UC San Diego

UC San Diego Previously Published Works

Title

Assessing an electronic self-report method for improving quality of ethnicity and race data in the Veterans Health Administration.

Permalink

https://escholarship.org/uc/item/7tg208v6

Journal

JAMIA Open, 6(2)

ISSN

2574-2531

Authors

Almklov, Erin
Cohen, Alicia J
Russell, Lauren E
et al.

Publication Date

2023-04-06

DOI

10.1093/jamiaopen/ooad020

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

JAMIA Open, 6(2), 2023, ooad020
https://doi.org/10.1093/jamiaopen/ooad020
Research and Applications

Research and Applications

Assessing an electronic self-report method for improving quality of ethnicity and race data in the Veterans Health Administration

Erin Almklov[1,2], Alicia J. Cohen[3,4,5], Lauren E. Russell[6], Maria K. Mor[7,8], Michael J. Fine[7,9], Leslie R. M. Hausmann[7,9], Ernest Moy[6], Donna L. Washington[10,11], Kenneth T. Jones[6], Judith A. Long[12,13], and James Pittman[1,2,14]

[1]VA Center of Excellence for Stress and Mental Health, San Diego, California, USA, [2]VA San Diego Healthcare System, San Diego, California, USA, [3]VA Center of Innovation in Long Term Services and Supports (LTSS-COIN), VA Providence Healthcare System, Providence, Rhode Island, USA, [4]Department of Family Medicine, Alpert Medical School of Brown University, Providence, Rhode Island, USA, [5]Department of Health Services, Policy, and Practice, Brown University School of Public Health, Providence, Rhode Island, USA, [6]VHA Office of Health Equity, Washington, District of Columbia, USA, [7]VA Center for Health Equity Research and Promotion, VA Pittsburgh Healthcare System, Pittsburgh, Pennsylvania, USA, [8]Department of Biostatistics, University of Pittsburgh School of Public Health, Pittsburgh, Pennsylvania, USA, [9]Division of General Internal Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA, [10]VA HSR&D Center for the Study of Healthcare Innovation, Implementation and Policy, VA Greater Los Angeles Healthcare System, Los Angeles, California, USA, [11]Division of General Internal Medicine and Health Services Research, Department of Medicine, University of California, Los Angeles Geffen School of Medicine, Los Angeles, California, USA, [12]Corporal Michael J. Crescenz VA Medical Center, VA Center for Health Equity Research and Promotion (CHERP), Philadelphia, Pennsylvania, USA, [13]Division of General Internal Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA and [14]Department of Psychiatry, University of California San Diego, San Diego, California, USA

Corresponding Author: Erin Almklov, PhD, VA San Diego Healthcare System, 3350 La Jolla Village Dr, San Diego, CA 92161, USA; erin.almklov@va.gov

Received 3 October 2022; Revised 18 January 2023; Editorial Decision 21 March 2023; Accepted 29 March 2023

ABSTRACT

Objective: Evaluate self-reported electronic screening (*eScreening*) in a VA Transition Care Management Program (TCM) to improve the accuracy and completeness of administrative ethnicity and race data.

Materials and Methods: We compared missing, declined, and complete (neither missing nor declined) rates between (1) *TCM-eScreening* (ethnicity and race entered into electronic tablet directly by patient using eScreening), (2) *TCM-EHR* (Veteran-completed paper form plus interview, data entered by staff), and (3) *Standard-EHR* (multiple processes, data entered by staff). The TCM-eScreening ($n=7113$) and TCM-EHR groups ($n=7113$) included post-9/11 Veterans. Standard-EHR Veterans included all non-TCM Gulf War and post-9/11 Veterans at VA San Diego ($n=92\,921$).

Results: *Ethnicity*: TCM-eScreening had lower rates of missingness than TCM-EHR and Standard-EHR (3.0% vs 5.3% and 8.6%, respectively, $P<.05$), but higher rates of "decline to answer" (7% vs 0.5% and 1.2%, $P<.05$). TCM-EHR had higher data completeness than TCM-eScreening and Standard-EHR (94.2% vs 90% and 90.2%, respectively, $P<.05$). *Race*: No differences between TCM-eScreening and TCM-EHR for missingness (3.5% vs 3.4%, $P>.05$) or data completeness (89.9% vs 91%, $P>.05$). Both had better data completeness than Standard-EHR ($P<.05$), which despite the lowest rate of "decline to answer" (3%) had the highest missingness (10.3%) and lowest overall completeness (86.6%). There was strong agreement between TCM-eScreening and TCM-

Published by Oxford University Press on behalf of the American Medical Informatics Association 2023.
This work is written by (a) US Government employee(s) and is in the public domain in the US.

EHR for ethnicity (Kappa = .92) and for Asian, Black, and White Veteran race (Kappas = .87 to .97), but lower agreement for American Indian/Alaska Native (Kappa = .59) and Native Hawaiian/Other Pacific Islander (Kappa = .50) Veterans.
Conculsions: eScreening is a promising method for improving ethnicity and race data accuracy and completeness in VA.

Key words: electronic screening, Veterans, health equity, racial and ethnic disparities

LAY SUMMARY
The Veterans Health Administration (VHA) depends on accurate ethnicity and race data to identify and address disparities in Veteran health and healthcare. This investigation evaluated self-reported electronic screening (*eScreening*) in a VA Transition Care Management Program (TCM) as a method for improving the completeness and accuracy of ethnicity and race data in the VHA electronic health record (EHR). We compared 3 methods of collecting ethnicity and race data: (1) Veteran-reported and entered data with eScreening (TCM-eScreening), (2) Veteran-reported, staff-entered data (TCM-EHR), and (3) standard methods for EHR data collection (Standard-EHR). We found that both methods of Veteran-reported data collection (TCM-eScreening and TCM-EHR) had significantly lower rates of missing data compared to Standard-EHR. Further, there was overall strong agreement between Veteran-reported methods for ethnicity and most racial groups. eScreening appears to be a promising method for improving ethnicity and race data accuracy and completeness.

INTRODUCTION

Although the Veterans Health Administration (VHA) is committed to improving the equity of Veterans' health and healthcare,[1,2] ethnic and racial disparities remain in Veteran healthcare experiences, healthcare utilization, and clinical outcomes.[1,3–7] Accurate administrative ethnicity and race data are critical for detecting, understanding, and ultimately eliminating such disparities.

Healthcare systems, such as the VHA, are generally well-positioned to collect patient demographic data as they tend to have information systems for data collection and reporting, staff who are used to collecting registration and admissions data, and an organizational culture that is familiar with the tools of quality improvement.[8] However, a 2019 U.S. Government Accountability Office (GAO) report[9] noted weaknesses in the completeness and accuracy of VHA ethnicity and race data, calling for evaluation and corrective action. The completeness of VHA's ethnicity and race data is generally better than its public and private sector counterparts.[7] Yet, prior work has demonstrated elevated levels of missing ethnicity and race data in the VHA electronic health records (EHR), as well as concerns that ethnicity and race data are sometimes recorded based on staff "observation" rather than Veteran self-report[10–16]—the gold standard for accurate collection of ethnicity and race data.[17–22] The explicit option for "observation" in VHA may result in undercounting of certain population groups[23–26] and lead staff to incorrectly believe that observation is comparable to Veteran self-report of ethnicity and race. Furthermore, the default recorded source for ethnicity and race data is "self-reported," meaning that unless staff manually correct the recorded data source it is impossible to track which data are based on Veteran self-report versus staff observation. Ultimately, both missing data and data based on staff observation—which can lead to misclassification and inaccurate data capture—may create significant difficulties in accurately identifying ethnic and racial health disparities among Veterans within VHA.[9]

Research suggests that patient-administered data collection via electronic tools can reduce certain sources of error and biases.[27] Additional benefits include time savings, broader reach, better use of organizational resources, and potential flexibility in data collection setting.[27–29] Self-administered questionnaires are also associated with higher rates of disclosure for potentially sensitive information than face-to-face or telephone interviews[30–33] with at least one study suggesting the highest levels of reporting noted for audio-computer self-administration questionnaires and computer-assisted self-completion interviews.[27]

The VHA utilizes several health information technology resources for Veterans to support their healthcare management including eScreening, a VHA developed web-based electronic screening platform used to collect self-reported health and demographic information, including ethnicity and race[34] (see Figure 1). eScreening was developed by the VA Center of Excellence for Stress and Mental Health (CESAMH) at the VA San Diego Healthcare System (VASDHS). Human-centered design was used in the development, testing, and integration of the software with the VA information technology systems, with input from clinical, research, informatics, and technology subject matter experts. eScreening was further refined with Veteran and staff feedback.[35] Using eScreening, Veterans can self-administer questions via secure webform, like other commonly used web-based data capture tools such as Qualtrics. eScreening supports a variety of patient self-report data elements and is unique among all other available systems in VHA because it can easily be customized for any program's specific needs. Highlighted features of the eScreening program include (1) the ability for Veterans to enter screening information directly without the involvement of a clinician and outside of the clinical session; (2) immediate scoring of measures; (3) instant patient feedback via printable handout with patient-friendly summary results; (4) an editable note generated in the EHR system; (5) electronic data storage with the ability to easily monitor health-related symptoms over time; (6) generating aggregate clinic data for managers; and (7) secure real-time alerts to

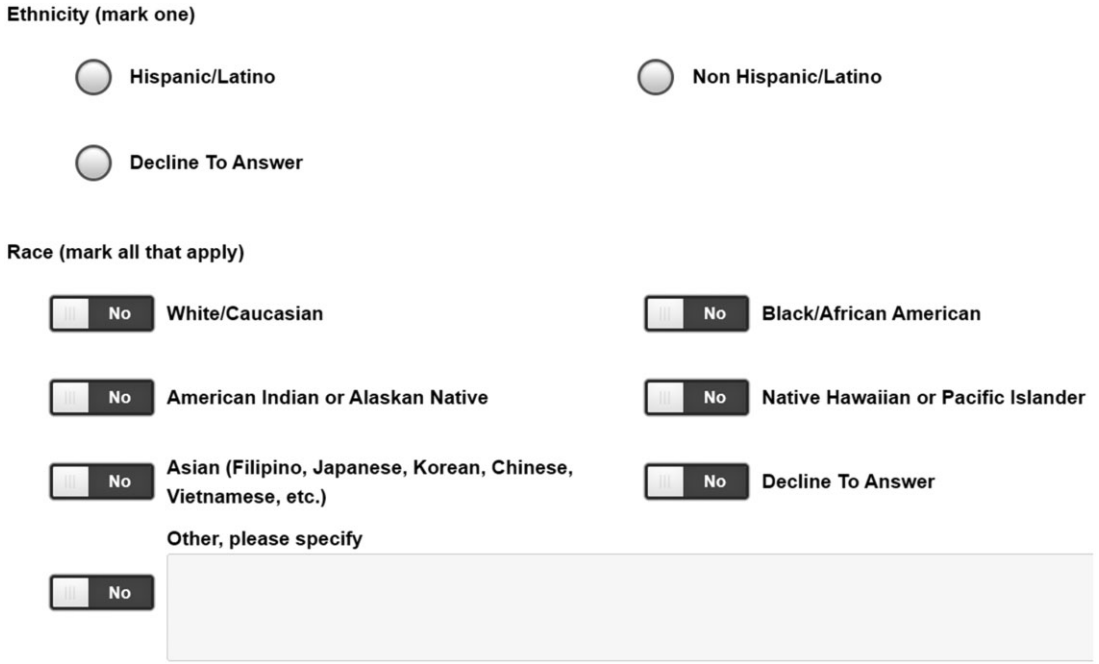


Figure 1. eScreening ethnicity and race questions.

clinicians for evaluating and triaging patients. The platform also connects directly to the VHA EHR, thus allowing clinical care teams to review responses in real time.

Given the successful implementation of eScreening at VA San Diego Healthcare System, expansion to additional sites, and plans to scale-up eScreening nationally, our team leveraged the opportunity to assess eScreening as a method for improving the completeness and accuracy of ethnicity and race data in the VHA EHR. Specifically, we aimed to (1) evaluate the effectiveness of eScreening in improving the completeness of ethnicity and race data within VHA, and (2) examine the concordance of Veteran-reported, Veteran-entered ethnicity and race data in eScreening with staff-entered ethnicity and race data in the EHR.

MATERIALS AND METHODS

Since eScreening was originally launched in 2013, it has been used in the VA San Diego Healthcare System Transition Care Management (TCM) clinic—a national program designed to support the reintegration needs of transitioning service members and Post-9/11 Veterans. Post-9/11 Veterans presenting to the VA San Diego Healthcare System enrollment office and identified as eligible for the TCM program complete an eScreening assessment via electronic tablet.

We conducted a retrospective analysis comparing the completeness and concordance of Veteran-reported, Veteran-entered ethnicity and race data via eScreening compared with standard EHR data collection methods between 2015 (when eScreening data collection began) and 2020 in the VA San Diego Healthcare System. Both eScreening and EHR ethnicity and race data were collected using standard Office of Management and Budget (OMB) ethnicity (Hispanic or Latino, or Not Hispanic or Latino) and race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White) categories. Additional demographic information was not collected as part of this project, though other eScreening modules include questions about other demographic information, such as gender identity, sexual orientation, relationship status, age, education, employment, income, language, military history, etc.

Data collection methods

We extracted ethnicity and race data from 3 sources—the TCM-eScreening, the TCM-EHR fields, and the standard-EHR fields (see Figure 2). In this study, eScreening was only used to collect ethnicity and race data during a Veteran's initial visit.

We extracted TCM-eScreening data from the eScreening program database which stores raw self-report data and is located on the Amazon Web Services government cloud. At the time of data collection, eScreening was a new technology, and data were not pushed directly into the medical record system. However, ethnicity and race data reported in eScreening is now captured in structured EHR data and eScreening could be used to populate the official EHR's ethnicity and race data fields.

We extracted TCM-EHR ethnicity and race data from the Corporate Data Warehouse (CDW), a national repository of VHA EHR data from Veterans Affairs (VA) clinical and administrative systems. As part of the standard process for enrolling for healthcare services in VHA, all Veterans presenting at the VA San Diego Healthcare System, enrollment office and identified as being eligible for the TCM program are asked to fill out a paper form (VA 10-10EZ)[36]—the official form for Veterans to apply for VHA health benefits—followed by an interview with enrollment staff to assess eligibility and explain VHA health benefits. The VA 10-10EZ form captures demographic information including Veteran-reported ethnicity and race. Data from the demographic form along with any additional information gleaned during the interview are entered into the EHR by enrollment staff, including the official EHR ethnicity and race fields.

Standard EHR ethnicity and race data were also extracted from the CDW. Capture of ethnicity and race data at VA San Diego Healthcare System outside of the TCM clinic include either Veteran-

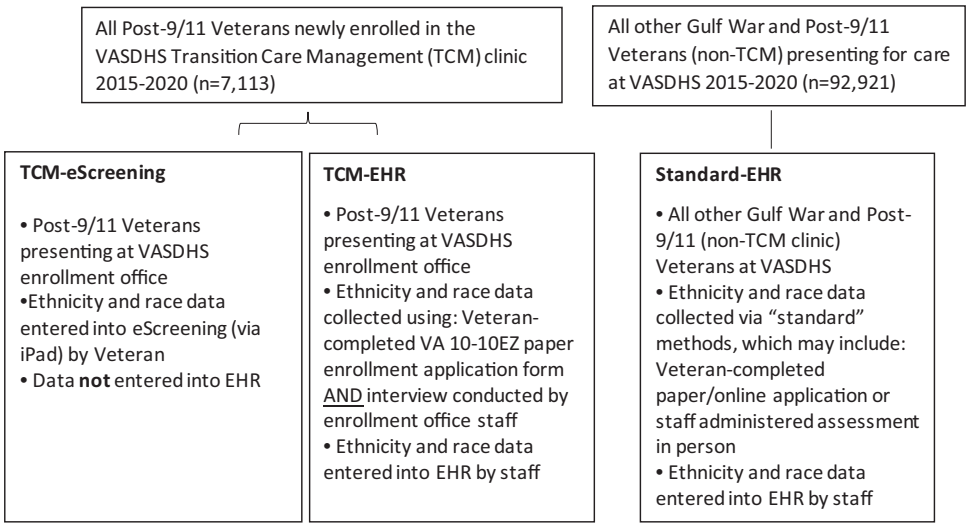


Figure 2. Schematic of 3 methods of ethnicity and race data collection at VA San Diego Healthcare System.

completed online or paper (in person or by mail) 10-10EZ forms, or staff-administered questions in person at a VA San Diego Healthcare System, enrollment office or in the emergency department. All data are entered into the EHR by enrollment staff and were extracted from the CDW.

Analytic plan

We linked Veteran-level eScreening ethnicity and race data (TCM-eScreening) to CDW data (TCM-EHR) using unique identifiers (eg, Vista Internal Entry Number, which is a designator that uniquely identifies patients within a VHA facility) to compare the completeness and concordance of ethnicity and race data between these 2 data collection methods. We also compared the completeness of ethnicity and race data for TCM-eScreening and TCM-EHR to Standard-EHR data collection.

For each method, ethnicity was documented as: Hispanic or Latino, not Hispanic or Latino, declined to answer, or missing. In the EHR, there was an additional category of "unknown." For the purposes of these analyses, we coded "unknown" as missing. For each method, race was documented as: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, declined to answer, or missing. In TCM-eScreening, there was also an additional category of "other." Race was categorized as "more than one race" for Veterans who selected more than one race category. In the EHR, if no race was recorded or race was recorded as "unknown," it was categorized as missing.

A major difference between methods is in the "decline to answer" response. In eScreening (TCM-eScreening), "decline to answer" is offered as an explicit response option for questions on ethnicity and race. Conversely, "decline to answer" is not an offered response option on the standard VA enrollment form (10-10EZ) used for TCM-EHR and Standard-EHR data collection, however, when staff are entering ethnicity and race data into the EHR, there is a response field that allows staff to record ethnicity and/or race as "decline to answer."

We analyzed ethnicity and race data separately. We used descriptive statistics to characterize the percent of Veterans with complete ethnicity and race data for each data collection method. We compared the rates of "missing," "decline to answer," and "complete" (neither missing/unknown nor declined) responses across the 3 groups: TCM-eScreening, TCM-EHR, and Standard-EHR. Specifically, we performed chi-square tests of independence to examine the relationship between data collection method (TCM-eScreening, TCM-EHR, Standard-EHR) and data completeness (complete, missing, decline to answer). Significant chi-square tests were followed with Bonferroni-corrected *post hoc* *z*-tests for independent proportions to determine which pairwise comparisons were significant at $P < .05$.

Table 1. Accuracy of TCM-eScreening with TCM-HER

	Kappa	Description
Ethnicity (Hispanic/Latino)	.92	Very good
Asian	.87	Very good
White	.91	Very good
Black or African American	.97	Very good
American Indian or Alaska Native	.59	Moderate
Native Hawaiian or Other Pacific Islander	.50	Moderate

To assess concordance, we describe the agreement among complete TCM-eScreening (Veteran-reported, Veteran-entered, considered the gold-standard) and TCM-EHR (Veteran-reported, staff entered) records. Categories used to assess ethnicity concordance were Hispanic/Latino and Not Hispanic/Latino. Race concordance categories were Yes and No responses for each of the following: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, and White or Caucasian. Declined to answer and missing/unknown responses were not included. We calculated Cohen's kappa for a measure of the extent of agreement beyond chance (see Table 1). Agreement statistics were not calculated for the non-matched Standard-EHR sample as this represented a different (Gulf War and Post-9/11, non-TCM clinic) population of Veterans. All descriptive statistics were generated using SPSS 26.

RESULTS

TCM-eScreening and TCM-EHR cohorts included the same 7113 Post-9/11 Veterans newly enrolled in the TCM clinic from 2015 to 2020. The Standard-EHR cohort included a non-matched sample of

all other Gulf War and Post-9/11 Veterans presenting for care at VA San Diego Healthcare System, but not enrolled in the TCM clinic, during the same 2015–2020 period ($n = 92\,921$). Table 2 shows the frequency and percentage of reported ethnicity and race categories by data collection method.

Ethnicity

There were significant differences among the 3 data collection methods (TCM-eScreening, TCM-EHR, Standard-EHR) in the rates of data completeness for ethnicity ($\chi^2 = 1842.49$, $P < .01$, Figure 2). TCM-eScreening had significantly lower rates of missingness compared with TCM-EHR (3% vs 5.3%; $P < .05$) and Standard-EHR (3% vs 8.6%; $P < .05$). However, TCM-eScreening had significantly higher rates of "decline to answer" compared with TCM-EHR (7% vs 0.5%; $P < .05$) and Standard-EHR (7% vs 1.2%; $P < .05$). TCM-EHR had significantly higher completeness when compared to TCM-eScreening and Standard-EHR (94.2% vs. 90% and 90.2%, respectively, $P < .05$) (Figure 3).

Race

There were significant differences among the data collection methods (TCM-eScreening, TCM-EHR, Standard-EHR) in the rates of

Table 2. Reported ethnicity and race by data collection method

	TCM-eScreening[a] ($n = 7113$)	TCM-EHR[a] ($n = 7113$)	Standard-EHR[b] ($n = 92\,921$)
Ethnicity			
Not Hispanic or Latino, *n* (%)	4704 (66.1)	5124 (72.0)	67 535 (72.7)
Hispanic or Latino, *n* (%)	1700 (23.9)	1575 (22.1)	16 308 (17.6)
Decline to answer, *n* (%)	496 (7.0)	39 (0.5)	1098 (1.2)
Missing/unknown, *n* (%)	213 (3.0)	375 (5.3)	7980 (8.6)
Race			
White, *n* (%)	3559 (50.0)	3943 (55.4)	51 044 (54.9)
Black or African American, *n* (%)	1110 (15.6)	1239 (17.4)	16 371 (17.6)
American Indian or Alaska Native, *n* (%)	96 (1.3)	100 (1.4)	984 (1.1)
Native Hawaiian or Other Pacific Islander, *n* (%)	76 (1.1)	191 (2.7)	1720 (1.9)
Asian, *n* (%)	832 (11.7)	766 (10.8)	8275 (8.9)
Other, *n* (%)	214 (3.0)	n/a	n/a
More than one race, *n* (%)	508 (7.1)	233 (3.3)	2093 (2.3)
Decline to answer, *n* (%)	470 (6.6)	396 (5.6)	2827 (3.0)
Missing/unknown, *n* (%)	248 (3.5)	245 (3.4)	9607(10.3)

Note: "other" was not an option for TCM-EHR and Standard-EHR.
EHR: electronic health record; TCM: Transition Care Management.
[a]All Post-9/11 Veterans newly enrolled in the VA San Diego Healthcare System Transition Care Management (TCM) clinic 2015–2020.
[b]All other Gulf War and Post-9/11 Veterans (non-TCM clinic) presenting for care at VA San Diego Healthcare System 2015–2020.

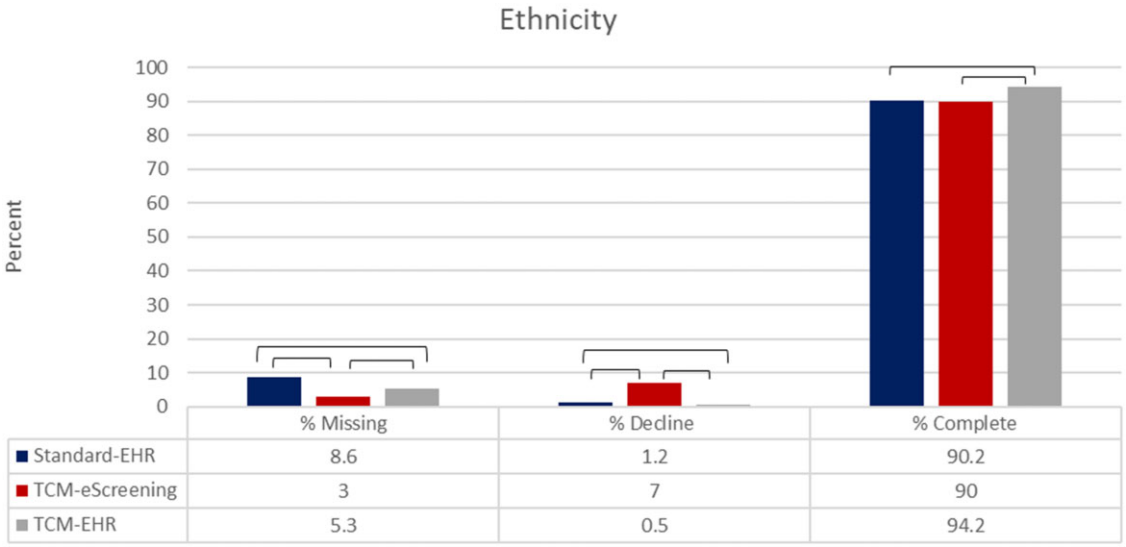
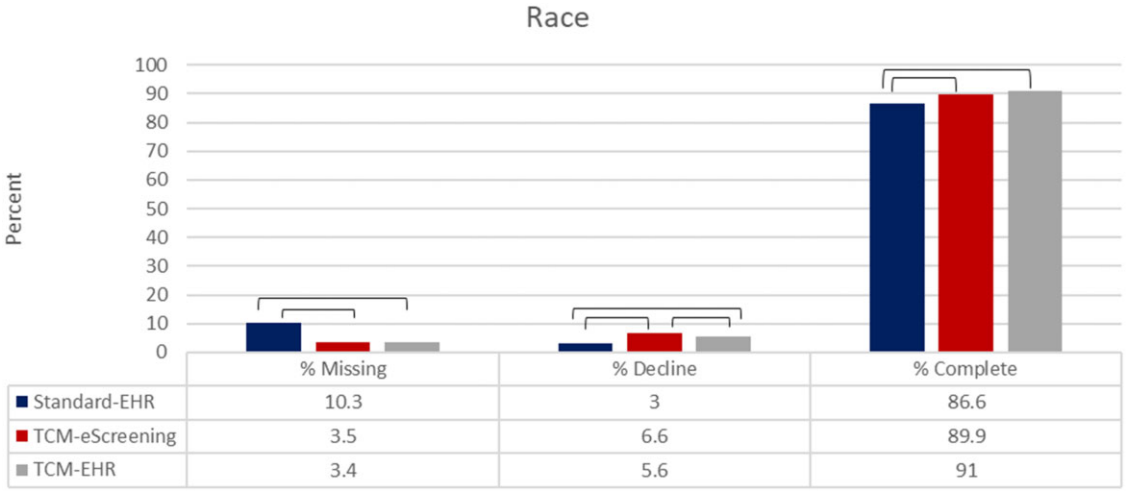


Figure 3. Completeness of ethnicity data by data collection method. Chi-square test of independence between data collection method (TCM-eScreening, TCM-EHR, Standard-EHR) and ethnicity data completeness (complete, missing, decline to answer) was significant ($\chi^2 = 1842.49$, $P < .01$). Brackets show significant differences for pair-wise comparisons ($P < .05$). EHR: electronic health record; TCM: Transition Care Management.



	% Missing	% Decline	% Complete
Standard-EHR	10.3	3	86.6
TCM-eScreening	3.5	6.6	89.9
TCM-EHR	3.4	5.6	91

Figure 4. Completeness of race data by data collection method. Chi-square test of independence between data collection method (TCM-eScreening, TCM-EHR, Standard-EHR) and race data completeness (complete, missing, decline to answer) was also significant ($\chi^2 = 982.64$, $P < .01$). Brackets show significant differences for pair-wise comparisons ($P < .05$). EHR: electronic health record; TCM: Transition Care Management.

data completeness for race ($\chi^2 = 982.64$, $P < .01$). While there were small but significant differences ($P < .05$) in rates of "decline to answer" between TCM-eScreening (6.6%) and TCM-EHR (5.6%), there were no significant differences between TCM-eScreening and TCM-EHR for rates of either missingness (3.5% vs 3.4%) or overall data completeness (89.9% vs 91%). Both TCM-eScreening and TCM-EHR had significantly better data completeness than Standard-EHR ($P < .05$), which despite having the lowest, rate of "decline to answer," (3%) had the highest rate of missingness (10.3%) and the lowest rate of overall completeness (86.6%) (Figure 4).

Concordance

Comparing TCM-eScreening with TCM-EHR, we found overall strong agreement for ethnicity (Kappa = .92) and Asian, White, and Black or African American race (Kappa's = .87, .91, and .97, respectively), with moderate agreement for American Indian or Alaska Native (Kappa = .59) and Native Hawaiian or Other Pacific Islander (Kappa = .50) Veterans.

DISCUSSION

In response to a GAO report calling for improved practices to better identify and address racial and ethnic disparities within the VHA, this investigation aimed to compare the completeness and concordance of Veteran-reported, Veteran-entered ethnicity and race data collected via a self-administered electronic screening platform (TCM-eScreening), versus Veteran-reported, staff-entered ethnicity and race data (TCM-EHR) and standard methods for EHR ethnicity and race data collection (Standard-EHR) within the VA San Diego Healthcare System. We found that both methods of Veteran-reported data collection (TCM-eScreening and TCM-EHR) were superior to Standard-EHR in terms of overall completeness of race data, and TCM-EHR was also superior to Standard-EHR for overall completeness of ethnicity data. While TCM-eScreening had significantly lower rates of missingness for ethnicity data than Standard-EHR, TCM-eScreening had significantly higher rates of "decline to answer," resulting in comparable rates of overall ethnicity data completeness.

TCM-eScreening was the only method of data collection with an explicit "decline to answer" response option. When given an option to "decline to answer" in the eScreening demographic module, 7% of Veterans declined to provide their ethnicity and 7% declined to provide their race, which was significantly higher than rates of declining to answer these questions in TCM-EHR and Standard-EHR groups where "declined to answer" was not an offered response option on the intake form but staff could subsequently document in the EHR that the Veteran declined. This raises the question as to whether VA should consider excluding "decline to answer" as an explicit response category when collecting data on ethnicity and race, since this is not one of the standard OMB categories and its inclusion results in lower rates of overall data completeness. Research also suggests that some Hispanic/Latino patients may not identify with any of the OMB-defined race categories, raising questions about the usefulness of the 2-question format of collecting ethnicity and race as separate fields, as well as how patients interpret the constructs of "race" and "ethnicity."[37–39] Potential changes to the existing OMB fields might include combining the ethnicity and race questions into a single question, which is consistent with an initial set of recommendations proposed by a Federal Interagency Technical Working Group on Race and Ethnicity Standards convened by OMB in 2022.[40] This change might also help reduce missingness, since race and ethnicity would be asked as a single question and allow people to identify as multiple ethnicities and races ("check all that apply"). Another potential change that had been previously proposed for the 2020 Census and was reiterated in the OMB Interagency Technical Working Group's initial recommendations would be to add a response category for Middle Eastern and North African. Alternatively, as suggested by others,[41] the existing OMB categories may need to be replaced by some other epistemology for describing human similarities and differences, such as taking into account the lived experiences of racial and ethnic minoritized communities.

Additional barriers to collection of ethnicity and race data noted in non-VA healthcare settings include patients lacking clarity on why these data are being collected or how they will be used, concerns about confidentiality, as well as concerns about potential bias or discrimination based on responses.[42,43] Future work should seek

to further understand implications of and patient preferences around including or not including an explicit "decline to answer" response option when collecting ethnicity and race data, reasons why Veterans may decline to answer questions about their ethnicity and race, and ways to help mitigate potential concerns. eScreening modules could also consider including an explanation about why information on ethnicity and race is being collected and how it will be used to enhance patient care services, as this may lead to more complete and accurate responding.

There was overall strong agreement between TCM-eScreening and TCM-EHR methods for ethnicity and most racial groups, with the notable exceptions of only moderate agreement for Veterans who identified as American Indian or Alaska Native and Native Hawaiian or Other Pacific Islander. These findings are largely consistent with earlier studies comparing Veteran ethnicity and race data self-reported via surveys compared with administrative data sources including the VA EHR, VA Defense Identity Repository (VADIR), and Medicare data.[7,44,45] Notably, American Indian or Alaska Native and Native Hawaiian or Other Pacific Islander Veterans may be more likely to identify with more than one race category, which might impact coding and agreement. For instance, of those who identified with American Indian or Alaska Native and Native Hawaiian or Other Pacific Islander, 63% and 52%, respectively, also identified with another race category and were coded as more than one race for TCM-eScreening, which had a higher percentage of Veterans coded as more than one race than did TCM-EHR. The relatively lower agreement for certain race categories might also reflect staff bias or error. Further, the fact that TCM-eScreening was the only data collection method with the option of "other" as a race category, which 3% of Veterans selected, might also have impacted the current findings.

These findings highlight discrepancies between Veteran-reported (gold-standard) and standard VHA EHR data collection methods, which underscores the need for collecting self-reported data, particularly among certain racial and ethnic groups. Our finding that the 2 different sources of Veteran-reported ethnicity and race data in our study (TCM-eScreening and TCM-EHR)—collected in the same population of Veterans during the same encounter—were largely concordant but not identical was unexpected. Notably, for the TCM-eScreening, Veterans entered their reported ethnicity and race and other demographic information directly into an electronic tablet; while for the TCM-EHR, Veterans filled out a paper form followed by an interview with enrollment staff during which time staff entered Veteran-reported data on ethnicity and race into the EHR. Having an additional step of staff recording Veteran-reported data (as in the TCM-EHR group) rather than Veterans being able to record their responses directly may have introduced error. It would be surprising that the same Veterans would have answered the same questions differently; but perhaps the different format led to Veterans changing answers. As noted previously, TCM-eScreening had a higher percentage of Veterans coded as more than one race than did TCM-EHR, and TCM-eScreening was also the only method with "other" as a response option. Overall, to accurately measure and ultimately eliminate health disparities, VA facilities should select and implement a single method for Veteran-reported, Veteran-entered ethnicity and race data collection that directly populates these fields in the VHA EHR. The chosen method should serve as the VA standard to decrease missingness, improve completeness, and improve accuracy. Potentially promising options include eScreening, which now connects with the VA EHR, or incorporating other electronic Veteran-reported methods such an online (10-10EZ) form into online patient portals or in-person check-in kiosks to directly populate fields in the EHR.

While electronic self-administered screening systems have largely been investigated for their use in collecting patient-reported outcome data, such as health-related symptoms,[44,46] their use in collecting sensitive demographic information is limited. Existent research, however, indicates that health information technology may have the potential to improve the collection and exchange of self-reported ethnicity and race data,[17] and that patient-facing tools give patients the opportunity to fill out or review their information directly. A previous study demonstrated improvement in ethnicity and race data quality after using a custom patient portal application on a tablet computer to allow patients to review their demographic information.[47] Prior work has found individuals often prefer electronic screening over other forms of data collection for sensitive subject matters.[46] Electronic (eg, computer-based) self-administered questionnaires may also result in more complete item response rates than paper and pencil methods.[27] Furthermore, these electronic self-report platforms are feasible and acceptable to a variety of patients and populations.[35,48] Thus, expanding on previous literature, the current findings indicate that electronic self-report screening systems, like eScreening, can be effective tools to accurately and completely collect ethnicity and race data which has important research, clinical, and evidence-based public policy implications.

Limitations of this investigation include the lack of a matched sample for the Standard-EHR cohort, slight differences in category response options among the data collection methods, data collection from only one VHA site, and the inability to calculate agreement statistics for the non-matched Standard-EHR sample as this represented a different (Gulf War and Post-9/11, non-TCM clinic) population of Veterans. Further, we are only able to speculate as to why the concordance was low for some of the smaller race categories. Notwithstanding these limitations, the findings suggest that implementing a standardized self-administered method of collecting ethnicity and race data, such as eScreening, can improve the ability of healthcare systems to monitor and address disparities affecting ethnic and racial minority patients. Further, excluding "decline to answer" as an explicit response category when collecting data on ethnicity and race may be considered, as its inclusion results in lower rates of overall data completeness and patients would still retain the ability to leave the data field(s) blank if desired. Overall, to improve health equity, healthcare organizations such as VHA should prioritize ensuring complete and accurate ethnicity and race data and use these data to guide healthcare improvement efforts. eScreening and the like represent promising tools to aid in such efforts and to support healthcare systems in advancing health equity.

FUNDING

This project was funded by VA Health Services Research & Development grant FOP 21-144 and supported by the VA Center of Excellence for Stress and Mental Health (CESAMH). This material also is the result of work supported with resources of the VA San Diego Healthcare System. AJC was additionally supported by VA Health Services Research & Development grant CDA 20-037.

AUTHOR CONTRIBUTIONS

EA contributed to the project design, data management and analysis, data interpretation, and drafting and revising the article. AJC and LER contributed to the project design, data interpretation, and

drafting and revising the article. MM contributed to data management and manuscript revisions. MJF, LRMH, EM, DLW, KJ, and JAL contributed to manuscript revisions. JP contributed to the project design and coordination, data collection, management and interpretation, and drafting and revising the article.

ACKNOWLEDGMENTS

This work was supported by many offices at the United States Department of Veterans Affairs, including the VA Center of Excellence for Stress and Mental Health and the VHA Office of Health Equity, Health Services Research & Development, Center for Health Equity and Resource Promotion, and Health Equity-Quality Enhancement Research Initiative National Partnered Evaluation Center, VA Greater Los Angeles. We also thank Juliette McClendon, PhD, Naomi Tomoyasu, PhD, and Joy Toyama, DrPH for their diligent work on this project.

DISCLOSURES

The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of Veteran Affairs or any of the institutions with which the authors are affiliated.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

As this project was completed as a quality improvement project (QIP), IRB approval and informed consent were not required.

CONFLICT OF INTEREST STATEMENT

None declared.

DATA AVAILABILITY

The datasets supporting the conclusions of this article are available from the corresponding author upon reasonable request.

REFERENCES

1. Saha S, Freeman M, Toure J, Tippens KM, Weeks C, Ibrahim S. Racial and ethnic disparities in the VA health care system: a systematic review. *J Gen Intern Med* 2008; 23 (5): 654–71.
2. Uchendu US. Institutional journey in pursuit of health equity: Veterans Health Administration's Office of Health Equity. *Am J Public Health* 2014; 104 (Suppl 4): S511–13.
3. Atkins D, Kilbourne A, Lipson L. Health equity research in the Veterans Health Administration: we've come far but aren't there yet. *Am J Public Health* 2014; 104 (S4): S525–6.
4. Samuel CA, Landrum MB, McNeil BJ, Bozeman SR, Williams CD, Keating NL. Racial disparities in cancer care in the Veterans affairs health care system and the role of site of care. *Am J Public Health* 2014; 104 (S4): S562–71.
5. Washington DL, Villa V, Brown A, Damron-Rodriguez J, Harada N. Racial/ethnic variations in veterans' ambulatory care use. *Am J Public Health* 2005; 95 (12): 2231–7.
6. Hausmann LRM, Gao S, Mor MK, Schaefer JH, Fine MJ. Understanding racial and ethnic differences in patient experiences with outpatient health care in Veterans Affairs Medical Centers. *Med Care* 2013; 51 (6): 532–9.
7. Hernandez SE, Sylling PW, Mor MK, *et al*. Developing an algorithm for combining race and ethnicity data sources in the Veterans Health Administration. *Mil Med* 2020; 185 (3–4): e495–500.
8. *Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement*. Washington, D.C.: National Academies Press; 2009. http://www.nap.edu/catalog/12696. Accessed April 6, 2022.
9. *VA Health Care: Opportunities Exist for VA to Better Identify and Address Racial and Ethnic Disparities*. U.S. Government Accountability Office (GAO-20-83); 2019. https://www.gao.gov/assets/gao-20-83.pdf. Accessed April 6, 2022.
10. Mays VM, Ponce NA, Washington DL, Cochran SD. Classification of race and ethnicity: implications for public health. *Annu Rev Public Health* 2003; 24 (1): 83–110.
11. McAlpine DD, Beebe TJ, Davern M, Call KT. Agreement between self-reported and administrative race and ethnicity data among medicaid enrollees in Minnesota: self-reported and administrative race and ethnicity data. *Health Serv Res* 2007; 42 (6 Pt 2): 2373–88.
12. Klinger EV, Carlini SV, Gonzalez I, *et al*. Accuracy of race, ethnicity, and language preference in an electronic health record. *J Gen Intern Med* 2015; 30 (6): 719–23.
13. Boehmer U, Kressin NR, Berlowitz DR, Christiansen CL, Kazis LE, Jones JA. Self-reported vs administrative race/ethnicity data and study results. *Am J Public Health* 2002; 92 (9): 1471–2.
14. Kressin NR, Chang BH, Hendricks A, Kazis LE. Agreement between administrative data and patients' self-reports of race/ethnicity. *Am J Public Health* 2003; 93 (10): 1734–9.
15. Kressin NR. Race/ethnicity identification: vital for disparities research, quality improvement, and much more than "meets the eye. *Med Care* 2015; 53 (8): 663–5.
16. Stroupe KT, Tarlov E, Zhang Q, Haywood T, Owens A, Hynes DM. Use of medicare and DOD data for improving VA race data quality. *J Rehabil Res Dev* 2010; 47 (8): 781–95.
17. Ulmer C, McFadden B, Nerenz DR, eds. *Institute of Medicine (US) Subcommittee on Standardized Collection of Race/Ethnicity Data for Healthcare Quality Improvement. Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement*. Washington (DC): National Academies Press (US); 2009. http://www.ncbi.nlm.nih.gov/books/NBK219756/. Accessed April 10, 2022.
18. Higgins PC, Taylor EF. *Measuring Racial and Ethnic Disparities in Health Care: Efforts to Improve Data Collection*. Washington, DC: Mathematica Policy Research, Incorporated; 2009.
19. OMB (Office of Management and Budget). *Recommendations from the Interagency Committee for the Review of the Racial and Ethnic Standards to the Office of Management and Budget Concerning Changes to the Standards for the Classification of Federal Data on Race and Ethnicity*. 1997. https://www.federalregister.gov/documents/1997/07/09/97-17664/recommendations-from-the-interagency-committee-for-the-review-of-the-racial-and-ethnic-standards-to. Accessed April 10, 2022.
20. Wei II, Virnig BA, John DA, Morgan RO. Using a Spanish surname match to improve identification of Hispanic women in Medicare administrative data. *Health Serv Res* 2006; 41 (4 Pt 1): 1469–81.
21. Dorsey R, Graham G. New HHS data standards for race, ethnicity, sex, primary language, and disability status. *JAMA* 2011; 306 (21): 2378–9.
22. Jarrín OF, Nyandege AN, Grafova IB, Dong X, Lin H. Validity of race and ethnicity codes in medicare administrative data compared with gold-standard self-reported race collected during routine home health care visits. *Med Care* 2020; 58 (1): e1–8.
23. Buescher PA, Gizlice Z, Jones-Vessey KA. Discrepancies between published data on racial classification and self-reported race: evidence from the 2002 North Carolina Live Birth Records. *Public Health Rep* 2005; 120 (4): 393–8.
24. Hahn RA, Truman BI, Barker ND. Identifying ancestry: the reliability of ancestral identification in the United States by self, proxy, interviewer, and funeral director. *Epidemiology* 1996; 7 (1): 75–80.
25. West CN. Race and ethnicity: comparing medical records to self-reports. *J Natl Cancer Inst Monogr* 2005; 2005 (35): 72–4.

26. Williams D. *The Quality of Racial Data. Closing the Gap*. Office of Minority Health Resource Center; 1998.
27. Bowling A. Mode of questionnaire administration can have serious effects on data quality. *J Public Health (Oxf)* 2005; 27 (3): 281–91.
28. Holzner B, Giesinger JM, Pinggera J, *et al.* The Computer-based Health Evaluation Software (CHES): a software for electronic patient-reported outcome monitoring. *BMC Med Inform Decis Mak* 2012; 12 (1): 126.
29. Weiner S, Horton L, Green T, Butler S. Feasibility of tablet computer screening for opioid abuse in the emergency department. *West J Emerg Med* 2015; 16 (1): 18–23.
30. Cullen D, Woodford A, Fein J. Food for thought: a randomized trial of food insecurity screening in the emergency department. *Acad Pediatr* 2019; 19 (6): 646–51.
31. Palakshappa D, Goodpasture M, Albertini L, Brown CL, Montez K, Skelton JA. Written versus verbal food insecurity screening in one primary care clinic. *Acad Pediatr* 2020; 20 (2): 203–7.
32. Gottlieb L, Hessler D, Long D, Amaya A, Adler N. A Randomized trial on screening for social determinants of health: the iScreen study. *Pediatrics* 2014; 134 (6): e1611–8.
33. Miller DP, Foley KL, Bundy R, *et al.* Universal screening in primary care practices by self-administered tablet vs nursing staff. *JAMA Netw Open* 2022; 5 (3): e221480.
34. Pittman JOE, Floto E, Lindamer L, Baker DG, Lohr JB, Afari N. VA eScreening program: technology to improve care for post-9/11 veterans. *Psychological Services* 2017; 14 (1): 23–33.
35. Almklov E, Afari N, Floto E, Lindamer L, Hurst S, Pittman JOE. Post-9/11 Veteran satisfaction with the VA eScreening program. *Mil Med* 2020; 185 (3–4): 519–29.
36. VA 10-10EZ. https://www.va.gov/Vaforms/medical/pdf/10-10EZ-fillable.pdf. Accessed April 6, 2022.
37. Markus HR. Pride, prejudice, and ambivalence: toward a unified theory of race and ethnicity. *Am Psychol* 2008; 63 (8): 651–70.
38. Robbin A. The problematic status of statistics on race and ethnicity: An "imperfect representation of reality. *J Gov Inf* 1999; 26 (5): 467–83.
39. Bhalla R, Yongue BG, Currie BP. Standardizing race, ethnicity, and preferred language data collection in hospital information systems: results and implications for healthcare delivery and policy. *J Healthc Qual* 2012; 34 (2): 44–52.
40. U.S. Office of Management and Budget Interagency Technical Working Group on Race and Ethnicity Standards [Internet]. [cited 2023Mar29]. https://spd15revision.gov.
41. Hayes-Bautista DE, Bryant M, Yudell M, *et al.* Office of management and budget racial/ethnic categories in mortality research: a framework for including the voices of racialized communities. *Am J Public Health* 2021; 111 (S2): S133–40.
42. Baker DW, Hasnain-Wynia R, Kandula NR, Thompson JA, Brown ER. Attitudes toward health care providers, collecting information about patients' race, ethnicity, and language. *Med Care* 2007; 45 (11): 1034–42.
43. Berry CA, Kaplan S, Mijanovich T, Mayer A. Moving to patient reported collection of race and ethnicity data: implementation and impact in ten hospitals. *Int J Health Care Qual Assur* 2014; 27 (4): 271–83.
44. Gardner W, Klima J, Chisolm D, *et al.* Screening, triage, and referral of patients who report suicidal thought during a primary care visit. *Pediatrics* 2010; 125 (5): 945–52.
45. Washington D, Toyama J. GAO Project Phase 1 Report: Determination of Completeness and Accuracy of VHA Race and Ethnicity Data. Office of Health Equity-Quality Enhancement Research Initiative Partnered Evaluation Center, Veterans Health Administration, Department of Veterans Affairs; 2020 (Final Report for VA response to U.S. Government Accountability Office report #GAO-20-83).
46. Lawrence ST, Willig JH, Crane HM, *et al.* Routine, self-administered, touch-screen, computer-based suicidal ideation assessment linked to automated response team notification in an HIV primary care setting. *Clin Infect Dis* 2010; 50 (8): 1165–73.
47. Polubriaginof FCG, Ryan P, Salmasian H, *et al.* Challenges with quality of race and ethnicity data in observational databases. *J Am Med Inform Assoc* 2019; 26 (8–9): 730–6.
48. Brinkman SD, Reese RJ, Norsworthy LA, *et al.* Validation of a self-administered computerized system to detect cognitive impairment in older adults. *J Appl Gerontol* 2014; 33 (8): 942–62.