



Published in final edited form as:

Alzheimers Dement. 2011 January ; 7(1): 110–123. doi:10.1016/j.jalz.2010.11.008.

Reducing case ascertainment costs in U.S. population studies of Alzheimer’s disease, dementia, and cognitive impairment—Part 2

Denis A. Evans^{a,b}, Francine Grodstein^{c,d}, David Loewenstein^f, Jeffrey Kaye^{g,h,i}, and Sandra Weintraub^{j,k}

^a Rush Institute on Healthy Aging, Rush University Medical Center, Chicago, IL, USA R, Verghese J, Wang C, Hall CB, Lipton RB. Within-person across-neuropsychological test variability and indecent dementia. *JAMA* 2008;300:823–30

^b Department of Internal Medicine, Rush University Medical Center, Chicago, IL, USA

^c Department of Epidemiology, School of Public Health, Harvard University, Boston, MA, USA

^d Brigham and Women’s Hospital, Boston, MA, USA

^e Department of Psychiatry and Behavioral Sciences, Miller School of Medicine, University of Miami, Miami, FL, USA

^f Wien Center for Alzheimer’s Disease and Memory Disorders, Mount Sinai Medical Center, Miami, FL, USA

^g Oregon Center for Aging and Technology, Oregon Health and Science University, Portland, OR, USA

^h Layton Aging and Alzheimer’s Disease Center, Oregon Health and Science University, Portland, OR, USA

ⁱ Portland Veterans Affairs Medical Center, Portland, OR, USA

^j Department of Psychiatry and Behavioral Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

^k Cognitive Neurology and Alzheimer’s Disease Center, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

Abstract

Dementia of the Alzheimer’s type (DAT) is a major public health threat in developed countries where longevity has been extended to the eighth decade of life. Estimates of prevalence and incidence of DAT vary with what is measured, be it change from a baseline cognitive state or a clinical diagnostic endpoint, such as Alzheimer’s disease. Judgment of what is psychometrically “normal” at the age of 80 years implicitly condones a decline from what is normal at the age of 30. However, because cognitive aging is very heterogeneous, it is reasonable to ask “Is ‘normal for age’ good enough to screen for DAT or its earlier precursors of cognitive impairment?” Cost containment and accessibility of ascertainment methods are enhanced by well-validated and reliable methods such as screening for cognitive impairment by telephone interviews. However, focused assessment of episodic memory, the key symptom associated with DAT, might be more effective at distinguishing normal from abnormal cognitive aging trajectories. Alternatively, the futuristic “Smart Home,” outfitted with unobtrusive sensors and data storage devices, permits the moment-to-moment recording of activities so that changes that constitute risk for DAT can be identified before the emergence of symptoms.

1. Introduction

With the extension of longevity in the United States and elsewhere, dementia with aging has become a major public health threat. In epidemiological studies of dementia and precursor conditions, what is measured has a large effect on reports of prevalence and incidence. What is measured also affects the cost of the research, especially with regard to case ascertainment. This article, which complements the one by Weir et al [1], focuses on the choice of desired endpoints of ascertainment and how to measure cognition.

Specifically, Evans (section 2) weighs the arguments both for assessing cognitive change against a personal baseline and for identifying markers of clinical diagnoses such as Alzheimer's disease (AD) or mild cognitive impairment (MCI). Because dementia of the Alzheimer's type (DAT) is not an acute event and pathophysiological changes in the brain can precede overt symptoms by many years [2–9], it would be preferable to detect accelerated cognitive decline before a frank diagnosis so that preventative measures can be taken. However, clinical diagnoses are practical and easier to conceptualize in the healthcare setting. Grodstein (section 3) points out that cost containment and accessibility are enhanced by wellvalidated and reliable telephone interviews that provide a global cognitive screen. Loewenstein (section 4), by contrast, sees value in the focused assessment of episodic memory and its components, which targets the key symptom associated with AD. This approach, especially if tests are shortened, might prove to be more effective than other cognitive measurements to distinguish normal from diseased aging. Kaye (section 5) provides a view toward the future fueled by technological advances, as he espouses a model of a scalable "Smart Home." Unobtrusive sensors and data storage devices permit the moment-to-moment recording of activities so that changes that might constitute a risk for cognitive decline can be identified before the emergence of symptoms. Finally, Weintraub (section 6) provides a synthesis of all of this material and suggests that our views of what is "normal for age" are challenged as we struggle to face the pending healthcare crisis of the next several decades.

2. Directly measured individual change in cognitive function as a study outcome

This section focuses on using directly measured individual cognitive decline as a study outcome, and a contrast is made with using AD and MCI diagnostic endpoints as research study outcomes. Both approaches are complementary, and neither is inherently superior to the other. Rather, the two approaches are fully compatible; each has strengths, and coordinated use of both may be preferable to relying on either one alone. A central goal of research in this area is to facilitate prevention by identifying at-risk subjects and potential risk factors earlier when the disease process is more susceptible to intervention. Focus on the intermediate syndrome of MCI is of undoubted value in this regard. Rigorous study of the entire continuum of directly measured individual cognitive decline may well be a logical next step.

Perhaps the greatest conceptual advantage of directly measuring individual decline in cognitive function as an outcome is the potential to identify risk factors early, facilitating prevention of progression to MCI and AD. As the goal of preventing fully manifest AD is increasingly seen as vitally important, the need to recognize the early, or preclinical, stages has been strongly emphasized [10,11]. Other conceptual advantages are (1) the greater validity of assessing decline in cognition directly by measuring it repeatedly rather than by attempting to infer cognitive decline indirectly from a single measurement without previous baseline, and (2) the potential reduction in bias. Practical advantages include greater statistical power and reduced cost of data collection.

In studies of cognitive impairment or dementia, the object of interest is each subject's loss of cognition from a former state, and not that subject's level of cognition as compared with an absolute, "normative" value. Accurate inference of the degree of cognitive loss from measurement of cognition at a single point (cross-sectional assessment) is challenging because cognitive abilities vary greatly among individuals and there is great random variation in measurement of cognition. Even assessment for incident dementia is typically cross-sectional because it usually consists of making a diagnosis of dementia among subjects who previously did not meet criteria for the condition. Direct measurement of individual change in cognition for each subject is preferable, but requires that cognitive function be assessed sequentially at a minimum of three time points and that appropriate statistical models be used to adequately identify change from an initial cognitive level. Typically, mixed models (random-effects models) are used [12].

Another conceptual advantage of directly measured individual change in cognitive function over a diagnostic endpoint as an outcome is its potential lack of bias. This advantage may be especially important in assessing whether apparent individual characteristics such as race/ethnicity or gender are related to dementia or cognition. Examiners cannot be effectively masked to these characteristics, which are typically obvious. The possibility of subtle, but important examiner bias in reaching a diagnosis is exceedingly difficult to eliminate. It is unlikely to be substantially reduced by providing information from the examiner to a panel of expert clinicians who actually assign a diagnosis on the basis of this information. The possibility of such bias affecting the analytic result is especially strong when the characteristic of interest coincides with a difference in average lifelong level of cognitive performance. For example, meaningful assessment of the possibility of differences in risk for dementia or cognitive loss between those belonging to African American race/ethnicity and Americans belonging to white race/ethnicity has been challenging because African Americans usually score lower on tests of cognition, most likely for cultural and educational reasons [13–17]. Race/ethnicity, education, and other factors distorting measurement of cross-sectional cognitive performance most likely do not affect directly measured individual change in cognitive function at older ages. In older age, most educational and cultural factors that affect measurement of cognition are fairly stable, and for each person, the factors distorting measurement of cognition at one time point are likely to be similar to those affecting subsequent time points. To the extent that the distorting factors are constant over time, they will not affect the measurement of individual changes in performance between these points.

Greater statistical power is achievable because analyses of directly measured change in cognitive function can include the entire spectrum of cognition. All subjects across this spectrum with the requisite measurements will have a value for cognitive decline. However, with studies of clinical syndromes, such as AD or MCI, the limiting factor determining statistical power is typically the smaller number of subjects developing these conditions. Further, in most studies, use of directly measured individual decline in cognitive function as an outcome results in reduced costs. Sequential administration of cognitive tests by trained technicians as required to measure this outcome typically is less costly as compared with clinical evaluation by skilled clinicians, often specialists, as required for diagnosis of the clinical syndromes of AD and MCI.

Conversely, using the clinical syndromes of AD and MCI as outcomes also has advantages. Because many criteria for AD diagnosis include functional loss, a sense of the burden of the disease is typically inherent in the disease definition. In diagnosing AD, a clinician judges that the subjects in the study have functional impairment from the condition, although the level of impairment required for diagnosis may vary among different examiners. An advantage is that incorporating functional impairment emphasizes the public health burden

of AD. A disadvantage is that it becomes difficult to rigorously study functional impairment as an AD consequence in cases when it is inherent in the disease definition [18]. From the perspective of AD consequences, including functional impairment as a disease criterion is a self-fulfilling prophecy; one must have the consequence of functional impairment to have a clinical diagnosis of probable AD. Further, communications to general clinical audiences are facilitated if they can be summarized in terms of particular clinical syndromes. Change in cognitive function is difficult to summarize easily in terms that are clearly understandable intuitively. Decline in cognition also does not provide specificity as to the conditions responsible for the decline; the extent to which the observed decline in cognition is because of AD, vascular dementia, another condition, or a combination of conditions is not certain.

These approaches to studying decline in cognition and the conditions responsible seem strongly complementary. The major advantages of using directly measured individual change in cognitive function as an outcome in studies of cognitive loss or dementia—especially reduced bias, direct assessment of cognitive decline rather than inference of decline from a single measurement, and increased statistical power—are essentially most applicable to analytic studies, those evaluating the relations between a putative risk factor and cognitive decline. In contrast, the advantages of studying the clinical syndromes of AD and MCI, especially ease of communication, apply most directly to either more descriptive studies, such as those seeking to measure the prevalence or incidence of AD or MCI in a population or healthcare setting, or to studies of the consequences of AD or MCI such as behaviors or healthcare utilization.

Conclusions regarding either risk factors or consequences are more secure when the results of the two approaches coincide in direction and, to the extent it can be accurately assessed, magnitude. However, if the results found with one approach differ substantially from those found with the other, an explanation must be sought and may well be informative.

3. Measuring cognition in very large populations

At a population level, average differences in cognition or cognitive change among individuals are often fairly subtle, thus a large sample size ensures precise estimates of differences. This section addresses the measurement of cognitive function in very large populations, at very low cost. The main focus is the telephone cognitive battery used for repeated measures of cognitive function in approximately 20,000 subjects of the Nurses' Health Study (NHS). The NHS cognitive assessment battery will be described briefly and its properties considered. Finally, the strengths and weaknesses of using a telephone instrument to measure cognition will be weighed. The NHS battery is made up of tests of global function, episodic memory, category fluency, working memory, and attention. Specifically, the battery includes the Telephone Interview for Cognitive Status (TICS) [19], an adapted telephone version of the Mini-Mental State Examination. Delayed recall of the TICS 10-word list, which constitutes the TICS-modified, is also tested. In addition, the battery incorporates immediate and delayed recall trials of the East Boston Memory Test [20], to assess paragraph recall. Thus, there are several measures of verbal anterograde memory impairment, which is a strong predictor of AD. Category fluency is tested by having the subject name as many animals as possible in one minute. An oral adaptation of the Trail Making Test, Parts A and B, which has been well-validated and compares favorably with in-person administration of the original Trail Making test, is administered [21]. Finally, subjects are given the Digit Span Backwards Test. With the inclusion of the Trail Making Test, the NHS battery covers six of the seven domains and/or functions (excluding naming) in the Uniform Dataset of the National Institute on Aging Alzheimer Disease Centers' neuropsychological battery [22].

The NHS battery was validated against a more extensive battery used in a sample of 61 subjects from the Religious Orders Study (ROS) who received both batteries. The ROS battery, administered in person, comprises 21 cognitive tests [23]. For each participant, a global score (average performance across all the tests) was obtained for each battery. The two global scores had a correlation value of 0.81. This finding suggests that cognitive testing by telephone captures cognitive performance in a manner similar to face-to-face testing, and that a short battery can effectively assess overall cognition.

A clinical validation study compared findings from the NHS battery with the clinical diagnosis of dementia (F.D. Grodstein, unpublished, 2005). Cognitive function was assessed by telephone in approximately 80 women. Later, informant interviews were conducted using the telephone Dementia Questionnaire [24]. With information from the informant interviews, a neurologist from the Massachusetts Alzheimer Disease Research Center determined dementia status in the 80 women. Those who scored ≥ 31 points on the TICS had an eightfold increased risk of dementia diagnosis. For scores on the verbal memory tests, there was a 12-fold increased risk of dementia diagnosis for women with a poor episodic-memory score as compared with those who performed well. All of these findings were statistically significant. This study is a strong clinical validation of the effectiveness of telephone tests in evaluating cognition.

Finally, perhaps the most important validation of telephone cognitive testing is the ability to detect established risk factors for cognitive impairment or decline. One of the most established risk factors for cognitive impairment is apolipoprotein E genotype. Approximately 4000 women in the NHS were genotyped [25]. Cognitive impairment was defined as a score falling within the worst 10% of the distribution of cognitive scores. For women with an $\epsilon 4$ allele, the odds ratio was 1.5, or a 50% higher risk of cognitive impairment. For those who were homozygous for $\epsilon 4$, the odds ratio was 2.6, with a highly significant trend of increased risk of cognitive impairment with increased number of $\epsilon 4$ alleles. These results are similar to those that would be expected in any study that assessed cognition in person. To further test ability to measure specific cognitive domains, the relationship between the $\epsilon 4$ allele and verbal memory was considered. Mean scores across four tests of verbal memory in women who had an $\epsilon 4$ allele were compared with those from women who did not [25]; women with an $\epsilon 4$ allele scored significantly worse than those with no $\epsilon 4$ allele, by 0.15 standard units. For comparison, ROS investigators conducted a similar analysis using their verbal memory tests from an in-person interview [26]. They found a mean difference of 20.1 standard units in verbal memory score for women with versus those without an $\epsilon 4$ allele—almost identical to the NHS findings. Again, this suggests that the NHS telephone cognitive battery and in-person cognitive assessments yield equivalent findings for risk factor relationships.

Other groups have also addressed the utility of a brief telephone cognitive battery for assessing cognition. For example, Wilson and Bennett considered one that tested episodic memory, working memory, semantic memory, and global cognitive function in 996 participants of the ROS [27]. They found very strong relationships between apolipoprotein E genotype and each cognitive system, and between each of these cognitive systems and AD pathology. This study also demonstrated the validity of a brief cognitive battery that can be administered by telephone.

Although validity is critical, test–retest reliability is important as well. The TICS was administered twice to the same group of 50 women in the NHS, within a gap of 31 days. There was a correlation of 0.7 in the scores, within acceptable standards for reliability measures in cognitive tests. This reliability test was performed in generally highfunctioning subjects; therefore, separate data were not available for those with MCI or dementia.

However, Brandt and Folstein [19] have reported test–retest reliability of 0.97 for the TICS in 100 subjects with AD.

Because administering cognitive tests to large groups of subjects requires many telephone interviewers, interinterviewer scoring reliability was also tested. Among a large team of telephone interviewers, there can be some level of judgment in deciding on how to score a given response within a cognitive test, and it is important that the scores are similarly scored regardless of the interviewer. In one study (F.D. Grodstein, unpublished, 1998), 10 telephone interviewers scored the same set of five cognitive interviews. A very high correlation of test scores across interviewers was found, .0.95 for each cognitive test. One important aspect of achieving high inter-interviewer reliability is to maintain strict and conservative scoring rules, diminishing the possibility of subjective judgment in the scoring.

In conclusion, an important advantage of telephone cognitive testing is its ease of administration and low cost. Further benefits include high participation because participants do not need to travel or to set aside significant amounts of time to participate. High participation enhances the validity of results, and also increases generalizability, such that larger segments of the population can participate instead of only a select group that is willing or able to participate in extensive in-person examinations.

There are disadvantages as well. The telephone method does not allow for certain types of cognitive tests. For example, visuospatial perception and naming, both vulnerable to early changes associated with DAT, cannot be tested. In addition, a potential disadvantage of telephoning is the possibility of hearing impairment interfering with performance, especially in the older subjects and especially in men, in whom hearing impairment can be more common than in women [28]. Some encouraging results were obtained in a pilot study comparing telephone cognitive tests in men aged ≥ 85 years from the Health Professionals Follow-up Study with scores obtained by separate but similar cohorts on tests administered in person (F.D. Grodstein, unpublished, 2002). First, there was 94% participation in the cognitive assessment, which is extremely high for this age group. Second, regarding the validity of performance, scores on the telephone tests were similar to those obtained from in-person testing of the oldest old. For example, on the test of category fluency (animal naming), the mean score among the men from the Health Professionals Follow-up Study was 15.6, as compared with the mean of 15.1 for the Consortium to Establish a Registry for Alzheimer’s Disease neuropsychological battery norm for the oldest-old men and women [29]. In the Digit Span Backwards Test, the mean was 5.5 in the Health Professionals Follow-up Study, versus 5.6 in men and women from the 901 Study [30].

The similarity of results between these men who were given a telephone cognitive battery and comparable subjects who were given in-person testing not only suggests that hearing ability may not meaningfully interfere with performance on the telephone, but also indicates that other issues (e.g., writing down word lists, background distractions) are not important limitations because one would not expect to find equivalent performance if there were substantial problems with administration of the telephone tests. In addition, in this same pilot study from the Health Professionals Follow-up Study, hearing ability was evaluated; as expected, 59% of participants reported at least some hearing problems. However, this did not affect participation, nor did it seem to affect performance. Interviewers asked each participant about their hearing ability and adjusted the volume of their voices on the basis of the subjects’ needs; importantly, 11 of 15 items in the telephone cognitive battery can be repeated if the participant has difficulty in hearing the question. On the TICS, the mean score was 31.2 among the men with hearing problems and 31.4 among those without; on category fluency, the mean was 15.8 among those with hearing problems and 15.2 among those without; on the 10-word delayed recall, the mean was 3.3 among the men with hearing

problems and 3.5 among those without; and on the East Boston Memory Test, the mean was 8.7 among the men with hearing problems and 8.2 among those without. Therefore, these data do not indicate that hearing ability severely influences results on a telephone cognitive battery.

In brief, a telephone instrument to measure cognitive function can provide meaningful information regarding cognitive status in large populations, at low cost and with good validity and test-retest reliability.

4. Approaches for detecting the earliest memory impairments in the elderly population: Implications for clinical and epidemiological studies

Researchers interested in the assessment of cognitive impairment in the elderly population face substantial challenges. There is considerable interest in using the most sophisticated methods to investigate subtle neuropsychological impairments associated with the earliest stages of neurodegenerative disease [31]. However, at the same time, the complexity of current research questions and the need to capture many different aspects of disease, from clinical signs [32] to biomarkers [33], have led to pressure to use shorter cognitive batteries to allow for the inclusion of additional types of measures. In practical terms, this has often meant the simple adaptation of many instruments that have long been used in dementia research, without evidence that they have the requisite psychometric properties or provide adequate sensitivity and/or specificity for conditions such as MCI or even more subtle forms of cognitive impairment.

Neuropsychological performance for any individual is determined by a complex array of variables including but not limited to cognitive reserve, premorbid strengths and weaknesses, engagement in testing, attentional capacity, sensory acuity, and fatigue, to name a few. Because any single cognitive measure can be affected by any one or a combination of these factors, neuropsychologists often examine an individual's pattern of performance across several different indices, comparing that performance with normative data, rather than relying on a single test in isolation [34]. Current trends, however, are to use limited numbers of measures and to use cut-offs for performance on the basis of the number of standard deviations that, for an individual, fall below a reference point set for patients of similar age, education, and cultural background [35].

The most sophisticated clinical and epidemiological studies of cognitive impairment in the elderly population assess different aspects of neuropsychological function at baseline and ensure that their measurements have sufficient range to track changes in different cognitive domains longitudinally. The optimal cognitive battery may, at minimum, include tests of (1) learning and retentive memory, (2) executive function, (3) language, and (4) visuospatial skills [31]. It is also desirable to have measures of both attention and processing speed because these are often impaired by a variety of brain disorders and, thus, serve as a more general marker of impairment.

Because memory dysfunction is a hallmark feature of AD and is present in many neurodegenerative disorders, this section will focus on selecting memory tests for studies of MCI and the trade-offs between efficiency and comprehensiveness of a cognitive battery.

Many neuropsychologists favor verbal list learning tasks for clinical work, longitudinal investigations of aging, clinical trials, and epidemiological studies. The advantage of this type of memory paradigm is that (1) there is more than one presentation of the to-be-remembered stimuli, thus minimizing the effect of poor attention on performance; (2) learning curves over repeated trials can be examined; (3) the processes of encoding,

acquisition, and storage of to-be-remembered information as well as specific types of memory can be distinguished; and (4) many measures also have parallel and alternate forms.

In choosing a list learning test, the number of to-be-remembered items should be sufficient to be challenging for high-functioning subjects (12 to 16 items is optimal). In populations where there may be cultural and/or language issues, alternative list learning tasks, such as, for example, using actual objects and making the test more challenging by requiring the subject to alternate between learning trials and distractor tasks may be beneficial [36].

There are many other memory tasks that are also well justified in the cognitive assessment of older adults including, but not limited to, paired associates learning, immediate and delayed recall of simple and increasingly complex geometric designs, spatial object location, face-name association, and paragraph recall.

For neuropsychologists, the most widely used paragraph recall test is the Logical Memory subtest on the Wechsler Memory Scale [37,38], which assesses immediate and delayed recall for two passages, each containing 25 elements. In one newer version of the test, the second passage is repeated to assess learning over more than one trial. Although Logical Memory takes no longer than 10 minutes to administer, the pressures to abbreviate the testing have been sufficiently powerful, such that a standard in the field today is to use a single paragraph to assess memory function. In fact, delayed recall for a single story passage is used in several significant national clinical research initiatives as the only memory measure in addition to a Clinical Dementia Rating [39] clinical score to derive a diagnosis of amnesic MCI. However, a potential problem with this approach is that factors such as auditory difficulties or inattention can greatly affect performance on this cognitive test. Conversely, persons with seemingly intact story memory, especially those with above-average premorbid function, might still have underlying difficulties with other aspects of memory. The solution to this problem is to have tests that are not as prone to attentional effects, have repeated trials, and are more sensitive to early, mild impairment. It is likely that these issues will only become more acute as investigators attempt to identify MCI at even earlier stages.

In their proposed revision of research criteria for AD, Dubois et al [10] contend that using paradigms that increase encoding specificity at acquisition and that assess failure to benefit from cueing at recall are superior to episodic memory tests using free recall alone. Buschke et al [40], using the Selective Reminding procedure, first observed that probing with the same semantic cues used for learning and retrieval was superior to testing free recall alone and was also superior to paired associate learning and Logical Memory in distinguishing mild dementia from normal aging. More recently, Buschke developed the Memory Capacity Test, which requires an individual to learn 16 items from different semantic categories with cued recall and then to learn 16 new items using the same semantic categories with cued recall. Cued recall for the second list and 30-minute delay of the second list of the Memory Capacity Test was particularly useful in identifying amyloid-positive normal elderly individuals with high cognitive reserve who were at greater risk for AD because of their increased amyloid load [41,42].

Loewenstein et al have focused on determining the extent to which vulnerability to semantic interference may identify cases of early AD [36]. The modified three-trial Fuld Object Memory Evaluation paradigm was extended by having subjects recall a second list of items that are all semantically similar to the original to-be-remembered targets (i.e., ring versus bracelet, key versus lock) [43]. Reduced recall for the second list as compared with the first list was thought to occur because of competition from the previously presented targets on the first list (proactive interference), whereas reduced recall for the first list after recall of

the second list was thought to be related to retroactive interference. The Semantic Interference Test (SIT) [36] evidenced high sensitivity and specificity in distinguishing normal elderly subjects from those with MCI and early dementia. Moreover, vulnerability to proactive interference was most associated with those subjects with MCI who progressed to dementia over a period of 2 to 3 years [44]. Recently, the SIT has been validated for use in epidemiological investigations [45].

Studies by Buschke et al and Loewenstein et al focus on semantic information processing deficits that may be specific for early AD and support the need to develop memory tests that optimize attention to the to-be-remembered stimuli and emphasize lack of encoding specificity.

For epidemiological studies, briefer memory tests may be required. The Florida Brief Memory Screen (FBMS) [46], which takes approximately 3 to 4 minutes to administer, was recently developed to identify those in need of further evaluation. The primary objective is for the subject to attend to and register the to-be-remembered targets. Free recall is then assessed. The FBMS has been found to be highly reliable and correctly classify 100% of patients with AD, 82.6% of individuals with amnesic MCI, and 87.5% of normal elderly controls without cognitive impairment [46]. Importantly, the FBMS scores were generally independent of age, education, and primary language (English versus Spanish). Although preliminary, these results suggest that a one-trial memory measure may have potential as a screening measure for amnesic MCI and dementia. For aforementioned reasons, the FBMS is not viewed as being sufficiently broad or having an adequate number of trials to be used as a stand-alone memory measure any more than the three words of the Mini-Mental State Examination [47] or the five words of the Montreal Cognitive Assessment Screen [48] would serve on their own to assess memory.

Because tests are developed to assess cognitive impairment in its earliest stages, the increased sensitivity may involve a trade-off of decreased specificity. Because false-positive and false-negative classification errors are related to underlying base rates of true impairment in a particular population, this trade-off needs to be evaluated with regard to the research questions being posed and the costs associated with misidentification.

Although there are many pressures to abbreviate memory measures, recent work with cued recall and semantic interference paradigms suggests that assessments may have to become even more detailed and comprehensive to capture subtle cognitive changes. A new era has commenced, where it is possible to obtain imaging and other biomarkers that point to cerebral compromise years before clinical symptoms are manifested [41]. Instead of a focus on how to reduce existing tests, there needs to be continued development of measures that (1) can use the person as his/her own control in evaluating decrements in performance, (2) reveal vulnerabilities in the cognitive system that cannot be masked by high education and cognitive reserve, and (3) have acceptable levels of specificity. The dementia field deserves no less.

5. Toward the future: Smart homes and cognitive assessment

A major challenge in clinical research is the identification of several kinds of events, such as those that are either rare, irregularly occurring, and evanescent (e.g., falls, episodic sleepiness, transient neurological events), or those that evolve slowly over time and may have poor demarcation with regard to onset and transition to new clinical states (e.g., depression, cognitive impairment). Currently, survey and assessment methodologies tend to rely on sparsely spaced queries either by questionnaire or in-person examinations. These depend on recall of events or a brief snapshot of function that assumes that observations recorded during the examination are representative of the persons' typical state of function

over relatively long periods. Consequently, it is difficult to identify data or events with detailed temporal precision and intra-individual specificity. They also are of limited ecological validity. A way to improve this state of affairs is to bring the locus of assessment into the home, providing an opportunity to record in real-time events as they occur. Ideally, this is done with minimal obtrusiveness, “in the background” of daily activity. Recent advances in ubiquitous computing and sensor network methodologies along with other convergent technologies and methods such as the techniques of ecological momentary assessment, telemedicine, and activity-monitoring provide a background for achieving these goals. The fully instantiated version of this approach to assessment has further evolved from the concept of the “Smart Home.”

The term “smart home” is often used to describe the general concept of a home that is outfitted with technology that automates the detection of events and performs activities or functions for the person living inside that structure without the resident needing to consciously engage with the technology. In the building industry, this home automation (sometimes called “domotics”) is most commonly focused on environmental engineering (e.g., cooling/heating, lighting, home entertainment systems, alarms).

In the healthcare or biomedicine sphere, the concept is extended to describe a home where technology is placed or used to automate data capture or to provide interventions relevant to health [49–51]. This ranges from automated assessment of motion to detect relevant events such as walking, falls, or general activity levels (e.g., through passive infrared motion sensors located around the home); activities of daily living (e.g., instrumenting a pillbox or medication dispenser); physiological function (e.g., vital signs monitored using an embedded scale in the floor or bed); and degree of socialization (e.g., using motion sensors to measure the amount of time spent outside the home, or with telephone monitoring, number of telephone calls). Ideally, these technologies “talk” to one another and form an intelligent sensor net that acts in the background to unobtrusively identify ongoing patterns of health or intermittent acute events. These data are then available for analysis using various methodologies that may directly record events or infer them through prediction models using the dense continuous data streams generated through these assessments.

To realize the power of transitioning from the current assessment approach (infrequently acquired, conducted in an artificial setting, scheduled at the examiners’ convenience, time limited, and recall-biased) to a more continuous, naturalistic, always-on paradigm, one needs to consider how such a goal may be achieved. Several key elements must be taken into account. These include first changing the principal locus of assessment from clinics or “appointments” (even when conducted as a home visit) to assessments that are integrated to the daily life of the home and community of the individual. Second, events should be recorded in real time or near real time. Timing may depend on how important the immediacy of the data is to the research questions. In the case of intervention studies, one may wish to know trends immediately so as to intervene or provide feedback quickly. In risk factor studies, this timing may not be as critical on the reporting side. However, an advantage of pervasive computing technology is that it can improve on current methods by having data that are inherently timestamped, thereby allowing for unusual temporal precision in connecting one risk factor to another. Third, the recording of activity or events of interest should be minimally obtrusive, occurring ideally in the background of a person’s usual daily routines. This quality has been referred to as “ambient assessment.” Finally, the recording of these data should be considered to be a continuous data stream.

As noted previously, the embodiment of this kind of approach comes from the tradition of the smart home. Applications based on this approach in the healthcare domain have, until recently, used it with the relatively limited purpose of serving as a demonstration apartment,

or in a limited number of free-standing homes to show feasibility of various technologies for assessing or maintaining function toward health applications. In these laboratory settings, the apartment or home has been outfitted with technologies for monitoring of function for a few hours or days [49–51]. Obviously, this does not represent a real-world setting and does not easily generalize to typical older persons or to a chronically ill population. The other highly relevant thread that applies to the home-based assessment model taking advantage of technology is in the classic application of telemedicine [52]. Here, there is quite a long experience, but most of the work in this area has focused on single domain or condition assessment such as physiological monitoring for diabetes or heart failure or mood assessment for depression. There has not been until recently a more integrated approach that attempts to fuse multiple data streams (physiological, cognitive, behavioral, environmental, etc.) in a continuous manner on a large scale.

As a means to achieve this goal, the Oregon Center for Aging and Technology, a National Institute on Aging-supported Roybal Center, has developed an approach to continuous in-home assessments geared toward the elderly population for assessment of key functions that together lead to loss of independence, or if maintained, successful aging [53,54]. Thus, the goals of the assessment platform are to be able to continuously assess primarily those functions that are critical to home independence such as cognitive function, psychological status, health events (e.g., falls, hospitalizations or emergency room visits, medication changes) and mobility. Figure 1 shows the layout of an individual home and an example of the technology installed for monitoring of these functions in a continuous ambient manner.

The backbone of this system is a network of passive infrared motion sensors that are strategically placed about the home to provide several key measures of function. These include total activity across several different time scales (e.g., activity per day, night, week, weekend, season, after or before a health event such as a fall or illness) and the location where the activity occurs. These same inexpensive sensors can be field-restricted and placed in series in a location trafficked throughout the day or night in the home to provide a measure of walking speed, a metric tied to many meaningful outcomes in gerontology including cognitive function, independent living, morbidity, and even mortality [55]. An instrumented pillbox can time-stamp the daily opening of the correct pill compartment, and thus measure the ability to adhere to medication regimens [56,57].

The home personal computer is a key element in this setup. It provides a multilevel assessment of function ranging from psychomotor speed, such as log-in typing speeds or keyboarding function or assessments of mouse movements, to “incidental” measures of cognitive function through assessment of simple trends in daily computer use ranging from time on the computer to time to respond to e-mails [58,59]. More structured assessment can be obtained through other means such as game playing performance over time where the games are designed to tap specific cognitive domains [60]. Although more obtrusive, one can still ask subjects to take formal cognitive tests online; doing this conveniently at home provides the opportunity to maximize the comfort level of test administration as well as easily vary the frequency of test administration. In addition, the home personal computer provides an opportunity to query volunteers directly throughout the year on a frequent basis. Experience suggests that a weekly e-mail is reliably answered to such recurring questions as, for example, “Did you fall? Did you have overnight visitors? Have you been feeling down or blue?”

Of course, this computer use requires that volunteers are taught to use computers if they do not know how to in the first place, a common need for the current elderly generation. However, because the next generation of elderly individuals—the Baby Boomers—are now approximately 75% online [61], this training need will recede. Even with growing numbers

of elderly individuals over time having increasing capacity for taking advantage of home-based technology, issues surrounding attitudes and beliefs about the appropriate use of these monitoring systems in everyday life need to be addressed. What types of data and under what circumstances elderly individuals will want to have this real-time information about themselves sent to various stakeholders, such as their doctors or their family members, remains to be clarified. Preliminary data suggest that the acceptance of unobtrusive in-home monitoring is closely tied to the perceived utility of the data generated and, accordingly, concerns about the privacy or security of this information are balanced by the value that these data may provide to the various stakeholders in the care ecosystem of the elderly individuals that can support their independence or optimize health [62].

It is thus important that we begin to understand how to best use this information and harness the power of these approaches as soon as possible. The input/output potential of the personal computer itself will be transformed, such that the distinctions among a desktop, a laptop, a tablet, a smartphone, or even a heads-up display in your mirror or embedded in appliances will become muted. Automated speech recognition technology will allow for more immediate analysis of spoken responses. This will extend beyond the simple capture of test question response accuracy to an entire new science of assessing affect and social interactions in real time. Ultimately, we will need to learn new ways to analyze these rich multidimensional data streams where the metrics for outcomes may become very different (Fig. 2) [63]. Thus, we may find that, instead of asking people how their memory function is or how they have felt in the past week, we now can more precisely measure real instances of activity patterns and how the individual felt about any changes (Fig. 3) [64]. Finally, the ability to perform this kind of automated assessment across hundreds or thousands of households holds the promise for a new kind of epidemiology that is conducted in real time and provides an unprecedented ability to connect activities that are currently difficult or impossible to capture by relying on recall of events.

6. Discussion

In 2007, life expectancy at birth for the total population in the United States reached 77.9 years [65]. This stands in sharp contrast to 47.3 years in 1900. The greatest risk factor for DAT is age and, thus, increased longevity is a double-edged sword. The issues discussed in this article lie at the heart of ascertainment of dementia in support of epidemiological studies. The various sections share several themes, but each also highlights some unique challenges faced in ascertaining cognitive decline and dementia in the community at large. All emphasize the need for brief, cost-effective measures that could enhance the ability to collect data from large, geographically distributed, and demographically diverse populations. Three themes emerged that can be emphasized as follows: (1) A suitable target of measurement. Should the outcome be an individual “slope of cognitive decline” or a distinct diagnostic endpoint such as MCI or AD? (2) A suitable instrument. What is the optimum measure of cognitive change and dementia? What form should it take? Is an overall, global cognitive score, a profile of performance along several cognitive dimensions, a single in-depth probe of episodic memory, or a functional measure with highest ecological validity most optimal? (3) If tests are administered, should they be conducted in person by a clinician and/or a technician or in the comfort of one’s home by telephone, Internet, or “in-the-background” through electronic monitoring? All sections emphasize the importance of early detection so that interventions, when they become available, could be introduced at a time point when they might prevent or slow decline.

At the crux of all these considerations lie our assumptions about what is “normal” for cognitive aging. Virtually every cross-sectional study of cognitive aging has observed age-related decline in scores on so-called “fluid” cognitive measures such as episodic

memory and mental speed of processing tests [66]. This finding is so well-established that we have come to expect that “normal” for an 80-year-old can represent a loss of 50% in a raw test score as compared with “normal” for a 30-year-old. In fact, normative tables for standard neuropsychological measures, such as the Wechsler Memory Scale [38], show this phenomenon, schematically illustrated in Fig. 4. For example, an 80-year-old can earn a scaled score of 10 (average) with a raw score of 13/50 on the Delayed Recall portion of the Logical Memory subtest. However, this represents an almost 50% reduction from a raw score of 24/50 that would earn a 30-year-old the same scaled score. Therefore, clinical judgments about the normalcy of cognitive test performance in older individuals assume loss as a given rather than as an end product of a long developmental process with variable speed of occurrence.

This bias is magnified when trying to determine whether an average-for-age score is acceptable for an 80-year-old who, as a 30-year-old, was well above average for age. Thus, two 80-year-olds can obtain the same test score, but one may actually be declining and the other may not. The nature of cognitive test scores, as discussed later in the text, makes them difficult to interpret without adjusting for several sociodemographic variables. To address this conundrum, Evans (section 2) makes the case for the importance of capturing intra-individual rates of change as well as diagnostic endpoints in epidemiological studies of age-related cognitive change. The individual is her/his own control and no automatic assumptions of “normal” decline need to be integrated into decision-making. Results from the Alzheimer’s Disease Neuroimaging Initiative have shown that cerebrospinal fluid and neuroimaging biomarkers associated with Alzheimer’s pathology precede the onset of symptoms by many years [67]. Therefore, a steep “personal slope” of cognitive change over time, even though test scores remain within normal limits for age, might signal greater vulnerability to dementia in future years. Several strategies for estimating decline against a standard of a peak prior level of ability have been suggested and exemplified in the work of Rentz et al [42,68–70]. However, in individuals with limited education or illiteracy, it may be more difficult to derive such an estimate.

Methods and measures of ascertainment were the topics covered in sections 3 to 5. A cognitive measurement that captures a single time point is problematic because, unlike some physiological health measures, the distinction between normal and abnormal can be blurred depending on age, education, racial/ethnic background, and other sociodemographic factors. Thus, there is no cognitive value akin to blood pressure that is either absolutely normal or abnormal. Another problem is that the baseline is unknown and, as aforementioned, can only be estimated. The lack of a baseline cognitive marker is a major handicap to a clinician and to researchers attempting to determine whether there has been a change. Annual health checks typically provide indices of physical health, such as blood pressure, cholesterol levels, bone density values, and cancer screenings, which can then be tracked and monitored to detect deviations from baseline values and from normal. But, to date, virtually no older individuals have had a baseline cognitive screening. In an unprecedented acknowledgement of the importance of cognitive health and change as an index of illness, Medicare has ruled that, as of January 2011, all beneficiaries be screened for cognitive impairment as part of an annual wellness benefit. This should be viewed as a major first step in the right direction to highlight the importance of cognitive and/or brain as well as physical and/or body health. However, from a clinical perspective, cognitive screening is most likely to identify dementia, and not risk for dementia, especially in previously high-functioning individuals. Thus, it could be recommended that neuropsychological assessment, or a “whole brain” health check, be carried out at a time when individuals are healthy and functioning well, in their 30s, 40s, and 50s, to establish a baseline against which cognitive decline could be detected. Such a practice will also make the detection of dementia less reliant on functional

activity changes, which can be very difficult to assess without knowledge of functional baselines and which are a lot less easy to characterize than objective cognitive test scores.

Inter-individual variability in indices of cognitive (and other) performance is a universal phenomenon, but the variability increases exponentially as the age scale is ascended. In a study comparing physicians ranging in age from 28 to 92 years on a computerized cognitive battery, it was reported that average mean scores declined but standard deviations increased dramatically from the youngest to the oldest age groups [71]. This is consistent with the view that aging is associated with increased heterogeneity and gives rise to the expectation that some older individuals may retain their cognitive abilities at a level that is, at a minimum, average for their much younger cohorts. The ability to measure individual rates of decline can capture the different cognitive aging trajectories that individuals will take over time, from “super” or very little change, to “usual” or average change, to “abnormal” or accelerated change.

Grodstein (section 3) offered evidence that a telephone-based test of mental state could provide a cost-effective method of ascertainment of cognitive functioning with good compliance for conducting repeated assessments. With the exception of naming and visuospatial perception, the telephone interview can survey domains vulnerable to AD, including attention, episodic memory, and word fluency. The instruments had good construct and criterion validity when compared with an in-person cognitive battery, and had strong inter-rater reliability. Furthermore, reduced hearing acuity, increasingly common among older individuals, did not constitute a significant deterrent to telephone screening for many participants. Telephones are plentiful in the United States and provide an acceptable, “low-tech” option for wide-scale assessment.

An alternative approach to a global measure was offered by Loewenstein (section 4), which suggested that measurement should focus on the cognitive domain most sensitive to AD, namely episodic memory, and its component operations. Mental operations entailed in encoding and retrieval have been validated against structural and functional brain imaging variables associated with “normal aging”, MCI, and early AD [72–76]. Not only does the ability to acquire, retain, and retrieve episodic information decay with AD but there is increased sensitivity to semantic interference also [36]. Thus, mechanisms of proactive and retroactive interference serve to erode memories, and tests based on the careful probing of these mechanisms can differentiate between healthy and diseased memory [36,77–79]. Furthermore, methods that control for encoding to ensure that the information to be remembered has been attended to and processed have also yielded important insights into normal and abnormal memory in aging [80–83]. The trade-off between a more domain-focused and a more global cognitive index approach to measurement is that the former requires clinician involvement and more time for test administration and scoring. The SIT [43] is offered as an instrument derived from principles of the cognitive/neuropsychological approach.

Kaye (section 5) considered a revolutionary approach to healthmonitoring and ascertainment of cognitive impairment, the “Smart Home.” Making optimum use of today’s technology (keeping in mind “big brother” concerns), moment-to-moment monitoring of activity patterns, appliance use, and other surrogate/functional indices of cognitive function, offers an unprecedented opportunity to focus on individual patterns and habits and detect deviations at a behaviorally microscopic level. Increased frequency of falls, missed medication dosages, and decreased use of the stove may signal underlying cognitive deficits. Although intra-individual variability has been explored as a construct in cognitive aging [84,85], there are only a handful of studies that have shown that this factor could serve to distinguish those who are cognitively normal from those experiencing decline [86–89].

“Smart home” technology could permit early detection of subtle variability before symptoms cross the line into the abnormal range. Although the technologies to deploy such systems are readily available, the feasibility of this approach remains to be tested and implemented on a large scale.

In conclusion, ascertainment of dementia at the population level creates many challenges. On the basis of the issues discussed in this article, how we define our endpoints and measurements will have a significant effect on estimating prevalence and incidence. The looming healthcare crisis of aging and dementia will require that a cognitive health check be as important to the primary care setting as cancer screenings and checks on cardiovascular, orthopedic, and sensory motor health.

Acknowledgments

This combined effort was supported by National Institute of Aging grants R01AG011101, R01AG024059, R01AG024215, P30AG008017, P30AG024978, and P30AG013854. The research was also supported by the Intel Corporation. The authors have no conflicts to disclose. The sponsors had neither a role in the analysis or interpretation of these data, nor in the content of the article. Appropriate approval procedures were used concerning human subjects.

References

1. Weir DR, Wallace RB, Langa KM, Plassman BL, Wilson RS, Bennett DA, et al. Reducing case ascertainment costs in U.S. population studies of Alzheimer’s disease, dementia, and cognitive impairment—Part 1. *Alzheimers Dement* 2011;7:94–109. [PubMed: 21255747]
2. Bennett DA, Schneider JA, Arvanitakis Z, Kelly JF, Aggarwal NT, Shah R, et al. Neuropathology of older persons without cognitive impairment from two community-based clinical-pathologic studies. *Neurology* 2006;66:1837–44. [PubMed: 16801647]
3. Hulette CM, Welsh-Bohmer KA, Murray MG, Saunders AM, Mash DC, McIntyre LM. Neuropathological and neuropsychological changes in “normal” aging: evidence for preclinical Alzheimer disease in cognitively normal individuals. *J Neuropathol Exp Neurol* 1998;57:1168–74. [PubMed: 9862640]
4. Haroutunian V, Purohit DP, Perl DP, Marin D, Khan K, Lantz M, et al. Neurofibrillary tangles in nondemented elderly subjects and mild Alzheimer disease. *Arch Neurol* 1999;56:713–8. [PubMed: 10369312]
5. Davis DG, Schmitt FA, Wekstein DR, Markesbery WR. Alzheimer neuropathologic alterations in aged cognitively normal subjects. *J Neuropathol Exp Neurol* 1999;58:376–88. [PubMed: 10218633]
6. Knopman DS, Parisi JE, Salviati A, Floriach-Robert M, Boeve BF, Ivnik RJ, et al. Neuropathology of cognitively normal elderly. *J Neuropathol Exp Neurol* 2003;62:1087–95. [PubMed: 14656067]
7. Resnick SM, Sojkova J, Zhou Y, An Y, Ye W, Holt DP, et al. Longitudinal cognitive decline is associated with fibrillar amyloid-beta measured by [11C] PiB. *Neurology* 2010;74:807–15. [PubMed: 20147655]
8. Storandt M, Mintun MA, Head D, Morris JC. Cognitive decline and brain volume loss as signatures of cerebral amyloid-beta peptide deposition identified with Pittsburgh compound B: cognitive decline associated with Abeta deposition. *Arch Neurol* 2009;66:1476–81. [PubMed: 20008651]
9. Hall CB, Lipton RB, Sliwinski M, Stewart WF. A change point model for estimating the onset of cognitive decline in preclinical Alzheimer’s disease. *Stat Med* 2000;19:1555–66. [PubMed: 10844718]
10. Dubois B, Feldman HH, Jacova C, DeKosky ST, Barberger-Gateau P, Cummings J, et al. Research criteria for the diagnosis of Alzheimer’s disease: revising the NINCDS ADRDA criteria. *Lancet Neurol* 2007;6:734–46. [PubMed: 17616482]
11. Petersen RC, Knopman DS, Boeve BF, Geda YE, Cha R, Roberts R, et al. Predicting Alzheimer’s disease. *Alzheimers Dement* 2009;5 (Suppl 1):130. [PubMed: 19328442]
12. Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics* 1982;38:963–74. [PubMed: 7168798]

13. Fillenbaum GG, Hughes DC, Heyman A, George LK, Blazer DG. Relationship of health and demographic characteristics to Mini-Mental State Examination score among community residents. *Psychol Med* 1988;18:719–26. [PubMed: 3263663]
14. Teresi JA, Golden RR, Cross P, Gurland B, Kleinman M, Wilder D. Item bias in cognitive screening measures: comparisons of elderly white, Afro-American, Hispanic and high and low education subgroups. *J Clin Epidemiol* 1995;48:473–83. [PubMed: 7722601]
15. Shadlen MF, Siscovick D, Fitzpatrick AL, Dulberg C, Kuller LH, Jackson S. Education, cognitive test scores, and black-white differences in dementia risk. *J Am Geriatr Soc* 2006;54:898–905. [PubMed: 16776783]
16. Manly JJ, Jacobs DM, Sano M, Bell K, Merchant CA, Small SA, et al. Effect of literacy on neuropsychological test performance in nondemented, education-matched elders. *J Int Neuropsychol Soc* 1999;5:191–202. [PubMed: 10217919]
17. Manly JJ, Touradji P, Tang MX, Stern Y. Literacy and memory decline among ethnically diverse elders. *J Clin Exp Neuropsychol* 2003;5:680–90. [PubMed: 12815505]
18. Seshadri S, Beiser A, Au R, Wolf PA, Evans DA, Wilson RS, et al. Operationalizing diagnostic criteria for Alzheimer's disease and other age-related cognitive impairment—Part 2. *Alzheimers Dement* 2011;7:35–52. [PubMed: 21255742]
19. Brandt, J.; Folstein, MF. Telephone Interview for cognitive status: professional manual. Lutz, FL: Psychological Assessment Resources, Inc; 2003.
20. Albert M, Smith LA, Scherr PA, Taylor JO, Evans DA, Funkenstein HH. Use of brief cognitive tests to identify individuals in the community with clinically diagnosed Alzheimer's disease. *Int J Neurosci* 1991;57:167–78. [PubMed: 1938160]
21. Abraham E, Axelrod BN, Ricker JH. Application of the Oral Trailmaking Test to a mixed clinical sample. *Arch Clin Neuropsychol* 1996;11:697–701.
22. Weintraub S, Salmon D, Mercaldo N, Ferris S, Graff-Radford NR, Chui H, et al. The Alzheimer's Disease Centers' Uniform Data Set (UDS): the neuropsychologic test battery. *Alzheimer Dis Assoc Disord* 2009;23:91–101. [PubMed: 19474567]
23. Wilson RS, Beckett LA, Barnes LL, Schneider JA, Bach J, Evans DA, Bennett DA. Individual differences in rates of change in cognitive abilities of older persons. *Psychol Aging* 2002;17:179–93. [PubMed: 12061405]
24. Kawas C, Segal J, Stewart WF, Corrada M, Thal LJ. Avalidation study of the Dementia Questionnaire. *Arch Neurol* 1994;51:901–6. [PubMed: 8080390]
25. Kang JH, Logroschino G, De Vivo I, Hunter D, Grodstein F. Apolipoprotein E, cardiovascular disease and cognitive function in aging women. *Neurobiol Aging* 2005;26:475–84. [PubMed: 15653176]
26. Wilson RS, Schneider JA, Barnes LL, Beckett LA, Aggarwal NT, Cochran EJ, et al. The apolipoprotein E e4 allele and decline in different cognitive systems during a 6-year period. *Arch Neurol* 2002;59:1154–60. [PubMed: 12117364]
27. Wilson RS, Bennett DA. Assessment of cognitive decline in old age with brief tests amenable to telephone administration. *Neuroepidemiology* 2005;25:19–25. [PubMed: 15855801]
28. Dillon, CF.; Gu, Q.; Hoffman, HJ.; Ko, CW. NCHS data brief. Hyattsville, MD: National Center for Health Statistics; 2010. Vision, hearing, balance and sensory impairments in Americans aged 70 year and older: United States, 1999–2006.
29. Beeri MS, Schmeidler J, Sano M, Wang J, Lally R, Grossman H, et al. Age, gender, and education norms on the CERAD neuropsychological battery in the oldest old. *Neurology* 2006;67:1006–10. [PubMed: 17000969]
30. Whittle C, Corrada MM, Dick M, Ziegler R, Kahle-Wroblewski K, Paganini-Hill A, et al. Neuropsychological data in nondemented oldest old: the 901 Study. *J Clin Exp Neuropsychol* 2007;29:290–9. [PubMed: 17454349]
31. Nelson AP, O'Conner MG. Mild cognitive impairment: a neuropsychological perspective. *CNS Spectr* 2008;13:56–64. [PubMed: 18204415]
32. Black R, Greenberg B, Ryan JM, Posner H, Seeburger J, Amatniek J, et al. Scales as outcome measures for Alzheimer's disease. *Alzheimers Dement* 2009;5:324–39. [PubMed: 19560103]

33. Craig-Schapiro R, Fagan AM, Holtzman DM. Biomarkers of Alzheimer's disease. *Neurobiol Dis* 2009;35:128–40. [PubMed: 19010417]
34. Jacobson MW, McEvoy LK, Dale A, Fennema-Notestine C. Cognitive phenotypes, brain morphometry and the detection of cognitive decline in preclinical AD. *Behav Neurol* 2009;21:29–37. [PubMed: 19847043]
35. Luis CA, Loewenstein DA, Acevedo A, Barker WW, Duara R. Mild cognitive impairment: directions for future research. *Neurology* 2003;61:438–44. [PubMed: 12939414]
36. Loewenstein DA, Acevedo A, Luis C, Crum T, Barker WW, Duara R. Semantic interference deficits and the detection of mild Alzheimer's disease and mild cognitive impairment without dementia. *J Int Neuropsychol Soc* 2004;10:91–100. [PubMed: 14751011]
37. Wechsler, D. *The Wechsler Memory Scale-Revised*. San Antonio, TX: The Psychological Corporation; 1987.
38. Wechsler, D. *The Wechsler Memory Scale*. 3. San Antonio, TX: The Psychological Corporation; 1997.
39. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* 1993;43:2412–4. [PubMed: 8232972]
40. Buschke H, Silwinsky MJ, Kuslansky G, Lipton RB. Diagnosis of dementia by the Double Memory Test: encoding specificity improves diagnostic sensitivity and specificity. *Neurology* 1997;48:989–99. [PubMed: 9109889]
41. Brooks LG, Loewenstein DA. Assessing the progression of mild cognitive impairment to Alzheimer's disease: current trends and future directions. *Alzheimers Res Ther* 2010;2:28. [PubMed: 20920147]
42. Rentz DM, Locasio JJ, Becker JA, Moran EK, Eng E, Buckner RL, et al. Cognition, reserve, and amyloid deposition in normal aging. *Ann Neurol* 2010;67:353–64. [PubMed: 20373347]
43. Loewenstein DA, Acevedo A, Schram L, Ownby R, White G, Mogosky B, et al. Semantic interference in mild Alzheimer's disease: preliminary findings. *Am J Geriatr Psychiatry* 2003;11:252–5. [PubMed: 12611756]
44. Loewenstein DA, Acevedo A, Agron J, Duara R. Vulnerability to proactive semantic interference and progression to dementia older adults with mild cognitive impairment. *Dement Geriatr Cogn Disord* 2007;24:363–8. [PubMed: 17911981]
45. Snitz B, Loewenstein DA, Saxton K, Ganguli M. A novel approach to assessing memory at the population level: vulnerability to semantic interference. *Int Psychogeriatr* 2010;22:785–94. [PubMed: 20105350]
46. Loewenstein DA, Acevedo A, Agron J, Ownby R, Barker W, Strauman S, et al. The utility of a brief memory screen in the diagnosis of mild memory impairment in the elderly: preliminary study. *Am J Geriatr Psychiatry* 2009;17:437–40. [PubMed: 19390301]
47. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–98. [PubMed: 1202204]
48. Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 2005;53:695–9. [PubMed: 15817019]
49. Stefanov DH, Zeungnam B, Bang W-C. The smart house for older persons and persons with physical disabilities: structure, technology arrangements, and perspectives. *IEEE Trans Neural Syst Rehabil Eng* 2004;12:224–50.
50. Chan M, Est'ève DA, Escriba C, Campo E. Review of smart homes—present state and future challenges. *Comput Methods Programs Biomed* 2008;91:55–81. [PubMed: 18367286]
51. Demiris G, Hensel BK, Geissbuhler A, Kulikowski C. Technologies for an aging society: a systematic review of "smart home" applications. *IMIA Yearbook of Medical Informatics* 2008. *Methods Inf Med* 2008;47(Suppl 1):33–40.
52. Bashshur, RL.; Shannon, GW. *History of telemedicine: evolution, context, and transformation*. New Rochelle, NY: Mary Ann Liebert; 2009.
53. Kaye, J.; Hayes, T.; Zitzelberger, T.; Yeagers, J.; Pavel, M.; Jimison, H., et al. Deploying wide-scale in-home assessment technology. In: Mihailidis, A.; Boger, J.; Kautz, H.; Normie, L., editors.

- Technology and Aging: Selected Papers from the 2007 International Conference on Technology and Aging. Amsterdam, the Netherlands: IOS Press; 2008. p. 19-26.
54. Hayes T, Pavel M, Kaye J. Gathering the evidence: supporting large-scale research deployments. *Intel Technol J* 2009;13:148–67.
 55. Hayes T, Abendroth F, Adami A, Pavel M, Zitzelberger T, Kaye J. Unobtrusive assessment of activity patterns associated with mild cognitive impairment. *Alzheimers Dement* 2008;4:395–405. [PubMed: 19012864]
 56. Hayes TL, Hunt JM, Adami A, Kaye JA. An electronic pillbox for continuous monitoring of medication adherence. *Conf Proc IEEE Eng Med Biol Soc* 2006;1:6400–3. [PubMed: 17946369]
 57. Hayes T, Larimer N, Kaye J. Medication adherence in healthy elders: small cognitive changes make a big difference. *J Aging Health* 2009;21:567–80. [PubMed: 19339680]
 58. Pavel, M.; Jimison, HB.; Hayes, TL.; Kaye, J.; Dishman, E.; Wild, K., et al. Continuous, unobtrusive monitoring for the assessment of cognitive function. In: Hofer, S.; Alwin, D., editors. *Handbook of cognitive aging: interdisciplinary perspectives*. Thousand Oaks, CA: Sage Publications, Inc; 2008. p. 524-43.
 59. Hatt WJ, Vanbaak EA, Jimison HB, Hagler S, Hayes TL, Pavel M, et al. The exploration and forensic analysis of computer usage data in the elderly. *Conf Proc IEEE Eng Med Biol Soc* 2009;1:1216–9. [PubMed: 19963743]
 60. Jimison H, Pavel M, McKanna J, Pavel J. Unobtrusive monitoring of computer interactions to detect cognitive status in elders. *IEEE Trans Inf Technol Biomed* 2004;8:248–52. [PubMed: 15484429]
 61. Rainie, L. Baby boomers and the internet. Washington, DC: Pew Internet and American Life Project, Pew Research Center; 2009. Available at: <http://www.pewinternet.org/Presentations/2009/Baby-Boomers-andthe-internet.aspx>
 62. Wild K, Boise L, Lundell J, Foucek A. Unobtrusive in-home monitoring of cognitive and physical health: reactions and perceptions of older adults. *J Appl Gerontol* 2008;27:181–200. [PubMed: 19165352]
 63. Kaye J. Home-based technologies: a new paradigm for conducting dementia prevention trials. *AlzheimersDement* 2008;4(1 Suppl 1):S60–6.
 64. Hayes T, Larimer N, Kaye J. Medication adherence in healthy elders: small cognitive changes make a big difference. *J Aging Health* 2009;2:567–80. [PubMed: 19339680]
 65. U.S.D.O.H.A.H. Services. Health, United States, 2009: with special feature on medical technology. Hyattsville, MD: National Center for Health Statistics; 2010. Website updated tables for life expectancy: <http://www.cdc.gov/nchs/hus/updatedtables.htm>
 66. Birren, JE.; Schaie, KW., editors. *Handbook of the psychology of aging*. 6. New York, NY: Elsevier; 2006.
 67. Jack CR Jr, Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, et al. Hypothetical model of dynamic biomarkers of the Alzheimer’s pathological cascade. *Lancet Neurol* 2010;9:119–28. [PubMed: 20083042]
 68. Rentz DM, Huh TJ, Faust RR, Budson AE, Scinto LF, Sperling RA, et al. Use of IQ-adjusted norms to predict progressive cognitive decline in highly intelligent older individuals. *Neuropsychology* 2004;18:38–49. [PubMed: 14744186]
 69. Rentz DM, Huh TJ, Sardinha LM, Moran EK, Becker JA, Daffner KR, et al. Intelligence quotient-adjusted memory impairment is associated with abnormal single photon emission computed tomography perfusion. *J Int Neuropsychol Soc* 2007;13:821–31. [PubMed: 17697413]
 70. Rentz DM, Sardinha LM, Huh TJ, Searl MM, Daffner KR, Sperling RA. IQ-based norms for highly intelligent adults. *Clin Neuropsychol* 2006;20:637–48. [PubMed: 16980251]
 71. Weintraub S, Powell D, Whitla D. Successful cognitive aging: individual differences among physicians on a computerized test of mental state. *J Geriatr Psychiatry* 1994;28:15–34.
 72. Apostolova LG, Morra JH, Green AE, Hwang KS, Avedissian C, Woo E, et al. Automated 3D mapping of baseline and 12-month associations between three verbal memory measures and hippocampal atrophy in 490 ADNI subjects. *Neuroimage* 2010;51:488–99. [PubMed: 20083211]

73. Dickerson BC, Salat DH, Bates JF, Atiya M, Killiany RJ, Greve DN, et al. Medial temporal lobe function and structure in mild cognitive impairment. *Ann Neurol* 2004;56:27–35. [PubMed: 15236399]
74. Dickerson BC, Sperling RA. Functional abnormalities of the medial temporal lobe memory system in mild cognitive impairment and Alzheimer's disease: insights from functional MRI studies. *Neuropsychologia* 2008;46:1624–35. [PubMed: 18206188]
75. Sperling RA, Bates JF, Chua EF, Cocchiarella AJ, Rentz DM, Rosen BR, et al. fMRI studies of associative encoding in young and elderly controls and mild Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 2003;74:44–50. [PubMed: 12486265]
76. Machulda MM, Senjem ML, Weigand SD, Smith GE, Ivnik RJ, Boeve BF, et al. Functional magnetic resonance imaging changes in amnesic and nonamnesic mild cognitive impairment during encoding and recognition tasks. *J Int Neuropsychol Soc* 2009;15:372–82. [PubMed: 19402923]
77. Binetti G, Magni E, Padovani A, Cappa SF, Bianchetti A, Trabucchi M. Release from proactive interference in early Alzheimer's disease. *Neuropsychologia* 1995;33:379–84. [PubMed: 7792005]
78. Ebert PL, Anderson ND. Proactive and retroactive interference in young adults, healthy older adults, and older adults with amnesic mild cognitive impairment. *J Int Neuropsychol Soc* 2009;15:83–93. [PubMed: 19128531]
79. Hanseeuw BJ, Seron X, Ivanoiu A. Increased sensitivity to proactive interference in amnesic mild cognitive impairment is independent of associative and semantic impairment. *Brain Cogn* 2010;72:325–31. [PubMed: 19906479]
80. Grober E, Merling A, Heimlich T, Lipton RB. Free and cued selective reminding and selective reminding in the elderly. *J Clin Exp Neuropsychol* 1997;19:643–54. [PubMed: 9408795]
81. Jacobs DM, Sano M, Dooneief G, Marder K, Bell KL, Stern Y. Neuropsychological detection and characterization of preclinical Alzheimer's disease [see comments]. *Neurology* 1995;45:957–62. [PubMed: 7746414]
82. Masur DM, Sliwinski M, Lipton RB, Blau AD, Crystal HA. Neuropsychological prediction of dementia and the absence of dementia in healthy elderly persons [see comments]. *Neurology* 1994;44:1427–32. [PubMed: 8058143]
83. Tounsi H, Deweer B, Ergis AM, Van der Linden M, Pillon B, Michon A, et al. Sensitivity to semantic cuing: an index of episodic memory dysfunction in early Alzheimer disease. *Alzheimer Dis Assoc Disord* 1999;13:38–46. [PubMed: 10192641]
84. Hultsch DF, MacDonald SW, Dixon RA. Variability in reaction time performance of younger and older adults. *J Gerontol B Psychol Sci Soc Sci* 2002;57:P101–15. [PubMed: 11867658]
85. Hultsch, DF.; Strauss, E.; Hunter, MA.; MacDonald, SWS. Intraindividual variability, cognition and aging. In: Craik, FIM.; Salthouse, TA., editors. *The handbook of aging and cognition*. 3. New York, NY: Psychology Press; 2008. p. 491-556.
86. Darby D, Maruff P, Collie A, McStephen M. Mild cognitive impairment can be detected by multiple assessments in a single day. *Neurology* 2002;59:1042–6. [PubMed: 12370459]
87. Gorus E, De Raedt R, Lambert M, Lemper JC, Mets T. Reaction times and performance variability in normal aging, mild cognitive impairment, and Alzheimer's disease. *J Geriatr Psychiatry Neurol* 2008;21:204–18. [PubMed: 18838743]
88. Kliegel M, Sliwinski M. MMSE cross-domain variability predicts cognitive decline in centenarians. *Gerontology* 2004;50:39–43. [PubMed: 14654726]
89. Holtzer R, Verghese J, Wang C, Hall CB, Lipton RB. Within-person across-neuropsychological test variability and indecent dementia. *JAMA* 2008;300:823–30. [PubMed: 18714062]

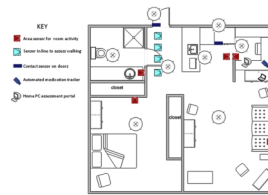


Fig. 1.

Home layout showing a typical array of sensors unobtrusively installed in existing homes to monitor activity and function on a continuous basis. Many other technologies or devices can also be added to this sensor net such as bed activity mats, telephone monitors, or physiological measuring devices (scales, oximeters, glucometers, etc.) depending on need. See key for explanation of symbols. For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.

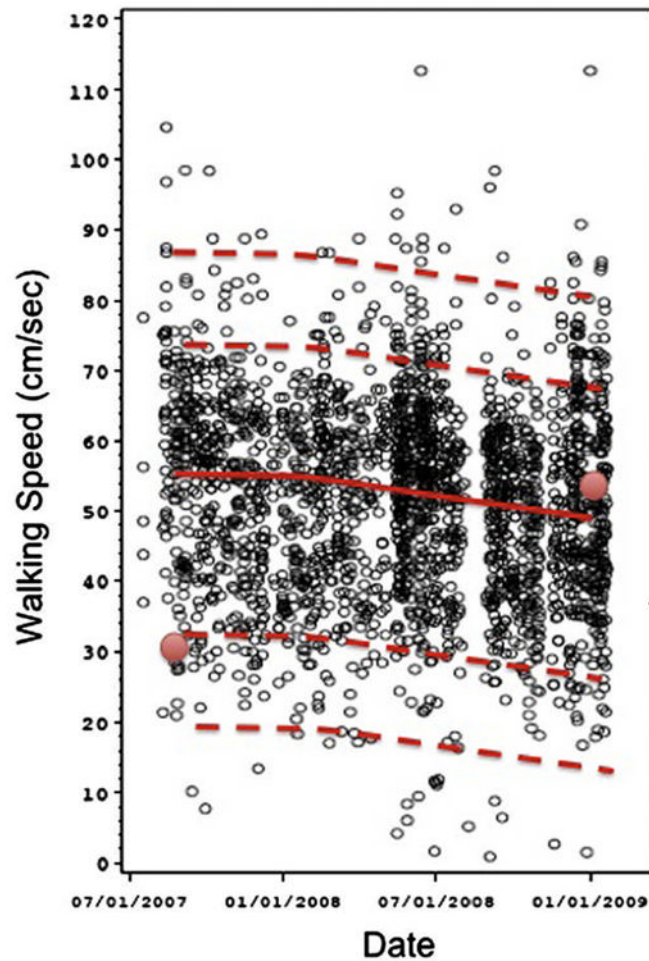


Fig. 2.

Daily walking speed measurements (cm/s) for an individual derived from an in-home sensor line operating over a period of 16 months showing the natural variability in this metric associated with cognitive decline. The automated measure provides a more comprehensive view of walking compared to current single point speed measures represented in the figure as the two red data points obtained with a stopwatch at widely spaced intervals. Uniquely, this methodology allows for the calculation of trends over time (red line) as well as defining naturally occurring populations of fast or slow walks (dashed lines) for an individual. For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.

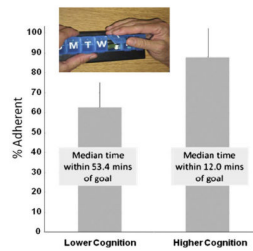


Fig. 3.

The bars represent the ability of two groups of elderly individuals to adhere to a twice daily medication-taking regimen assessed using a medication tracking device (inset) that automatically records the time of day when a pill compartment is opened. Those scoring greater than the upper limit of the 95% confidence interval on the Alzheimer's Disease Assessment Scale cognitive subscale (higher scores represent poorer performance) were significantly less able to adhere to the medication regimen. These results suggest the ability to detect subtle deficits in cognitive function using performance on a daily task that is unobtrusively monitored in real time. Data are from Hayes et al [64]. For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.

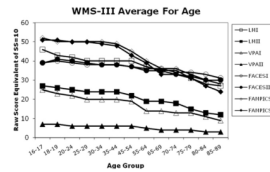


Fig. 4.

The graph represents a translation of data from normative tables of the subtests of the Wechsler Memory Scale—third edition [38]. To generate the data points, the raw score equivalent of a scaled score of 10 was plotted for each subtest for each age group included in the normative tables. If, for a particular test, the raw score was a range rather than a single value, the average of the scores was computed to generate the graph. The trajectory is unequivocally downward over the age groups from youngest to oldest. However, each subtest has its own aging trajectory, with some showing sharper decline than others over the same period.