–4.2)</b>
T (n=354)	44	(12.4)		<b>11.7</b>	<b>(8.3–15.0)</b>	<b>2.3</b>	<b>(1.0–3.7)</b>



Table 3. (continued)

	Discordance		p value	Adjusted Prevalence without Survey Screen <sup>f</sup>		Adjusted Prevalence with Survey Screen <sup>f</sup>	
	Unadjusted Prevalence			%	(95% CI)	%	(95% CI)
	N	(%)					
U (n =232)	31	(13.4)		<b>12.5</b>	<b>(8.2–16.7)</b>	<b>2.0</b>	<b>(0.8–3.3)</b>
<b>Survey Screen<sup>d</sup></b>			<0.0005				
No past year alcohol use <sup>e</sup>	10	(0.3)		–	–	0.3	(0.1–0.4)
Low-risk drinking	27	(1.3)		–	–	<b>1.1</b>	<b>(0.7–1.6)</b>
Possible alcohol misuse	56	(9.3)		–	–	<b>8.9</b>	<b>(6.6–11.2)</b>
Alcohol misuse	362	(67.4)		–	–	<b>69.2</b>	<b>(65.0–73.5)</b>
Severe misuse	106	(46.5)		–	–	<b>46.0</b>	<b>(38.7–53.3)</b>

<sup>a</sup>Includes mental health and substance use disorders within past year: VA administrative documentation of ICD 295-298.xx, 300-304.xx, 305.0, 305.2, 305.3, 305.4, 305.5, 305.6, 305.7, 305.8, 306.xx, 307.1, 308-309, 311.xx and 312.3x

<sup>b</sup>PM=performance measure

<sup>c</sup>The 21 VA networks were ranked in order of the prevalence of overall discordance and labeled based on ranks with the letters A-U. Bold indicates that a network had significantly increased discordance compared to Network A, the network with the lowest discordance.

<sup>d</sup>AUDIT-C scores of 0 indicate no alcohol use in the past year. AUDIT-C scores 1–2 for women and 1–3 for men indicate low-level use; AUDIT-C scores 3–4 for women and 4 for men indicate possible alcohol misuse. AUDIT-C scores 5–7 indicate alcohol misuse and AUDIT-C scores 8–12 indicate severe alcohol misuse for both men and women.

<sup>e</sup>Referent group

<sup>f</sup>Bolded adjusted prevalence of discordance indicates a significant difference from the referent group in multivariable models.

random measurement error, changes in patients' drinking between two screens, or regression to the mean. However, the observed discordance between clinical and survey AUDIT-Cs cannot be accounted for by these factors. Discordance due to random measurement error and changes in drinking would be expected to occur in a normal distribution around the screening threshold similar to that observed across clinical screening scores (Fig. 2 Panel a). However, the distribution of discordance across survey screening scores was not normal (Fig. 1, Panel b). Furthermore, discordance due to changes in drinking should increase as the time between screens increases, but no such association was observed. Discordance cannot solely reflect regression to the mean or decreased drinking at repeat screening because there was no association between the order of screens and discordance. Furthermore, randomized controlled trials suggest that repeated screening leads to lower reported consumption on later screens.<sup>26–28</sup> This bias would be expected to result in lower AUDIT-C scores on the survey, which tended to follow the clinical screen, the opposite of the observed association.

Social desirability bias likely contributed to the observed discordance. Over twice as many patients who screened positive on confidential mailed surveys had discordant results compared to patients who had positive clinical screens (61% vs. 24%). If the latter is an estimate of the magnitude of "expected" discordance, 37% (95% CI 34–41%) of patients who screened positive on surveys had "excess" discordance. Patients may under-report alcohol consumption on clinical alcohol screening due to stigma or a desire to avoid discussing their drinking with providers.<sup>29</sup> The fact that Black patients were more likely than White patients to have discordant screening results might reflect greater social desirability bias among these patients, although it could also reflect bias due to differences in the way the AUDIT-C was interpreted and/or administered across racial/ethnic subgroups.

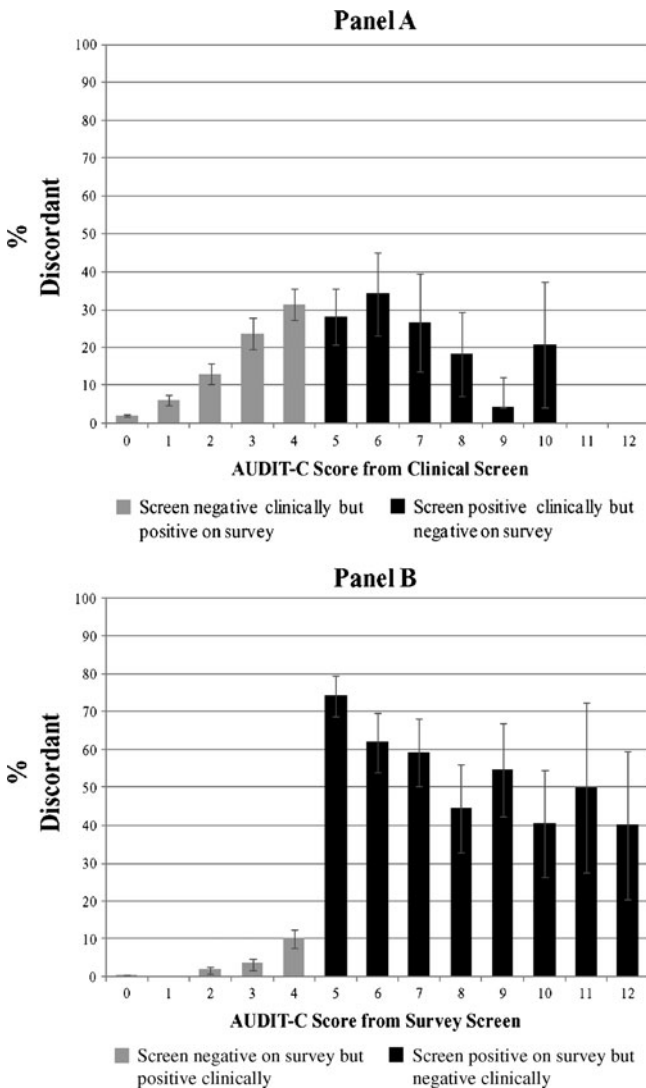
Factors other than social desirability bias likely contributed to the magnitude of the observed discordance for several reasons. First, the AUDIT-C was validated in studies that used interviewers to administer screens,<sup>10</sup> and in which patients

were aware providers would receive screening results.<sup>8</sup> Second, the variation in discordance across VA networks suggests that institutional factors contributed to the observed discordance. Anecdotal reports suggest that considerable variability exists across sites in the privacy of screening. Differences in training and/or decisions about who conducts screening might also contribute to variability; medical assistants or nurses may be more likely to follow screening instructions verbatim than primary care or mental health providers assessing alcohol use as part of the medical history.

Several limitations of this study should be noted. In order to study a large diverse national sample without alerting providers that the quality of screening was being evaluated, this study used AUDIT-Cs from confidential mailed surveys as a comparison standard. While in-depth interviews are often the ideal comparison standard, recruiting patients and providers for such studies biases results.<sup>30</sup> In addition, primary data collection would have delayed assessment of the quality of screening. Finally, use of secondary data allowed us to evaluate regional variation in the quality of alcohol screening in a cost-efficient manner. For these reasons, the AUDIT-C from a mailed survey was the best available comparison standard for this translational study of alcohol screening implementation.

Other limitations relate to the study sample. Patients who returned the VA's outpatient satisfaction survey are older and drink less than non-respondents,<sup>30</sup> potentially under-estimating discordance since patients who screened positive for alcohol misuse had higher discordance. The study sample was also too small to evaluate facility-level variation. VA patients differ in important ways from other outpatients, and the VA health care system is currently atypical in that it uses performance incentives to achieve high rates of alcohol screening. Finally, this study did not collect data on alcohol screening procedures across VA networks, so it could not determine whether differences in implementation of alcohol screening account for the observed differences in discordance across VA networks.

Nevertheless, this study has important implications for health care systems implementing routine alcohol screening,



**Figure 2.** Among patients with each AUDIT-C score based on clinical screens (Panel a) or survey screens (Panel b), the percent with discordant results on the other screen (N=6,861). Alcohol screening results are considered discordant if a patient screened positive on the clinical or survey screen but not the other. Panel a: The percent of patients with each AUDIT-C score (0–12 points) on the clinical screen who had discordant survey screen results. Panel b: The percent of patients with each AUDIT-C score (0–12 points) on the survey screen who had discordant clinical screen results.

Almost 7% of the total study sample screened positive for alcohol misuse on surveys but were missed by clinical screening, and 1.5% of the total sample had severe alcohol misuse that was missed. No known prior studies have evaluated the validity of alcohol screening when integrated into routine clinical care. This study found that use of validated questionnaires does not—by itself—ensure the quality of screening and suggests that the quality of clinical alcohol screening should be monitored, even when well-validated screening questionnaires are used. While it is unknown whether a similar issue affects screening for other health risk behaviors and mental health conditions, this issue merits evaluation. Self-administered measures of alcohol screening, whether on paper, online as in electronic health risk assessments (eHRA),<sup>31</sup> or by interactive voice recording on the telephone (IVR),<sup>32</sup> may be the most valid approaches to

implementing alcohol screening. If alcohol screening is administered by clinicians they may need focused training to prepare them to screen in a valid manner.<sup>33,34</sup>

This study is another demonstration of the challenge of developing effective performance measures for preventive care.<sup>35</sup> The current VA alcohol screening performance measure that sets targets for rates of alcohol screening creates incentives for documentation of screening results, but does not provide incentives for high quality screening that identifies patients with alcohol misuse. Furthermore, setting very high targets for screening could contribute to lower quality screening by encouraging providers to document screening when they do not have time to ask screening questions verbatim in a private setting. Future research must address the need for performance measures that create incentives for providers not only to screen, but to identify patients with alcohol misuse.

The VA has recently succeeded in achieving high rates of annual alcohol screening. Brief alcohol interventions are effective for patients with alcohol misuse,<sup>36</sup> and patients with alcohol use disorders benefit from referral or repeated brief interventions, with and without lab monitoring or medications.<sup>37–40</sup> This study suggests that mandating clinical use of a validated alcohol screening questionnaire does not ensure high quality screening. Three out of every five patients who screened positive for alcohol misuse on confidential mailed surveys were not identified by clinical screening. Put another way, 6,821 patients with alcohol misuse would be missed out of every 100,000 patients screened in VA. Moreover, significant variation across VA networks, after accounting for differences in patient characteristics, suggests organizational influences on the quality of alcohol screening. Together these findings indicate a need to focus on monitoring and improving the quality of alcohol screening in order to identify as many patients as possible who could benefit from brief alcohol interventions.

**Acknowledgements:** The authors greatly appreciate the VA Office of Quality and Performance for sharing their data for use in this study and feedback on several drafts of the manuscript.

The research reported here was supported by Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development Service. The manuscript was specifically produced with support from the Veterans Affairs Substance Use Disorders Quality Enhancement Research Initiative (SUB 98-000). Data for this study were provided by the VA Office of Quality and Performance.

The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the University of Washington.

**Conflict of Interest:** None disclosed.

**Corresponding Author:** Katharine A. Bradley, MD, MPH; Health Services Research & Development (HSR&D), Veterans Affairs (VA) Puget Sound Health Care System, 1100 Olive Way, Suite 1400, Seattle, WA 98101, USA (e-mail: Katharine.Bradley@va.gov).

## REFERENCES

1. Saitz R. Clinical practice. Unhealthy alcohol use. *N Engl J Med*. 2005;352(6):596–607.
2. Room R, Babor T, Rehm J. Alcohol and public health. *Lancet*. 2005;365(9458):519–530.

3. **Whitlock EP, Polen MR, Green CA, Orleans T, Klein J.** Behavioral counseling interventions in primary care to reduce risky/harmful alcohol use by adults: A summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med.* 2004;140:557-568.
4. **Kaner E, Beyer F, Dickinson H, et al.** Effectiveness of brief alcohol interventions in primary care populations. *Cochrane Database Syst Rev.* 2007;(2):CD004148.
5. **Solberg LI, Maciosek MV, Edwards NM.** Primary care intervention to reduce alcohol misuse ranking its health impact and cost effectiveness. *Am J Prev Med.* 2008;34(2):143-152.
6. **Bradley KA, Williams EC, Achtmeyer CE, Volpp B, Collins BJ, Kivlahan DR.** Implementation of evidence-based alcohol screening in the Veterans Health Administration. *Am. J. Manag. Care.* 2006;12(10):597-606.
7. **Lapham GT, Achtmeyer CE, Williams EC, Hawkins EJ, Kivlahan DR, Bradley KA.** Increased documented brief alcohol interventions with a performance measure and electronic decision support. *Med. Care.* In Press.
8. **Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA.** The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol use disorders identification test. *Arch Intern Med.* 1998;158(16):1789-1795.
9. **Bradley KA, Bush KR, Epler AJ, et al.** Two brief alcohol-screening tests From the Alcohol Use Disorders Identification Test (AUDIT): validation in a female Veterans Affairs patient population. *Arch Intern Med.* 2003;163(7):821-829.
10. **Bradley KA, DeBenedetti AF, Volk RJ, Williams EC, Frank D, Kivlahan DR.** AUDIT-C as a brief screen for alcohol misuse in primary care. *Alcohol Clin Exp Res.* 2007;31(7):1208-1217.
11. **Frank D, DeBenedetti AF, Volk RJ, Williams EC, Kivlahan DR, Bradley KA.** Effectiveness of the AUDIT-C as a screening test for alcohol misuse in three race/ethnic groups. *J Gen Intern Med.* 2008;23(6):781-787.
12. **Aertgeerts B, Buntinx F, Ansoms S, Fevery J.** Screening properties of questionnaires and laboratory tests for the detection of alcohol abuse or dependence in a general practice population. *Br J Gen Pract.* 2001;51(464):206-217.
13. **Rumpf HJ, Hapke U, Meyer C, John U.** Screening for alcohol use disorders and at-risk drinking in the general population: psychometric performance of three questionnaires. *Alcohol Alcohol.* 2002;37(3):261-268.
14. **Hawkins EJ, Kivlahan DR, Williams EC, Wright SM, Craig T, Bradley KA.** Examining quality issues in alcohol misuse screening. *Subst Abuse.* 2007;28(3):53-65.
15. **Goulet JL, Erdos J, Kancir S, et al.** Measuring performance directly using the veterans health administration electronic medical record: a comparison with external peer review. *Med Care.* 2007;45(1):73-79.
16. **Wright SM, Craig T, Campbell S, Schaefer J, Humble C.** Patient satisfaction of female and male users of Veterans Health Administration services. *J Gen Intern Med.* 2006;21(Suppl 3):S26-S32.
17. **Bradley KA, McDonell MB, Bush K, Kivlahan DR, Diehr P, Fihn SD.** The AUDIT alcohol consumption questions: reliability, validity, and responsiveness to change in older male primary care patients. *Alcohol Clin Exp Res.* 1998;22(8):1842-1849.
18. **Dybek I, Bischof G, Grothues J, et al.** The reliability and validity of the Alcohol Use Disorders Identification Test (AUDIT) in a German general practice population sample. *J Stud Alcohol.* 2006;67(3):473-481.
19. **Bradley KA, Williams EC, Achtmeyer CE, et al.** Measuring performance of brief alcohol counseling in medical settings: a review of the options and lessons from the Veterans Affairs (VA) health care system. *Subst Abuse.* 2007;28(4):133-149.
20. **Bradley KA, Kivlahan DR, Zhou XH, et al.** Using alcohol screening results and treatment history to assess the severity of at-risk drinking in Veterans Affairs primary care patients. *Alcohol Clin Exp Res.* 2004;28(3):448-455.
21. **Au DH, Kivlahan DR, Bryson CL, Blough D, Bradley KA.** Alcohol screening scores and risk of hospitalizations for GI conditions in men. *Alcohol Clin Exp Res.* 2007;31(3):443-451.
22. **Harris AH, Bryson CL, Sun H, Blough D, Bradley KA.** Alcohol screening scores predict risk of subsequent fractures. *Subst Use Misuse.* 2009;44:1055-1069.
23. **Bryson CL, Au DH, Sun H, Williams EC, Kivlahan DR, Bradley KA.** Alcohol screening scores and medication nonadherence. *Ann Intern Med.* 2008;149(11):795-804.
24. **Stata Statistical Software: Release Special Edition 11.1** [computer program]. College Station TX: Stata Corporation; 2007.
25. **Kinder LS, Bryson CL, Sun H, Williams EC, Bradley KA.** Alcohol screening scores and all-cause mortality in male Veterans Affairs patients. *J Stud Alcohol Drugs.* 2009;70(2):253-260.
26. **McCambridge J, Day M.** Randomized controlled trial of the effects of completing the Alcohol Use Disorders Identification Test questionnaire on self-reported hazardous drinking. *Addiction.* 2008;103(2):241-248.
27. **Kypri K, Langley JD, Saunders JB, Cashell-Smith ML.** Assessment may conceal therapeutic benefit: findings from a randomized controlled trial for hazardous drinking. *Addiction.* 2007;102(1):62-70.
28. **Moos RH.** Context and mechanisms of reactivity to assessment and treatment. *Addiction.* 2008;103(2):249-250.
29. **Davis CG, Thake J, Vilhena N.** Social desirability biases in self-reported alcohol consumption and harms. *Addict. Behav.* Nov 10 2009.
30. **Bradley KA, Bush KR, McDonell MB, Malone T, Fihn SD.** Screening for problem drinking: comparison of CAGE and AUDIT. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *J Gen Intern Med.* 1998;13(6):379-388.
31. **Reid RJ, Fishman PA, Yu O, et al.** Patient-centered medical home demonstration: a prospective, quasi-experimental, before and after evaluation. *Am. J. Manag. Care.* 2009;15(9):e71-e87.
32. **Helzer JE, Rose GL, Badger GJ, et al.** Using interactive voice response to enhance brief alcohol intervention in primary care settings. *J Stud Alcohol Drugs.* 2008;69(2):251-258.
33. **Babor TF, Stephens RS, Marlatt A.** Verbal report methods in clinical research on alcoholism: response bias and its minimization. *J Stud Alcohol.* 1987;48:410-424.
34. **Del Boca FK, Darkes J.** The validity of self-reports of alcohol consumption: state of the science and challenges for research. *Addiction.* 2003;98(Suppl 2):1-12.
35. **Walter LC, Davidowitz NP, Heineken PA, Covinsky KE.** Pitfalls of converting practice guidelines into quality measures: lessons learned from a VA performance measure. *JAMA.* 2004;291(20):2466-2470.
36. **Kaner E, Bland M, Cassidy P, et al.** Screening and brief interventions for hazardous and harmful alcohol use in primary care: a cluster randomised controlled trial protocol. *BMC Public Health.* 2009;9:287.
37. **Brown RL, Saunders LA, Bobula JA, Mundt MP, Koch PE.** Randomized-controlled trial of a telephone and mail intervention for alcohol use disorders: three-month drinking outcomes. *Alcohol Clin Exp Res.* 2007;31(8):1372-1379.
38. **Willenbring ML, Olson DH.** A randomized trial of integrated outpatient treatment for medically ill alcoholic men. *Arch Intern Med.* 1999;159(16):1946-1952.
39. **Willenbring ML.** Medications to treat alcohol dependence: adding to the continuum of care. *JAMA.* 2007;298(14):1691-1692.
40. **Anton RF, O'Malley SS, Ciraulo DA, et al.** Combined pharmacotherapies and behavioral interventions for alcohol dependence: the COMBINE study: a randomized controlled trial. *JAMA.* 2006;295(17):2003-2017.