The need for validation of large administrative databases: Veterans Health Administration ICD-9CM coding of exudative age-related macular degeneration and ranibizumab usage

Paul Latkany · Mona Duggal · Joseph Goulet · Hyung Paek · Michael Rambo · Philip Palmisano · Woody Levin · Joseph Erdos · Amy Justice · Cynthia Brandt

Received: 23 March 2010 / Accepted: 19 May 2010 / Published online: 9 July 2010
© Springer Science+Business Media, LLC 2010

Abstract We performed a validation study by chart review of data for exudative age-related macular degeneration (eAMD) and, because of the Veterans Administration (VA) therapy policy, ranibizumab usage in the largest electronic medical record system in the USA. We reviewed 5,854 distinct patients who visited an ophthalmology clinic within VA Connecticut from January 2006–December 2008. We randomly selected 98 of 138 distinct eAMD patients and 265 of 5,588 non-eAMD patients who did not receive ranibizumab. International Classification of Diseases, Ninth Revision, Clinical Modification coding of eAMD had an excellent positive predictive value of 97.8% (95% confidence interval (CI), 93.5–99.4%). The national Decision Support System (DSS) had an excellent positive predictive value of 100% (95% CI, 79.9–100%) for ranibizumab. However, the negative predictive value of the DSS dispensed ranibizumab decreased to 67.5 (95% CI, 62.1–72.4) because of a change in the way local values were stored that led to errors. Therefore, validation of clinical information over time in large databases is necessary.

Keywords Age-related macular degeneration · Validation study · Informatics · Ranibizumab

Introduction

The Veterans Administration (VA) has the largest electronic health records (EHR) system in the USA and affiliated territories and is utilized in over 1,400 facilities and contains over 874 million documents [1]. Within the VA EHR, the Veterans Health Information Systems and Technology Architecture (VistA) is the backbone suite of over 100 Health IT Systems including the Computerized Patient Record System (CPRS). The local VistA/CPRS is used to collect all clinical data (including progress and procedure notes) from local patients and most administrative patient-related data in the VA. In addition to textual documentation regarding direct clinical care, VistA has structured and coded data that may be entered by a variety of sources that are part of a patient's electronic health record.

The use of International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9CM) coded data for clinical investigation permits valid clinical questions to be answered [2] more economically than the expense associated with clinical trials. However, there are many possible errors that may occur with the use of ICD-9CM coded data [3]. Validating the accuracy of coded data is an optimal first step with its use. Most validation data studies have examined general medical or surgical care [3].

Diseases treated by subspecialties, outside of general medical or surgical care are important clinically and economically. The annual costs of the VA blind rehabilitation centers were \$250–300 million in 2003 [4]. Age-related macular degeneration (AMD) is one of the more common etiologies of the over one million veterans over 45 years of age with visual impairment. There are currently about 1.7

P. Latkany (✉) · H. Paek · P. Palmisano · W. Levin
VA Connecticut,
West Haven, CT, USA
e-mail: all27email@yahoo.com

M. Duggal · J. Goulet · M. Rambo · J. Erdos · A. Justice ·
C. Brandt
Yale University,
New Haven, CT, USA

million people with AMD in the USA with an estimated cost of treatment of exudative AMD (eAMD) of \$569 million annually [5, 6]. Despite the importance of AMD, the validation of AMD coded administrative data has not been reported as in other eye diseases [7, 8]. There are several reports examining the quality of general medical care within the VA that began with data validation as the necessary foundation [9]. Generally, only administratively coded data from the national VA system are available for analysis. In 2008, over five million people received care within the VA (more than the total population of Norway or New Zealand) [10]. The VA's disseminated policy which begun in 2007 was to utilize ranibizumab as the preferred therapy for eAMD. The objective of this work was to validate eAMD administrative data in the VA in order to perform further analysis of the quality of care of eAMD treatment within the VA using electronic medical record data.

Methods

Study setting and population

The VA Connecticut Healthcare System serves an area of more than 333,000 veterans in Connecticut and southern New England. There are two campuses located in West Haven and Newington and six community-based outpatient clinics. There were 5,854 people that were evaluated in the ophthalmology clinic from 2006 to 2008. All patients were considered to have eAMD by administrative data for study purposes if the patient had more than one ophthalmology visit (using the Veterans Aging Cohort Study analysis design of national administrative VA data) [11] with the ICD-9CM code: 362.52. Data for the control group were extracted from VistA for patients who had a visit to the eye clinic during the study period (Fig. 1). The local pharmacy administrative code was used to determine whether patients administratively had been given ranibizumab.

Patients were categorized into three groups: (1) patients who had received ranibizumab and had eAMD, (2) patients who had eAMD and did not receive ranibizumab, and (3) patients who had neither received ranibizumab nor had eAMD. Of the 138 patients who had eAMD ICD-9CM entered on two separate eye clinic visits but did not receive ranibizumab, 98 were randomly selected. Ranibizumab usage was identified in 108 patients by administrative code, and 20 additional patients were identified by chart review, and all of these patients had received an ICD-9CM eAMD code. Text searching of all notes of patients who had visited the ophthalmology clinic using the terms “rani” and “lucen” was also performed. There were 5,588 patients that had neither received ranibizumab nor had eAMD, and 265 patients were randomly selected from this group for chart review (Fig. 1).

Chart review

Chart review was considered the criterion standard for determination of presence or absence of eAMD and administration of ranibizumab. Since there is no access to the full text medical record nationally, the goal of the local chart review is to determine if the local variables which are uploaded to the national accessible databases can be validated by local chart review. Charts were reviewed for all patients who were identified to have received ranibizumab. Text searching of all progress notes of patients who had visited the ophthalmology clinic was performed to augment the chart review. In addition, the local pharmacy value that was uploaded to the national Decision Support System database was compared to the data abstracted from the chart review.

One of the authors (P.L.), who is a vitreoretinal surgeon, performed the initial chart review. The chart was evaluated by all three physicians, and terms consistent with eAMD in the free text field were considered confirmed presence of eAMD for the chart review. Four cases required adjudication, and all were resolved with discussion. The vitreoretinal surgeon

Fig. 1 Identification and analysis of groups who had visited the eye clinic from VistA data

clarified the text for the generalists. The ophthalmology clinic notes were categorized into presence or absence of eAMD. The categories for ranibizumab use were administered or not administered. The categorized review was then compared to the ranibizumab administrative and ICD-9CM data. All grading of medical texts were masked regarding the ICD-9CM and ranibizumab data. When a disagreement occurred regarding assignment, a discussion resolved all conflicts. This study was approved by the Institutional Review Boards of Yale University School of Medicine and the West Haven, CT VA.

Data sources

ICD-9CM coded data were obtained from VistA and the CPRS for VA Connecticut. The local VistA/CPRS data collects all clinical data (including progress and procedure notes). Providers assign ICD-9CM codes for outpatient visits while professional coders are used for inpatient hospitalizations. The DSS national database was the source for the values of what data were reported nationally. The local VISTA/CPRS was the source of the demographics information in Table 1.

Statistical analysis

Data were entered into VA Microsoft SQL Server 2008. Random selection of patients was generated through its NewID function.

Results were analyzed using Statistical Analysis System version 9.1 (SAS Institute, Cary, NC, USA). The demographics are reported as frequencies and not compared between groups. The skilled chart reviewers' restriction on available time permitted an analysis of all patients logged as having received ranibizumab and a large random sample of those who had eAMD without having received ranibizumab and those without eAMD. We estimated that at least 95 patients taking ranibizumab would be required to place a 95% confidence interval (CI) around the positive predictive value of at most ± 0.1 (NCSS PASS Kaysville, Utah, 2005). We had budgetary

and time constraints in terms of availability of physicians to perform chart review and had to sample the charts. The positive predictive value (PPV) and the negative predictive value (NPV) of the administrative data were calculated and corrected for the sampling scheme by weighting [12]. The positive and negative predictive values are the extent to which the presence or absence of eAMD ICD-9CM data correlates with the presence or absence in the text of the charts, respectively. The sensitivity and specificity of eAMD were calculated correcting for the sampling scheme [13]. The kappa statistic was weighted and used to assess agreement between the ICD-9CM data and chart review and allows them to be independent from the necessity of the establishment of a "criterion standard."

Results

The mean age of the patients who had eAMD without ranibizumab usage was 82.3 (SD 8.56), and there were 99% males. Of these, 54.1% had glaucoma, 11.2% had cerebral vascular disease (CVD), 84.7% had hypertension, and 32.6% had diabetes. The mean age of patients who did not have eAMD and did not receive ranibizumab was 73.2 (SD 12.07), 93.2% were male, 56.2% had glaucoma, 17.4% had CVD, 81.9% hypertension, and 39.2% had diabetes (Table 1).

All patients who had local eAMD ICD-9CM codes were also present nationally without any data management discrepancies identified.

In the sample of 265 patients who were examined in the ophthalmology clinic but did not have an associated eAMD ICD-9CM code, only four patients by chart review were identified having eAMD (the incorrectly coded patients included those who recently has post-cataract extraction and had a recent diagnosis of eAMD; Table 2). The sensitivity of eAMD by administrative code is 61.6% (95% CI, 48.2–80.0%) with a specificity of 100% (95% CI, 99.9–100%). The positive predictive values for eAMD are 97.8% (95% CI, 93.5–99.4).

Of 98 sampled patients who received an ICD-9CM code of eAMD, two were coded incorrectly. Both had

Table 1 Demographics

	eAMD	No eAMD
Age	82.3 (8.56 SD)	73.2 (12.07 SD)
Male	97/98	247/265
Glaucoma	53/98	149/265
CVD	11/98	46/265
Hypertension	83/98	217/265
Diabetes mellitus	32/98	104/265

CVD cardiovascular disease

Table 2 ICD eAMD validation

ICD eAMD	Chart review		
	YES	No	Total
Yes ICD eAMD	96 (TP)	2 (FP)	98
No ICD eAMD	4 (FN)	261 (TN)	265
Total	100	263	363

TP true positive, FP false positive, FN false negative, TN true negative (without Ranibizumab)

non-exudative AMD and had no text passages that confirmed the presence of exudative macular degeneration (Table 2).

Ranibizumab

The mean age of patients who had received ranibizumab through 2008 was 81.1 (7.24), 96.1% were males, 45.3% had glaucoma, 11.7% had CVD, 79.7% had hypertension, and 34.4% had diabetes. Ranibizumab was dispensed to 128 separate patients according to chart review. One patient who received ranibizumab had a choroidal neovascular membrane secondary to a prior laser scar for a branch retinal vein occlusion (Table 3). All patients were considered as having macular degeneration with eAMD for the analysis plan.

The ranibizumab NPV for the national Decision Support System (DSS) is poor because of a change in the way that local ranibizumab values were stored between the years 2007 and 2008. This impacted the results for ranibizumab in Table 3.

The positive predictive values for coding of eAMD and for utilization of ranibizumab are excellent (Table 3). The positive predictive values for ranibizumab administration through 2008 are 99.1% (95% CI, 94.90–100).

Discussion

The VA's electronic medical record system represents a framework for the future of how medical care will be delivered and stored in most of USA. Our study was the first analysis of the agreement between ICD-9CM data and chart review within the VA EMR for eAMD and usage of ranibizumab. While delving directly into large administrative databases for analysis is appealing, our study demonstrates the need for data validation as a necessary first step. Within the range of validation reports of ocular and other medical diseases, our results

place the VA eAMD ICD-9CM data in the more accurate grouping of these variables. The use of ICD-9CM data has been reported based on Medicare claims databases, and the range of the PPV of ICD-9CM data was from 0.53 for peripheral vascular disease to 0.94 for hip fracture and a sensitivity from 0.58 for peripheral vascular disease to 0.97 breast cancer [14]. Also, within Medicare's large claims database, variations in the accuracy of ICD-9CM data have been reported as dependent on the type of eye disease. The error in coding of cataract procedures of 802 Medicare claims was minimal with a positive predictive value of 0.99 (95% CI, 0.93–1.00) [15]; however, retinal detachment coding had a sensitivity of 149/161 (78%) [16, 17].

The accuracy of ICD9 data within the VA also depends on the disease examined. For instance within the VA, detection of central venous catheter insertion determined by administrative data alone had a sensitivity of less than 12% [18], but the positive predictive value of administrative data for serious infection-related hospital admission was 80% [19]. Internal vetting of VA data occurs through planned reviews, and for instance, the DSS and PBM pharmacy data have been reported to be similar.

AMD is an important public health problem and is one of the leading causes of blindness in the elderly [20]. There have been recent improvements in efficacy of treatment of eAMD. Since the year 2000, there has been a rapid evolution of treatment from using laser photocoagulation, to verteporfin, to the anti-vascular endothelial growth factors of bevacizumab and the FDA June 2006 approval of ranibizumab in eAMD. The VA ICD-9CM EMR data of eAMD appear to be very accurate and is likely secondary to the fact that the diagnosis is primarily made by subspecialists and once diagnosed is a lifelong disease. The validation of the administrative data within the VA ophthalmology clinic can be used as a foundation for larger analysis of the VA eAMD quality of care including analyzing optimal detection of eAMD and care utilization in the vast VA system of clinical care.

Use of ranibizumab within the VA was mandated as a preferred therapy over bevacizumab in 2007, and we chose to analyze the use of ranibizumab because of its strict indication of use. We reviewed an initial validated algorithm for 2007 data extraction when an expected increase in administrations for 2008 from our local site to the national database dropped to zero. We learned that new local data management processes were developed that necessitated a different algorithm for successful data analysis of 2008 local data. The initial validated algorithm for ranibizumab administration for uploaded national DSS data was subsequently invalid for a later time point because of a change in local data management. In addition, our findings led to a corrective investigation of DSS national

Table 3 Ranibizumab validation local and national

		Chart		
		Yes	No	Total
Ranibizumab (national)	Yes	20 (TP)	0 (FP)	20
	No	108 (FN)	102 (TN)	210
	Total	128	102	230
Ranibizumab (local)	Yes	108(TP)	1 (FP)	109
	No	20 (FN)	101(TN)	121
	Total	128	102	230

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

data reporting from the local site. In the VA, the administration of ranibizumab by national pharmacy data should be viewed as having excellent PPV. Since there is no maximum monthly number of administrations of ranibizumab per patient when indicated, a cumulative increase of ranibizumab use should likely be expected as evident from our local data. Errors of data management should be suspected when as in our data no reported administrations occurred from a clinic where previous injections occurred. Concerns may arise in big administrative databases when this validation cannot be performed, and validation over time may be required as we found in our study.

Possible reasons for high positive predictive values in our study of eAMD in the VA are that it is not a rare disease and has very few ICD-9CM codes associated with its diagnosis and the high expense associated with using ranibizumab leading to better documentation.

A limitation of our study is that it examined data from one region within the VA healthcare system limited to VA Connecticut where both ophthalmology and retinal specialists are available.

Validation of values in large databases permits accurate plans of further analysis. There is a high PPV of eAMD ICD-9CM coding and for ranibizumab treatment within VA CT as shown by manual chart review of the EMR and provides the opportunity for further study of the quality of eAMD care within the VA nationally.

References

1. Ashton CM, Menke TJ, Deykin D, et al. A state-of-the-art conference on databases pertaining to veterans' health. A resource for research and decision making. *Med Care*. 1996;34 Suppl 3:1–8.
2. O'Malley KJ, Cook KF, Price MD, et al. Measuring diagnoses: ICD code accuracy. *Health Serv Res*. 2005;40:1620–39.
3. Quan H, Parsons GA, Ghali WA. Validity of procedure codes in International Classification of Diseases, 9th revision, clinical modification administrative data. *Med Care*. 2004;42:801–9.
4. Stroupe KT, Stelmack JA, Tang XC, et al. Economic evaluation of blind rehabilitation for veterans with macular diseases in the Department of Veterans Affairs. *Ophthalmic Epidemiol*. 2008;15: 84–91.
5. Jager RD, Mieler WF, Miller JW. Age-related macular degeneration. *N Engl J Med*. 2008;358:2606–17.
6. Friedman DS, O'Colmain BJ, Muñoz B, et al. Eye diseases prevalence research group. Prevalence of age-related macular degeneration in the United States. *Arch Ophthalmol*. 2004;122:564–72.
7. Coleman AL, Greenland S. Glaucoma outcome studies using existing databases: opportunities and limitations. *J Glaucoma*. 1995;4:295–8.
8. Coleman AL, Morgenstern H. Use of insurance claims databases to evaluate the outcomes of ophthalmic surgery. *Surv Ophthalmol*. 1997;42:271–8.
9. Harris AH, Kivlahan DR, Bowe T, et al. Developing and validating process measures of health care quality: an application to alcohol use disorder treatment. *Med Care*. 2009;47:1244–50.
10. Kinsella K, He W. US Census Bureau, International population 'reports, P95/09-1, an aging world. Washington: US Government Printing Office; 2008.
11. Justice AC, Dombrowski E, Conigliaro J, et al. Veterans Aging Cohort Study (VACS): overview and description. *Med Care*. 2006;44:S13–24.
12. Fleiss JL, Levin B, Paik MC, Fleiss J. Statistical methods for rates and proportions. 3rd ed. New York: Wiley-Interscience; 2003. p. 223–5.
13. Zhou XH, Obuchowski NA, McClish DK. Statistical methods in diagnostic medicine. New York: Wiley; 2002. p. 309–12.
14. Fisher ES, Whaley FS, Krushat WM, et al. The accuracy of Medicare's hospital claims data: progress has been made, but problems remain. *Am J Public Health*. 1992;82:243–8.
15. Javitt JC, McBean AM, Sastry SS, DiPaolo F. Accuracy of coding in Medicare part B claims. Cataract as a case study. *Arch Ophthalmol*. 1993;111:605–7.
16. Javitt JC, Tielsch JM, Canner JK, et al. National outcomes of cataract extraction. Increased risk of retinal complications associated with Nd:YAG laser capsulotomy. The cataract patient outcomes research team. *Ophthalmology*. 1992;99:1487–97.
17. Tielsch JM, Legro MW, Cassard SD, et al. Risk factors for retinal detachment after cataract surgery. A population-based case-control study. *Ophthalmology*. 1996;103:1537–45.
18. Penz JF, Wilcox AB, Hurdle JF. Automated identification of adverse events related to central venous catheters. *J Biomed Inform*. 2007;40:174–82.
19. Schneeweiss S, Robicsek A, Scranton R, et al. Veteran's affairs hospital discharge databases coded serious bacterial infections accurately. *J Clin Epidemiol*. 2007;60:397–409.
20. Coleman HR, Chan CC, Ferris 3rd FL, et al. Age-related macular degeneration. *Lancet*. 2008;372:1835–45.