



Elizabeth Tarlov, RN, PhD; Kevin T. Stroupe, PhD

## **Advancing veterans' healthcare using electronic data: Lessons learned from researchers in the field**

The Department of Veterans Affairs (VA) provides nationwide access, lifetime coverage, and an integrated care structure to its enrollees. Those key aspects of VA healthcare—together with data contained in VA's electronic information systems supporting over eight million veterans—provide unique opportunities to study processes, outcomes, and costs of care. Recently, for example, VA data have been used to study outcomes associated with acute postoperative inpatient rehabilitation and care in specialized rehabilitation bed units after lower-limb amputation [1–2], medication adherence and relapse among patients discharged from a VA posttraumatic stress disorder (PTSD) treatment program [3], the provision and costs of assistive technology devices to veterans after stroke [4], and use of mental health services by veterans disabled by auditory disorders [5].

### **UNDERPINNINGS OF SINGLE-TOPIC ISSUE ON DATA QUALITY**

In 1998, the VA Health Services Research & Development Service funded the VA Information Resource Center (VIREC) to facilitate the use of VA administrative data for research. Since that time, VIREC has developed and disseminated research-relevant information about VA databases and information systems. Knowledge-building activities include data quality investigations conducted within VIREC. This single-topic issue recognizes the vast stores of information about VA data that VA investigators generate nationwide, in the normal course of their data analysis activities, and seeks to leverage that information to advance electronic-data-based research.

A call for abstracts was issued in the spring of 2009 for investigations of the quality and research utility of electronic data used in research to advance the care of veterans. VA investigators responded—many more worthy abstracts were submitted than could be included in one *JRRD* issue. Reflecting strong research-community support for VA research, 54 scholars readily accepted invitations to peer review and almost 100 individuals participated as authors, reviewers, or editors.

VA researchers have long been major contributors to the methodological literature in the area of administrative data use [6–17]. This issue advances that literature by focusing specifically on the quality and research utility of databases used in veterans research. We view data quality—roughly, the completeness and accuracy of the data collected and entered—as a necessary but not sufficient feature of a “research-useful” data set [18]. The research

utility of a data set entails additional factors such as coverage of the population of interest and degree of validity achievable when specific data elements are used as proxies for important but absent information. Research utility questions, therefore, relate to the inevitable challenges involved in the use of secondary data—i.e., data that were collected for a different purpose—rather than a deficiency in the data per se.

## **VITAL VETERANS RESEARCH AND VA HEALTH INFORMATION SYSTEMS**

The questions of data quality and research utility addressed in the included articles were encountered during research tackling key issues in the care and rehabilitation of wounded war veterans. Promoting high-quality care and providing the tools for independent living to veterans with spinal cord injury (SCI) [19], polytrauma [20], PTSD, substance-use disorder and other mental health conditions [21–23] and to others needing rehabilitative and assistive services [24–25] are high VA priorities. Investigating obesity and chronic kidney disease and its antecedents are important because these are among the most common chronic conditions afflicting veterans seeking care in the VA [26–27]. Understanding the healthcare needs and ensuring accessible care for women and other vulnerable populations [23,28–30] are among the VA's top research priorities. Identifying cost-effective approaches to veterans' healthcare [25,29,31] is critical to long-term viability of the VA healthcare system. Studies included in this issue were derived from research investigating these issues using data from sources inside and outside the VA.

The included articles also spotlight the great value that VA's health information systems have for research. Chief among the data sources mined for research is the electronic medical record (EMR), portions of which are extracted to create the VA National Patient Care Database, providing clinical and health services utilization information. Condition-based registries maintained for clinical and administrative purposes and other case-reporting systems

provide additional clinical and treatment detail and so are also rich data sources. Data originating outside the VA are critical for learning about care received and costs of care for veterans obtaining services in the private sector and help to fill a gap in information for many studies. In addition to the National Patient Care Database [19–21,24,30], data sources evaluated include the Veterans Health Administration Corporate Data Warehouse [26], the VA's National Spinal Cord Dysfunction Registry [19], the Minimum Data Set (long-term care resident assessment) [24], Medicare's End-Stage Renal Disease Medical Evidence Report [27], Medicaid and Medicare claims [29–30], and the Medicare Current Beneficiary Survey [31].

## **SPOTLIGHT ON DATA QUALITY AND RESEARCH UTILITY STUDIES**

Several studies examined the completeness and/or validity of these data resources. Some common observations emerge from among this set of studies. First, findings suggest substantial variability in record completeness across VA facilities and regional networks [20,22]. Both the inconsistency itself and the inferior quality found in some locations have clear implications for patient care quality as well as validity of performance measurement and research conclusions. Because all locations are within the same healthcare system, the variability is unlikely to be attributable to heterogeneity in mission, goals, economic incentives, or data-entry system. Although the reasons for the inconsistency in quality are unclear and probably multifactorial, the recognition of variability nonetheless presents opportunities to improve systemwide performance.

Second, data collected for purposes peripheral to direct patient care or billing seem to be particularly prone to inaccuracy [19–20,24,27,30]. Several studies examined data collection systems that require manual data entry by clinical staff, separate from routine medical record documentation [19,24,27]. Not surprisingly, accurate and complete data entry will receive lower priority when not essential for patient care. In other cases, lack of an administrative

mandate for data collection, which otherwise may provide incentives and technical controls, may contribute to incomplete data.

Unlike data quality, the research utility of a data set depends somewhat on the skill and experience of the researcher. A thorough understanding of the data and time and effort spent constructing and evaluating the validity of operational definitions, for example, can substantially alter measurement of important variables. B. Smith et al. exemplify this truth as they share results of extensive experience in employing VA electronic data to identify and characterize veterans with SCI [19]. The point is also illustrated by Frayne et al., who tested a number of algorithms to construct a length-of-stay measure and found that estimates of mental illness-related disparity depended on the choice of algorithm [23]. Several articles investigated methods of case identification using administrative data [19–22,24]. Carlson et al. examined the feasibility of using external cause-of-injury codes to identify subsets of the injured patient population [20]. Harris et al. determined that use of VA clinic stop codes alone is inadequate to identify patients receiving outpatient treatment for substance use disorder but that adding diagnosis codes to the case definition greatly improved ascertainment [22]. Berlowitz et al. evaluated whether administrative data could provide more accurate information than the Minimum Data Set about the presence of comorbid conditions among residents in long-term and rehabilitation care settings [24]. Washington et al. reported their experience in combining data from several sources to create a sampling frame of women veterans [28], and Bangerter et al. investigated the accuracy of an algorithm to identify patients with PTSD in administrative data [21]. M. Smith et al. described techniques used to link patient-level information in the EMR and the National Prosthetics Patient Database (NPPD) and concluded that use of the NPPD can improve assessment of resource utilization, but they caution researchers against use of the NPPD to identify the universe of patients receiving specific prosthetics services [25]. Hendricks et al. explored the value of including Medicaid data in veterans studies and shared insights gained from extensive

experience in use of VA and Medicaid data for research [29]. All these studies contradict a common impression that use of secondary data for research is straightforward and uncomplicated.

As the VA continues to respond to the health-care needs of U.S. veterans and, in particular, to the unique and evolving needs of those returning from Afghanistan and Iraq, research that identifies those needs, evaluates the VA's response, and assesses patient outcomes is essential. VA electronic data will continue to be a vital tool for that research and analysis. The broad goal of the studies reported in this issue is uniform: to achieve excellence in the conduct of current and future research using VA data. We hope and anticipate that the results presented in this single-topic issue, through an immediate positive impact on research going forward, will advance VA healthcare toward the ultimate goal of each veteran's optimum health and well-being.

**Elizabeth Tarlov, RN, PhD;\* Kevin Stroupe, PhD**

Edward Hines Jr. VA Hospital—VA Information Resource Center, Hines, IL

\*Email: [elizabeth.tarlov@va.gov](mailto:elizabeth.tarlov@va.gov)

## REFERENCES

1. Kurichi JE, Small DS, Bates BE, Prvu-Bettger JA, Kwong PL, Vogel WB, Bidelsbach DE, Stineman MG. Possible incremental benefits of specialized rehabilitation bed units among veterans after lower extremity amputation. *Med Care*. 2009;47(4):457–65. [PMID: 19238103] DOI:10.1097/MLR.0b013e31818b08c6
2. Stineman MG, Kwong PL, Kurichi JE, Prvu-Bettger JA, Vogel WB, Maislin G, Bates BE, Reker DM. The effectiveness of inpatient rehabilitation in the acute postoperative phase of care after transtibial or transfemoral amputation: Study of an integrated health care delivery system. *Arch Phys Med Rehabil*. 2008; 89(10):1863–72. [PMID: 18929014] DOI:10.1016/j.apmr.2008.03.013
3. Lockwood A, Steinke DT, Botts SR. Medication adherence and its effect on relapse among patients discharged from a Veterans Affairs posttraumatic stress disorder treatment program. *Ann Pharmacother*.

- 2009;43(7):1227–32. [PMID: 19584387]  
DOI:10.1345/aph.1M017
4. Winkler SL, Vogel B, Hoenig H, Ripley DC, Wu S, Fitzgerald SG, Mann WC, Reker DM. Cost, utilization, and policy of provision of assistive technology devices to veterans poststroke by Medicare and VA. *Med Care*. 2010;48(6):558–62. [PMID: 20125048]  
DOI:10.1097/MLR.0b013e3181bd4a11
  5. Kendall CJ, Rosenheck R. Use of mental health services by veterans disabled by auditory disorders. *J Rehabil Res Dev*. 2008;45(9):1349–60. [PMID: 19319759]  
DOI:10.1682/JRRD.2007.11.0185
  6. Berlowitz DR, Brandeis GH, Moskowitz MA. Using administrative databases to evaluate long-term care. *J Am Geriatr Soc*. 1997;45(5):618–23. [PMID: 9158586]
  7. Best WR, Khuri SF, Phelan M, Hur K, Henderson WG, Demakis JG, Daley J. Identifying patient pre-operative risk factors and postoperative adverse events in administrative databases: Results from the Department of Veterans Affairs National Surgical Quality Improvement Program. *J Am Coll Surg*. 2002;194(3):257–66. [PMID: 11893128]  
DOI:10.1016/S1072-7515(01)01183-8
  8. Borzecki AM, Wong AT, Hickey EC, Ash AS, Berlowitz DR. Identifying hypertension-related comorbidities from administrative data: What's the optimal approach? *Am J Med Qual*. 2004;19(5):201–6. [PMID: 15532912]  
DOI:10.1177/106286060401900504
  9. Frayne SM, Miller DR, Sharkansky EJ, Jackson VW, Wang F, Halanych JH, Berlowitz DR, Kader B, Rosen CS, Keane TM. Using administrative data to identify mental illness: What approach is best? *Am J Med Qual*. 2010;25(1):42–50. [PMID: 19855046]
  10. Kerr EA, McGlynn EA, Van Vorst KA, Wickstrom SL. Measuring antidepressant prescribing practice in a health care system using administrative data: Implications for quality measurement and improvement. *Jt Comm J Qual Improv*. 2000;26(4):203–16. [PMID: 10749005]
  11. Kurichi JE, Stineman MG, Kwong PL, Bates BE, Reker DM. Assessing and using comorbidity measures in elderly veterans with lower extremity amputations. *Gerontology*. 2007;53(5):255–59. [PMID: 17435390]  
DOI:10.1159/000101703
  12. Lafrance JP, Miller DR. Defining acute kidney injury in database studies: The effects of varying the baseline kidney function assessment period and considering CKD status. *Am J Kidney Dis*. 2010. [Epub ahead of print] [PMID: 20673605]
  13. Phibbs CS, Bhandari A, Yu W, Barnett PG. Estimating the costs of VA ambulatory care. *Med Care Res Rev*. 2003;60(3 Suppl):54S–73S. [PMID: 15095546]  
DOI:10.1177/1077558703256725
  14. Reker DM, Hamilton BB, Duncan PW, Yeh SC, Rosen A. Stroke: Who's counting what? *J Rehabil Res Dev*. 2001;38(2):281–89. [PMID: 11392661]
  15. Rosenheck R, Fontana A, Stolar M. Assessing quality of care: Administrative indicators and clinical outcomes in posttraumatic stress disorder. *Med Care*. 1999;37(2):180–88. [PMID: 10024122]  
DOI:10.1097/00005650-199902000-00008
  16. Swartztrauber K, Anau J, Peters D. Identifying and distinguishing cases of parkinsonism and Parkinson's disease using ICD-9 CM codes and pharmacy data. *Mov Disord*. 2005;20(8):964–70. [PMID: 15834854]  
DOI:10.1002/mds.20479
  17. Wray NP, Ashton CM, Kuykendall DH, Hollingsworth JC. Using administrative databases to evaluate the quality of medical care: A conceptual framework. *Soc Sci Med*. 1995;40(12):1707–15. [PMID: 7660184]  
DOI:10.1016/0277-9536(94)00275-X
  18. Wang RY, Strong DM. Beyond accuracy: What data quality means to data consumers. *J Manage Inform Syst*. 1996;12(4):5–33.
  19. Smith BM, Evans CT, Ullrich P, Burns S, Guihan M, Miskevics S, LaVela SL, Rajan S, Weaver FM. Using VA data for research in persons with spinal cord injuries and disorders: Lessons from SCI QUERI. *J Rehabil Res Dev*. 2010;47(8):679–88. [PMID: 20125048]  
DOI:10.1682/JRRD.2009.08.0117
  20. Carlson KF, Nugent SM, Grill J, Sayer NA. Accuracy of external cause-of-injury coding in VA poly-trauma patient discharge records. *J Rehabil Res Dev*. 2010;47(8):689–98. [PMID: 20125048]  
DOI:10.1682/JRRD.2009.08.0118
  21. Bangerter A, Gravely A, Cutting A, Clothier B, Spoot M, Sayer N. Brief report: Comparison of methods to identify Iraq and Afghanistan war veterans using Department of Veterans Affairs administrative data. *J Rehabil Res Dev*. 2010;47(8):815–22. [PMID: 20125048]  
DOI:10.1682/JRRD.2009.08.0115
  22. Harris AHS, Reeder RN, Ellerbe L, Bowe T. Are VHA administrative location codes valid indicators

- of specialty substance use disorder treatment? *J Rehabil Res Dev*. 2010;47(8):699–708.  
DOI:10.1682/JRRD.2009.07.0106
23. Frayne SM, Berg E, Holmes TH, Laungani K, Berlowitz DR, Miller DR, Pogach L, Jackson VW, Moos R. Mental illness-related disparities in length of stay: Algorithm choice influences results. *J Rehabil Res Dev*. 2010;47(8):709–18.  
DOI:10.1682/JRRD.2009.08.0112
  24. Berlowitz DR, Hickey EC, Saliba D. Can administrative data identify active diagnoses for long-term care resident assessment? *J Rehabil Res Dev*. 2010;47(8):719–24.  
DOI:10.1682/JRRD.2009.08.0123
  25. Smith MW, Su P, Phibbs CS. Matching prosthetics order records in VA National Prosthetics Patient Database to healthcare utilization databases. *J Rehabil Res Dev*. 2010;47(8):725–38.  
DOI:10.1682/JRRD.2009.07.0098
  26. Noël PH, Copeland LA, Perrin RA, Lancaster AE, Pugh MJ, Wang C-P, Bollinger MJ, Hazuda HP. VHA Corporate Data Warehouse height and weight data: Opportunities and challenges for health services research. *J Rehabil Res Dev*. 2010;47(8):739–50.  
DOI:10.1682/JRRD.2009.08.0110
  27. Fischer MJ, Stroupe KT, Hynes DM, Blemur P, Sohn M-W, Browning MM, Huo Z, O'Hare AM, Kaufman JS. Validation of erythropoietin use data on Medicare's End-Stage Renal Disease Medical Evidence Report. *J Rehabil Res Dev*. 2010;47(8):751–62.  
DOI:10.1682/JRRD.2009.08.0108
  28. Washington DL, Sun S, Canning M. Creating a sampling frame for population-based veteran research: Representativeness and overlap of VA and Department of Defense databases. *J Rehabil Res Dev*. 2010;47(8):763–72.  
DOI:10.1682/JRRD.2009.08.0127
  29. Hendricks A, Gardner J, Frakt A, Gilden D, Prentice J, Wolfsfeld L, Pizer S. What can Medicaid data add to research on VA patients? *J Rehabil Res Dev*. 2010;47(8):773–80.  
DOI:10.1682/JRRD.2009.07.0107
  30. 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):781–96.  
DOI:10.1682/JRRD.2009.08.0122
  31. Jonk Y, O'Connor H, Schult T, Cutting A, Feldan R, Cowper Ripley D, Dowd B. Using the Medicare Current Beneficiary Survey to conduct research on Medicare-eligible veterans. *J Rehabil Res Dev*. 2010;47(8):797–814.  
DOI:10.1682/JRRD.2009.10.0174
- This article and any supplementary material should be cited as follows:  
Tarlov E, Stroupe KT. Advancing veterans' health-care using electronic data: Lessons learned from researchers in the field. *J Rehabil Res Dev*. 2010;47(8):vii–xii.  
DOI:10.1682/JRRD.2010.08.0148

