Published 2018. This article is a U.S. Government work and is in the public domain in the USA
DOI: 10.1111/1475-6773.13081
SPECIAL ISSUE: LINKING VA AND NON-VA DATA TO ADDRESS US VETERAN
HEALTH SERVICES ISSUES

HSR Commentary: Linking VA and Non-VA Data to Address Important US Veteran Health Services Research Issues

*Denise M. Hynes , Matthew L. Maciejewski , and
David Atkins *

Objective. This commentary summarizes the methods and topics addressed in the special issue of HSR focused on linkage of United States Department of Veterans Affairs (VA) and non-VA datasets. The issue illustrates that researchers are increasingly linking diverse datasets as a valuable method for obtaining outcomes, treatments, and covariates to evaluate and examine health care delivery that includes non-VA services. The issue serves as a reference to VA and non-VA investigators alike who employ data linkage methods to address high-impact clinical and health policy evaluations that span different care systems and different datasets.

Key Words. Data science, linkage methods, Veterans

The theme of this special issue of HSR is linkage of United States Department of Veterans Affairs (VA) and non-VA datasets. As the largest integrated health care system in the United States with an interconnected electronic health record going back more than 20 years, the VA has a deep and rich repository of clinical information that the health system records for patient care and aggregates for performance assessment and quality improvement. For almost as long, VA researchers have used these data to study a wide variety of health services questions, from examining causes of quality variations to evaluating the impact of new programs. Researchers have learned that there can be some limitations to working with VA data alone, since many Veterans using the VA also receive some portion of their care from other health care providers and encounter-based VA data may not capture everything that is relevant to a patient's health. The papers that follow in this special issue illustrate that researchers are increasingly linking diverse datasets as a valuable method for obtaining outcomes, treatments, and covariates that would otherwise be unavailable. We think these lessons are broadly applicable to both the

challenges and benefits of data linkage as VA and non-VA researchers address questions that span different care systems and different datasets.

The most prevalent use of data linkage is linking VA clinical data to Medicare claims files. As Vietnam Veterans age, a large proportion of VA patients are now eligible for Medicare, and the use of Medicare services increases as they age. For clinical and research purposes, VA has deterministically linked Medicare and Medicaid claims to VA records to assess Veterans' care outside of the VA health system. Deterministic matching, often used in linkage of one claims dataset to another, is possible when there are one or more unique patient identifiers (such as Social Security Number, first name or last name) that are included in both datasets to facilitate exact linkages. Deterministic matching can be done through a series of iterative steps and also can include concealing the linkage keys with cryptographic hashing (Quantin et al. 1998; Dusetzina et al. 2014). Most papers in this special issue linked VA and non-VA data using deterministic matching. Probabilistic matching is another method that has been used in other contexts, such as linkage of clinical trial registries to Medicare claims (Hammill et al. 2009), and was used by Carlson and colleagues. Probabilistic matching is the linkage of two datasets when multiple variables are used to determine the likelihood that the same person or events are represented in both datasets and the degree of matching can be measured by linkage scores and the application of decision rules. The choice of using a deterministic or probabilistic approach for data linkage depends upon whether direct identifiers of good quality are available, as well as the goals for the project (Roos, Wajda, and Nicol 1986).

All of the papers in this issue relied in part on a VA national database known as the Corporate Data Warehouse, the provenance for which is the VA electronic health record system. Understanding this context is critical to the validity of linkage with other data sources. These data reflect the nuances of the VA clinical system activities that focus on workload reporting and patient care management, instead of diagnosis and procedure coding with the intention of submitting bills for reimbursement by a third-party payer. Data

Address correspondence to Denise M. Hynes, M.P.H., Ph.D., R.N., Center to Improve Veteran Involvement in Care, Veterans Affairs Portland VA Health Care System, Portland, OR; Veterans Affairs Information Resource Center, Hines, IL; and also College of Public Health and Human Sciences, Oregon State University, Corvallis, OR; e-mail: hynesd@oregonstate.edu. Matthew L. Maciejewski, Ph.D., is with the Center for Health Services Research in Primary Care, Durham Veterans Affairs Health Care System, Durham, NC; and also Department of Department of Population Health Sciences, Duke University, Durham, NC. David Atkins, M.D., M.P.H., is with the Department of Veterans Affairs, Office of Research and Development, Washington, DC.

linkages and analyses of the types represented in this special issue have been supported by the VA Information Resource Center (VIREC), which has facilitated understanding of the data by generating and disseminating documentation on the utility and quality of VA data for research (Tarlov and Stroupe 2010). Efforts to apply standardized data models to transform VA and linked Medicare data, such as ongoing work by VIREC and the VA Informatics and Computing Infrastructure based on the Observational Medical Outcomes Partnership data model and quality tools, may enable linkage and use of a broader range of data in the future (Lynch and Viernes 2018). Such data transformation efforts will be necessary as Veterans increasingly seek health care from community providers that contract with VA, thus requiring data linkages of electronic health record systems to support a fuller assessment of health and health care services.

The papers in this special issue used a range of non-VA datasets, several presented for the first time in a research application. Medicare data were the most commonly linked data (used in Wang et al. 2018; Thorpe et al. 2018; Vaughan Sarrazin et al. 2018; Trivedi et al. 2018; Reddy et al. 2018; Nelson et al. 2018a, b; Vanneman et al. 2018; Liu et al. 2018; Chui et al. 2018; Hebert et al. 2018; Lei et al. 2018), because these data have been available to VA researchers within the VA under an agreement with Centers for Medicare and Medicaid Services via the VIREC for more than fifteen years and have well-established documentation and file structures (Hynes et al. 2007). Other non-VA datasets used by authors in this special issue included US Renal Data System (USRDS) (Wang et al. 2018), state data from California birth records (Shaw et al. 2018), and Oregon prescription drug monitoring program (PDMP) data (Carlson et al. 2018). These papers leveraged the linked data they created to address a range of topics, including VA and non-VA health care use and costs (Wang et al. 2018; Vaughan Sarrazin et al. 2018; Hebert et al. 2018; Vanneman et al. 2018; Liu et al. 2018; Lei et al. 2018), medication use (Thorpe et al. 2018; Carlson et al. 2018; and Chiu et al. 2018), and the impact of homelessness on health care use and costs (Nelson et al. 2018, and Trivedi et al. 2018). One study focused on health care use in VA and VA community care (Rosen et al. 2018). Clinical conditions addressed included mental health (Vanneman et al. 2018), dementia (Thorpe et al. 2018; Lei et al. 2018), end-stage kidney disease (Wang et al. 2018), pain management (Carlson et al. 2018; Chiu et al. 2018), high-risk pregnancy (Shaw et al. 2018), and infectious disease (Nelson et al. 2018). Two studies linked Medicare data to examine the impact of the VA's patient aligned care team intervention (Trivedi et al. 2018; Reddy et al. 2018).

Two papers demonstrated the value of linking VA and non-VA datasets in creative ways. The paper by Hebert and colleagues linked 15 years of VA health care use and Medicare claims to a large survey of Veterans conducted in 1999 to examine how become age-eligible for Medicare changes Veterans' use of eight categories of VA and Medicare services over many years and compare survival between Veterans more or less reliant on VA services. Since Veterans could self-select the specific services obtained in VA or Medicare and their timing and unobserved confounding were likely an issue, the authors implemented an instrumental variable analysis to address these validity threats. As nonelderly Veterans increasingly access community-based care, an analysis modeled on this paper would be valuable to understand time trends in Veteran demand for specific types of services.

The paper by Carlson and colleagues conducted a novel linkage of the VA Department of Defense Identity Repository that was deterministically linked with VA health care use data and then with information from the Oregon Health Authority (OHA) PDMP via probabilistic matching to examine the frequency with which post-9/11 Veterans seen in Oregon in 2014–2016 were obtaining opioids and sedative-hypnotics. This paper is especially relevant because opioid use and misuse have become an epidemic in the United States and the VA initiated the Opioid Safety Initiative in 2014 to attempt to combat it among Veterans (Lin et al. 2017). Using the linked VA and OHA PDMP data afforded Carlson and colleagues the ability to address this issue among a younger, lesser-studied population and found that more than one-third of Veterans who received opioids or sedative-hypnotics from the VA had also received these prescriptions from non-VA prescribers at some point during the study period. This work highlights the value of using probabilistic linkage of state PDMP data to obtain accurate prescription histories in high-risk populations.

We conclude this introduction by noting that this special issue originally was envisioned by our dear colleague, Jim Burgess, who passed away in June 2017. Jim was passionate about conducting rigorous research to address high-priority policy questions and supporting the next generation of health services researchers through mentoring. Just prior to his untimely passing, he was to become Co-Editor-in-Chief of HSR and had begun envisioning this special issue and planning with us in his role on the VIREC Steering Committee. His absence is sorely felt at HSR, the community of VA investigators, VIREC, and the larger community of health services researchers with whom he was actively engaged. We hope that this collection of manuscripts lives up to his vision for the special issue and expect it will serve as a reference to VA and

non-VA investigators alike who employ data linkage methods to address high-impact clinical and health policy evaluations.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: This work was supported by the U.S. Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development Service, project numbers SDR 02-237 and SDR 98-004; RCS 98-352 (DMH); RCS 10-391 (MM).

Disclosures: Dr. Maciejewski owns Amgen stock due to his spouse's employment.

Disclaimer: 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 United States government.

REFERENCES

- Carlson, K. F., T. A. Gilbert, B. J. Morasco, D. Wright, J. V. Otterloo, A. Herrndorf, and L. J. Cook. 2018. Linkage of VA and State Prescription Drug Monitoring Program Data to Examine Concurrent Opioid and Sedative-Hypnotic Prescriptions Among Veterans. *Health Services Research* 53 (6 Pt 2): 5285-308. <https://doi.org/10.1111/1475-6773.13025>
- Chui, P. W., L. A. Bastian, E. DeRycke, C. A. Brandt, W. C. Becker, and J. L. Goulet. 2018. Dual Use of Department of Veterans Affairs and Medicare Benefits on High-Risk Opioid Prescriptions in Veterans Aged 65 Years and Older: Insights From the VA Musculoskeletal Disorders Cohort. *Health Services Research* 53 (6 Pt 2): 5402-18. <https://doi.org/10.1111/1475-6773.13060>
- Dusetzina, S. B., S. Tyree, A. M. Meyer, A. Meyer, L. Green, and W. R. Carpenter. 2014. *Linking Data for Health Services Research: A Framework and Instructional Guide*. Rockville, MD: Agency for Healthcare Research and Quality (US).
- Hammill, B. G., A. F. Hernandez, E. D. Peterson, G. C. Fonarow, K. A. Schulman, and L. H. Curtis. 2009. "Linking Inpatient Clinical Registry Data to Medicare Claims Data Using Indirect Identifiers." *American Heart Journal* 157 (6): 995-1000.
- Hebert, P. L., A. S. Batten, E. Gunnink, A. Reddy, E. S. Wong, S. D. Fihn, and C. Liu. 2018. Reliance on Medicare Providers by Veterans After Becoming Age-Eligible for Medicare is Associated With the Use of More Outpatient Services. *Health Services Research* 53 (6 Pt 2): 5159-80. <https://doi.org/10.1111/1475-6773.13033>
- Hynes, D. M., K. Koelling, K. Stroupe, N. Arnold, K. Mallin, M. W. Sohn, F. M. Weaver, L. Manheim, and L. Kok. 2007. "Veterans' Access to and Use of Medicare and Veterans Affairs Health Care." *Medical Care* 45 (3): 214-23.

- Lei, L., S. G. Cooley, C. S. Phibbs, B. Kinosian, R. M. Allman, A. P. Porsteinsson, and O. Intrator. 2018. Attributable Cost of Dementia: Demonstrating Pitfalls of Ignoring Multiple Health Care System Utilization. *Health Services Research* 53 (6 Pt 2): 5331-51. <https://doi.org/10.1111/1475-6773.13048>
- Lin, L. A., A. S. B. Bohnert, R. D. Kerns, M. A. Clay, D. Ganoczy, and M. A. Ilgen. 2017. "Impact of the Opioid Safety Initiative on Opioid-Related Prescribing in Veterans." *Pain* 158 (5): 833-9.
- Liu, C., A. Batten, E. S. Wong, S. D. Fihn, and P. L. Hebert. 2018. Fee-for-Service Medicare-Enrolled Elderly Veterans are Increasingly Voting With Their Feet to Use More VA and Less Medicare, 2003–2014. *Health Services Research* 53 (6 Pt 2): 5140-58. <https://doi.org/10.1111/1475-6773.13029>
- Lynch, K., and B. Viernes. 2018. "Conducting a Research Study using OMOP Data." V. I. a. C. Infrastructure [accessed on October 15, 2018]. Available at https://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/2487-notes.pdf
- Nelson, R. E., M. Jones, C. Liu, M. H. Samore, M. E. Evans, V. W. Stevens, T. Reese, and M. A. Rubin. 2018a. The Impact of Healthcare-Associated Methicillin-Resistant *Staphylococcus Aureus* Infections on Postdischarge Health Care Costs and Utilization across Multiple Health Care Systems. *Health Services Research* 53 (6 Pt 2): 5419-37. <https://doi.org/10.1111/1475-6773.13063>
- Nelson, R. E., Y. Suo, W. Pettey, M. Vanneman, A. E. Montgomery, T. Byrne, J. D. Fargo, and A. V. Gundlapalli. 2018b. Costs Associated With Health Care Services Accessed Through VA and in the Community Through Medicare for Veterans Experiencing Homelessness. *Health Services Research* 53 (6 Pt 2): 5352-74. <https://doi.org/10.1111/1475-6773.13054>
- Quantin, C., H. Bouzelat, F. A. Allaert, A. M. Benhamiche, J. Faivre, and L. Dusserre. 1998. "Automatic Record Hash Coding and Linkage for Epidemiological Follow-Up Data Confidentiality." *Methods of Information in Medicine* 37 (3): 271-7.
- Reddy, A., E. Wong, A. Canamucio, K. Nelson, S. D. Fihn, J. Yoon, and R. M. Werner. 2018. Association Between Continuity and Team-Based Care and Health Care Utilization: An Observational Study of Medicare-Eligible Veterans in VA Patient Aligned Care Team. *Health Services Research* 53 (6 Pt 2): 5201-18. <https://doi.org/10.1111/1475-6773.13042>
- Roos Jr, L. L., A. Wajda, and J. P. Nicol. 1986. "The Art and Science of Record Linkage: Methods that Work with Few Identifiers." *Computers in Biology and Medicine* 16 (1): 45-57.
- Rosen, A. K., T. H. Wagner, W. B. Pettey, M. Shwartz, Q. Chen, J. Lo, W. J. O'Brien, and M. E. Vanneman. 2018. Differences in Risk Scores of Veterans Receiving Community Care Purchased by the Veterans Health Administration. *Health Services Research* 53 (6 Pt 2): 5438-54. <https://doi.org/10.1111/1475-6773.13051>
- Shaw, J. G., V. R. Joyce, S. K. Schmitt, S. M. Frayne, K. A. Shaw, B. Danielsen, R. Kimerling, S. M. Asch, and C. S. Phibbs. 2018. Selection of Higher Risk Pregnancies Into Veterans Health Administration Programs: Discoveries From Linked Department of Veterans Affairs and California Birth Data. *Health Services Research* 53 (6 Pt 2): 5260-84. <https://doi.org/10.1111/1475-6773.13041>

- Tarlov, E., and K. T. Stroupe. 2010. "Advancing Veterans' Healthcare Using Electronic Data: Lessons Learned from Researchers in the field." *Journal of Rehabilitation Research and Development* 47 (8): vii-xi.
- Thorpe, C. T., W. F. Gellad, M. K. Mor, J. P. Cashy, J. R. Pleis, C. H. Van Houtven, L. J. Schleiden, J. T. Hanlon, J. D. Niznik, R. L. Carico, C. B. Good, and J. M. Thorpe. 2018. Effect of Dual Use of Veterans Affairs and Medicare Part D Drug Benefits on Antihypertensive Medication Supply in a National Cohort of Veterans With Dementia. *Health Services Research* 53 (6 Pt 2): 5375-401. <https://doi.org/10.1111/1475-6773.13055>
- Trivedi, A. N., L. Jiang, E. E. Johnson, J. C. Lima, M. Flores, and T. P. O'Toole. 2018. Dual Use and Hospital Admissions Among Veterans Enrolled in the VA's Homeless Patient Aligned Care Team. *Health Services Research* 53 (6 Pt 2): 5219-37. <https://doi.org/10.1111/1475-6773.13034>
- Vanneman, M. E., C. S. Phibbs, S. K. Dally, A. N. Trivedi, and J. Yoon. 2018. The Impact of Medicaid Enrollment on Veterans Health Administration Enrollees' Behavioral Health Services use. *Health Services Research* 53 (6 Pt 2): 5238-59. <https://doi.org/10.1111/1475-6773.13062>
- Vaughan Sarrazin, M., G. E. Rosenthal, and C. L. Turvey. 2018. Empirical-Based Typology of Health Care Utilization by Medicare Eligible Veterans. *Health Services Research* 53 (6 Pt 2): 5181-200. <https://doi.org/10.1111/1475-6773.12995>
- Wang, V., C. J. Coffman, K. M. Stechuchak, T. S. Berkowitz, P. L. Hebert, D. Edelman, A. M. O'Hare, H. J. Weidenbacher, and M. L. Maciejewski. 2018. Comparative Assessment of Utilization and Hospital Outcomes of Veterans Receiving VA and Non-VA Outpatient Dialysis. *Health Services Research* 53 (6 Pt 2): 5309-30. <https://doi.org/10.1111/1475-6773.13022>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix SA1: Author Matrix.