Published in final edited form as:

J Am Med Dir Assoc. 2017 December 01; 18(12): 1010–1018.e1. doi:10.1016/j.jamda.2017.07.014.

Consensus Approaches to Identify Incident Dementia in Cohort Studies: Systematic Review and Approach in the SAGES Study

Eyal Y. Kimchi, MD, PhD^{a,*}, Tammy T. Hsieh, MD, MPH^{b,*}, Ray Guo, MA^c, Bonnie Wong, PhD^a, Margaret O'Connor, PhD^d, Edward R. Marcantonio, MD, SM^e, Eran Metzger, MD^f, Jason Strauss, MD^g, Steven E. Arnold, MD^a, Sharon K. Inouye, MD, MPH^{c,e,†}, and Tamara G. Fong, MD, PhD^{c,d,†}

^aDepartment of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA

^bDepartment of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA

^cAging Brain Center, Institute for Aging Research, Hebrew SeniorLife, Boston, MA

^dDepartment of Neurology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

^eDepartment of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

^fDepartment of Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

^gDepartment of Psychiatry, Cambridge Health Alliance, Harvard Medical School, Boston, MA

Abstract

OBJECTIVES—To survey the current methods used to ascertain dementia and MCI in longitudinal cohort studies, in order to categorize differences in approaches and to identify key components of expert panel methodology in current use.

METHODS—We searched PubMed for the past 10 years, from March 6, 2007–March 6, 2017 using a combination of controlled vocabulary and keyword terms to identify expert panel consensus methods used to diagnose MCI or dementia in large cohort studies written in English. From these results, we identified a framework for reporting standards, and describe as an exemplar the clinical consensus procedure used in an ongoing study of elective surgery patients (the Successful AGing after Elective Surgery study).

RESULTS—Thirty-one articles representing unique cohorts were included. Among published methods, membership of experts panel varied significantly. There was more similarity in what

Corresponding Author/Request for reprints: Tamara G. Fong, MD, PhD, Aging Brain Center, Institute for Aging Research, Hebrew SeniorLife, 1200 Centre Street, Boston, MA 02135, tfong@bidmc.harvard.edu, 617-667-4074.

*contributed equally as first-authors

†contributed equally as co-senior authors

Conflict of Interest/Disclosure Statement

The authors have no conflict of interest to report.

types of information was used to ascertain disease status. However, information describing the diagnostic decision process and resolution of disagreements was often lacking.

CONCLUSION—Methods used for expert panel diagnosis of MCI and dementia in large cohort studies are widely variable, and there is a need for more standardized reporting of these approaches. By describing the procedure in which our expert panel achieved consensus diagnoses, we hope to encourage the development and publication of well-founded and reproducible methods for diagnosis of MCI and dementia in longitudinal studies.

Keywords

consensus; dementia; diagnosis; mild cognitive impairment

Introduction

In longitudinal studies of aging, tracking progression of cognitive changes over time often requires detection of the onset of clinically significant mild cognitive impairment (MCI) or dementia. The reference standard diagnosis of MCI or dementia requires direct evaluation by a clinician, integrating information obtained through history from patient and family members, physical examination, laboratory testing, and neuroimaging, to determine if the findings meet established clinical criteria – such as the National Institute on Aging-Alzheimer’s Association (NIA-AA) criteria for all-cause dementia,¹ the Diagnostic and Statistical Manual (DSM-5) or the International Classification of Disease (ICD-10) criteria for major neurocognitive disorders.² However, in large epidemiologic or cohort studies, this approach may not be feasible due to expense and/or logistical constraints. An alternative method utilizes expert panels for diagnosing MCI or dementia in cohort studies. However, there is currently no standardized approach and limited evidence on preferred methodology.³

For our own large-scale, prospective epidemiologic study of cognitive outcomes following major surgery, we wanted to implement rigorous methods for a clinical consensus procedure for cognitive outcomes following accepted approaches by expert panels. Thus, we first conducted a systematic literature review on expert panel consensus diagnosis methods for MCI and dementia used in cohort studies to provide detailed information on each of the approaches, and to identify key components of the expert panel methodology. Next, we provide a description of the clinical consensus procedure we developed for the Successful AGing after Elective Surgery Study (SAGES),⁴ and describe its alignment with and extension of the other approaches. Finally, based on lessons learned, suggested reporting standards are presented.

Materials and Methods

Literature Review

A systematic literature review was conducted according to PRISMA guidelines.^{5,6} To identify relevant papers on the expert panel diagnosis of MCI or dementia, a PubMed search using the search strategy ([dementia] AND [diagnosis] AND ([expert panel] OR [consensus methods] OR [consensus diagnosis])) was performed by one of the authors (TGF). The search was limited to papers published in the past 10 years (March 6, 2007–March 6, 2017),

in human subjects, and written in English. Studies were required to be original investigations that included diagnosis of MCI or dementia as an outcome measure, for example, cohort studies of dementia prevalence or incidence. We specifically excluded studies where a clinical diagnosis of MCI or Alzheimer's disease was supported by biomarkers, such as in the Alzheimer's Disease Neuroimaging Initiative (ADNI) or Australian Imaging, Biomarker & Lifestyle Flagship Study of Ageing (AIBL) studies⁷, or studies which use pathologic results from autopsy to confirm diagnoses, such as in the Nun study⁸ or the Rush Religious Orders Study.⁹ In addition, we excluded review articles, papers that described consensus guidelines intended for diagnostic use in clinical practice, letters, and case reports (Figure 1).

We identified a total of 1168 articles. After exclusion of articles based on title and abstract (Figure 1), 56 remaining full-text articles were reviewed. It is important to note that the overall goal of this review was to identify varying approaches to consensus methods, *not* to provide an exhaustive review of all studies of incident dementia. In addition, following Institute of Medicine (IOM) standards for systematic reviews,¹⁰ three articles describing the Aging, Demographics, and Memory Study (ADAMS), Women's Health Initiative Memory Study (WHIMS) and Honolulu Asian-American Study (HAAS) cohort, were added based on hand-review of references, since these cohorts are frequently referenced by other studies and are considered to be important among longitudinal studies on aging.

Information on the consensus method for diagnosing MCI or dementia was abstracted for each study by one of three study authors (EYK, EM, TTH) into a table capturing key categories of interest. Data were confirmed and cross-checked by another author for inclusion in the table (TGF). For several papers identified by the literature search, details on the expert panel methods were obtained from or verified with additional articles, since some papers were follow-up analyses of long-standing longitudinal studies, with more detailed methods having been published earlier.

Results

Systematic Review

We categorized expert panel methods in 31 studies (Table 1) in detail. Specifically, we examined the composition of the expert panel, the information available to the panel, and the decision process. For the decision process, we examined who performed the initial review, which criteria were used and how they were applied, and how disagreements were handled. Variability in the detail of the consensus methods was notable, with the number or expertise of the panel members not defined in nearly half of the studies (15/31, 48%). For those studies reporting such details, the number of panel members ranged from 2 to 7. A third of studies with panel descriptions explicitly included 5 or more experts (5/15, 33%). Around half of the studies (15/31, 48%) noted the composition of their expert panels, typically one or more members from the following disciplines: neurology, geriatric medicine, neuropsychology, geriatric psychiatry, nursing, gerontology, or neuroscience. All studies included cognitive or neuropsychological testing data for panel review. However, other measures were not universally included; for example, functional measures were included in

87% (27/31), proxy measures in 93% (28/31), and direct patient assessments in 77% (24/31).

The decision process used by each expert panel was variably reported. Once a case was brought to the panel, the initial evaluation was performed individually in 21 (68%) studies, by the entire group in 3 (10%) studies, or not specified in 5 (16%) studies. Most panels (22/31, 71%), relied on published diagnostic criteria, for example Diagnosis and Statistical Manual – version 5 (DSM-5),² National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer’s Disease and Related Disorders Association (NINCDS-ADRDA),¹¹ or National Institute on Aging – Alzheimer’s Association Diagnostic Guidelines (NIA-AA),¹ but some studies used their own definitions. Specific details on how diagnostic criteria were met by the available data were not provided in most cases. The decision process was specified as “majority vote” or “consensus” in 21 of 31 studies reviewed (70%), but few studies provided specific criteria required to achieve consensus. Furthermore, handling of disagreements was not well described, with only 6 of 31 studies (20%) reporting methods used to reach a final decision. One exception to this is in WHIMS, which provides a useful example reporting details of their consensus approach. The authors specifically describe how 2 adjudicators independently evaluated the data and assigned a classification, which was then matched with the diagnostic classification made by a field physician; all 3 adjudicators needed to be in agreement for consensus diagnosis. If the adjudicators disagreed, they discussed the case and if consensus was still not achieved, the 3 adjudicators plus a geriatric psychologist familiar with the neuropsychological measures discussed the case until a consensus was reached.¹²

Clinical Consensus Procedure for the Successful Aging after Elective Surgery (SAGES) Study

SAGES is an ongoing prospective cohort study of older adults undergoing major elective non-cardiac surgery, which examines the contribution of delirium to long-term outcomes, including cognitive decline. Incident dementia or mild cognitive impairment are secondary outcome measures. The study design and methods have been described previously.⁴ The expert panel procedure for classification of MCI or dementia is summarized in Table 2. In brief, a seven-member interdisciplinary expert panel chaired and coordinated by a non-voting neurologist (TGF) reviewed detailed demographic and clinical data, which provided a framework for interpreting the cognitive, functional, and proxy-rated outcome measures. Keeping these factors in mind, the next step in the individual diagnostic procedure was to review specific data that was used to score the NIA-AA criteria.¹ These criteria provided a conceptual “checklist” for establishing a research diagnosis of dementia or MCI (eTable 1). In the case of dementia, for example, the NIA-AA requirement for a symptom “that interferes with usual activities, is a decline from a prior level of functioning, and is not explained by delirium or other psychiatric factor” to be present was evaluated by review of the following data: impairment in memory or other cognitive patient-reported Activities of Daily Living (ADLs)/Instrumental Activities of Daily Living (IADLs) (disruption in usual activities), proxy Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) (decline from a prior level of functioning), Confusion Assessment Method (CAM) (presence or absence of delirium), and Short Form-12/Geriatric Depression Scale (presence or absence

of other psychiatric factor). For evidence of cognitive decline, the expert panel chose to focus the adjudication process on selected neuropsychological tests that represented domains largely affected by MCI and dementia diagnosis, including the Hopkins Verbal Learning Test-Discrimination Index (HVLTDI) and HVLTPercent Retention, Boston Naming Test (BNT) category fluency, and Trails B. Cut points for MCI were defined as impairment more than 2 standard deviations (SD) below mean (age-adjusted norm) on HVLTDI or percent retention, or impairment more than 1.5 SD below mean (age-adjusted norm) on two or more of the selected tests.^{13,14} For dementia, cognitive performance had to be more than 2 SD below mean (age-adjusted norm) on two or more of the selected tests.^{15,16} When longitudinal data was available, trends over time were considered in diagnosis.

After completing initial individual review for each case, panel members entered their preliminary diagnoses and any comments about the case to assist adjudication into a centralized database developed in REDCap.¹⁷ Panel members were advised to consider factors that could potentially affect cognitive performance, such as the presence of a delirium, depression, anxiety, medications, vision/hearing impairment, medical comorbidities, and developmental disabilities. Table 2 provides details of the SAGES clinical consensus protocol, which included a modified Delphi consensus process,¹⁸ along with the number of patients undergoing each step. The SAGES procedure is summarized and compared descriptively with other studies from the systematic review (Table 3).

Discussion

Methods to identify incident dementia and related cognitive disorders are crucial for many large-scale longitudinal studies, which aim to enhance our understanding of the cognitive aging process. The ability to compare and reproduce methods across studies is important to enhance and speed the advancement of scientific understanding in this important area. In our systematic review, we have demonstrated that the methods and the level of detail in reporting of the methods are highly variable across studies. The variability in approaches combined with inconsistency in reporting makes comparison of findings across studies quite challenging. Moreover, since none of the reported methods in this review have been validated against a reference standard, such as cerebrospinal fluid Alzheimer's biomarkers, amyloid positron emission tomography (PET) imaging, or pathological confirmation, comparison of the relative diagnostic efficacy of these approaches was beyond the scope of the present study.

Some of the studies from the systematic review used a two stage approach, where the first stage utilized a brief cognitive screen on all study participants, followed by a clinical examination by skilled clinicians (often neurologists) of those participants who show evidence of cognitive impairment by the screen. Inclusion of a clinical examination, however, also does not guarantee accuracy of a diagnosis, and for many large cohort studies, direct clinical assessments are not always feasible. For these reasons, large cohort studies often apply epidemiologic tools and an expert panel to approximate a clinical diagnosis because using multiple pieces of information can improve classification of the presence or absence of the disease³. Our review indicates that these methods are often highly variable, and to date no minimum standards have been established for reporting¹⁹.

Based on our findings and prior work,³ we have proposed some summary recommendations for reporting of expert consensus procedures (Table 4) across 4 domains: expert panel constitution; data provided to panel; decision process; and validity of diagnosis. In terms of the panel constitution, Bertens et al³ have recommended assuring the same panel members throughout the study, using an odd number of members if voting is part of the process (to avoid ties), and defining how panels will resolve disagreements prior to initiating the study. We recommend including members from 4 expert disciplines if possible – neurology, geriatrics, psychiatry, and neuropsychology – to cover the breadth of the clinical expertise for optimal evaluation of the study data. Ensuring diversity of expertise among panel members improves accuracy of the modified Delphi approach, compared with ratings of a single clinician.²⁰ Neuropsychologists are especially helpful to assist in interpreting the neuropsychological testing data, which can be extremely useful in complex cases. The inclusion of a non-voting coordinator for the process can help to minimize bias in results. In terms of data provided to the panel, well organized, clearly formatted reports providing all relevant demographic, background, clinical, and proxy-rated information to the panel along with embedded instructions maximizes the efficiency and effectiveness of the process. For neuropsychological testing, provision of both raw and adjusted scores, along with population norms, can greatly assist in interpretation.

Criteria for entering the adjudication process and for review by the full panel should be pre-specified for the decision process. As a first step, a blinded review by individual panel members is recommended. When full consensus is not achieved, then the case goes to second stage review by the full panel. Panel members utilize pre-specified decision rules to classify diagnoses. Clinical judgment is required to consider the influence of factors such as comorbidity, low educational level, or developmental disabilities; and to thoughtfully evaluate cognitive performance without relying solely on cut-points. Establishing the rules for handling disagreements and to achieve consensus should be completed *á priori*. The diagnostic accuracy for this approach, a modified Delphi method--which requires expert panels to first individually review a uniform set of data, followed by a group discussion and final group decision--has been shown to be superior to individual diagnosis, and is considered to provide sufficiently accurate diagnoses in cases where histopathological data are not available.¹⁷

Enhancing the validity of diagnoses can be achieved through the following steps. Use of accepted, published diagnostic criteria (e.g., NIA-AA criteria) maximizes accuracy, and evaluation of inter-rater agreement across panel members in the blinded ratings enhances reliability and reproducibility of diagnoses. Diagnoses can be further supported by longitudinal trends and clinical data (e.g., neuroimaging). True validation requires comparison with a reference standard, such as cerebrospinal fluid Alzheimer's biomarkers, amyloid PET imaging, or pathological confirmation.

In creating the SAGES clinical consensus procedure, we incorporated as many of these features (Table 4) as possible. The SAGES clinical consensus procedure was a structured and iterative diagnostic process, first undertaken individually by the seven members of the panel, and subsequently affirmed by group consensus adjudication using a modified Delphi method.¹⁸ This process takes advantage of a large panel with an odd-number of members,

which provides a wider range of experience and expertise in approaching the data, particularly when the diagnosis may not be straightforward. As with other studies, SAGES uses published criteria for diagnosis, and the data available to our expert panel is explicitly aligned to address each component of the criteria (eTable 1), which is a strength of our process. Finally, the SAGES adjudication process facilitates consideration of other important factors such as developmental disability, comorbidity, or delirium. Although our SAGES method is not validated against a reference standard, it provides a means to approximate the NIA-AA criteria in an epidemiologic study.

Several caveats about this study are worthy of mention. First, as with any systematic review, some articles may have been missed. For a review of study methods, as opposed to study results, a particular challenge is in the ability of the search to identify studies, when specific details of the methods are not always part of the title, abstract or keywords. While we acknowledge this important limitation, it is unlikely that we missed a large number of additional methodologies that would differ substantially from the 31 approaches identified in the current review^{3,6}. In addition, while we present SAGES as a representative example, it is not a validated approach. We did not intend for it to serve as a reference standard for these methods, only to provide a useful framework to demonstrate more detailed reporting of key steps in the expert panel process.

Conclusion

Methods used for expert panel diagnosis of MCI/dementia in large cohort studies are highly variable and often poorly described. Based on the studies reviewed here, the description of the expert panel methods should include at minimum both the number and expertise of the panel members, data available to the panel, definition of how established criteria are met, and details of the diagnostic decision making process, including the handling of disagreements. By describing the procedure in which our SAGES expert panel achieved consensus diagnoses, we hope to encourage development and publication of well-described and reproducible methods for the diagnosis of incident MCI/dementia in clinical and epidemiologic studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by Grants No. P01AG031720 (SKI), K07AG041835 (SKI), R24AG054259 (SKI) and R01AG044518 (SKI) from the National Institute on Aging, and 3 UL1 TR001102-04S2 (TGF) from the National Center for Advancing Translational Sciences. Dr. Kimchi was supported in part by Grant No. T32NS048005 from the National Institute of Neurological Disorders and Stroke and Dr. Marcantonio is supported in part by Grant No. K24AG035075 from the National Institute on Aging. Dr. Inouye holds the Milton and Shirley F. Levy Family Chair. The authors gratefully acknowledge the patients, family members, nurses, physicians, and research staff who made the SAGES study possible. This work is dedicated to the memory of Joshua Bryan Inouye Helfand.

References

1. Mckhann GM, Knopman DS, Chertkow H, et al. The Diagnosis Of Dementia Due To Alzheimer's Disease: Recommendations From The National Institute On Aging-Alzheimer's Association Workgroups On Diagnostic Guidelines For Alzheimer's Disease. *Alzheimer's & Dementia : The Journal Of The Alzheimer's Association*. May; 2011 7(3):263–269.
2. Association AP. Diagnostic And Statistical Manual Of Mental Disorders: DSM-5. Washington, DC: American Psychiatric Association; 2013.
3. Bertens LCM, Broekhuizen BDL, Naaktgeboren CA, et al. Use Of Expert Panels To Define The Reference Standard In Diagnostic Research: A Systematic Review Of Published Methods And Reporting. *Plos Medicine*. 2013; 10(10):E1001531. [PubMed: 24143138]
4. Schmitt EM, Marcantonio ER, Alsop DC, et al. Novel Risk Markers And Long-Term Outcomes Of Delirium: The Successful Aging After Elective Surgery (SAGES) Study Design And Methods. *Journal Of The American Medical Directors Association*. Nov; 2012 13(9):818 E811–810.
5. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA Statement For Reporting Systematic Reviews And Meta-Analyses Of Studies That Evaluate Health Care Interventions: Explanation And Elaboration. *Annals Of Internal Medicine*. Aug 18; 2009 151(4):W65–94. [PubMed: 19622512]
6. Lilford RJ, Richardson A, Stevens A, et al. Issues In Methodological Research: Perspectives From Researchers And Commissioners. *Health Technology Assessment*. 2001; 5(8):1–57.
7. Ellis KA, Bush AI, Darby D, et al. The Australian Imaging, Biomarkers And Lifestyle (AIBL) Study Of Aging: Methodology And Baseline Characteristics Of 1112 Individuals Recruited For A Longitudinal Study Of Alzheimer's Disease. *International Psychogeriatrics / IPA*. Aug; 2009 21(4): 672–687.
8. Snowdon DA, Kemper SJ, Mortimer JA, Greiner LH, Wekstein DR, Markesbery WR. Linguistic Ability In Early Life And Cognitive Function And Alzheimer's Disease In Late Life: Findings From The Nun Study. *JAMA : The Journal Of The American Medical Association*. 1996; 275(7):528–532. [PubMed: 8606473]
9. Bennett DA, Wilson RS, Schneider JA, et al. Natural History Of Mild Cognitive Impairment In Older Persons. *Neurology*. Jul 23; 2002 59(2):198–205. [PubMed: 12136057]
10. Finding What Works In Health Care: Standards For Systematic Reviews. Washington, DC: National Academies Press (US); 2011.
11. 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*. Jul; 1984 34(7):939–944. [PubMed: 6610841]
12. Shumaker SA, Legault C, Kuller L, et al. Conjugated Equine Estrogens And Incidence Of Probable Dementia And Mild Cognitive Impairment In Postmenopausal Women: Women's Health Initiative Memory Study. *JAMA : The Journal Of The American Medical Association*. 2004; 291(24):2947–2958. [PubMed: 15213206]
13. Jak AJ, Bondi MW, Delano-Wood L, et al. Quantification Of Five Neuropsychological Approaches To Defining Mild Cognitive Impairment. *The American Journal Of Geriatric Psychiatry : Official Journal Of The American Association For Geriatric Psychiatry*. May; 2009 17(5):368–375. [PubMed: 19390294]
14. Tuokko H, Garrett DD, McDowell I, Silverberg N, Kristjansson B. Cognitive Decline In High-Functioning Older Adults: Reserve Or Ascertainment Bias? *Aging & Mental Health*. Jul; 2003 7(4):259–270. [PubMed: 12888437]
15. Strauss, E., Speen, O. A Compendium Of Neuropsychological Tests: Administration, Norms, And Commentary. Oxford University Press; 2006.
16. Shapiro AM, Benedict RH, Schretlen D, Brandt J. Construct And Concurrent Validity Of The Hopkins Verbal Learning Test-Revised. *Clin Neuropsychol*. Aug; 1999 13(3):348–358. [PubMed: 10726605]
17. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research Electronic Data Capture (Redcap)--A Metadata-Driven Methodology And Workflow Process For Providing Translational Research Informatics Support. *J Biomed Inform*. Apr; 2009 42(2):377–381. [PubMed: 18929686]

18. Dalkey N, Helmer O. An Experimental Application Of The Delphi Method To The Use Of Experts. *Management Science*. 1963; 9:458–467.
19. Mayeux R, Reitz C, Brickman AM, et al. Operationalizing Diagnostic Criteria For Alzheimer's Disease And Other Age-Related Cognitive Impairment-Part 1. *Alzheimer's & Dementia : The Journal Of The Alzheimer's Association*. Jan; 2011 7(1):15–34.
20. Gabel MJ, Foster NL, Heidebrink JL, et al. Validation Of Consensus Panel Diagnosis In Dementia. *Archives Of Neurology*. 2010; 67(12):1506–1512. [PubMed: 21149812]
21. Gross AL, Jones RN, Habtemariam DA, et al. Delirium And Long-Term Cognitive Trajectory Among Persons With Dementia. *Archives Of Internal Medicine*. Sep 24; 2012 172(17):1324–1331. [PubMed: 23403619]
22. Schmitt EM, Saczynski JS, Kosar CM, et al. The Successful Aging After Elective Surgery Study: Cohort Description And Data Quality Procedures. *Journal Of The American Geriatrics Society*. Dec; 2015 63(12):2463–2471. [PubMed: 26662213]
23. Tremont G, Papandonatos GD, Kelley P, Bryant K, Galioto R, Ott BR. Prediction Of Cognitive And Functional Decline Using The Telephone-Administered Minnesota Cognitive Acuity Screen. *Journal Of The American Geriatrics Society*. Mar; 2016 64(3):608–613. [PubMed: 27000332]
24. Gerstenecker A, Triebel KL, Martin R, Snyder S, Marson DC. Both Financial And Cognitive Decline Predict Clinical Progression In MCI. *Alzheimer Disease And Associated Disorders*. Jan-Mar; 2016 30(1):27–34. [PubMed: 26900988]
25. Paddick SM, Kisoli A, Samuel M, et al. Mild Cognitive Impairment In Rural Tanzania: Prevalence, Profile, And Outcomes At 4-Year Follow-Up. *The American Journal Of Geriatric Psychiatry : Official Journal Of The American Association For Geriatric Psychiatry*. Sep; 2015 23(9):950–959. [PubMed: 25579049]
26. Hendrie HC, Hake A, Lane K, et al. Statin Use, Incident Dementia And Alzheimer Disease In Elderly African Americans. *Ethnicity & Disease*. Aug 07; 2015 25(3):345–354. [PubMed: 26673814]
27. Steffens DC, Welsh-Bohmer KA, Burke JR, et al. Methodology And Preliminary Results From The Neurocognitive Outcomes Of Depression In The Elderly Study. *Journal Of Geriatric Psychiatry And Neurology*. Dec; 2004 17(4):202–211. [PubMed: 15533991]
28. Cavallari M, Hsieh TT, Guttmann CR, et al. Brain Atrophy And White-Matter Hyperintensities Are Not Significantly Associated With Incidence And Severity Of Postoperative Delirium In Older Persons Without Dementia. *Neurobiology Of Aging*. Jun; 2015 36(6):2122–2129. [PubMed: 25824618]
29. Roberts RO, Geda YE, Knopman DS, et al. The Mayo Clinic Study Of Aging: Design And Sampling, Participation, Baseline Measures And Sample Characteristics. *Neuroepidemiology*. 2008; 30(1):58–69. [PubMed: 18259084]
30. Tschanz JT, Welsh-Bohmer KA, Lyketsos CG, et al. Conversion To Dementia From Mild Cognitive Disorder: The Cache County Study. *Neurology*. Jul 25; 2006 67(2):229–234. [PubMed: 16864813]
31. Radford K, Mack HA, Robertson H, et al. The Koori Growing Old Well Study: Investigating Aging And Dementia In Urban Aboriginal Australians. *International Psychogeriatrics*. 2014; 26(6):1033–1043. [PubMed: 24507414]
32. Dong Y, Gan DZQ, Tay SZ, et al. Patterns Of Neuropsychological Impairment In Alzheimer's Disease And Mixed Dementia. *Journal Of The Neurological Sciences*. Oct 15; 2013 333(1–2):5–8. [PubMed: 23978422]
33. Rosenblatt A, Samus QM, Steele CD, et al. The Maryland Assisted Living Study: Prevalence, Recognition, And Treatment Of Dementia And Other Psychiatric Disorders In The Assisted Living Population Of Central Maryland. *Journal Of The American Geriatrics Society*. 2004; 52(10):1618–1625. [PubMed: 15450036]
34. Katz MJ, Lipton RB, Hall CB, et al. Age And Sex Specific Prevalence And Incidence Of Mild Cognitive Impairment, Dementia And Alzheimer's Dementia In Blacks And Whites: A Report From The Einstein Aging Study. *Alzheimer Disease And Associated Disorders*. 2012; 26(4):335–343. [PubMed: 22156756]

35. Nose M, Kodama C, Ikejima C, et al. Apoe4 Is Not Associated With Depression When Mild Cognitive Impairment Is Considered. *International Journal Of Geriatric Psychiatry*. Feb; 2013 28(2):155–163. [PubMed: 22700491]
36. Canadian Study Of Health And Aging. Study Methods And Prevalence Of Dementia. *CMAJ: Canadian Medical Association Journal*. 1994; 150(6):899–913. [PubMed: 8131123]
37. Mathuranath PS, George A, Ranjith N, et al. Incidence Of Alzheimer's Disease In India: A 10 Yearsfollow-Up Study. *Neurology India*. Nov-Dec;2012 60(6):625–630. [PubMed: 23287326]
38. Handels RLH, Aalten P, Wolfs CAG, et al. Diagnostic And Economic Evaluation Of New Biomarkers For Alzheimer's Disease: The Research Protocol Of A Prospective Cohort Study. *BMC Neurology*. 2012; 12:72–72. [PubMed: 22883691]
39. Annweiler C, Fantino B, Schott AM, Krolak-Salmon P, Allali G, Beauchet O. Vitamin D Insufficiency And Mild Cognitive Impairment: Cross-Sectional Association. *European Journal Of Neurology*. Jul; 2012 19(7):1023–1029. [PubMed: 22339714]
40. Kukull WA, Higdon R, Bowen JD, et al. Dementia And Alzheimer Disease Incidence: A Prospective Cohort Study. *Archives Of Neurology*. 2002; 59(11):1737–1746. [PubMed: 12433261]
41. Dotson VM, Beydoun MA, Zonderman AB. Recurrent Depressive Symptoms And The Incidence Of Dementia And Mild Cognitive Impairment. *Neurology*. 2010; 75(1):27–34. [PubMed: 20603482]
42. Kumar R, Dear KBG, Christensen H, et al. Prevalence Of Mild Cognitive Impairment In 60- To 64-Year-Old Community-Dwelling Individuals: The Personality And Total Health Through Life 60+ Study. *Dementia And Geriatric Cognitive Disorders*. 2005; 19(2–3):67–74. [PubMed: 15572874]
43. Lopez OL, Jagust WJ, Dekosky ST, et al. Prevalence And Classification Of Mild Cognitive Impairment In The Cardiovascular Health Study Cognition Study: Part 1. *Archives Of Neurology*. 2003; 60(10):1385–1389. [PubMed: 14568808]
44. Cook SE, Marsiske M, Mccoy KJM. The Use Of The Modified Telephone Interview For Cognitive Status (TICS-M) In The Detection Of Amnesic Mild Cognitive Impairment. *Journal Of Geriatric Psychiatry And Neurology*. 2009; 22(2):103–109. [PubMed: 19417219]
45. Steenland NK, Auman CM, Patel PM, et al. Development Of A Rapid Screening Instrument For Mild Cognitive Impairment And Undiagnosed Dementia. *Journal Of Alzheimer's Disease : JAD*. 2008; 15(3):419–427. [PubMed: 18997295]
46. Smith K, Flicker L, Lautenschlager NT, et al. High Prevalence Of Dementia And Cognitive Impairment In Indigenous Australians. *Neurology*. Nov 04; 2008 71(19):1470–1473. [PubMed: 18799785]
47. Galasko D, Salmon D, Gamst A, et al. Prevalence Of Dementia In Chamorros On Guam: Relationship To Age, Gender, Education, And APOE. *Neurology*. May 22; 2007 68(21):1772–1781. [PubMed: 17515539]
48. Sachdev PS, Chen X, Joscelyne A, Wen W, Altendorf A, Brodaty H. Hippocampal Size And Dementia In Stroke Patients: The Sydney Stroke Study. *Journal Of The Neurological Sciences*. Sep 15; 2007 260(1–2):71–77. [PubMed: 17482210]
49. Satizabal CL, Beiser AS, Chouraki V, Chêne G, Dufouil C, Seshadri S. Incidence Of Dementia Over Three Decades In The Framingham Heart Study. *The New England Journal Of Medicine*. 2016; 374(6):523–532. [PubMed: 26863354]
50. Langa KM, Plassman BL, Wallace RB, et al. The Aging, Demographics, And Memory Study: Study Design And Methods. *Neuroepidemiology*. 2005; 25(4):181–191. [PubMed: 16103729]
51. White L, Petrovitch H, Ross G, et al. Prevalence Of Dementia In Older Japanese-American Men In Hawaii: The Honolulu-Asia Aging Study. *JAMA : The Journal Of The American Medical Association*. 1996; 276(12):955–960. [PubMed: 8805729]
52. Van Norden AG, De Laat KF, Gons RA, et al. Causes And Consequences Of Cerebral Small Vessel Disease. The RUN DMC Study: A Prospective Cohort Study. Study Rationale And Protocol. *BMC Neurology*. 2011; 11:29. [PubMed: 21356112]
53. Tombaugh TN. Trail Making Test A And B: Normative Data Stratified By Age And Education. *Archives Of Clinical Neuropsychology*. 2004; 19(2):203–214. [PubMed: 15010086]
54. Trenerry, M. Visual Search And Attention Test: Professional Manual. Odessa, FL: Psychological Assessment Resources; 1990.

55. Brandt, J., Benedict, R. Hopkins Verbal Learning Test - Revised: Professional Manual. Odessa, FL: Psychological Assessment Resources; 2001.
56. Wechsler, D. Wechsler Adult Intelligence Scale-Revised Manual. New York, NY: Psychological Corporation, A Harcourt Assessment Company; 1989.
57. Benton A. Development Of A Multilingual Aphasia Battery. Progress And Problems. J Neurol Sci. 1969; 9:39–48. [PubMed: 5820858]
58. Benton A. Verbal Fluency. Neuropsychologica. 1968; 6:53–60.
59. Goodglass, HEK. The Assessment Of Aphasia And Related Disorders. 2. Malvern, PA: Lea & Febiger; 1983.
60. Randolph, C. Repeatable Battery For The Assessment Of Neuropsychological Status. San Antonio: The Psychological Corporation; 1998.
61. Teng EL, Chui HC. The Modified Mini-Mental State (3MS) Examination. The Journal Of Clinical Psychiatry. Aug; 1987 48(8):314–318. [PubMed: 3611032]
62. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies Of Illness In The Aged. The Index Of Adl: A Standardized Measure Of Biological And Psychosocial Function. Jama. Sep 21.1963 185:914–919. [PubMed: 14044222]
63. Lawton MP, Brody EM. Assessment Of Older People: Self-Maintaining And Instrumental Activities Of Daily Living. The Gerontologist. Autumn;1969 9(3):179–186. [PubMed: 5349366]
64. Ware J Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: Construction Of Scales And Preliminary Tests Of Reliability And Validity. Medical Care. Mar; 1996 34(3):220–233. [PubMed: 8628042]
65. Plassman BL, Langa KM, McCommon RJ, et al. Incidence Of Dementia And Cognitive Impairment, Not Dementia In The United States. Annals Of Neurology. Sep; 2011 70(3):418–426. [PubMed: 21425187]
66. Inouye SK, Van Dyck CH, Alessi CA, Balkin S, Siegel AP, Horwitz RI. Clarifying Confusion: The Confusion Assessment Method. A New Method For Detection Of Delirium. Annals Of Internal Medicine. Dec 15; 1990 113(12):941–948. [PubMed: 2240918]
67. Jorm AF. A Short Form Of The Informant Questionnaire On Cognitive Decline In The Elderly (IQCODE): Development And Cross-Validation. Psychological Medicine. Feb; 1994 24(1):145–153. [PubMed: 8208879]
68. Yesavage JA. Geriatric Depression Scale. Psychopharmacology Bulletin. 1988; 24(4):709–711. [PubMed: 3249773]
69. Tierney MC, Moineddin R, McDowell I. Prediction Of All-Cause Dementia Using Neuropsychological Tests Within 10 And 5 Years Of Diagnosis In A Community-Based Sample. Journal Of Alzheimer's Disease : JAD. 2010; 22(4):1231–1240. [PubMed: 20930315]
70. Albert MS, Dekosky ST, Dickson D, et al. The Diagnosis Of Mild Cognitive Impairment Due To Alzheimer's Disease: Recommendations From The National Institute On Aging-Alzheimer's Association Workgroups On Diagnostic Guidelines For Alzheimer's Disease. Alzheimer's & Dementia : The Journal Of The Alzheimer's Association. May; 2011 7(3):270–279.

Figure 1.
Flowchart for article selection for inclusion in systematic review of study methods.

Table 1

Summary of expert panels used in longitudinal dementia or MCI studies

	Cohort (n) /panel (n)	Study objective	Panel composition (n)		Information available to panel ²	Longitudinal data ³	Decision Process				Ref	
			Members	Disciplines represented ¹			Identified for panel adjudication ⁴	Initial Review ⁵	Criteria ⁶	Decision process ⁷		Handling of disagreements
Successful Aging after Elective Surgery	566 (112)	Functional and cognitive outcomes after delirium	7	G, N, NP, P	C, F, P	Y	Neuropsych cutpoint	I	P	C	Discussion to 5 of 7 agreement	21,22
Minnesota Cognitive Acuity Screen - telephone	97 (97)	Utility of MCAS-t to identify MCI over time	?	?	C, F, P, D	Y	Full cohort	G	P	Majority vote	?	23
Cognitive Observations in Seniors	143 (51)	MCI progression	?	N, NP, RN	C, F, P, D	Y	Completed two study visits	?	P	?	?	24
Tanzania	1198 (296)	Dementia prevalence, community-based	3–4	MD, RN	C, F, P, D	Y	Performance on screening measure	I	P	?	Single consultant (P)	25
Indianapolis-Ibadan Dementia Project	974 (974)	Effect of statin use on MCI and dementia incidence in community-based sample	?	N, NP, P	C, F, P, D	Y	Full cohort	G	P	?	?	26
Neurocognitive Outcomes of Depression in the Elderly	587 (63)	Neurocognitive outcomes of depression	5–6	N, NP, P, O	C	N	Subjective cognitive complaints	I	P	?	?	27
Shanghai Aging Study	3541 (3541)	Prevalence, incidence and risk factors for dementia and MCI	5	N, NP	C, F, P, D	N	Full cohort	?	P	?	If not 100% agreement, marked as discordant	28
Mayo Clinic Study of Aging (Olmstead County)	2719 (2719)	Incident dementia in MCI	3	N, RN, NP	C, F, P, D (24.6% by phone)	Y	Full cohort	?	P	C	Discussion	29
Cache County Study on Memory, Health,& Aging	4463 (4463)	Conversion rate to dementia	2+	NP, P, MD	C, F, P, D	Y	Full cohort	I	C	C	?	30
Koori Growing Old Well Study	336 (153)	Aging, cognitive decline, and dementia in Aboriginal Australians	3	P, NP	C, F, P, D	Y	Screening cutpoint	I	P	C	?	31
National University Health System Memory Clinic	151 (151)	Cognitive profile in mixed dementia vs. AD	?	?	C, F, D	N	Full cohort	?	P	C	?	32
Maryland Assisted Living Study	200 (200)	Prevalence, recognition, and treatment of dementia and other psychiatric disorders in the assisted living population	?	G, NP, P, RN	C, F, P, D	Y	Full cohort	I	P,S	C	?	33
Einstein Aging Study	1944 (1944)	Incidence MCI, AD, and dementia	?	?	C, F, P, D	Y	Full cohort	I	C, P	C	?	34
Tone Town Japan	738 (738)	Relationship between ApoE 4 and depression	?	P, NP	C, F, P, D	N	Full cohort	G	S	C	?	35
Canadian Study of Health and Aging	10263 (2420)	Prevalence of dementia in Canada	?	MD, NP, RN,	C, F, P, D	Y	Screening cutpoint	I	P	C	?	36
Cognition in older adults in Trivandrum study	2466 (1066)	Incidence rates of Alzheimer's disease (AD) in India	?	N, NP, P	C, F, P, D	Y	Screening cutpoint	I	P	C	?	37

	Cohort (n) /panel (n)	Study objective	Panel composition (n)		Information available to panel ²	Longitudinal data ³	Decision Process			Ref		
			Members	Disciplines represented ¹			Identified for panel adjudication ⁴	Initial Review ⁵	Criteria ⁶		Decision process ⁷	Handling of disagreements
Handels, et al	241 (241)	Emerging and novel techniques for early AD diagnosis	?	?	C, F, P, D	Y	Full cohort	I	S	C	Discussion	38
Gait and Alzheimer Interaction Tracking study	95 (95)	Association between vitamin D levels and MCI	?	G, N, NP	C, P	N	Full cohort	I	S	C	?	39
Adult Changes in Thought study	2581 (415)	Dementia and AD incidence	?	MD, NP, RN	C, P, D	Y	Screening cutpoint	I	P	C	?	40
Baltimore Longitudinal Study on Aging	1239 (1239)	Depression and incidence of AD and MCI	?	?	C, F, P, D	Y	Full cohort	I	P	C	?	41
Personality and Total Health through Life Study	224 (117)	MCI prevalence	4	MD, P	C, F, P, D	N	Screening cutpoint	I	S	C	?	42
Cardiovascular Health Study - Cognition Study	2470 (2470)	Prevalence and classification of MCI	?	?	C, F, P, D	Y	Full cohort	I	S	?	?	43
Cook, et al	71 (71)	Use of TICS-M in detecting MCI	5	NP, O	C, F, P	N	Full cohort	I	P	C	Discussion, re-vote, majority rules	44
Steenland et al.	178 (178)	Rapid screen for MCI or dementia	2	N	C, F, P	N	Full cohort	I	S	C	Discussion	45
Kimberley indigenous Australian	363 (165)	Prevalence of dementia and cognitive impairment	2	G, P	C, F, P, D	N	Screening cutpoint, stratified	I	P	?	?	46
Guam	2029 (1058)	Dementia prevalence	3	N, G	C, F, P, D	N	Screening cutpoint, stratified	I	P	C	?	47
Sydney Stroke Study	165 (165)	Hippocampal size and dementia after stroke	4+	N, NP, P, G	C, F, P, D	N	Full cohort	I	S	?	?	48
Framingham Heart Study	5205 (5205)	Dementia incidence	2	N, NP	C, F, P, D	Y	Full cohort	I	P	C	?	49
Aging Demographics, and Memory Study	1770 (856)	Prevalence and outcomes of dementia and CIND	?	N, NP, P, MD	C, F, P	Y	Cognitive cutpoints, stratified	I	P	C	?	50
Women's Health Initiative Memory Study	10426 (1029)	Effect of estrogen on incidence of dementia and MCI in older women	3–4	N, GP	C, F, P, D	Y	Cognitive cutpoint, examination by field physician, stratified	I	P	C	Individual revision of diagnosis, discussion, discussion with additional experts	12
Honolulu Asia Aging Study	3734 (426)	Prevalence of dementia in	3	N, G	C, F, P, D	Y	Cognitive cutpoint	?	P	?	?	51

AD = Alzheimer's Disease; CIND = Cognitive Impairment, Not Demented; MCI = Mild Cognitive Impairment; MCI-A = Mild Cognitive Impairment, Amnesic subtype;

¹Disciplines represented: neurology (N), geriatrics (G), psychiatry (P), neuropsychology (NP), nursing (RN), other physician (MD), other (O). Other included cognitive neuroscientist

²Information available to panel: C=Cognitive testing; F=Functional measure; P=Proxy measure; D=Direct patient assessment

³Longitudinal data – “yes” if study included at least two assessments, separated in time

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

⁴Diagnostic criteria: C=cutpoint, P=published (for example, ICD-10, DSM-IVTR, NINDS-AIRE, NIA-AA), S=study defined, ?=not documented

Table 2**Description of the Clinical Consensus Protocol used in SAGES**

Study objective	Functional and cognitive outcomes after delirium
Total Cohort (n)	566
Cases Reviewed by Panel (n)	112
Panel Composition	
Members (n)	7
Clinical Expertise	Neurology, neuropsychology, geriatrics, geriatric psychiatry
Selection of Cases for Panel Review	Data were screened quarterly. Decline in cognitive performance greater than 1.5 standard deviations ⁵² from the age-adjusted mean on any one neuropsychological test, or greater than one SD on two or more tests, at least one of which needed to be a measure of memory (HVLTDI or HVLTDI-percent retention)
Information Available to Panel	
Demographic	Age, primary language, highest education, Vision, Alcohol, Medical diagnosis. Wechsler Test of Adult Reading (WTAR) ⁵³ at baseline, a test of single word reading used as a measure of premorbid intelligence
Cognitive testing	3MS, VSAT ⁵⁴ , HVLTDI ⁵⁵ , Digit Span Forward and Backward ⁵⁶ , Category Fluency (animal naming) ⁵⁷ , Phonemic F-A-S Fluency Tasks ⁵⁸ , BNT ⁵⁹ , RBANS Digit Symbol Substitution Test ⁶⁰ , TMT A and B ⁵³ , and intersecting pentagons from the 3MS ⁶¹ .
Functional data	ADLs ⁶² , IADLs ⁶³ , SF-12 ⁶⁴ , CAM ⁶⁵ , IQCODE ⁶⁷ , and GDS ⁶⁸ , obtained from both the patient and an informant.
Interview Quality Assessment	Interviewer confidence; sensory or memory impairment, physical illness, pain, stress, poor phone connection, presence of another person
Decision Process	
Criteria	NIA-AA
Initial Review (Individual, Blinded) n=578 cases	In the first round of the clinical consensus process, each panel member was asked to review specific data identified as analogs for the NIA-AA criteria ¹ . These criteria were used as a conceptual "checklist" in making a diagnosis of dementia or MCI (Table xx). Panel members were to independently assign one of three possible diagnoses for each participant at each time point: No Evidence of Disease ⁶⁹ , Mild Cognitive Impairment (MCI), or Dementia
Second Review (Group) N=112 cases	A consensus conference was typically held within a week after individual panel ratings were completed. Consensus agreement was defined as agreement by five of seven panel members. Any cases from the individual review that did not meet this underwent a second round of review at a consensus conference; any panel member could also request a case to be brought to consensus conference for discussion and review regardless of consensus criteria
Handling of Disagreements	Panel members had access to the same data provided to them in the initial review. As each case was reviewed, the preliminary diagnosis and comments from individual panel members regarding why each diagnosis was made in the initial review were presented to the group. Panel members had the opportunity to discuss diagnostic considerations and challenges with each other. Discussion continued until consensus reached.
Consensus	Five of seven panel members reached agreement on the diagnosis.

3MS=Modified Mini-Mental State Examination; ADL =Basic Activities of Daily Living Scale; BNT=Boston Naming Test; CAM = Confusion Assessment Method; GDS = Geriatric Depression Scale; HVLTDI =Hopkins Verbal Learning Test-Revised; IADLs = Instrumental Activities of Daily Living Scale; IQCODE = Informant Questionnaire on Cognitive Decline in the Elderly; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status; SF-12 = Short Form-12 Health Survey; TMT = Trail-Making Tests A and B; VSAT= Visual Search and Attention Test

Table 3

SAGES study within Framework of Other Studies

Key features	Systematic Review (N=30)*	SAGES Study
Expert Panel		
Panel number	2–6, when specified	7
Panel composition – interdisciplinary domains represented	Up to 4, when specified	4
Consistency of members over time	Not specified	Yes
Prespecified Standard Criteria		
Criteria for Entry to Adjudication Process	Full cohort or following specific cognitive screening criteria	Specific Neuropsych cutpoints
Criteria for Proceeding to Clinical Consensus	Often not specified	Less than 5 of 7 agreement on individual diagnosis
Criteria for MCI or dementia	Published clinical criteria (70%), or study defined	Standard Clinical Criteria
Consensus Decision Process		
Initial review	Individual review (67%), group review (10%), or not specified (17%)	Blinded, individual
Handling disagreements	Most studies (80%) did not specify	Group discussion
Re-vote and achieving final consensus	Most studies (83%) did not specify	Yes
Quality Assurance Considerations		
Non-voting chair	No	Yes
Electronic rating system (e.g., REDCap) to provide detailed instructions and comment fields	No	Yes

Table 4

Key Recommendations for Reporting of Expert Consensus Procedures in MCI and dementia

Expert Panel Constitution
Consistent membership for duration of study
Number of experts (odd number, if voting part of process)
Expertise – minimum 4 disciplines (i.e., neuropsychology, neurology, geriatrics, geriatric psychiatry)
Non-voting expert to coordinate process
Data Provided to Panel
Demographic, background, and clinical information; include source of data (patient vs. proxy)
Neuropsychological testing (include raw and adjusted scores, population norms, expert interpretation)
Guidelines for assigning diagnosis included in real time with each review
Decision Process
Initial, blinded individual review
Group review for those not achieving consensus (e.g., 5 of 7 ratings)
Follow pre-specified decision rules to classify diagnosis.
Use clinical judgment to consider any co-morbid or reversible conditions, including history of learning disability, high or low education, low IQ
Utilize experts to interpret objective data (usually neuropsychological performance) and do not rely solely on cutpoints
Pre-specify criteria for review by full panel
Ensure data needed to fulfill diagnostic criteria are reviewed
Pre-defined procedure for handling of disagreements
Validity of Diagnosis
Evaluate inter-rater agreement during initial blinded review
Support diagnosis with neuroimaging, longitudinal trends or other clinical data
Comparison to reference standard: Confirm diagnosis with autopsy, CSF, or PET neuroimaging