Department of Veterans Affairs

Veterans Health Administration multiple sclerosis surveillance registry: The problem of case-finding from administrative databases

William. J. Culpepper II, MA;[1–3]*Mary Ehrmantraut, RN, MS;[2–3] Mitchell T. Wallin, MD, MPH;[3–4] Kathleen Flannery, PharmD;[2] Douglas D. Bradham, DrPH[1–3]

[1]*Department of Epidemiology and Preventive Medicine, Healthcare Outcomes Research, University of Maryland School of Medicine, Baltimore, MD;* [2]*Veterans Integrated Service Network 5 Health Services Research and Development Service Center, Research Office, Baltimore Department of Veterans Affairs Medical Center (VAMC), Baltimore, MD;* [3]*Multiple Sclerosis Center of Excellence East, Baltimore VAMC, Baltimore, MD;* [4]*Department of Neurology, Washington, DC, VAMC, Washington, DC*

Abstract—Establishment of a national multiple sclerosis (MS) surveillance registry (MSSR) is a primary goal of the Department of Veterans Affairs (VA) MS Center of Excellence. The initial query of Veterans Health Administration (VHA) databases identified 25,712 patients (labeled "VHA MS User Cohort") from fiscal years 1998 to 2002 based on International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) code; service-connection for MS; and/or disease-modifying agent (DMA) use. Because of ICD-9-CM limitations, the initial query was overinclusive and resulted in many non-MS cases. Thus, we needed a more rigorous case-finding method. Our gold standard was chart review of the Computerized Patient Record System for the mid-Atlantic VA medical centers. After chart review, we classified patients as not having MS or having MS/possible MS. We also applied a statistical algorithm to classify cases based on service-connection for MS, DMA use, and/or at least one healthcare encounter a year with MS coded as the primary diagnosis. We completed two analyses with kappa coefficient and sensitivity analysis. The first analysis (efficacy) was limited to cases with a definitive classification based on chart review ($n = 600$). The kappa coefficient was 0.85, sensitivity was 0.93, and specificity was 0.92. The second analysis (effectiveness) included unknown cases that were classified as MS/possible MS ($N = 682$). The kappa coefficient was 0.82, sensitivity was 0.93, and specificity was 0.90. These findings suggest that the database algorithm reliably eliminated non-MS cases from the initial MSSR population and is a reasonable case-finding method at this intermediate stage of MSSR development.

Key words: case finding, chart review, database, diagnosis, kappa, multiple sclerosis, sensitivity, specificity, surveillance registry, Veterans Health Administration.

INTRODUCTION

Multiple sclerosis (MS) is the most common progressive neurological disease of young adults [1–3]. It is of major interest to the Department of Veterans Affairs

---

(VA) because of the associated high rates of healthcare use. In late 2002, the Veterans Health Administration (VHA) funded two MS Centers of Excellence (MSCoEs) to conduct research and education for improving the clinical care of veterans with MS. On the East Coast, the MSCoE is located at the VA medical center (VAMC) in Baltimore, Maryland; on the West Coast, the MSCoE is shared between the VAMCs in Portland, Oregon, and Seattle, Washington. A primary goal of the MSCoEs was development of an MS surveillance registry (MSSR). The MSSR would provide researchers with a pool of veterans for future studies of unresolved epidemiological, clinical, and quality-of-care issues in MS diagnosis, treatment, and management. Ultimately, the MSSR would facilitate a better systemwide understanding of the needs of MS patients. This mandate seemed feasible. However, case finding from extant VHA data presents unique challenges [4–7], especially when the diagnosis cannot be made based on specific, objective medical tests and/or procedures. This challenge is especially true for diagnoses that are made clinically, such as MS. Another problem is reliance on diagnostic coding schemes with insufficient taxonomic detail to allow adequate distinction between confirmed clinical diagnoses and suspected diagnoses that must be "ruled out." For example, the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) provides only one 3-digit code (340) for suspected MS, evaluations to rule out MS, and clinically definite MS.

We initially identified cases through a retrospective query of extant VHA data sources based on ICD-9-CM MS diagnostic code or disease-modifying agent (DMA) use. This query identified 25,712 veterans from fiscal year (FY) 1998 to FY2002 who had at least one healthcare encounter in which MS was coded, had a service-connected disability for MS, or used a DMA. This initial data set was labeled the "VHA MS User Cohort" [8]. We recognized that refinement of the case-finding process was needed because of the limitations of the ICD-9-CM and resultant overidentification of cases.

This article describes our initial attempt to refine the case-finding process to yield a cohort of veterans with a high probability of having MS or possible MS. We did not intend to derive a clinical diagnosis by applying the McDonald et al. criteria [9] to the extant data, since many critical data elements were missing from the computerized patient record (e.g., magnetic resonance imaging, cerebrospinal fluid). Rather, our intent was to eliminate cases that clearly were not MS because other cases would eventually be formally evaluated by an MS specialist participating in the MSSR and/or through screening for inclusion in future studies.

METHODS

We identified 25,712 veterans with our initial query of the extensive VHA databases. In this query, we defined a case as any inpatient or outpatient veteran with one or more ICD-9-CM Code 340 for MS from FY1998 to FY2002 (25,290 unique veterans). Additionally, we queried Veterans Benefits Administration (VBA) databases. Specifically, we used the VBA Compensation and Pension Mini-Master File to identify veterans who were service-connected for MS. Service-connection for MS requires definitive evidence of clinical–MS signs upon examination during or within 7 years of active military service [10]. Finally, we examined VHA Pharmacy Benefits Management (Chicago, Illinois) data to identify veterans who had been prescribed an MS-specific DMA from FY1998 to FY2002. These last two queries yielded 422 additional unique veterans. The final result was the initial VHA MS User Cohort of 25,712 veterans. Details of this methodology can be found in the MSCoE's 2003 annual report [8].

Chart Review Sample

We used chart review as the gold standard for classification of MS diagnostic status. For this study, we reviewed charts in the Computerized Patient Record System (CPRS) from only the 682 VHA MS User Cohort cases (2.6%) in Veteran Integrated Service Network (VISN) 5. VISN 5 encompasses the VAMCs in Baltimore, Maryland; Washington, DC; and Martinsburg, West Virginia. The flowchart in **Figure 1** describes the inclusion of VISN 5 cases in the chart review sample.

A nurse practitioner who was experienced in MS care and research was the primary reviewer and used the McDonald diagnostic criteria to guide the classification of cases as [9]—

1. MS or possible MS (some symptoms consistent with MS such as clinically isolated syndrome and other possible diagnoses not yet ruled out).

2. Definitely not MS.

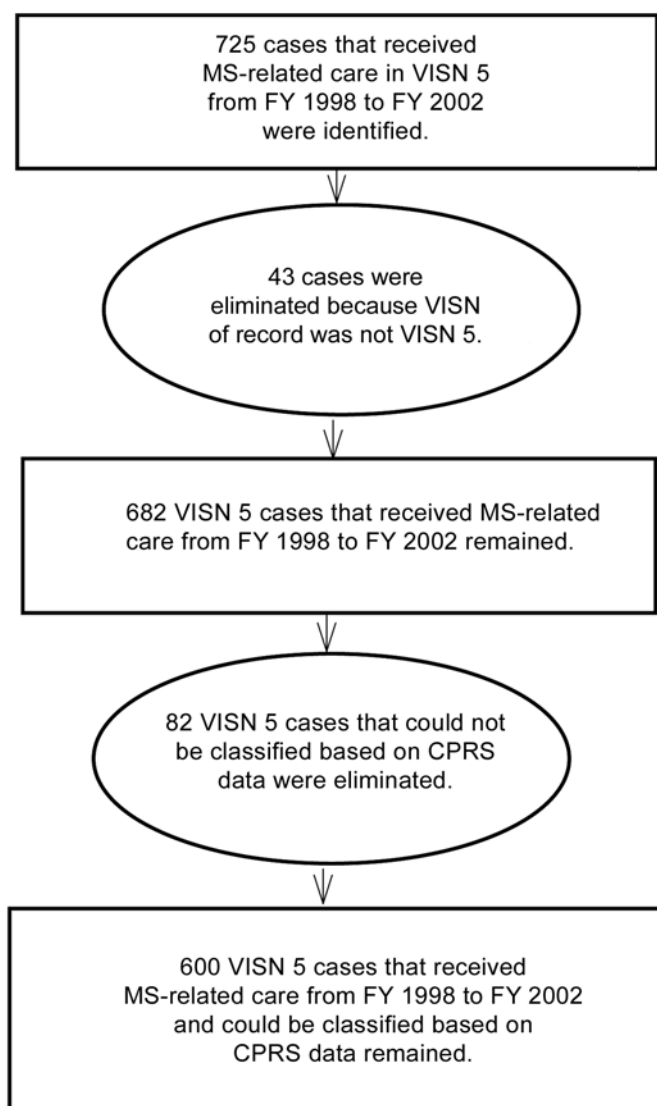3. Unknown (not enough data to determine).

19

CULPEPPER et al. VHA multiple sclerosis surveillance registry

Figure 1.
Flowchart describing selection of cases for chart review. CPRS = Computerized Patient Record System, FY = fiscal year, MS = multiple sclerosis, VISN = Veterans Integrated Service Network.

A total of 50 cases (7%) required adjudication. For these cases, the MS specialist in our group independently reviewed the chart in the CPRS. A consensus was reached in each case. When a classification was in doubt, we conservatively classified the case as possible MS. Because our focus was on eliminating cases that clearly were not MS as opposed to making formal diagnoses, we did not conduct independent ratings on a subset of cases for interrater reliability.

Cases that were classified as not MS had no clear alternative diagnosis. The most common diagnosis among these cases was "other unknown and unspecified morbidity/mortality" (ICD-9-CM Code 799.9) and occurred in 32 cases (10.8%). Parkinson's disease was found in only 16 cases (5.4%). No other diagnosis accounted for more than 2 percent of cases classified as not MS.

Statistical Algorithm Development

Since chart review was impractical for the nearly 26,000 cases in the initial VHA MS User Cohort, we developed a statistical algorithm that we could apply to the VHA MS User Cohort to identify the cases that were not MS and leave the remaining cases (i.e., MS/possible MS) in the MSSR.

The statistical algorithm classified cases based on the frequency of healthcare encounters with MS coded as the primary diagnosis, service-connection for MS [10–11], and DMA use (Avonex®, Rebif®, Betaseron®, and Copaxone®). We assumed that patients who had more frequent healthcare encounters with MS coded as the primary diagnosis were more likely to have MS. We chose a cutoff of at least one healthcare encounter each year with MS coded as the primary diagnosis. This cutoff was based on a comparison of the distribution of the number of ICD-9-CM Code 340 diagnoses for the cases classified in the chart review as not MS vs MS/possible MS (**Table 1**). Thus, a patient with MS who entered the VHA MS User Cohort in FY1999 and did not have a service-connected disability for MS or use a DMA would need four or more healthcare encounters (i.e., at least one each year) with ICD-9-CM Code 340 to be classified as MS/possible MS. Because the statistical algorithm is based on the extant VHA databases, we refer to it as the "database algorithm."

The database algorithm classified a case as not MS if the patient did not have, on average, at least one healthcare encounter each year with MS coded as the primary diagnosis, did not have a service-connected disability for MS, and did not use a DMA. If any of these criteria were met, the case was classified as MS/possible MS.

Because our objective was to develop a tool that would reliably identify the initial MSSR population, we empha-

Table 1.
Number of healthcare encounters each year with multiple sclerosis (MS) coded as primary diagnosis.

MS Classification by Chart Review	Median	Mean ± SD
Not MS	1	1.6 ± 1.5
MS/Possible MS	18	44.2 ± 80.4
SD = standard deviation.		

sized the elimination of cases that were clearly not MS. Thus, retaining cases for which diagnostic classification could not be made based on the available medical record data was important. These cases were classified as unknown based on chart review. We believed that this approach, while including a few more false-positive cases, was preferable to eliminating false-negative cases (cases who had MS) from the MSSR. We plan to reclassify cases each year when new clinical event data become available, which is in keeping with the clinical nature of MS diagnosis.

Reliability Analysis

We used the kappa (κ) coefficient [11] and traditional sensitivity analysis (sensitivity and specificity) [12] to evaluate how well the database algorithm correctly classified MS diagnostic status compared with chart review. Each case was dichotomized as not MS or MS/possible MS based on chart review and the database algorithm.

The κ measures the correspondence between two binary ratings [11]. Most frequently, κ is used for measuring interrater or intertest agreement. In this study, the chart review classification and the database algorithm classification were treated as separate raters or as versions of the same test. We felt that this κ application, while untraditional, was more conservative than using the Spearman correlation or simply reporting percent agreement.

Sensitivity analyses (sensitivity and specificity) are traditional epidemiologic tools for evaluating how well a test identifies cases relative to an established gold standard [12]. In this study, the test was the database algorithm classification and the gold standard was the chart review classification of MS. Not MS was the target condition because our primary objective was to eliminate as many non-MS cases as possible from the MSSR.

Two separate analyses were performed. The first analysis ($n = 600$) removed the 82 cases that were classified as unknown based on the chart review. The focus of this analysis was the evaluation of the efficacy of the database algorithm, i.e., how well the database algorithm performs under ideal conditions. The second analysis ($N = 682$) included the 82 unknown cases and classified them as MS/possible MS. Here, the focus was assessment of the effectiveness or real-world utility of the database algorithm.

RESULTS

In **Table 2**, the complete results of the chart review classification and the database algorithm classification are cross-referenced.

The κ from the first analysis (i.e., efficacy, $n = 600$) was 0.85 (95% confidence interval [CI]: 0.80–0.89; $p < 0.001$) and indicated very good agreement between the two classification methods. **Table 3** summarizes the results from this sensitivity analysis. These results also show that the database algorithm performed very well, with an overall agreement of 92 percent with the chart review (554 of 600 correctly classified compared with chart review).

The κ from the second analysis (i.e., effectiveness, $N = 682$), in which the unknown cases were included as MS/possible MS, was 0.82 (95% CI: 0.78–0.86; $p < 0.001$). **Table 4** summarizes the results from the sensitivity analysis in this larger patient sample. The database algorithm performed as well as the chart review, with an overall agreement of 91 percent (621 of 682 correctly classified compared with chart review).

DISCUSSION

Results of this study indicate that the database algorithm reliably distinguishes between not MS and MS/possible MS cases from extant VHA databases. Overall, the database algorithm correctly classified a little better than 9 out of every 10 cases compared with chart review. Chart review of the VISN 5 cases revealed that a diagnostic classification could not be made in 12 percent (82) of the cases because of insufficient information and that 43.3 percent of cases were classified as not MS. We hypothesize that this high percentage is due mostly to the imprecision of ICD-9-CM MS coding and the clinical nature of MS diagnosis.

Several caveats need comment before a discussion of the application and implications of the database algorithm. First, inter- and intrarater reliabilities were not performed because one individual conducted chart reviews and adjudication was required in only 50 (7%) of the cases. This lack of reliability testing would have been a major flaw if our intent had been to make formal clinical MS diagnoses. However, our aim was to eliminate as many non-MS cases as possible. This approach was deemed most appropriate for the initial MSSR population, since the cases will undergo further routine evaluation as part of the MSSR project and through screening for possible participation in future studies.

A second caveat concerns how well the VISN 5 results represent the rest of the VHA. The extent of this limitation is unknown because the only way to estimate

21

CULPEPPER et al. VHA multiple sclerosis surveillance registry

Table 2.
Results of Computerized Patient Record System chart review and database algorithm multiple sclerosis (MS) classifications. A nurse practitioner completed chart review and classified each case as MS/possible MS, unknown, or not MS. Database algorithm classified a case as MS/possible MS if the patient had at least one healthcare encounter each year with MS coded as primary diagnosis, was service-connected (SC) for MS, or used a disease-modifying agent (DMA). If none of these criteria was met, the case was classified as not MS.

Database Algorithm	Chart Review			Total
	MS/Possible MS	Unknown[*]	Not MS	
SC for MS				
n	52	37	1	90
% Within Chart Review	57.8	41.1	1.1	100.0
% Within Database Algorithm	17.1	45.1	0.3	13.2
DMA				
n	82	5	0	87
% Within Chart Review	94.3	5.7	0	100.0
% Within Database Algorithm	27.0	6.1	0	12.8
SC and DMA				
n	90	8	0	98
% Within Chart Review	91.8	8.2	0	100.0
% Within Database Algorithm	29.6	9.8	0	14.4
≥ 1 Healthcare Encounter/Year for MS				
n	56	17	21	94
% Within Chart Review	59.6	18.1	22.3	100.0
% Within Database Algorithm	18.4	20.7	7.1	13.8
Not MS				
n	24	15	274	313
% Within Chart Review	7.7	4.8	87.5	100.0
% Within Database Algorithm	7.9	18.3	92.6	45.9
Total				
n	304	82	296	682
% Within Chart Review	44.6	12.0	43.4	100.0

[*]Insufficient data in chart to allow classification. These cases were classified as MS/possible MS for analysis.

Table 3.
Sensitivity analysis comparing multiple sclerosis (MS) diagnostic classification (No. of cases) between database algorithm[*] and chart review[†] with unknown cases excluded. Sensitivity (274/296) = 0.93, specificity (280/304) = 0.92, positive predictive value (274/298) = 0.92, and negative predictive value (280/302) = 0.93.

		Chart Review		Total
		Not MS	MS	
Database Algorithm	Not MS	274	24	298
	MS	22	280	302
Total		296	304	600

[*]MS classification: patients who used disease-modifying agent, were service-connected for MS, and/or had at least one healthcare encounter each year with MS coded as primary diagnosis.
[†]MS classification: patients with MS and possible MS.

Table 4.
Sensitivity analysis comparing multiple sclerosis (MS) diagnostic classification (No. of cases) between database algorithm[*] and chart review[†] with unknown cases included. Sensitivity (274/296) = 0.93, specificity (347/386) = 0.90, positive predictive value (274/313) = 0.88, and negative predictive value (347/369) = 0.94.

		Chart Review		Total
		Not MS	MS	
Database Algorithm	Not MS	274	39	313
	MS	22	347	369
Total		296	386	682

[*]MS classification: patients who used disease-modifying agent, were service-connected for MS, and/or had at least one healthcare encounter each year with MS coded as primary diagnosis.
[†]MS classification: patients with MS, possible MS, and 82 unknown cases.

representativeness would be with similar data and analyses from other VISNs. However, part of our rationale in eliminating non-MS cases was to minimize the potential impact of basing the database algorithm on data from a single VISN. Obtaining sufficient funding for evaluation of one or two other VISNs has been difficult. Thus, the extent to which our conservative approach offsets the limitation of a single VISN's worth of data is also unknown.

Despite these caveats, we feel comfortable applying the database algorithm to the entire VHA MS User Cohort. In this study, we applied the database algorithm to the VHA MS User Cohort through a computer program that reviewed and classified each of the 25,712 cases as not MS if the patient had less than one healthcare encounter a year with MS coded as the primary diagnosis, did not have a service-connected disability for MS, and did not use a DMA. Cases were classified as MS/possible MS if the patient had more than one healthcare encounter a year with MS coded as the primary diagnosis, had a service-connected disability for MS, or used a DMA. **Figure 2** displays the transition from the original VHA MS User Cohort to the more refined MSSR that followed application of the database algorithm. The database algorithm eliminated 11,159 cases (43.4%) that were not MS. The remaining 14,553 cases (56.6%) with a high probability of MS/possible MS became the initial MSSR population. These proportions are nearly identical to the chart review sample (45.9% not MS and 54.1%
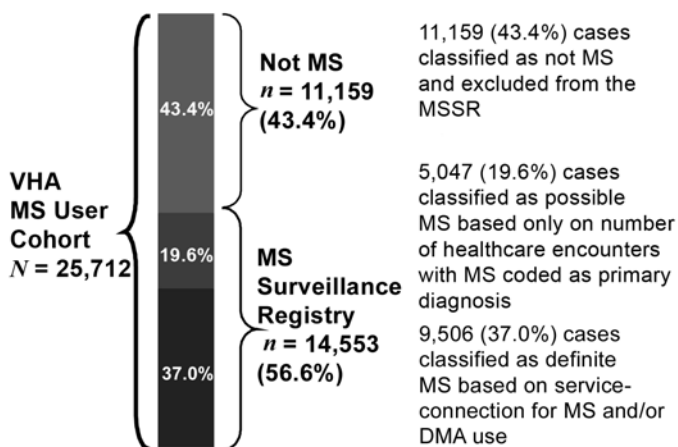

Figure 2.
Application of database algorithm to Veterans Health Administration (VHA) Multiple Sclerosis (MS) User Cohort to derive the MS surveillance registry (MSSR). DMA = disease-modifying agent.

MS/possible MS) and further support the reliability and robustness of the database algorithm's ability to identify MS/possible MS cases from extant VHA data sources.

Additionally, the database algorithm allows discrimination between possible and definite MS cases. The criteria for definite MS cases are service-connected disability for MS and/or DMA use. In the chart review sample, 99.6 percent of the 275 cases classified as MS based on service-connection and/or DMA use were classified as MS/possible MS by chart review. In contrast, cases that met the criteria based only on the number of healthcare encounters with MS coded as the primary diagnosis but not on service connection or DMA use were classified as possible MS (**Figure 2**). In the chart review sample, 77.6 percent of the 94 cases classified as possible MS based on number of healthcare encounters with MS coded as primary diagnosis (by the algorithm) were also classified as MS/possible MS based on chart review. The lower agreement on possible MS cases reflects the ambiguity involved in MS diagnosis.

As this MSCoE project proceeds, we expect to further refine the MSSR through screening for participation in various studies and clinical trials, ongoing clinical evaluation, and the annual addition of more clinical event data that we will refine by reapplying the database algorithm.

Other efforts for improving MS diagnostic accuracy in the VHA are underway. The MSCoEs are currently field-testing a CPRS MS-specific template that when implemented VHA-wide will help validate and document MS diagnosis for questionable MSSR and future MS cases. Additionally, several studies have recently been approved or are in review that will screen MSSR cases for inclusion in the studies. Diagnostic information from these efforts will be fed into the MSSR so that we may further refine and define the MSSR. We anticipate that over the next several years, we will achieve our ultimate goal of further reducing the MSSR to only those veterans with confirmed MS or possible MS according to current diagnostic criteria [8].

ACKNOWLEDGMENTS

23

CULPEPPER et al. VHA multiple sclerosis surveillance registry

REFERENCES

1. Courtney SW. Multiple sclerosis: the process and medical treatment. 4th ed. Cherry Hill (NJ): Multiple Sclerosis Association of America; 2002.

2. Courtney SW. All about multiple sclerosis. 2nd ed. Cherry Hill (NJ): Multiple Sclerosis Association of America; 2003.

3. Ebers GC, Koopman WJ, Hader W, Sadovnik AD, Kremenchutzky M, Mandalfino P, Wingerchuk DM, Baskerville J, Rice GP. The natural history of multiple sclerosis: a geographically based study. 8: Familial multiple sclerosis. Brain. 2000;123(Pt 3):641–49. [PMID: 10686184]

4. Van der Heyden J, Tafforeau J, Van Oyen H, Demarest S. Measurement of the use of curative health services: health interview survey versus national registers. Arch Public Health. 2003;61(4):177–90.

5. Faciszewski T, Broste SK, Fardon D. Quality of data regarding diagnoses of spinal disorders in administrative databases. A multicenter study. J Bone Joint Surg Am. 1997;79(10):1481–88. [PMID: 9378733]

6. Kashner TM. Agreement between administrative files and written medical records: a case of the Department of Veterans Affairs. Med Care. 1998;36(9):1324–36. [PMID: 9749656]

7. Petersen LA, Wright S, Normand SL, Daley J. Positive predictive value of the diagnosis of acute myocardial infarction in an administrative database. J Gen Intern Med. 1999;14(9):555–58. [PMID: 10491245]

8. Multiple Sclerosis Center of Excellence. Description of VHA MS Cohort VHA utilization and VHA expenditures: preliminary analyses [monograph on the Internet]. Seattle, Washington: Department of Veterans Affairs; c2004 [cited 2006 May 5]. Available from: http://www.va.gov/ms/

9. McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, McFarland HF, Paty DW, Polman CH, Reingold SC, Sandberg-Wollheim M, Sibley W, Thompson A, Van den Noort S, Weinshenker BY, Wolinsky JS. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the Diagnosis of Multiple Sclerosis. Ann Neurol. 2001;50(1):121–27. [PMID: 11456302]

10. Wallin MT, Page WF, Kurtzke JF. Multiple sclerosis in U.S. veterans of Vietnam era and later military service: race, sex, and geography. Ann Neurol. 2004;55(1):65–71. [PMID: 14705113]

11. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20:37–46.

12. Rothman KJ, Greenland S. Modern epidemiology. 2nd ed. New York (NY): Lippincott Williams & Wilkins; 1998.