Reliability of Medical History Reporting in Older Adults With and Without Cognitive Impairment

Journal of Central Nervous System Disease
Volume 11: 1–9

DOI: 10.1177/1179573519843874

Nicholas Curcio[1], Kristin Wilmoth[1], Christian LoBue[1,3] and C Munro Cullum[1,2,3,4]

[1]Department of Psychiatry, The University of Texas Southwestern Medical Center, Dallas, TX, USA. [2]Department of Neurology and Neurotherapeutics, The University of Texas Southwestern Medical Center, Dallas, TX, USA. [3]Department of Neurological Surgery, The University of Texas Southwestern Medical Center, Dallas, TX, USA. [4]Peter O'Donnell Jr. Brain Institute, The University of Texas Southwestern Medical Center, Dallas, TX, USA.

ABSTRACT

BACKGROUND: Clinical diagnosis of cognitive disorders depends on accurate reporting of medical history, yet little is known about the reliability and the validity of such reports, particularly in older patients with and without cognitive impairment.

METHODS: In 2 studies, we examined the reliability and the validity of reported histories of select medical events in adults with and without cognitive impairment from a large national cohort.

RESULTS: Information from subjects ($N_1 = 3664$), obtained from 2 time points, 6 to 12 months apart, was consistent across most medical events, regardless of the diagnostic group (range = 97.6%-100% agreement; Cohen κ range = 0.712-0.945), with few exceptions. Validity analyses ($N_2 = 382$) revealed that 3 of 5 medical events assessed showed substantial agreement between self-report information and clinician diagnosis.

CONCLUSIONS: These data represent some of the first to demonstrate the reliability and the validity of reported select medical events in older adults with and without cognitive impairment.

KEYWORDS: Cognitive impairment, Alzheimer disease, self-report, reliability, validity medical events

RECEIVED: May 11, 2018. **ACCEPTED:** March 21, 2019.

TYPE: Original Research

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: Funding for this study was provided in part by the Texas Institute for Brain Injury and Repair at the University of Texas (UT) Southwestern Medical Center. Additional support came from the UT Southwestern Alzheimer's Disease Center's National Institute of Aging (NIA) grant, P3012300-19, and the Texas Alzheimer's Research and Care Consortium. The National Alzheimer's Coordinating Center (NACC) database is funded by NIA/National Institutes of Health grant U01 AG016976.The NACC data are contributed by the NIA-funded Alzheimer Disease Centers: P30 AG019610 (Principal Investigator [PI] Eric Reiman, MD), P30 AG013846 (PI Neil Kowall, MD), P50 AG008702 (PI Scott Small, MD), P50 AG025688 (PI Allan Levey, MD, PhD), P50 AG047266 (PI Todd Golde, MD, PhD), P30 AG010133 (PI Andrew Saykin, PsyD), P50 AG005146 (PI Marilyn Albert, PhD), P50 AG005134 (PI Bradley Hyman, MD, PhD), P50 AG016574 (PI Ronald Petersen, MD, PhD), P50 AG005138 (PI Mary Sano, PhD), P30 AG008051 (PI Thomas Wisniewski, MD), P30 AG013854 (PI M. Marsel Mesulam, MD), P30 AG008017 (PI Jeffrey Kaye, MD), P30 AG010161 (PI David Bennett, MD), P50 AG047366 (PI Victor Henderson, MD, MS), P30 AG010129 (PI Charles DeCarli, MD), P50 AG016573 (PI Frank LaFerla, PhD), P50 AG005131 (PI James Brewer, MD, PhD), P50 AG023501 (PI Bruce Miller, MD), P30 AG035982 (PI Russell Swerdlow, MD), P30 AG028383 (PI Linda Van Eldik, PhD), P30 AG053760 (PI Henry Paulson, MD, PhD), P30 AG010124 (PI John Trojanowski, MD, PhD), P50 AG005133 (PI Oscar Lopez, MD), P50 AG005142 (PI Helena Chui, MD), P30 AG012300 (PI Roger Rosenberg, MD), P30 AG049638 (PI Suzanne Craft, PhD), P50 AG005136 (PI Thomas Grabowski, MD), P50 AG033514 (PI Sanjay Asthana, MD, FRCP), P50 AG005681 (PI John Morris, MD), P50 AG047270 (PI Stephen Strittmatter, MD, PhD).

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: C Munro Cullum, Departments of Psychiatry, The University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75390, USA. Email: Munro.Cullum@UTSouthwestern.edu

Introduction

Obtaining medical history information often relies on self-report using interview and/or questionnaire methods. In older patients with suspected or known cognitive impairment, the accuracy of such information may be called into question. This is a particularly important issue, as the likelihood of cognitive impairment and medical diseases both increase dramatically with age.[1,2] In addition to various types of dementia, memory impairment can also result from neurological and cardiac events, such as traumatic brain injury (TBI),[3–6] stroke,[7] seizures,[8] heart failure,[9] cardiac arrest,[10] and atrial fibrillation.[11] Thus, it is important in any clinical evaluation to obtain an accurate report of a patient's medical history, as this can be essential to identify the nature and course of cognitive impairment, as well as providing insight into a patient's prognosis and possible treatment recommendations.

Despite the importance of accurate medical history reporting, little is known about its reliability and validity, particularly among older and/or cognitively impaired individuals. Research by Okura et al[12] investigated the agreement of patients' self-reported medical conditions compared with that found in medical records within a large cognitively intact, diverse sample of community-dwelling adults aged 45 years and older (N = 2037). The authors found self-reports of congestive heart failure (κ = 0.46) were considerably less consistent than other vascular and neurological events (κs = 0.75-0.80). In another investigation of 2380 community-dwelling elderly patients, Kriegsman et al[13] found that such patients provided fairly accurate reports for many chronic diseases documented in their medical records, such as cardiac disease and diabetes (κs = 0.69 and 0.85, respectively), yet provided inaccurate reports of some other diseases such as atherosclerosis and arthritis (κs = 0.38 and 0.31,

respectively). Wilmoth et al[14] demonstrated that subjects aged 50 years and older in a large national database reliably self-reported TBIs across 2 time points and that subjects' level of cognitive impairment did not predict reporting consistency.

Taken together, these findings suggest that some reports of medical events may be unreliable, even in cognitively intact individuals, and very little is known about medical history reporting among older individuals with varying degrees of cognitive impairment. Although prior studies have focused mainly on medical history reporting in healthy adult samples, little research has been conducted on (a) the stability of reporting across time and (b) the concordance of individual and provider/informant reports in patients with cognitive impairment. The purpose of this investigation was to examine the reliability and the validity of various reported medical history events in individuals with normal cognition (NC), mild cognitive impairment (MCI), and Alzheimer disease (AD) using a large national database.

Study 1

Methods

Data were obtained from the National Alzheimer's Coordinating Center (NACC) Uniform Data Set (UDS) form versions 1 and 2, which has compiled sociodemographic and clinical information from subjects volunteering for research at all NIA-funded Alzheimer Disease Centers (ADCs) in the United States since September 2005.[15] Written informed consents were obtained from participants at each ADC and approved by the ADC's Institutional Review Board (IRB). Research using the NACC database was approved by the University of Washington (UW) IRB (36178). The University of Washington IRB does not require tracking of external IRB numbers because NACC only distributes data sets that are Health Insurance Portability and Accountability Act limited, at a minimum. The maintenance of the NACC is covered by the UW IRB and is limited to the Coordinating Center itself. The individual ADCs that collect the human subjects data have their own approvals that they maintain.

Data received from NACC are either completely deidentified or contain limited data elements that would not allow the researcher to readily identify individuals in the data set. The maintenance of the (potentially) identifiable data in the NACC is covered by the UW IRB approval, and the secondary use of deidentified data would likely not require an IRB approval from external institutions as it is not identifiable to researchers (ie, it does not fit the definition of "human subjects research" per regulations). The NACC database is funded by NIA/National Institutes of Health grant U01 AG016976.

This analysis used data from 33 different ADCs. Sociodemographic and clinical data were obtained for initial and follow-up visits completed between September 2005 and June 2015. Subjects were included if they were aged 50 years or older, completed an initial and follow-up visit within 6 to 12 months, and were classified as being an NC or having a clinical diagnosis of MCI or possible/probable AD at both time points. Clinical diagnoses were made by experienced clinicians at each UDS visit using standard Petersen diagnostic criteria for MCI[16] and National Institute of Neurological and Communicative Disorders and Stroke (NINCDS)/Alzheimer's Disease and Related Disorders Association (ADRDA) criteria for AD.[17]

As part of the standard UDS, subjects and accompanying informants are asked at each ADC visit whether they have experienced different neurological and cardiac events, among other medical history questions. Each medical event question was recorded as having never occurred ("absent"), having occurred within the past year or still requiring active management that is consistent with available information ("recent/active"), or having occurred more than a year prior to the visit but was resolved with no current treatment underway ("remote/inactive"). As our intent was to focus on remote reporting of select medical history events, and because events occurring during the follow-up period would adversely impact test-retest reliability, individuals reporting "recent/active" events or those having follow-up visits after 12 months were excluded. Subjects with missing data were also excluded.

We chose to focus on neurological and cardiac events because of their potential impact on cognitive functioning. The neurological events examined included stroke, transient ischemia attack, seizure, and TBI with either <5 minutes loss of consciousness (LOC) or $\geqslant 5$ minutes LOC. Cardiac events included were heart attack/cardiac arrest, atrial fibrillation, angioplasty/endarterectomy/stent implantation, cardiac bypass procedure, pacemaker implantation, and congestive heart failure. Other medical diagnoses such as hypertension and diabetes were excluded due to the ongoing nature of medical treatment and medication adherence, which may confound remote history reporting.

Analyses. Descriptive and frequency analyses were conducted for age, sex, education, and race (see Table 1). Chi-square and 1-way analyses of variance (ANOVAs) examined group differences in demographic factors with the significance level set at .05. To characterize the overall functional status of the sample, the Clinical Dementia Rating Scale Sum of Boxes (CDR-SOB) scores for each participant were included and averaged across diagnostic groups.

Self-report reliability was examined using Cohen κ for each neurological and cardiac event. Within each cognitive group, inconsistencies in reporting were defined as participants who endorsed the medical event at time 1 but not time 2 or vice versa (ie, yes/no or no/yes). Confidence intervals (95.00% CIs) are provided for each κ coefficient, with the lower CI providing the reader with an estimate of the lowest estimate that might be found with similar samples. Kappa coefficients greater than or equal to 0.60, which also contain a lower bound value of the 95% CI greater than or equal to 0.60, were deemed to have met

Table 1. Demographic characteristics and dementia rating for study 1 normal cognition, MCI, and AD subjects.

	NORMAL COGNITION, N = 1765	MCI, N = 627	AD, N = 1272	*P*
Age, mean (SD)	72.6 (9.0)	73.8 (8.6)	74.6 (9.6)	<.001[a]
Education, mean (SD)	15.6 (2.9)	15.1 (3.2)	14.2 (3.5)	<.001[a]
CDR-SOB score, mean (SD)	0.1 (0.3)	1.3 (1.0)	6.2 (3.93)	—
White, %	83.0	81.3	85.1	<.001[b]
Gender: n Female (%)	1153 (65.3)	296 (47.2)	703 (55.3)	<.001[b]

Abbreviation: MCI, mild cognitive impairment; AD, Alzheimer disease; CDR-SOB, Clinical Dementia Rating Scale Sum of Boxes.
[a]Analysis of variance; [b]Pearson χ^2.

Table 2. Demographic characteristics and dementia rating for study 2 normal cognition, MCI, and AD Subjects.

	NORMAL COGNITION, N = 243	MCI, N = 53	AD, N = 86	*P*
Age, mean (SD)	70.3 (9.1)	72.3 (7.6)	70.2 (9.6)	>.05[a]
Education, mean (SD)	16.4 (2.6)	16.5 (2.7)	14.9 (3.8)	<.001[a]
CDR-SOB score, mean (SD)	0.0 (0.26)	1.6 (1.2)	5.4 (3.8)	—
White, %	74.1	81.1	81.4	<.05[b]
Gender: n Female (%)	153 (63.0)	24 (45.3)	40 (46.5)	<.05[b]

Abbreviation: MCI, mild cognitive impairment; AD, Alzheimer disease; CDR-SOB, Clinical Dementia Rating Scale Sum of Boxes.
[a]Analysis of variance; [b]Pearson χ^2.

criteria for substantial agreement.[18] Analyses were conducted using IBM SPSS Statistics Package version 22.

Results

Our initial query from the NACC UDS yielded data from 31 792 participants. After applying the exclusion criteria, there were 1765 subjects identified as NC, 627 with MCI, and 1272 with AD included in the analyses (see Table 2). Diagnostic groups differed significantly by sex and race, $\chi^2 = 72.31$ and 21.94, respectively ($P < .001$), and they also differed with respect to level of education and age, $F(2, 3600) = 69.05$ and $F(2, 3611) = 17.51$, $Ps < .001$. A Tukey post hoc test revealed that the NC group was significantly younger than the MCI and AD groups, but the MCI and AD groups did not differ significantly for age ($P > .05$). Consistent with expectations, participants with AD displayed the greatest functional deficits as reflected by CDR sum of boxes ratings, while still considered relatively mild for AD, and the MCI group displayed greater functional deficits than the NC group ($mean_{CDR\text{-}SOB} = 6.2$, 1.3, and 0.1, respectively).

Data were examined across the initial and follow-up visits to examine consistency in medical history reporting. Cohen κ analyses revealed significant agreement for all medical events for each diagnostic group, with few exceptions, $Ps < .001$. The κ coefficients and CIs pertaining to each neurological and cardiac event across diagnostic groups, the relative percentage agreement between baseline and follow-up visits, and the number of participants who reported remote events at both time points 1 and 2 for each event can be found in Tables 3 and 4.

All κ coefficient CIs met criteria for substantial agreement (lower bound κ CI ⩾ 0.60),[18] with 3 exceptions: the seizure condition for both the NC and the AD groups, and the congestive heart failure condition for the NC group. Generally, a more robust agreement was seen within cardiac events overall compared with neurological events. Among neurological events, TBI with less than 5 minutes LOC showed the lowest observed percentage agreements across all cognition groups, and TBI with ⩾5 minutes LOC revealed the highest percentage agreement within the AD group. About 60.10% of medical events revealed a higher number of reporters at follow-up than baseline.

Discussion

Older individuals and assisting caregivers appear to be reliable in reporting of most of the remote neurological and cardiac events queried, regardless of the presence or the absence of cognitive impairment. However, there was a trend across all conditions for a greater number of participants having reported a remote medical event at the follow-up visit than baseline. A closer examination of participants' reports revealed that no participant within any diagnostic group reported the event at baseline but then denied the event at time point 2 (yes/no) for

Table 3. Test-retest reliability of neurological event reporting.

	NORMAL COGNITION	MCI	AD
Stroke, n	1762	625	1266
Consistent 1 and 2, n (%)	1758 (99.77)	618 (98.88)	1257 (99.29)
Reported at time 1	26	34	54
Reported at time 2	30	41	63
Inconsistent reporters	4	7	9
κ (CI)	0.93 (0.86-1.00)	0.90 (0.82-0.97)	0.92 (0.87-0.97)
TIA, n	1752	624	1260
Consistent 1 and 2, n (%)	1748 (99.77)	620 (99.36)	1253 (99.44)
Reported at time 1	43	20	43
Reported at time 2	47	24	50
Inconsistent reporters	4	4	7
κ (CI)	0.95 (0.91-1.00)	0.91 (0.81-1.00)	0.92 (0.87-0.98)
Seizure, n	1760	626	1263
Consistent 1 and 2, n (%)	1752 (99.55)	624 (99.68)	1256 (99.44)
Reported at time 1	14	12	12
Reported at time 2	14	12	17
Inconsistent reporters	8	2	7
κ (CI)	0.71 (0.52-0.90)	0.92 (0.80-1.03)	0.76 (0.58-0.93)
TBI LOC <5 min, n	1747	621	1258
Consistent 1 and 2, n (%)	1716 (98.22)	607 (97.75)	1228 (97.62)
Reported at time 1	97	46	85
Reported at time 2	102	46	77
Inconsistent reporters	31	14	30
κ (CI)	0.84 (0.79-0.89)	0.84 (0.75-0.92)	0.80 (0.73-0.87)
TBI LOC ≥ 5 min, n	1751	623	1261
Consistent 1 and 2, n (%)	1740 (99.37)	616 (98.88)	1257 (99.68)
Reported at time 1	39	27	27
Reported at time 2	40	24	27
Inconsistent reporters	11	7	4
κ (CI)	0.86 (0.78-0.94)	0.86 (0.75-0.96)	0.92 (0.85-1.00)

Abbreviation: MCI, mild cognitive impairment; AD, Alzheimer disease; TIA, transient ischemic attack; TBI, traumatic brain injury; LOC, loss of consciousness; CI, confidence interval.
*All *Ps* $<.001$.

stroke, transient ischemic attack, or any of the 6 cardiac events. For participants who reported such events at baseline, 100% reliability in reporting was observed at follow-up.

Within each diagnostic group, remote history of a seizure(s) and TBI yielded several yes/no and no/yes inconsistencies. The observed contrast in yes/no inconsistencies for the seizure(s) event could be due to the large variability in presentation and subjective experience and reporting of seizure phenomena, which can overlap with other conditions (eg, migraines and psychiatric disorders). Participants may have been more reliable

Table 4. Test-retest reliability of cardiac event reporting.

	NORMAL COGNITION	MCI	AD
Heart attack/cardiac arrest, n	1759	625	1270
Consistent 1 and 2, n (%)	1758 (99.94)	622 (99.52)	1266 (99.69)
Reported at time 1	55	36	50
Reported at time 2	56	39	54
Inconsistent reporters	1	3	4
κ (CI)	0.99 (0.97-1.01)	0.96 (0.91-1.00)	0.96 (0.92-1.00)
Atrial fibrillation, n	1754	624	1268
Consistent 1 and 2, n (%)	1750 (99.77)	622 (99.68)	1267 (99.92)
Reported at time 1	24	11	16
Reported at time 2	28	13	17
Inconsistent reporters	4	2	1
κ (CI)	0.92 (0.85-1.00)	0.92 (0.80-1.03)	0.97 (0.91-1.03)
Angioplasty, n	1761	626	1271
Consistent 1 and 2, n (%)	1747 (99.20)	623 (99.52)	1266 (99.60)
Reported at time 1	56	31	42
Reported at time 2	70	34	47
Inconsistent reporters	14	3	5
κ (CI)	0.89 (0.83-0.94)	0.95 (0.90-1.01)	0.94 (0.89-1.00)
Cardiac bypass, n	1761	626	1271
Consistent 1 and 2, n (%)	1759 (99.89)	625 (99.84)	1269 (96.46)
Reported at time 1	52	30	45
Reported at time 2	54	31	47
Inconsistent reporters	2	1	2
Kappa	0.98 (0.95-1.01)	0.98 (0.95-1.02)	0.98 (0.95-1.01)
Pacemaker, n	1762	627	1271
Consistent 1 and 2, n (%)	1762 (100.00)	627 (100.00)	1271 (100.00)
Reported at time 1	2	1	4
Reported at time 2	2	1	4
Inconsistent reporters	0	0	0
κ (CI)	1.00 (1.00-1.00)	1.00 (1.00-1.00)	1.00 (1.00-1.00)
Congestive heart failure, n	1760	627	1271
Consistent 1 and 2, n (%)	1757 (99.83)	627 (100.00)	1271 (100.00)
Reported at time 1	6	5	6
Reported at time 2	9	5	6
Inconsistent reporters	3	0	0
κ (CI)	0.80 (0.58-1.02)	1.00 (1.00-1.00)	1.00 (1.00-1.00)

Abbreviation: MCI, mild cognitive impairment; AD, Alzheimer disease; CI, confidence interval.
*All *Ps* < .001.

to report seizure(s) if they required hospitalization, chronic medication, and/or if it was subsequently verified by a medical professional.

Regarding TBI, differences in reporting at both time points could be due to a variety of reasons. Determining the length of LOC after injury, which may have occurred years or even decades earlier, can prove difficult, particularly if post-traumatic amnesia was present. In addition, a witness to the injury, which caused the TBI, is often necessary to estimate the duration of LOC. Recent findings within the literature suggest poor validity of self/parent-reported TBIs that have been verified via medical records. Mckinlay et al[19] recently demonstrated in a longitudinal study that even when young adults were carefully cued to recall remote documented instances of TBI, many instances of TBI, including those severe enough to require hospitalization, were not accurately recalled. Interestingly, the agreement among reporting for all TBI events was similar across the 3 groups of normal, MCI, and AD individuals, suggesting that the presence of cognitive impairment does not substantially affect such reporting in the NACC population.

Within the present study, both TBI events revealed relatively strong reliability across diagnostic groups, with κ coefficients ranging from 0.802 to 0.836 for TBI < 5 minutes LOC and 0.857 to 0.924 for TBI ⩾ 5 minuets LOC. However, a larger number of inconsistent responders were found within the TBI < 5 minutes LOC group. These findings are consistent with those by Wilmoth et al.[14] Given the variability of symptom reporting among milder concussions, it should not come as a surprise that there may be greater variability in event reporting than in more serious injuries, such as those requiring hospitalization. It is unclear as to why our agreement rates were much higher, although this may relate to methodological differences in subject characteristics, data collection, and analysis. In addition, in the NACC procedure for medical history reporting, information is obtained from patients presenting for evaluation, sometimes with an accompanying informant. As such, a study partner may have assisted with medical history reporting. Although the presence of accompanying caregivers more closely mimics clinical scenarios in which cognitively impaired patients are evaluated, the inability to know whether or not the informants provided information in individual cases may limit the generalizability of these findings.

Conversely, the implantation of a pacemaker was the only event to demonstrate 100% agreement across each diagnostic group. This is not surprising given the nature of the surgery, as people do not seem to forget this procedure, although the UDS does not offer information regarding when any of these remote events occurred, which could affect reporting reliability. Interestingly, the congestive heart failure and angioplasty conditions yielded stronger agreement rates for those classified as having MCI and AD compared with the NC group. Of note, the observed strength of agreement observed within the congestive heart failure condition differs from the results observed by Okura et al.[12] The reasons for these observed trends are unclear.

Study 2

Methods

In a separate sample, initial visit data were collected for participants who completed form version 3 of the UDS, the most recently updated protocol used within NACC ADCs, up until the June 2017 data freeze. This analysis used data from 23 different ADCs. Form version 3 of the NACC UDS, which was implemented in 2015, introduced a section pertaining to clinician-verified medical conditions while removing the assessment of participant self-reported medical histories at each follow-up visit. As such, to assess the validity of medical history reporting within this population, a separate sample from the one used in study 1 was necessary. Subjects were included if they were aged 50 years or older, completed an initial visit using form version 3 of the UDS, and were classified as having NC, MCI, or AD.

The presence of medical conditions that were assessed under similar conditions were included in analyses as follows: myocardial infarction, atrial fibrillation, angioplasty, congestive heart failure, and stroke. As fewer identical conditions were assessed across both clinician-assessed and patient health history sections on the UDS form version 3, a smaller list of medical events were included in study 2 compared with study 1. Only medical events that similarly assessed individual and clinician report were included in analyses. Chi-square analyses examined the concordance between self-report and clinician report among diagnostic groups (NC, MCI, and AD). Clinician-assessed medical conditions were based on a review of all available information, including new diagnoses made during the current ADC visit, previous medical records, procedures, laboratory tests, and the clinical examination.

Descriptive and frequency analyses were conducted for age, sex, education, and race (see Table 2). Chi-square and 1-way ANOVAs examined group differences in demographic factors with the significance level set at .05. To characterize the overall functional status of the sample, the Clinical Dementia Rating Scale Sum of Boxes (CDR-SOB) scores for each participant were included and averaged across diagnostic groups. Validity was assessed by agreement between self-report and clinician documentation of conditions using Cohen κ. Confidence intervals (95.00%) are provided for each κ coefficient, with the lower CI providing the reader with an estimate of the lowest estimate that might be found with similar samples. Kappa coefficients greater than or equal to 0.60 that also contain a lower bound value of the 95% CI greater than or equal to 0.60 were deemed to have met criteria for substantial agreement.[18]

Results

Our second query from the NACC UDS yielded 382 subjects after applying the exclusion criteria. Diagnostic groups differed significantly by sex and race, $\chi^2 = 10.34$ and 26.10, respectively ($Ps < .05$), as well as level of education, $F(2, 3600) = 69.05$,

Table 5. Concordance of reported and clinician-assessed medical histories.

Myocardial infarction, n	312
Consistent 1 and 2, n (%)	310 (99.36)
Only patient reported	2
Only Clinician reported	0
κ (CI)	0.83 (0.60 to 1.06)
Atrial fibrillation, n	312
Consistent 1 and 2, n (%)	309 (99.03)
Only patient reported	1
Only clinician reported	2
κ (CI)	0.92 (0.83 to 1.01)
Angioplasty, n	313
Consistent 1 and 2, n (%)	307 (98.08)
Only patient reported	6
Only clinician reported	0
κ (CI)	0.66 (0.40 to 0.91)
Congestive heart failure, n	312
Consistent 1 and 2, n (%)	312 (100.00)
Only patient reported	0
Only clinician reported	0
κ (CI)	1.00 (1.00 to 1.00)
Stroke, n	314
Consistent 1 and 2, n (%)	310 (98.73)
Only patient reported	3
Only clinician reported	0
κ (CI)	0.40 (−0.10 to 0.94)

Abbreviation: CI, confidence interval.
*All Ps $< .001$.

$P < .001$, but not age $F(2, 379) = 1.10$, $P = .38$. Again, as expected, participants with AD displayed the greatest functional deficits as reflected by CDR sum of boxes scores, and the MCI group displayed higher scores than the NC group ($mean_{CDR\text{-}SOB} = 5.4$, 1.6, and 0.0, respectively).

Cohen κ analyses revealed significant agreement (lower bound κ CI ⩾ 0.60) for the following conditions: myocardial infarction, atrial fibrillation, and congestive heart failure. Stroke (κ CI = −0.10 to 0.94) and angioplasty (κ CI = 0.40-0.91) displayed poor rates of agreement (see Table 5). Diagnostic groups did not differ significantly by the proportion of inconsistent responders, $\chi^2(2, N = 382) = 1.77$, $P > .05$.

Discussion

Although significant concordance was observed in many of the medical conditions assessed by clinicians compared with patient report, no clear distinction was observed between the type of events that elicited a high level of agreement (eg, myocardial infarction) compared with events that elicited a lower level of agreement (eg, stroke). Whereas past research has observed differences in chronic disease reporting,[13] our results were limited to select medical events. Investigating the differences in reporting between chronic and acute conditions may provide a fruitful area of future research. Our results regarding myocardial infarction (κ = 0.83) are similar to those observed by Okura et al[12] (κ = 0.80). A possible explanation for the observed results within the stroke condition could be the manner of diagnosis. For example, if clinician(s) observed signs of a remote lacunar infarct on magnetic resonance imaging (MRI), it is possible the patient may not have noticed or reported clinical symptoms, although the MRI was later read as reflecting a stroke, which is not uncommon in clinical practice.

The medical literacy of individuals endorsing medical conditions found in their charts may also have influenced results. For example, patients may deny a previous myocardial infarction but endorse a "heart attack." The way information is conveyed to patients, including the specific terms used, may influence their understanding of their medical conditions.

Summary

In summary, these findings are some of the first to demonstrate the reliability and the validity of medical event history reporting in older adults with and without cognitive impairment, including those with mild AD. Given the necessity for valid medical history reporting and the extent to which health care providers and researchers rely on self-report information, these findings provide cautionary support for the ongoing use and inclusion of self and assisting caregiver report information within these populations.

Our results identified that for several different neurological and cardiac events, there was a high level of consistency in reporting between baseline and follow-up visits occurring 6 to 12 months later. Similar high rates of reporting agreement were found by Okura et al[12] in their sample of 2037 older Minnesota residents, noting that rates of agreement were highest for life-threatening events with a rapid onset (eg, stroke and myocardial infarction), as well as in chronic disorders which required ongoing care (eg, diabetes). From a psychological standpoint, it makes sense that potentially life-threatening or life-altering events would be remembered more vividly, and ongoing medical conditions would be reported with good consistency, as they were present at both time points. Combining results from both studies highlights the reliability of reporting major medical event information by older individuals/informants. In addition to supporting the results from the previous studies of healthy

older samples, our results extend these findings to aging individuals with different levels of cognitive impairment.

More complex and/or heterogeneous conditions appear to be reported less consistently and accurately based on these results. This observation may be due to numerous factors, such as the transient nature of such events, less severe nature, and/or the presence of overlapping symptoms with other conditions. Although TBI and other events yielded generally strong reliability, agreement between self-report and medical records could not be assessed due to the limitations of the available database. Along these lines, Mckinlay et al[19] found large inconsistencies regarding agreement between self-reported and medical record reports of TBI in a sample of 1265 individuals age 25 years and younger. Of note, most of the TBIs sustained in the study by Mckinlay were considered "mild," which likely contributed to lower rates of agreement, as definitions and diagnoses of mild TBI vary across settings. Within our study, we were limited to specifiers within the NACC database such as estimated duration of LOC (ie, < or >5 minutes), which reveal little information regarding the severity of the injury. Furthermore, no information was available regarding subjects' age at the time of TBI or whether repeated injuries were sustained. As such, it may be important to take into account the nature and the severity of the injury/event when gathering self-report information. Clinicians would be well advised to use a variety of collateral information to guide treatment decisions in cases involving such events. These findings also point to the need for continued research within these populations, as obtaining accurate clinical history is not only important for understanding the progression and treatment of cognitive impairment, but for treating the patient in general.

These studies have several limitations that may limit generalizability. First, all participants were seen at ADCs and thus are involved in volunteer research programs. As such, these individuals may not be representative of the general population and may have varying levels of medical literacy, and replication in community-based samples is needed. As previously noted, many ADC participants present with accompanying informants to complete the NACC UDS questionnaires and interview questions. Although this scenario more mimics the clinical setting wherein cognitively impaired patients often present with an informant or caregiver, we were limited to the acquired data from NACC, which do not indicate whether patient and/or caregiver provided the medical history information or offered confirmation of its veracity. It may be that healthy participants and those with MCI would tend to rely less on informants for medical history information/validation than those with dementia, although this remains a question that goes beyond the available data. Future studies may wish to examine the reliability of self-report data separately from patients and caregivers to provide additional measures of reporting reliability based on the source of information. Nevertheless, despite this methodological limitation, it is worth noting that similar high rates of reporting reliability were seen across these large samples of healthy aging subjects and those with MCI and dementia who were systematically and thoroughly evaluated and diagnosed using standard procedures across multiple sites nationally. Furthermore, in both studies, the samples were primarily white and highly educated. Last, our AD samples were only mildly impaired ($mean_{CDR\text{-}SOB1} = 6.2$, SD = 3.93; $mean_{CDR\text{-}SOB2} = 5.4$, SD = 3.8), as CDR-SOB scores ranging from 4.5 to 9.0 are considered "mild."[20] Thus, our findings may not be reflective of patients with more severe cognitive impairment, as medical history reporting may become less reliable with greater deficits, which merits further investigation. Last, our findings reflect the reliability and the validity of reporting only certain neurological and cardiac events/conditions, which may differ from reporting of other common diseases such as hypertension and diabetes.

Conclusions

Two studies examined the consistency of reported histories of select medical events in older adults with and without cognitive impairment from a large national cohort. Results demonstrated that retrospective recall of neurological and cardiac events appears to be reliable, even in those with cognitive impairment due to MCI and mild AD. Results also yielded substantial agreement between select clinician and patient-reported medical event histories. These data represent some of the first to document the reliability of reported medical histories in older adults with and without cognitive impairment.

Author Contributions

All authors were involved in the design, literature review, data acquisition, analysis, and interpretation for both studies involved in this manuscript. Additionally, all authors participated in drafting and revising the manuscript. Final approval was obtained from each author prior to journal submission.

ORCID iDs

Nicholas Curcio https://orcid.org/0000-0002-6299-6611
Kristin Wilmoth https://orcid.org/0000-0003-4192-8756

REFERENCES

1. Graham JE, Rockwood K, Beattie BL, et al. Prevalence and severity of cognitive impairment with and without dementia in an elderly population. *Lancet*. 1997;349:1793–1796. doi:10.1016/S0140-6736(97)01007-6.
2. Doraiswamy PM, Leon J, Cummings JL, Marin D, Neumann PJ. Prevalence and impact of medical comorbidity in Alzheimer's disease. *J Gerontol A Biol Sci Med Sci*. 2002;57:M173–M177.
3. Lye TC, Shores EA. Traumatic brain injury as a risk factor for Alzheimer's disease: a review. *Neuropsychol Rev*. 2000;10:115–129. doi:10.1016/j.neubiorev.2012.02.013.
4. Fleminger S, Oliver DL, Lovestone S, Rabe-Hesketh S, Giora A. Head injury as a risk factor for Alzheimer's disease: the evidence 10 years on; a partial replication. *J Neurol Neurosurg Psychiatry*. 2003;74:857–862.
5. Guo Z, Cupples LA, Kurz A, et al. Head injury and the risk of AD in the MIRAGE study. *Neurology*. 2000;54:1316–1323. doi:10.1212/WNL.54.6.1316.

6. Plassman BL, Havlik RJ, Steffens DC, et al. Documented head injury in early adulthood and risk of Alzheimer's disease and other dementias. *Neurology*. 2000;55:1158–1166.
7. Patel MD, Coshall C, Rudd AG, Wolfe CDA. Cognitive impairment after stroke: clinical determinants and its associations with long-term stroke outcomes. *J Am Geriatr Soc*. 2002;50:700–706. doi:10.1046/j.1532-5415.2002.50165.x.
8. Aldenkamp AP. Effect of seizures and epileptiform discharges on cognitive function. *Epilepsia*. 1997;38:S52–S55. doi:10.1111/j.1528-1157.1997.tb04520.x.
9. Sloan RS, Pressler SJ. Cognitive deficits in heart failure: re-cognition of vulnerability as a strange new world. *J Cardiovasc Nurs*. 2009;24:241–248. doi:10.1097/JCN.0b013e3181a00284.
10. Grubb NR, O'Carroll R, Cobbe SM, Sirel J, Fox KA. Chronic memory impairment after cardiac arrest outside hospital. *BMJ*. 1996;313:143–146. doi:10.1136/bmj.313.7050.143.
11. Knecht S, Oelschläge C, Duning T, et al. Atrial fibrillation in stroke-free patients is associated with memory impairment and hippocampal atrophy. *Eur Heart J*. 2008;29:2125–2132. doi:10.1093/eurheartj/ehn341.
12. Okura Y, Urban LH, Mahoney DW, Jacobsen SJ, Rodeheffer RJ. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. *J Clin Epidemiol*. 2004;57:1096–1103. doi:10.1016/j.jclinepi.2004.04.005.
13. Kriegsman DM, Penninx BW, van Eijk JT, Boeke AJ, Deeg DJ. Self-reports and general practitioner information on the presence of chronic diseases in community-dwelling elderly—a study on the accuracy of patients self-reports and on determinants of inaccuracy. *J Clin Epidemiol*. 1996;49:1407–1417. doi:10.1016/s0895-4356(96)00274-0.
14. Wilmoth K, LoBue C, Clem MA, et al. Consistency of traumatic brain injury reporting in older adults with and without cognitive impairment. *Clin Neuropsychol*. 2017; 32:524–529. doi:10.1080/13854046.2017.1378371.
15. Morris JC, Weintraub S, Chui HC, et al. The Uniform Data Set (UDS): clinical and cognitive variables and descriptive data from Alzheimer disease centers. *Alzheimer Dis Assoc Disord*. 2006;20:210–216. doi:10.1097/01.wad.0000213865.09806.92.
16. Winblad B, Palmer K, Kivipelto M, et al. Mild cognitive impairment – beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *J Intern Med*. 2004;256:240–246.
17. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*. 1984;34:939–944. doi:10.1212/WNL.34.7.939.
18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.
19. Mckinlay A, Horwood LJ, Fergusson DM. Accuracy of self-report as a method of screening for lifetime occurrence of traumatic brain injury events that resulted in hospitalization. *J Int Neuropsychol Soc*. 2017;22:717–723. doi:10.1017/S1355617716000497.
20. O'Bryant SE, Waring SC, Cullum CM, et al. Staging dementia using clinical dementia rating scale sum of boxes scores. *Popul (English Ed.)*. 2011;65:1091–1095. doi:10.1001/archneur.65.8.1091.