

Brief report: Comparison of methods to identify Iraq and Afghanistan war veterans using Department of Veterans Affairs administrative data

Ann Bangerter, BS;^{1*} Amy Gravely, MA;¹ Andrea Cutting, MA;¹ Barb Clothier, MA, MS;¹ Michele Spont, PhD;¹⁻³ Nina Sayer, PhD, LP^{1-2,4}

¹Center for Chronic Disease Outcomes Research, Minneapolis Department of Veterans Affairs Medical Center, Minneapolis, MN; Departments of ²Psychology, ³Psychiatry, and ⁴Medicine, University of Minnesota, Twin Cities Campus, Minneapolis, MN

Abstract—The Department of Veterans Affairs (VA) has made treatment and care of Operation Iraqi Freedom/Operation Enduring Freedom (OIF/OEF) veterans a priority. Researchers face challenges identifying the OIF/OEF population because until fiscal year 2008, no indicator of OIF/OEF service was present in the Veterans Health Administration (VHA) administrative databases typically used for research. In this article, we compare an algorithm we developed to identify OIF/OEF veterans using the Austin Information Technology Center administrative data with the VHA Support Service Center OIF/OEF Roster and veterans' self-report of military service. We drew data from two different institutional review board-approved funded studies. The positive predictive value of our algorithm compared with the VHA Support Service Center OIF/OEF Roster and self-report was 92% and 98%, respectively. However, this method of identifying OIF/OEF veterans failed to identify a large proportion of OIF/OEF veterans listed in the VHA Support Service Center OIF/OEF Roster. Demographic, diagnostic, and VA service use differences were found between veterans identified using our method and those we failed to identify but who were in the VHA Support Service Center OIF/OEF Roster. Therefore, depending on the research objective, this method may not be a viable alternative to the VHA Support Service Center OIF/OEF Roster for identifying OIF/OEF veterans.

Key words: administrative data, Department of Veterans Affairs, Global War on Terrorism, health services use, Iraq and Afghanistan war veterans, OIF/OEF, period of service, PTSD, Veterans Health Administration, VHA Support Service Center OIF/OEF Roster.

INTRODUCTION

More than 2 million servicemembers have served in Operation Iraqi Freedom (OIF) and Operation Enduring Freedom (OEF) and nearly 1.2 million have been discharged as veterans. Veterans from these conflicts have experienced multiple deployments and unique combat environments [1–3], along with high rates of psychiatric disturbance and physical injury [4–7]. Recognizing the importance of the healthcare needs and demands of OIF/OEF veterans, the Veterans Health Administration (VHA) has solicited research focused on them [8–9].

The Department of Veterans Affairs (VA) health services use SAS data sets found in the Austin Information Technology Center (AITC) contain national VHA-provided

Abbreviations: AITC = Austin Information Technology Center, DMDC = Defense Manpower Data Center, DSS = Decision Support System, EES = Environmental Epidemiology Service, FY = fiscal year, NPV = negative predictive value, OEF = Operation Enduring Freedom, OIF = Operation Iraqi Freedom, POS = period of service, PPV = positive predictive value, PTSD = posttraumatic stress disorder, SF = outpatient visit, VA = Department of Veterans Affairs, VHA = Veterans Health Administration.

*Address all correspondence to Ann Bangerter, BS; Center for Chronic Disease Outcomes Research, VA Medical Center, One Veterans Drive (152/2E), Minneapolis, MN 55417; 612-467-1384; fax: 612-727-5699. Email: ann.bangerter@va.gov
DOI:10.1682/JRRD.2009.08.0115

healthcare information on veterans, which researchers use to identify patient populations [10–11]. Using these data sets, VA researchers can identify periods of service (POS) for veterans who have served from the Spanish-American War to the gulf war. However, VA researchers cannot search the VA health services use SAS data sets by POS to identify OIF/OEF veterans because Congress has not designated OIF/OEF as a POS. Other methods are therefore needed to identify this priority group.

The VHA Support Service Center OIF/OEF Roster is arguably the gold standard for identification of OIF/OEF veterans [12–14]. The OIF/OEF Roster is cataloged by the Department of Defense's Defense Manpower Data Center (DMDC) and has been provided to VHA's Environmental Epidemiology Service (EES) since September 2003. This comprehensive database includes identifying information on VA-registered OIF/OEF veterans with military discharges starting October 1, 2001, who were (1) physically located within the OIF/OEF combat zones or areas of operation or (2) identified by their service branch as directly supporting the OIF/OEF mission outside designated combat zones.* This data set is not available to researchers without funded studies. Furthermore, the approval process, although appropriate given the need to ensure veteran privacy and data security, can take several months. Another method for identifying OIF/OEF veterans is self-reported POS on surveys or during interviews. This method, however, cannot be used for selection of OIF/OEF veterans before data collection.

In this article, we describe results from two studies that compare an algorithm we developed to identify OIF/OEF VA users that makes use of variables available in the Outpatient Visit (SF) file of the VA health services use database. We compared this OIF/OEF Algorithm with the OIF/OEF Roster (study 1) and then with veterans' self-report surveys (study 2). We describe the positive predictive value (PPV) and sensitivity of our OIF/OEF Algorithm and also present differences between veterans we identified using our OIF/OEF Algorithm against those we failed to identify according to the OIF/OEF Roster and self-report survey. Findings will help investigators determine the most appropriate approach for identifying OIF/OEF veterans for their research.

*Access to the OIF/OEF Roster description is provided on a Web site restricted to VA employees. Please contact the corresponding author for an electronic copy.

METHODS

Samples and Study Design

Study 1 included two samples: (1) a national sample of OIF/OEF veterans who used VA healthcare from fiscal year (FY) 2004 to FY2007 selected on the basis of the OIF/OEF Algorithm (outlined in the **Figure**) for a study examining OIF/OEF veteran reintegration problems and treatment interests and (2) OIF/OEF veterans included in the OIF/OEF Roster with military discharges through December 2007 who had VA health services use data from FY2004 to FY2007. Study 2 included a sample of veterans who had at least one diagnosis of posttraumatic stress disorder (PTSD) in VA health services use data during FY2008 or FY2009 for a study of PTSD treatment participation. In study 2, we compared those who self-reported OIF/OEF POS on the completed survey with those we identified as OIF/OEF veterans using the OIF/OEF Algorithm.

Measures

For study 1, we used VA health services use data to obtain demographic information and health services use information, including psychiatric diagnoses and use of a VA facility in the past 2 years. For study 2, we obtained age from VA health services use data. We used the self-report survey for all other demographic information and POS. The survey included a question about which conflicts the survey recipient served in. We classified survey respondents who indicated they had participated in OEF (2001–present), OIF (2003–present), or the Global War on Terrorism (2004–present) as self-reported OIF/OEF veterans.

OIF/OEF Algorithm Based on VA Health Services Use Data: Design and Data Source

OIF/OEF veterans are, by definition, combat veterans discharged after October 1, 2001, and eligible for VA services. Before the National Defense Authorization Act of 2008 was enacted on January 28, 2008, the VA granted OIF/OEF combat veterans discharged or released from Active Duty on or after January 28, 2003, eligibility to enroll for VA healthcare for 2 years from the date of discharge or release. That end date has subsequently been expanded to 5 years [15].

To identify OIF/OEF veterans, we restricted the following three variables in the SF file: (1) combat eligibility flag, (2) end date of VA healthcare eligibility, and (3) POS. It should be noted that the variable indicating the end date of VA healthcare eligibility will exist only if the combat

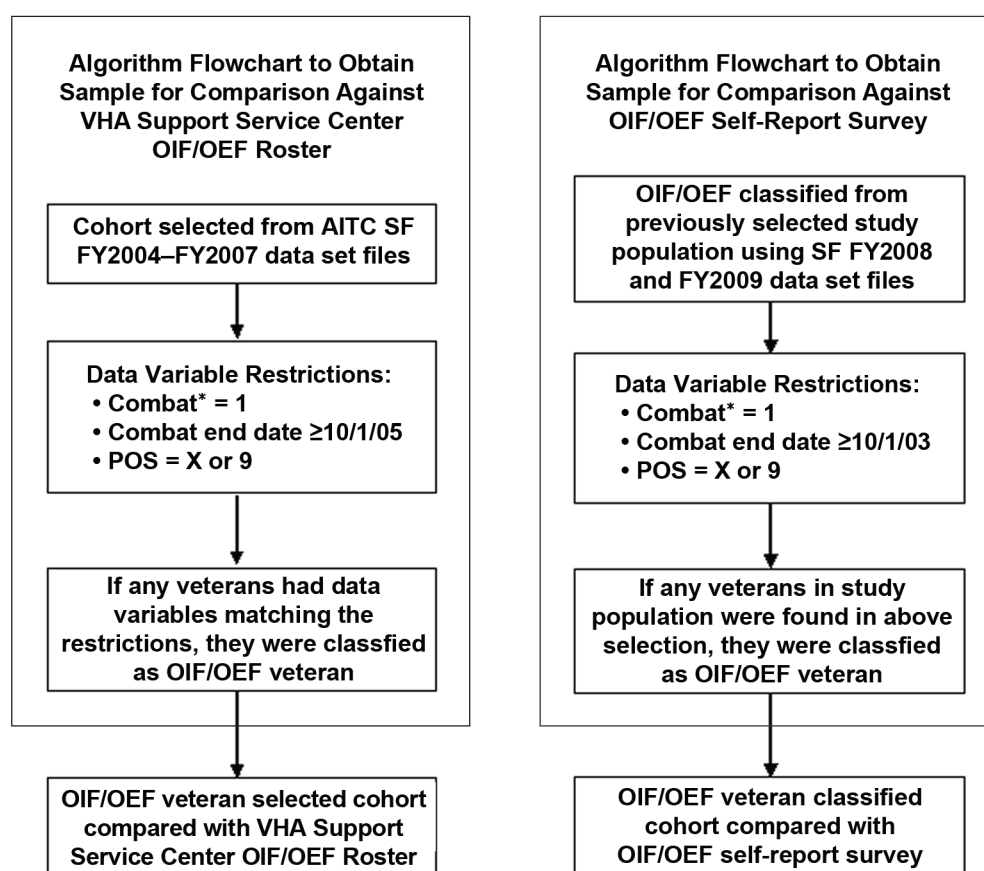


Figure.

Flowcharts for obtaining Operation Iraqi Freedom/Operation Enduring Freedom (OIF/OEF) veteran status using Department of Veterans Affairs health services use data compared with Veterans Health Administration (VHA) Support Service Center OIF/OEF Roster and self-reported survey data. *Indicates whether veteran served Active Duty in theater of combat operations during period of war, after gulf war, or in combat against hostile force during period of “hostilities” after November 11, 1998. AITC = Austin Information Technology Center, FY = fiscal year, POS = period of service, SF = outpatient visit, X = served in gulf war, 1 = served in time of conflict, 9 = served in other or none.

eligibility flag occurs in a veteran’s record. We defined the end date of VA healthcare eligibility differently in study 1 and study 2. Because the researchers prioritized surveying OIF veterans who would have been discharged after October 2003, we specified selection of those veterans with an eligibility end date after October 1, 2005, for study 1. For study 2, we used an earlier VA eligibility end date (October 1, 2003) to encompass a larger proportion of OEF veterans, since the war began in 2001. The **Figure** depicts the OIF/OEF Algorithm.

Statistical Analysis

The gold standards were the OIF/OEF Roster and self-report survey for study 1 and study 2, respectively.

We performed analyses with SAS 9.1 software (SAS Incorporation Inc; Cary, North Carolina).

For studies 1 and 2, we calculated sensitivity and PPV. For study 2, we also calculated specificity and negative predictive value (NPV). However, given that the OIF/OEF Roster only includes OIF/OEF veterans, we were unable to calculate specificity or NPV for study 1. To compare demographic characteristics for study 1 and study 2, we used Pearson chi-square tests for categorical variables and two-sample *t*-tests for continuous variables. For study 1, we also used the nonparametric Wilcoxon rank-sum test for number of visits because of the significant positive skew in its distribution.

RESULTS

Study 1: OIF/OEF Algorithm Versus VHA Support Service Center OIF/OEF Roster

Of the 181,612 potential OIF/OEF veterans we identified using of OIF/OEF Algorithm, we confirmed 92 percent (PPV, $n = 167,110$) as OIF/OEF when compared with the OIF/OEF Roster. However, our OIF/OEF Algorithm failed to identify 119,758 OIF/OEF veterans. The sensitivity was 58 percent ($167,110/286,868$) and the false negative rate was 42 percent. There were differences between those OIF/OEF veterans identified using our OIF/OEF Algorithm and those OIF/OEF veterans we failed to identify (**Table 1**). The small differences in age and sex reached statistical significance because of the large sample size. The differences in rate of psychiatric diagnoses and healthcare use, however, appear meaningful.

Table 1.

Comparison of Operation Iraqi Freedom/Operation Enduring Freedom (OIF/OEF) veterans selected using OIF/OEF Algorithm versus OIF/OEF veterans not selected.

Characteristic	Confirmed OIF/OEF Veteran*		p-Value
	Selected ($n = 167,110$)	Not Selected ($n = 119,758$)	
Male, n (%)	147,057 (88)	104,190 (87)	<0.001
Reservist, n (%)	90,240 (54)	55,089 (46)	<0.001
Diagnosis in VA Records, [†] n (%)			
Anxiety	17,463 (10.5)	8,168 (6.82)	<0.001
PTSD	54,528 (32.6)	18,371 (15.3)	<0.001
Mood Disorder	44,669 (26.7)	17,964 (15.0)	<0.001
Psychosis	1,855 (1.1)	886 (0.7)	<0.001
MH Visit Recorded, [‡] n (%)	18,766 (11.2)	9,233 (7.7)	<0.001
Number of VA Appoint- ments [†] (median)	10	2	<0.001
Age (mean \pm SD)	33.2 \pm 9.0	36 \pm 9.8	<0.001

*OIF/OEF veteran status confirmed using Veterans Health Administration Support Service Center OIF/OEF Roster.

[†]Based on at least one ICD-9-CM code in VA health services use data 2 years before survey. ICD-9-CM codes used for diagnoses are as follows:

Anxiety = 300.0, 300.2, 300.3, 308, 309.

PTSD = 309.81.

Mood Disorder = 296, 300.4, 311.

Psychosis = 295, 297, 298.

[‡]Based on clinical visit to any MH Clinic Stop ($500 \leq$ clinic stop < 600) at medical center or associated VA community-based outpatient clinic 1 year before survey.

ICD-9-CM = International Classification of Diseases-9th Revision-Clinical Modification, MH = mental health, PTSD = posttraumatic stress disorder, SD = standard deviation, VA = Department of Veterans Affairs.

We failed to identify most ($n = 88,974$, 74%) of the 119,758 OIF/OEF veterans using our OIF/OEF Algorithm because they were not classified as “combat” veterans in the SF file. The remaining 30,784 (26%) did not have a VA healthcare eligibility date or POS that fit our restrictions.

Study 2: OIF/OEF Algorithm Versus OIF/OEF Veteran Self-Report

Of the 7,952 veterans surveyed, 5,207 returned study material for a response rate of 65.5 percent. Among the 4,563 who self-reported a POS on the survey, 1,229 of those we identified using the OIF/OEF Algorithm also self-reported OIF/OEF POS. However, of the 4,563 veterans that self-reported POS, we failed to identify 331 (7.3%) OIF/OEF veterans and misidentified 24 (0.5%). Our OIF/OEF Algorithm had a PPV of 0.98, NPV of 0.90, sensitivity of 0.79, and specificity of 0.99 (**Table 2**). Thus, assigning OIF/OEF veteran status using our OIF/OEF Algorithm had a higher false negative rate (21%) than a false positive rate (1%), as detailed in **Table 2**.

Table 3 presents differences between those we identified and failed to identify using our OIF/OEF Algorithm. OIF/OEF veterans not identified using our OIF/OEF Algorithm were more likely than the identified OIF/OEF veterans to report having served in Afghanistan or the Global War on Terrorism but less likely to report having served in Iraq.

Table 2.

Sensitivity and specificity of Operation Iraqi Freedom/Operation Enduring Freedom (OIF/OEF) Algorithm as compared with self-reported survey data.

Selected Using OIF/OEF Algorithm	OIF/OEF Veteran*		Predictive Value (%)
	Yes	No	
Yes	TP = 1,229	FP = 24	Positive = 98
No	FN = 331	TN = 2,979	Negative = 90
Sensitivity (%)	79	—	—
Specificity (%)	—	99	—

*OIF/OEF status confirmed using self-reported survey data.

Note: Sensitivity = $TP / (TP + FN) = 1,229 / (1,229 + 331) = 79\%$.

Specificity = $TN / (FP + TN) = 2,979 / (24 + 2,979) = 99\%$.

Positive predictive value = $TP / (TP + FP) = 1,229 / (1,229 + 24) = 98\%$.

Negative predictive value = $TN / (FN + TN) = 2,979 / (331 + 2,979) = 90\%$.

FP rate (α) = $FP / (FP + TN) = 24 / (24 + 2,979) = 0.8\% = 1 - \text{specificity}$.

FN rate (β) = $FN / (TP + FN) = 331 / (1,229 + 331) = 21.0\% = 1 - \text{sensitivity}$.

FN = false negative, FP = false positive, TN = true negative, TP = true positive.

Table 3.

Comparison of Operation Iraqi Freedom/Operation Enduring Freedom (OIF/OEF) veterans selected using OIF/OEF Algorithm versus OIF/OEF veterans not selected.

Characteristic	Confirmed OIF/OEF Veteran*		<i>p</i> -Value
	Selected (<i>n</i> = 1,229)	Not Selected (<i>n</i> = 331)	
Male, <i>n</i> (%)	932 (76)	242 (73)	0.31
Conflicts, <i>n</i> (%)			
OIF (2003–present)	1,090 (89)	270 (82)	0.001
OEF (2001–present)	217 (18)	92 (28)	<0.001
Global War on Terrorism (2004–present)	169 (14)	75 (23)	<0.001
Race/Ethnicity, <i>n</i> (%)			
Caucasian	563 (46)	176 (53)	0.02
African American	196 (16)	78 (24)	0.001
Asian American	79 (7)	11 (3)	0.031
Hispanic (any)	415 (34)	70 (21)	<0.001
Income, <i>n</i> (%): >\$50,000	174 (14)	77 (23)	0.001
Education, <i>n</i> (%): High school only	283 (23)	59 (18)	0.042
Disabled, <i>n</i> (%)	506 (41)	170 (51)	<0.001
Marital Status, <i>n</i> (%)			
Single (never married)	338 (28)	78 (24)	0.09
Married/Partnered	678 (55)	181 (55)	0.09
Divorced/Widowed	207 (17)	71 (22)	0.09
Employed (full- or part-time), <i>n</i> (%)	750 (61)	203 (61)	0.97
Age [†] (mean ± SD)	32.5 ± 8.9	36.8 ± 11.1	<0.001

*OIF/OEF status confirmed using self-report survey.

[†]Age information is from Department of Veterans Affairs health services use data; all other variables are from survey data.

SD = standard deviation.

For those 331 OIF/OEF veterans that our OEF/OIF Algorithm failed to identify, 267 (80.6%) were not classified as “combat” veterans in the SF file. The remaining 64 did not have a POS that fit the criterion we used for that variable.

DISCUSSION

The PPV of our OIF/OEF Algorithm against the OIF/OEF Roster was 92 percent (167,110/181,612) and, as detailed in **Table 2**, the PPV of the self-reported survey was 98 percent. The PPV of our OIF/OEF Algorithm was acceptable for both studies. However, the false negative rate could be problematic depending on the sampling goals. Those OIF/OEF veterans identified by the OIF/OEF Algorithm (study 1) were more likely to have a mental health diagnosis and more VA appointments than those not identi-

fied. Therefore, using an algorithm based on VA health services use data to select OIF/OEF veterans may result in a nonrepresentative sample of OIF/OEF VA users. It is unknown whether veterans not identified using our OIF/OEF Algorithm have similar healthcare needs and are simply receiving healthcare services outside of the VA. Another possibility is that reduced VA appointment frequency may result in decreased opportunities to update combat status and diagnose all of a veteran’s medical conditions.

Study 2 suggests that the OIF/OEF Algorithm produces few false positives when the population of interest is OIF/OEF veterans diagnosed with PTSD within the VA. We would expect this group of veterans to use healthcare services more routinely and therefore have an accurate combat status in VA health services use data. Our OIF/OEF Algorithm could be a useful tool for developing a sampling frame if no other method to identify OIF/OEF veterans is available. However, as in study 1, we missed those without a combat indicator.

Starting in FY2008, a new method for OIF/OEF veteran identification was released. There is now an OIF/OEF flag available to researchers in the National Data Extracts of the Decision Support System (DSS), located in the AITC system. This DSS OIF/OEF veteran flag is created using the DMDC data from the OIF/OEF Roster that is sent to EES and does not rely on the combat indicator that we found to be unreliable. Additionally, using the DSS OIF/OEF veteran flag to select a cohort of OIF/OEF veterans requires only one data variable restriction as opposed to three different variable restrictions that we used in the OIF/OEF Algorithm. This method is limited to veterans who have used the VA healthcare system from FY2008 onward.

CONCLUSIONS

Our OIF/OEF Algorithm allows for accurate OIF/OEF veteran identification of a subgroup of veterans who are classified by combat eligibility status in VA health services use data. However, it fails to identify nearly 42 percent of OIF/OEF veterans on the OIF/OEF Roster and 21 percent who self-reported OIF/OEF veteran status. Therefore, our OIF/OEF Algorithm would not be appropriate if a researcher wanted to sample from the entire OIF/OEF veteran population that uses VA services, but it might be useful to identify the subgroup of OIF/OEF veterans who use the VA more frequently and have mental health diagnoses such as PTSD.

ACKNOWLEDGMENTS

Author Contributions:

Study concept and design: N. Sayer, M. Spont.

Acquisition of data: A. Bangerter, A. Cutting, N. Sayer, M. Spont.

Analysis and interpretation of data: B. Clothier, A. Gravely, A. Cutting, A. Bangerter, N. Sayer, M. Spont.

Drafting of manuscript: A. Bangerter, N. Sayer.

Critical revision of manuscript for important intellectual content:

N. Sayer, A. Bangerter, A. Gravely, A. Cutting, M. Spont, B. Clothier.

Statistical analysis: B. Clothier, A. Gravely, A. Cutting, A. Bangerter.

Obtained funding: N. Sayer, M. Spont.

Administrative, technical, or material support: A. Bangerter, A. Cutting, A. Gravely.

Study supervision: N. Sayer, M. Spont.

Financial Disclosures: The authors have declared that no competing interests exist.

Funding/Support: This material was based on work supported by the VA Health Services Research and Development Service (grants IAC 06-266 and RRP 07-315).

Additional Contributions: The views expressed in this article are those of the author(s) and do not necessarily represent the position or policy views of the VA or any other government agency.

Institutional Review: The data for this work came from two separately funded projects approved by the Subcommittee on Human Studies, Minneapolis VA Medical Center, Minneapolis, Minnesota.

REFERENCES

- Hosek J, Kavanagh J, Miller L. How deployments affect service members. Santa Monica (CA): RAND; 2006.
- Hoge CW, Castro CA, Messer SC, McGurk D, Cotting DI, Koffman RL. Combat duty in Iraq and Afghanistan, mental health problems, and barriers to care. *N Engl J Med*. 2004; 351(1):13–22. [PMID: 15229303] DOI:10.1056/NEJMoa040603
- Tanielian TL, Jaycox L; RAND Corporation. Invisible wounds of war: Psychological and cognitive injuries, their consequences, and services to assist recovery. Santa Monica (CA): RAND; 2008.
- Okie S. Traumatic brain injury in the war zone. *N Engl J Med*. 2005;352(20):2043–47. [PMID: 15901856] DOI:10.1056/NEJMp058102
- Gawande A. Casualties of war—Military care for the wounded from Iraq and Afghanistan. *N Engl J Med*. 2004; 351(24):2471–75. [PMID: 15590948] DOI:10.1056/NEJMp048317
- Seal KH, Metzler TJ, Gima KS, Bertenthal D, Maguen S, Marmar CR. Trends and risk factors for mental health diagnoses among Iraq and Afghanistan veterans using Department of Veterans Affairs Health Care, 2002–2008. *Am J Public Health*. 2009;99(9):1651–58. [PMID: 19608954] DOI:10.2105/AJPH.2008.150284
- Seal KH, Bertenthal D, Maguen S, Gima K, Chu A, Marmar CR. Getting beyond “Don’t ask; don’t tell”: An evaluation of US Veterans Administration postdeployment mental health screening of veterans returning from Iraq and Afghanistan. *Am J Public Health*. 2008;98(4):714–20. [PMID: 18309130] DOI:10.2105/AJPH.2007.115519
- Office of Research and Development program announcement: Quality Enhancement Research Initiative (QUERI) special solicitation for projects: Implementing research into practice to improve care delivery. Washington (DC): Department of Veterans Affairs; 2005.
- Afghanistan & Iraq [Internet]. Washington (DC): Department of Veterans Affairs; 2009 [cited 2009 Aug 3]. Available from: <http://www.research.va.gov/resources/pubs/docs/OIF-OEF-brochure.pdf>.
- Rosenheck RA, Fontana AF. Recent trends in VA treatment of post-traumatic stress disorder and other mental disorders. *Health Aff (Millwood)*. 2007;26(6):1720–27. [PMID: 17978391] DOI:10.1377/hlthaff.26.6.1720
- Jia H, Zheng YE, Cowper DC, Stansbury JP, Wu SS, Vogel WB, Duncan PW, Reker DM. Race/ethnicity: Who is counting what? *J Rehabil Res Dev*. 2006;43(4):475–84. [PMID: 17123187] DOI:10.1682/JRRD.2005.05.0086
- Seal KH, Bertenthal D, Miner CR, Sen S, Marmar C. Bringing the war back home: Mental health disorders among 103,788 US veterans returning from Iraq and Afghanistan seen at Department of Veterans Affairs facilities. *Arch Intern Med*. 2007;167(5):476–82. [PMID: 17353495] DOI:10.1001/archinte.167.5.476
- Cohen BE, Gima K, Bertenthal D, Kim S, Marmar CR, Seal KH. Mental health diagnoses and utilization of VA non-mental health medical services among returning Iraq and Afghanistan veterans. *J Gen Intern Med*. 2010;25(1): 18–24. [PMID: 19787409] DOI:10.1007/s11606-009-1117-3
- Haskell SG, Brandt CA, Krebs EE, Skanderson M, Kerns RD, Goulet JL. Pain among veterans of Operations Enduring Freedom and Iraqi Freedom: Do women and men differ? *Pain Med*. 2009;10(7):1167–73. [PMID: 19818028] DOI:10.1111/j.1526-4637.2009.00714.x
- National Defense Authorization Act of 2008, HR 4986, 110th Cong. 2nd Session (2008).

Submitted for publication August 7, 2009. Accepted in revised form July 6, 2010.

This article and any supplementary material should be cited as follows:

Bangerter A, Gravely A, Cutting A, Clothier B, Spont M, Sayer N. Brief report: Comparison of methods to identify Iraq and Afghanistan war veterans using Depart-

ment of Veterans Affairs administrative data. *J Rehabil Res Dev*. 2010;47(8):815–22.

DOI:10.1682/JRRD.2009.08.0115



