Technical Brief ■

Use of Administrative Records to Describe Longitudinal Patterns of Health Services Use among Veterans

BRENT C. WILLIAMS, MD, MPH, DAVID R. MEHR, MD, BRANT E. FRIES, PhD

Abstract Administrative records of the Department of Veterans Affairs (DVA) and the National Death Index were linked to create a four-year longitudinal database that describes the clinical status, hospital and nursing home use, and mortality for a nationwide cohort of persons admitted to DVA nursing homes ($n = 23,039$). Using Social Security Numbers as identifiers, the records of only 6% of these persons had logically inconsistent or implausible patterns. Nineteen percent of the remaining records had correctable logical errors. Information on the creation, consistency, and potential uses of this database may prove useful to health services researchers interested in describing longitudinal patterns of health care use across multiple settings within and outside the DVA.

■ *J Am Med Informatics Assoc.* 1994;1:325–327.

Large administrative databases are increasingly being used to describe the health services utilization patterns and, to a lesser extent, the clinical characteristics of large groups of persons.^{1–3} Few researchers, however, have succeeded in describing multiple-year patterns of the clinical status and health care use of large groups of persons, especially across multiple settings (e.g., hospitals and nursing homes). Substantial difficulties exist in linking records across time and among different payers, and administrative databases that include clinical information are rare. We

report the development and internal consistency of a unique database that describes the clinical and functional status and health services utilization patterns of a national cohort of users of Department of Veterans Affairs (DVA) nursing homes over a substantial period (three to four years).

We linked administrative records of the DVA and the National Death Index to create a four-year (October 1985 to October 1989) longitudinal database of health services use and mortality for a complete nationwide cohort of persons who resided in or were admitted to DVA-operated nursing homes between October 1986 and October 1987 [fiscal year (FY) 1987]. DVA data sources included the Patient Assessment File (PAF), which includes information on functional status (e.g., impairments in eating, toileting, and transferring) and clinical characteristics (e.g., terminal illness, needs for wound or ostomy care) of all persons in DVA-operated nursing homes at admission, at return from the hospital, and semiannually; and the Patient Treatment File (PTF), which includes information on diagnoses and dates of admission to and discharge from DVA hospitals, DVA-operated nursing homes, and community nursing homes under DVA contract. The presence of 23,039 patients in DVA-operated nursing homes during FY 1987 was established through the PAF. Using Social Security Numbers, records for individual patients were linked

Affiliations of the authors: University of Michigan (BCW, BEF), Ann Arbor, MI, and University of Missouri (DRM), Columbia, MO.

Presented in part at the Eighth Annual Department of Veterans Affairs Services Research and Development Meeting, Washington, DC, April 1992.

Supported by grant IIR 90-096 from Health Services Research and Development, Department of Veterans Affairs; and by a grant from the Claude D. Pepper Geriatric Research and Training Center at the University of Michigan (National Institute of Aging grant AG-08808).

Correspondence and reprints: Brent C. Williams, MD, MPH, University of Michigan Medical Center, Department of Internal Medicine, 3116 Taubman Center, 1500 E. Medical Center Drive, Ann Arbor, MI 48109-0376. e-mail: BWILLIAM@umich.edu

Received for publication: 3/22/93; accepted for publication: 3/9/94.

Table 1 ■

Error Patterns in Constructed Longitudinal Records for Persons in Department of Veterans Affairs (DVA)-operated Nursing Homes during Fiscal Year 1987 ($n = 23,029$)*

Error Type†	Number	%
Irreconcilable		
Dates of NH admission from PAF and PTF records for same NH stay differ by > 2 days	497	2.2
PAF semiannual assessment recorded, patient not in nursing home (PAF or PTF)	220	1.0
NH admission recorded in PTF while patient is still in hospital (PTF)	192	0.8
Overlapping NH stays from PTF records	181	0.8
Any event (NH or hospital admission, NH assessment) occurring > 3 days after death	123	0.5
Other	218	0.9
SUBTOTAL	1,431	6.2
Reconcilable‡		
Nursing home admission recorded in PAF while PTF indicates patient is still in hospital		
> 2 days prior to discharge in PTF	1,421	6.2
Within 2 days of discharge in PTF	1,284	5.6
Dates of birth different on several records	859	3.7
Dates of NH admission from PAF and PTF records for same NH stay differ by ≤ 2 days	286	1.2
Different NH admission dates among PAF records	175	0.8
Assessment recorded in PAF ≤ 2 days after NH discharge	169	0.7
Other	133	0.6
SUBTOTAL	4,327	18.8

*Within each class (reconcilable and irreconcilable), only one type of error was defined to occur for each patient.

†PAF = Patient Assessment File; PTF = Patient Treatment File; NH = DVA-operated nursing home or community nursing home under DVA contract.

‡Among records without irreconcilable error patterns.

to PTFs for FYs 1986 through 1989 and to PAFs for FYs 1987 through 1989, requiring a search of approximately 1.2-million PTFs and 75,000 PAFs per year. Mortality during the follow-up period was established using the National Death Index, whose accuracy has been established using this⁴ and other⁵⁻⁶ populations. The information was then reorganized to create a person-based data set that included information on up to 40 events (e.g., admissions to, transfers to or from, or discharges from DVA hospitals or nursing homes, and death) for each patient.

As they were constructed, the longitudinal profiles

were examined for any of 27 potential illogical or implausible patterns. These "error patterns" were classified as reconcilable (e.g., probably due to miscoding a date by a few days) or irreconcilable (e.g., records indicating simultaneous residence in a nursing home and a hospital). Information was retained for only one error pattern per patient—the first irreconcilable error pattern to occur, or the last reconcilable error pattern to occur when no irreconcilable error was found.

The frequencies of the error patterns are shown in Table 1. Only 6.2% of patients had irreconcilable errors; additional work probably would have reconciled some of these errors (e.g., those with inverted dates or wrong years). The most common irreconcilable errors were conflicting dates of nursing home admission between PAF and PTF records (2.2% of cases), and a PAF semiannual assessment occurring when no nursing home admission had been recorded in either the PAF or the PTF (1.0% of cases).

The most common reconcilable errors were nursing homes admission recorded in the PAF while the PTF indicated that the patient was still in the hospital (6.2% of cases had more than a two-day discrepancy), and conflicting dates of birth on separate records (3.7% of cases). In the first type of case, dates of nursing home admission were taken from the PTF, since the DVA uses this information administratively (unlike PAF dates of admission) and it is more likely to be accurate. In the second type of case, the same date of birth occurring on more than one record was used, even though other records included different dates of birth.

This study demonstrated the feasibility of creating a unique, longitudinal database using computerized records of the DVA. The database offered several advantages over previous efforts to describe the health services utilization patterns among the elderly. First, the database was person based rather than group based, in contrast to some previous research that has focused on identifying "transition rates" (or "transition probabilities") for large cohorts of elderly persons between hospital, nursing home, and community living, and death.⁷ Second, the database followed a large (nationwide) cohort of individuals for substantial periods of time. Most previous research has been limited to relatively brief (e.g., six to 12 months) follow-ups of smaller cohorts of individuals.⁸⁻¹⁰ Third, the database combined information on clinical status with information on health services utilization for members of the study population. Most previous research using administrative data has included extremely limited measures of health status, such as

discharge diagnosis or diagnosis-related group.^{1,11} Fourth, the structure of the database allowed the accurate description of back-to-back institutional stays. Most previous research in longitudinal health care patterns of nursing home residents has not accounted for the fact that many persons "discharged" from nursing homes are in fact only temporarily hospitalized and subsequently return to the nursing homes. Not accounting for back-to-back episodes of institutional care can substantially alter estimates of nursing home discharge destination and length of stay.¹²

The database has made possible the examination of a variety of important issues in health services research. For example, the clinical characteristics and length-of-stay patterns of a nationwide cohort of persons newly admitted to VA nursing homes have been described.¹³ A second application of the database has been to describe the patterns and predictors of health services utilization and death among a national cohort of persons who were discharged alive from VA nursing homes for a period of up to three years.¹⁴ Other applications of the database have included examination of the predictors of live discharge from a VA nursing home, and a comparison of the characteristics of a "current-stay" cohort of VA nursing home residents with those of non-VA nursing home residents as reported in the 1985 National Nursing Home Survey.¹⁵

The most important limitation of the data was the lack of information on the use of non-DVA health resources by members of the cohort. Thus, estimates of the amount and predictors of total health services use in this population may be biased.

We suggest that 1) computerized records can be used to describe longitudinal patterns of health services utilization and clinical status for large groups of persons; and 2) future efforts to examine longitudinal health service patterns of users of the DVA health system should attempt to include information on the use of non-DVA health services and on the reliability and validity of information contained in the PAF and PTF. Health services researchers interested in developing similar databases within or outside the DVA may contact the authors for further information.

The authors thank Clare Weipert for assisting with manuscript preparation.

References ■

1. Holloway JJ, Medendorp SV, Bromberg J. Risk factors for early readmission among veterans. *Health Serv Res.* 1990;25(Part II):213-37.
2. Wennberg JE, Freeman JL, Shelton RM, Bubolz TA. Hospital use and mortality among Medicare beneficiaries in Boston and New Haven. *N Engl J Med.* 1989;321:1168-73.
3. Shapiro E, Webster LM. Nursing home utilization patterns for all Manitoba admissions, 1974-1981. *Gerontologist.* 1984;24:610-5.
4. Williams BC, Demitrack LB, Fries BE. The accuracy of the National Death Index when personal identifiers other than Social Security number are used. *Am J Public Health.* 1992;82:1145-7.
5. Wentworth DN, Neaton JD, Rasmussen WL. An evaluation of the Social Security Administration Master Beneficiary Record File and the National Death Index in the Ascertainment of vital status. *Am J Public Health.* 1983;73:1270-4.
6. Boyle CA, Decoufle P. National source of vital status information: extent of coverage and possible selectivity in reporting. *Am J Epidemiol.* 1990;131:160-8.
7. Densen PM. The elderly and the health care system: another prospect. *Milbank Q.* 1987;65:614-38.
8. Lewis MA, Kane RL, Cretin S, Clark V. The immediate and subsequent outcomes of nursing home care. *Am J Public Health.* 1985;75:758-62.
9. Wolinsky FD, Coe RM, Mosely RR. The use of health services by elderly Americans: implications from a regression-based cohort analysis. In: Ward RA, Tobin SS (eds). *Health in Aging: Sociological Issues and Policy Directions.* New York: Springer Publishing Company 1987:106-32.
10. Evashwick C, Rowe G, Diehr P, Branch L. Factors explaining the use of health care services by the elderly. *Health Serv Res.* 1984;19:357-82.
11. Roos NP, Wennberg JE, McPherson K. Using diagnosis-related groups for studying variations in hospital admissions. *Health Care Financing Rev.* 1988;9:53-62.
12. Spence DA, Weiner JD. Nursing home length of stay patterns: results from the 1985 National Nursing Home Survey. *Gerontologist.* 1990;30:16-20.
13. Williams BC, Mehr DR, Fries BE. Length of stay in VA nursing homes: comparative characteristics of brief-, medium-, and long-stay residents. *J Aging Health.* 1993;5:208-28.
14. Williams BC, Fries BE, Mehr DR. Patterns and determinants of health services use and mortality after VA nursing home care. Submitted for publication, 1993.
15. Mehr DR, Fries BE, Williams BC. How different are VA nursing home residents? *J Am Geriatr Soc.* 1993;41:1095-101.