Published in final edited form as:

Alzheimer Dis Assoc Disord. 2017 ; 31(2): 168–172. doi:10.1097/WAD.0000000000000164.

Validity of death certificate and hospital discharge ICD codes for dementia diagnosis: the Multi Ethnic Study of Atherosclerosis

Akira Fujiyoshi^{1,2}, David R. Jacobs Jr², Alvaro Alonso², José A. Luchsinger³, Stephen R. Rapp⁴, and Daniel A. Duprez⁵

¹Department of Public Health, Shiga University of Medical Science, Shiga, Japan

²Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN, USA

³Departments of Medicine and Epidemiology, Columbia University Medical Center, New York, NY, USA

⁴Department of Psychiatry and Behavioral Medicine, Wake Forest University School of Medicine, Winston-Salem, NC, USA

⁵Department of Medicine, University of Minnesota, Minneapolis, MN, USA

INTRODUCTION

Large epidemiologic studies of dementia are made difficult by the complexity of case finding and definition. A simple and widely available approach would be to base the case definition on the International Classification of Diseases (ICD) codes, but additional data on their validity are needed. In the Multi-Ethnic Study of Atherosclerosis (MESA), we examined the concordance of dementia-related ICD codes with textual comments in corresponding medical records (here restricted to hospital records and death certificates). Our hypothesis was that ICD codes reflecting dementia diagnosis reasonably correspond to textual language in medical records.

METHODS

MESA focuses on the prevalence, risk factors, and progression of subclinical cardiovascular disease in a multi-ethnic cohort, as previously described.¹ In brief, 6814 community-dwelling men and women who identified themselves as white, black, Hispanic, or Chinese, aged 45 to 84 years, free of clinically apparent cardiovascular disease (CVD), and other serious illness and cognitively able to participate in the study were recruited from 2000 through 2002 from Baltimore City and Baltimore County, Maryland; Chicago, Illinois; Forsyth County, North Carolina; Los Angeles County, California; northern Manhattan and

Corresponding author: Akira Fujiyoshi, MD, PhD, fujiy001@umn.edu, Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, 1300 S. 2nd Street, Minneapolis, MN 55454, USA, Tel: 612-624-1818; Fax: 612-624-0315.

Conflicts of Interest

All authors have no conflict of interest.

the Bronx, New York; and St. Paul, Minnesota. For this study we used data through 2013 (released to investigators in June 2015).

MESA telephone interviewers inquired about interim hospital admissions and deaths every 9 months. Copies of all obtainable death certificates and ICD10 codes, face sheets and ICD9 codes from hospital records and some outpatient diagnoses were assembled by MESA staff at each center. MESA medical records were collected to adjudicate fatal and hospitalized cardiovascular events. An attempt was made to obtain at least ICD coding for each hospitalization or death. If the events were related to MESA outcomes the full medical records were requested and the clinic coordinator submitted copies of parts deemed relevant to the MESA adjudication committee for review. Although dementia was not a pre-specified outcome for MESA adjudication, medical records often mentioned it.

Procedure

The set of ICD-codes used to identify dementia cases (Supplemental Table) included non-specific codes such as “memory loss” (ICD9: 780.93) or “persistent mental disorders due to conditions classified elsewhere” (ICD9: 294.8, 294.9) in the list because studies in the US showed that use of these codes was not uncommon^{2,3}, even for dementia cases assessed by neurologists.⁴

We defined dementia as characterized by a significant decline in cognitive function compared to a previous level, not better accounted for by other mental disorders (such as major depressive disorder, schizophrenia) or secondary conditions (due to either infection, malignancy, trauma, or substance use). One clinician (AF) read medical records blinded to ICD codes, looking for phrases that would indicate, or contradict the conditions defined above. Dementia diagnosis textually documented in the medical records was considered the reference (i.e. true positive). We tabulated in categories 1) medical record text concordant with ICD dementia code, 2) medical record discordant with ICD code, 3) medical record indeterminate, and 4) medical record text not available for review. For categories 3 and 4, we also examined all medical records obtained in the Minnesota research clinic (location of the clinician reviewer) for cases where the records sent to the adjudication committee were indeterminate or inadequate (n=11, with 4 judged concordant, 5 judged discordant, and 2 remaining indeterminate), but this further review was not logistically possible in the other clinics.

RESULTS

We identified 306 potential cases of dementia by ICD code (52 based solely on death certificate). As shown in Table 1a, 224 cases were concordant, while 18 were discordant, and 64 were indeterminate (among which 31 had no available textual record).

Positive predictive value (PPV) of dementia identified by ICD code was 73.2% (224 concordant /306 with dementia according to ICD coding). Considering that MESA did not explicitly attempt to obtain medical records for all possible dementia cases and assuming the 64 participants in indeterminate categories 3 and 4 were missing information at random, then the most optimistic PPV would be 92.6% (224/242).

The pattern of ICD codes encountered in these MESA participants is detailed in Table 1a (one condition at a time, whether other conditions were present or not) and Table 1b (mutually exclusive breakdown by single defining ICD code condition or number of defining ICD code conditions). Half of the 306 dementia diagnoses were based on meeting 2 or more ICD code criteria (Table 1b), and medical text in these cases was concordant with the ICD code diagnosis in 84.3% (2 criteria met) with 3.7% discordant and 97.8% (3+ criteria met) with none discordant. In the other half of cases which met only a single criterion, medical text confirmed the ICD diagnosis in only 58.2%, with 9.2% discordant and 32.7% indeterminate. Among cases meeting a single criterion (Table 1b), substantially higher discordance rates occurred only in the rare categories “F. Cognitive deficits, late effects of cerebrovascular disease” (12 cases) and “G. General signs and symptoms: memory loss” (2 cases). Indeterminate records were common in these two categories and in the category “E. Other specified dementias and non-specific conditions” (16 cases). Considering any mention of an ICD criterion code group (Table 1a), concordance percentage was 80% or more for “A. Vascular dementia”, “B. Alzheimer’s dementia”, “C. Other dementias and chronic organic psychotic conditions”, and “D. Other persistent mental disorders”, corresponding to any ICD9 290 or 294 code plus 331.0, or any ICD10 F00, F01, F03, F04, or G30 code.

Few of the identified dementia cases had the dementia ICD code in the 1st or 2nd position; therefore the medical textual wording which justified our judgment of concordance was limited. We found it to be adequate to claim concordance with dementia, but insufficient for the specific dementia diagnosis in part because in 63% of cases (140/224) the wording in the medical record did not provide a specific dementia diagnosis. For the 18 discordant cases, the following alternative conditions (number of cases) were considered more likely: delirium or acute confusional state (7), cognitive decline induced by remote trauma (3), normal pressure hydrocephalus (2), alcohol-related or hepatic encephalopathy (2), brain metastasis from malignancy (1), infection of the central nervous system (1), other conditions (2).

DISCUSSION

We found a high PPV of ICD code dementia diagnosis in people who were not demented at MESA baseline. We made a conservative PPV estimate of 73.2% (i.e. medical text was concordant with the ICD code diagnosis), and a less conservative estimate of 92.6% under the liberal assumption that the records specific to the dementia ICD code were either not obtained from hospital or were retained in clinic, given the cardiovascular focus of the MESA adjudication effort.

Studies on validity of dementia ICD codes from clinical/administrative practice have been limited, and most of them were from European countries where a national registry or similar system exists.⁵⁻⁸ PPV of ICD codes from those studies, including a study on Alzheimer’s disease in the US⁹, ranged from 78.9 to 96.3 %, close to our estimate. We found that specific dementia diagnoses were less commonly documented in the medical records that were not focused on documenting dementia. Similar results were reported in the literature.^{4,6,8}

Inclusion of non-specific codes such as “memory loss” in epidemiological studies may be a challenge as it is likely to increase both true positive and false positive cases.⁴ Such trade-off

should be considered carefully in selecting code sets, but these codes were relatively rare in this MESA sample, occurring as the only defining criteria in 30 people, <10% of all MESA cases (categories E, F, and G in Table 1). Therefore inclusion or exclusion of these more questionable cases does not seem to be an important epidemiologic consideration in defining dementia by ICD code based on death and hospital records. Many of these questionable cases could have had mild cognitive impairment, which would alter our conclusion in a study that included outpatient records.

A limitation in MESA is that discovery of false negatives is problematic. Dementia cases that were never hospitalized or died without mention of dementia, lost to follow-up due to their dementia, or less advanced cases (who are less likely to be documented in medical records^{2,9}) are likely scenarios for false negativity. Additionally, our review of medical records was not complete (89.9%, 275 out of 306 potential cases). However, a strength of our study is the completeness of MESA follow-up and attempt to obtain at least ICD diagnoses from all participants hospitalized or dead, whether for a presumed MESA target event, or not. Another strength is that MESA is a study of community-dwelling people initially without overt CVD or severe cognitive loss.

In conclusion, in a study with near complete follow-up of hospitalizations and deaths, ICD code-based dementia diagnosis had a true positive rate of at least 73%. While our broad dementia category has its own limitation, assuming nondifferential misclassification of dementia by exposure, our case definition would enhance studies of dementia risk in cohort studies such as MESA and in “big data” such as CMS/Medicare or EMR generally.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Source of Funding:

This research was supported by contracts N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168 and N01-HC-95169 from the National Heart, Lung, and Blood Institute and by grants UL1-TR-000040 and UL1-TR-001079 from NCRR. The authors thank the other investigators, the staff, and the participants of the MESA study for their valuable contributions. A full list of participating MESA investigators and institutions can be found at <http://www.mesa-nhlbi.org>. AF is supported by Fulbright Program (RS 2014)

References

1. Bild DE, Bluemke DA, Burke GL, et al. Multi-Ethnic Study of Atherosclerosis: objectives and design. *Am J Epidemiol*. 2002; 156:871–881. [PubMed: 12397006]
2. Cho K, Gagnon DR, Driver JA, et al. Dementia Coding, Workup, and Treatment in the VA New England Healthcare System. *Int J Alzheimers Dis*. 2014; 2014:821894. [PubMed: 24701364]
3. Butler D, Kowall NW, Lawler E, et al. Underuse of diagnostic codes for specific dementias in the Veterans Affairs New England healthcare system. *J Am Geriatr Soc*. 2012; 60:910–915. [PubMed: 22587853]
4. Pippenger M, Holloway RG, Vickrey BG. Neurologists' use of ICD-9CM codes for dementia. *Neurology*. 2001; 56:1206–1209. [PubMed: 11342688]

5. Solomon A, Ngandu T, Soininen H, et al. Validity of dementia and Alzheimer's disease diagnoses in Finnish national registers. *Alzheimers Dement*. 2014; 10:303–309. [PubMed: 23849592]
6. van de Vorst IE, Vaartjes I, Sinneker LF, et al. The validity of national hospital discharge register data on dementia: a comparative analysis using clinical data from a university medical centre. *Neth J Med*. 2015; 73:69–75. [PubMed: 25753071]
7. Jin YP, Gatz M, Johansson B, et al. Sensitivity and specificity of dementia coding in two Swedish disease registries. *Neurology*. 2004; 63:739–741. [PubMed: 15326258]
8. Phung TK, Andersen BB, Hogh P, et al. Validity of dementia diagnoses in the Danish hospital registers. *Dement Geriatr Cogn Disord*. 2007; 24:220–228. [PubMed: 17690555]
9. Taylor DH Jr, Fillenbaum GG, Ezell ME. The accuracy of medicare claims data in identifying Alzheimer's disease. *J Clin Epidemiol*. 2002; 55:929–937. [PubMed: 12393082]

Table 1a
Judgments about agreement of medical record text and ICD coding for diagnosis of dementia

Univariate analysis of each criterion grouping (cases may occur in multiple criterion ICD code categories). Each criterion alone qualifies for a diagnosis of dementia according to ICD coding in hospital or death records.

Criterion definition	Concordance of ICD code with medical record text (considered to be the "true positive") (number)				Concordance of ICD code with medical record text (%)			
	ICD9 code	ICD10 code	Concordant with ICD code (N)	Discordant with ICD code (N)	Indeterminate (N)	Total (N)	Frequency distribution: % of total	Indeterminate (%)
All cases	See Supplemental table		224	18	64	306	100.0	5.9
A. Vascular dementia	290.4	F01	22	0	4	26	8.5	0.0
A. Criterion not met			202	18	60	280	91.5	6.4
B. Alzheimer's dementia	331.0	F00, G30	84	0	8	92	30.1	0.0
B. Criterion not met			140	18	56	214	69.9	8.4
C. Other dementias and chronic organic psychotic conditions	290, 294 (not 290.4, 294.8, 294.9)	F03	154	6	22	182	59.5	3.3
C. Criterion not met			70	12	42	124	40.5	9.7
D. Other persistent mental disorders	294.8 or 294.9	F04	120	7	23	150	49.0	4.7
D. Criterion not met			104	11	41	156	51.0	7.1
E. Other specified dementias and non-specific conditions	331.1, 331.2, 331.8, 331.9	G31 not G31.2	19	2	14	35	11.4	5.7
E. Criterion not met			205	16	50	271	88.6	5.9
F. Cognitive deficits: late effects of cerebrovascular disease	438.0	I69.91	3	1	1	5	1.6	20.0
F. Criterion not met			221	17	63	301	98.4	5.6
G. General signs and symptoms: memory loss	780.93	R41	6	6	7	19	6.2	31.6
G. Criterion not met			218	12	57	287	93.8	4.2
								19.9

Judgments about agreement of medical record text and ICD coding for diagnosis of dementia. Mutually exclusive combinations of criteria met

Table 1b

Criterion definition	Concordance of ICD code with medical record text (number)					Concordance of ICD code with medical record text (%)				
	ICD9 code	ICD10 code	Concordant with ICD code (N)	Discordant with ICD code (N)	Indeterminate (N)	Total (N)	Frequency distribution: % of total N	Concordant with ICD code (%)	Discordant with ICD code (%)	Indeterminate (%)
Exactly 1 criterion met			89	14	50	153	50.0	58.2	9.2	32.7
A. Vascular dementia	290.4	F01	5	0	2	7	2.3	71.4	0.0	28.6
B. Alzheimer's dementia	331.0	F00, G30	29	3	9	41	13.4	70.7	7.3	22.0
C. Other dementias and chronic organic psychotic conditions	290, 294 (not 290.4, 294.8, 294.9)	F03	32	5	13	50	16.3	64.0	10.0	26.0
D. Other persistent mental disorders	294.8 or 294.9	F04	21	0	4	25	8.2	84.0	0.0	16.0
E. Other specified dementias and non-specific conditions	331.1, 331.2, 331.8, 331.9	G31 not G31.2	1	1	14	16	5.2	6.3	6.3	87.5
F. Cognitive deficits: late effects of cerebrovascular disease	438.0	I69.91	0	1	1	2	0.7	0.0	50.0	50.0
G. General signs and symptoms: memory loss	780.93	R41	1	4	7	12	3.9	8.3	33.3	58.3
Any 2 criteria met*			91	4	13	108	35.3	84.3	3.7	12.0
Any 3-5 criteria met**			44	0	1	45	14.7	97.8	0.0	2.2

* 12 of the 108 with 2 criteria met had a code in categories E, F, or G; 4 of the 12 were confirmed and 3 were contradicted

** 14 of the 45 with 3-5 criteria met had a code in categories E, F, or G; all 14 had at least 2 codes in categories A-D and all were confirmed.