Mechanisms and Risk Factors of Cognitive Aging

MECHANISMS AND RISK FACTORS OF COGNITIVE AGING

ALAIN KOYAMA

A Dissertation Submitted to the Faculty of

The Harvard T.H. Chan School of Public Health

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Science

in the Department of Epidemiology

Harvard University

Boston, Massachusetts.

May 2015

Dissertation Advisor: Dr. Francine Grodstein Alain Koyama

Mechanisms and Risk Factors of Cognitive Aging

Abstract

This dissertation explores the role of biologic factors and a novel method of measuring cognitive function in investigating mechanisms and risk factors of cognitive aging. With a rapidly increasing aging population, the public health burden of dementia is expected to rise in the future. Therefore, it is important to both better understand the etiology and identify novel risk factors, in order to reduce the incidence of new cases and develop effective treatments. Blood-based biomarkers, such as amyloid-beta and sex hormones, can provide an objective measure of factors likely associated with dementia risk. Additionally, computerized cognitive testing can provide efficient and accurate measurement of cognitive function, yet is seldom used in epidemiologic studies.

Our first chapter involves the first systematic review and meta-analysis of prospective studies exploring the association between plasma amyloid-beta and incident Alzheimer's disease or dementia. While preclinical prediction of Alzheimer's disease is important for effective intervention, studies have considerably varied in design, assays and sample size, making it difficult to interpret the overall data and thus necessitating a systematic review and meta-analysis.

The second chapter explores the association between endogenous hormone levels and cognitive function in a population of older women. While many studies have investigated the role of hormone therapy in cognitive aging, fewer studies have investigated endogenous hormones, which may provide a more objective measure of hormonal status. We therefore present a prospective cohort study investigating the association between several endogenous sex hormones and their prohormones at baseline, and cognitive function over 20 years later.

The third chapter evaluates the feasibility and performance of a self-administered computerized cognitive battery. Although most epidemiologic studies of cognitive aging rely on traditional neuropsychological testing to measure cognitive outcomes, such methods can be prone to error, interviewer bias, and demand substantial time and cost. Therefore, we present results demonstrating for the first time the feasibility and performance of an unsupervised self-administered computerized cognitive battery in a population of older men.

Table of Contents

List of Figures

List of Tables

Acknowledgements

There are many individuals to whom I would like to express my sincere gratitude, without whom this dissertation would not have been possible. To my research advisor, Dr. Francine Grodstein, for her constant attention and guidance throughout all phases of my research. To my committee members, Dr. Olivia Okereke, Dr. Bernard Rosner, and Dr. Marc Weisskopf, who dedicated their busy schedules to provide me their valuable expertise and experience. To Dr. Mary Townsend, whose invaluable assistance allowed me to quickly begin my thesis work. To Dr. Kristine Yaffe and past colleagues for allowing me to gain valuable research experience prior to my dissertation. To friends and family for their support and encouragement. And to all those in public health who work in the pursuit of promoting health as a fundamental right of the human condition.

Plasma Amyloid β as a Predictor of Dementia and Cognitive Decline: A Systematic Review and Meta-analysis

Alain Koyama, Olivia I. Okereke, Ting Yang, Deborah Blacker Dennis J. Selkoe, Francine Grodstein

1.1 Abstract

Context: Preclinical prediction of Alzheimer's disease is important, critical to effective intervention. Plasma levels of amyloid β-peptides have been a principal focus of the growing literature on blood-based biomarkers, but studies to date have varied in design, assay methods and sample size, making it difficult to readily interpret the overall data.

Objective: To conduct a systematic review and meta-analysis of relevant prospective studies in order to determine if plasma amyloid β levels may predict development of dementia, Alzheimer's disease, and cognitive decline.

Data Sources: Prospective studies published between 1995 and 2011 indexed in the PubMed, EMBASE, and PsycInfo databases were searched.

Study Selection: Selected studies included those measuring at least one relevant plasma amyloid β species ($A\beta_{40}$, $A\beta_{42}$, $A\beta_{42}$:$A\beta_{40}$ ratio) and reporting an effect estimate for dementia, Alzheimer's disease, or cognitive change.

Data Extraction: Using a standardized extraction form, appropriate study parameters on subject information, exposure, and outcome were extracted. Random effects models were utilized to generate summary risk ratios and 95% confidence intervals, comparing the bottom versus top quantile for each plasma measure.

Results: Thirteen studies with a total of 10,303 subjects met inclusion criteria for meta-analysis.

Lower $A\beta_{42}$:$A\beta_{40}$ ratios were significantly associated with development of Alzheimer's disease (summary RR=1.60, 95% CI=1.04,2.46; p=0.03) and dementia (RR=1.67 95% CI=1.02,2.75; p=0.04). Significant heterogeneity was found for both summary estimates, which could not be explained by participants' age, sex distribution, the study's follow-up time, or year of publication. Plasma levels of $A\beta_{40}$ and $A\beta_{42}$ alone were not significantly associated with either outcome.

Conclusions: Overall, the literature indicates that plasma $A\beta_{42}$:$A\beta_{40}$ ratios predict development of Alzheimer's disease and dementia. However, significant heterogeneity in the meta-analysis underlines the need for substantial further investigation of plasma amyloid β levels as a preclinical biomarker.

1.2 Introduction

An enormous public health burden is caused by senile dementia, with Alzheimer's disease (AD) alone being the seventh leading cause of death in the United States and costing an estimated $172 billion annually[1]. Current therapies to treat AD are minimally effective and do not alter the disease process. It is widely believed that novel therapeutic agents expected to be developed in the coming years will be optimally administered preclinically, before patients develop full dementia. Thus, preclinical prediction of dementia through biomarkers is an important field, critical to effective intervention and disease modification[2]. Although the Alzheimer's Association and the National Institute on Aging recently established research guidelines for identifying preclinical dementia using neuroimaging and cerebrospinal fluid (CSF) proteins[3], a blood-based biomarker would be less invasive and more cost-effective than CSF or imaging-based methods. Moreover, a blood-based biomarker might also be used in a complementary role to CSF and imaging, as a first-step screen for high-risk individuals who would maximally benefit from these more invasive and expensive modalities.

Plasma levels of amyloid β-peptides have been a focus of the growing literature on blood-based biomarkers for dementia[4-17], but studies to date have varied substantially in their design, assay methods and sample size - making it difficult to interpret the overall data. Therefore, we performed a systematic review and meta-analysis to evaluate the scientific literature, asking whether plasma Aβ levels predict development of dementia, including AD, and cognitive decline.

1.3 Methods

1.3.1 Search Strategy

Following a pre-established protocol, a systematic review was conducted by two investigators with methodological expertise (A.K. and F.G.) using a Boolean search strategy on the electronic databases MEDLINE, EMBASE, and PsycInfo. Keywords shown in Figure 1.1 were used to search for the

exposure and outcomes of interest, as well as to confine our search to epidemiological studies. Studies were limited to those published after 1995, due to the lack of well-developed Aβ assays before this time. The bibliographies of all relevant articles and review papers were also hand-searched; abstracts from major scientific meetings were also examined by the authors, and experts in the field consulted for any further studies.

1.3.2 Inclusion Criteria

Study selection was carried out in two stages, using the same inclusion criteria. The first stage involved reviewing only the title and abstract of each article, and the second stage involved reviewing the full text. For an article to be included in either stage, it had to fulfill four criteria for study quality: a prospective cohort (including case-cohort or nested case-control designs); measurement of the relevant plasma amyloid β species ($A\beta_{40}$, $A\beta_{42}$, and/or $A\beta_{42}$:$A\beta_{40}$ ratio); report of the relative risk or equivalent effect estimates for incident AD, total dementia, and/or mean differences in cognitive decline for studies of that outcome; be adjusted for age at a minimum. All languages were included in the searches.

1.3.3 Data Extraction

Data extraction was performed using a standardized extraction form. We extracted the following variables from each study: year of publication; study design; country of study population; name of cohort, exposures measured and variable coding method; outcomes measured and standard for diagnosis; length of follow-up; sample size; demographics (mean age at baseline, gender, ethnicity); effect measures, respective *p* values and confidence intervals and/or standard errors; number of cases in each group; covariates used in modeling.

Figure 1.1 – Search strategy

MEDLINE:

"Amyloid beta-Protein"[Mesh] OR "Amyloid beta-Protein/blood"[MAJR] OR ("plasma"[tiab] AND ("beta amyloid"[tiab] OR "amyloid beta"[tiab] OR "abeta"[tiab])) AND ("Dementia"[MeSH] OR "Dementia"[tiab] OR "Alzheimer Disease"[MeSH] OR "Dementia, Vascular"[Mesh] OR "Dementias, Vascular"[tiab] OR "Vascular Dementia"[tiab] OR "vascular dementias"[tiab] OR "Lewy Body Disease"[MeSH] OR "Lewy Body Dementia"[tiab] OR "Senile Dementia"[tiab] OR "Senile Dementias"[tiab] OR "Presenile Dementia"[tiab] OR "Presenile Dementias"[tiab] OR "Alzheimer's Disease"[tiab] OR "Alzheimer Dementia"[tiab] OR "Alzheimer's Dementia"[tiab] OR "Alzheimer Dementias"[tiab] OR "Alzheimer's Dementias"[tiab] OR "cognitive impairment"[tiab] OR "cognitive decline"[tiab] OR "mci"[tiab]) AND (("Longitudinal Studies"[Mesh] OR "longitudinal"[tiab]) OR ("Prospective Studies"[Mesh] OR "prospective"[tiab]) OR("Cohort Studies"[Mesh] OR "cohort"[tiab]) OR "Epidemiologic Studies"[Mesh])

PsycINFO:

(DE "Dementia" OR DE "Dementia with Lewy Bodies" OR DE "Presenile Dementia" OR DE "Senile Dementia" OR DE "Vascular Dementia" OR DE "Alzheimers Disease" OR "Dementia" OR "Dementia with Lewy Bodies" OR "Presenile Dementia" OR "Senile Dementia" OR "Vascular Dementia" OR "Alzheimers Disease" OR DE "cognitive impairment" OR "mci" or "cognitive decline") AND (DE "Beta Amyloid" OR "amyloid beta" OR "beta amyloid" OR "abeta")

EMBASE:

'amyloid beta protein'/exp AND ('dementia'/exp OR 'multiinfarct dementia'/exp OR 'vascular dementia'/exp OR 'diffuse lewy body disease'/exp OR 'lewy body dementia'/exp OR 'alzheimer disease'/exp OR 'alzheimer dementia'/exp OR 'senile dementia'/exp OR 'dementia senilis'/exp OR 'presenile dementia'/exp OR 'dementia, presenile'/exp OR 'mild cognitive impairment'/exp OR 'cognitive defect'/exp) AND ('longitudinal study'/exp OR 'prospective study'/exp OR 'epidemiology'/exp) AND ([adult]/lim OR [aged]/lim) AND [humans]/lim AND [embase]/lim AND [1995-2011]/py

1.3.4 Data Synthesis

For the analyses, odds ratios, incidence rate ratios, hazard ratios, or risk ratios for dichotomous outcomes were considered as equivalent effect measures[18]. For the sake of simplicity, these effect measures will hereafter be referred to as risk ratios (RR). We focused on data regarding $A\beta_{42}$ and $A\beta_{42}$:$A\beta_{40}$, since these likely provide the most relevant information for risk prediction based on the existing literature. In addition, there is less biological rationale supporting the measurement of $A\beta_{40}$ alone as a predictor of dementia; therefore, we evaluated those studies secondarily. In studies reporting plasma amyloid β-protein as a categorical variable, we considered the highest quantile as the reference group for our meta-analysis and generated a summary effect estimate for the comparison of the bottom versus the top quantile. These categorical analyses were considered, a priori, as our primary analyses for several reasons. First, because absolute measures of Aβ can differ widely between current plasma Aβ assays[19], the categorical classification of Aβ is subject to less misclassification than a continuous variable. That is, while a continuous measure requires that each unit is appropriately estimated, an ordered categorical variable only requires that subjects are generally ranked correctly across three or four categories and thus yields less misclassification. Additionally, ordered categories are less susceptible to outliers of high levels of Aβ as well as very low levels that approach the detection limit of the assay, again resulting in less misclassification when using quantiles. Most importantly, in eventual clinical practice, it is most likely that Aβ will not be utilized as a continuous measure, but rather that threshold categories will be defined for different risk states. Finally, the majority of studies presented analyses of Aβ as a categorical variable. However, secondary analyses were also performed to derive a summary effect estimate from the incremental dose-response RR for each study, when available. The four studies reporting cognitive decline as an outcome[6, 11] were not included in the meta-analysis due to large variations in the methods by which cognition was assessed, but are reviewed here qualitatively.

For the dementia outcomes (total dementia and incident AD), both fixed and random effects models were used to generate summary risk ratios across relevant studies. As results were similar using both models, only DerSimonian and Laird random effects estimates are presented[20]. Heterogeneity was assessed using the I^2 statistic, and, if heterogeneity was found, we explored possible explanations using meta-regression models;[21] we tested mean age, gender percentage, year of publication and follow-up time in the meta-regression models. We also conducted meta-analyses excluding certain studies with which were meaningfully different from other investigations in terms of sex. We could not conduct stratified analyses according to follow-up time, since this would have yielded strata with an insufficient number of studies to provide meaningful information in a summary estimate. To assess study quality, since many studies reported results from multiple regression models with minimal and maximal control for potential confounding factors, we conducted two separate meta-analyses of the least and most adjusted risk ratios, and the pooled estimates for each were compared for significant differences. Besides evaluating maximal control of confounding factors, we did not conduct additional analyses examining study quality, since our inclusion criteria (see above) already addressed many primary issues of study quality; that is, given our assessment of study quality as part of inclusion criteria, attempts to further stratify studies by quality within the meta-analysis would have resulted in strata with insufficient studies to yield meaningful summary estimates. Publication bias was assessed by means of the Egger test[22], and was found to be nonsignificant for our primary meta-analyses of $A\beta_{42}$ and $A\beta_{42}$:$A\beta_{40}$. All calculations were performed using STATA version 11 (StataCorp 2007, College Station, TX).

1.4 Results

1.4.1 Study Selection

After the initial keyword search, there were 424 results from Pubmed, 82 from Embase, and 252 from PSYCInfo, for a total of 758 studies (Figure 1.2). After compilation of all studies into Endnote version X3 (Thomson Reuters, 2009, New York, NY), removal of duplicates resulted in 726 distinct studies. Two investigators (A.K. and F.G.) independently reviewed the remaining articles and after the first stage of study selection, 25 studies were identified for further consideration. After reviewing the full text of these articles, 14 publications remained which met all of our inclusion criteria.

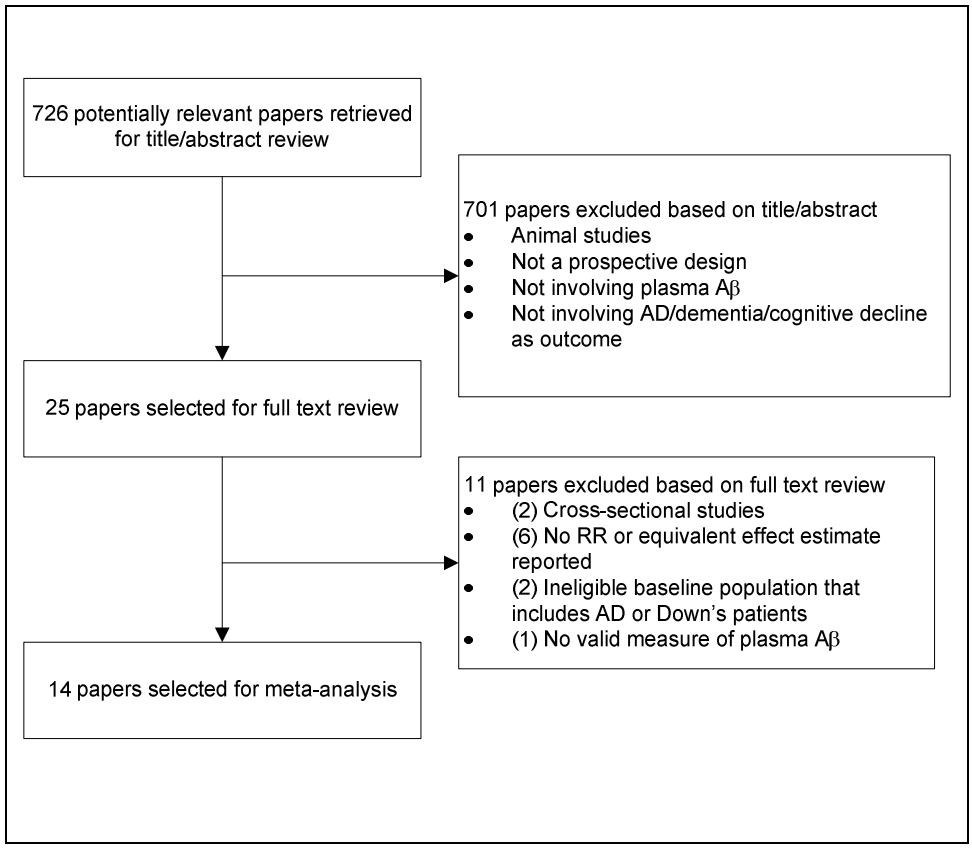


Figure 1.2 – Study Selection

1.4.2 Description of Studies

Fourteen publications met inclusion criteria for meta-analysis. Of these, two publications were each included as two distinct studies in the meta-analysis rather than as one, because results in each of these publications were presented for two separate subcohorts[14, 15]. In addition, two publications utilized

the same cohort, and only the more recent was included as it contained a larger sample[9, 10]. All studies in the meta-analysis were published in 2003 or after. All studies were prospective cohorts, although two were case-cohort studies. The total subject pool was largely female (48-69% across the studies), with one study comprised exclusively of females[11] and another of males[15].

1.4.3 Plasma Amyloid β-Protein and Dementia

Six studies reported risk ratios for the association between $A\beta_{42}$ levels and risk of dementia, and all used ordered categories of $A\beta_{42}$ levels (Table 1.1). Of these, five reported increased risks of developing dementia for lower levels of $A\beta_{42}$ in the least adjusted models, although only two were statistically significant[14]. The pooled risk ratio estimate across the studies was modest, and not statistically significant (summary RR=1.37; 95% CI=0.95,1.98; p=0.10) (Figure 1.3).

Among the six studies investigating $A\beta_{42}$:$A\beta_{40}$ ratio, five reported an increased risk of developing dementia for the lowest $A\beta_{42}$:$A\beta_{40}$ ratios[7, 14, 16] compared to the highest quantile; four of these found statistically significant increased relative risks. All studies reported the $A\beta_{42}$:$A\beta_{40}$ ratio in quantiles. A pooled analysis yielded a statistically significant RR of 1.67 (95% CI=1.02,2.75; p=0.04).

In all these meta-analyses of plasma Aβ and dementia, the pooled estimate did not change significantly when using results from the most adjusted models, or when excluding the study which only included men. Significant heterogeneity was found in the each of the above meta-analyses, which was not explained by age or gender distribution of the populations studied, or by follow-up time or year of publication. In a secondary analysis of $A\beta_{42}$ as a continuous measure, results were consistent with those reported here for the quantile comparisons, although as expected, with increased misclassification of Aβ level in the continuous variable, the summary RRs were generally weaker.

Table 1.1 – Baseline characteristics of studies

Study	Design	Outcomes	No. at risk	No. of Events		Mean Age	Sex (female)	Follow-up time (yrs)	Adjusted For	Dementia		AD	
				Dem.	AD					RR (Aβ42)	RR ($A\beta_{42}$:$A\beta_{40}$)	RR (Aβ42)	RR ($A\beta_{42}$:$A\beta_{40}$)
Abdullah 2009	cohort	AD	203	--	14	76.8	48.3%	2.1 (median)	age, sex, education	--	--	4.47 (1.19,9.85)	3.91 (1.36,11.24)
Graff-Radford 2007	cohort	AD, cognitive change	563	--	17	78.0 (median)	62.0%	3.7 (median)	age, APOE	--	--	1.18 (0.50,2.75)	3.08 (1.12,8.3)
Lambert 2009	case-cohort	AD, dementia	8414	233	154	74.6	60.4%	4	age, gender, education, site	1.33 (0.97,1.92)	2.00 (1.41,2.78)	--	2.08 (1.30,3.23)
Mayeux 2003	cohort	AD	451	--	86	76.2	68.9%	5	age, education, APOE, Aβ40 level, BMI	--	--	0.4 (0.21,0.77)	--
Schupf 2008	cohort	AD	1125	--	104	76.9	68.3%	4.6 (mean)	age, sex, APOE, education, ethnicity, BMI, cohort membership	--	--	0.29 (0.12,0.71)	1.11 (0.59,2.00)
Seppalla 2010a	cohort	Dementia	197	9[c]	--	70.0[a]	55.0%[a]	3	unknown	3.12 (1.25,7.79)	3.26 (1.31,8.11)	--	--
Seppalla 2010b	cohort	Dementia	60	7[c]	--	70.0[a]	55.0%[a]	6	unknown	4.77 (1.14,19.98)	8.4 (1.83,83.57)	--	--
Sundelof 2008a	cohort	AD, dementia	1045	146	82	71.0	0.0%	11.2 (median)	age, APOE	0.7 (0.44,1.13)	0.67 (0.42,1.08)	0.85 (0.48,1.50)	0.89 (0.54,1.48)
Sundelof 2008b	cohort	AD, dementia	680	74	46	77.6	0.0%	5.3 (median)	AD: age dementia: age, APOE, diabetes	1.58 (0.87,2.86)	1.06 (0.59,1.90)	2.5 (1.04,6.03)	1.27 (0.61,2.66)
van Oijen 2006[b]	case-cohort	AD, dementia	6713	392	289	68.6	61.0%	8.6 (mean)	age, sex	1.16 (0.82,1.64)	2.13 (1.49,3.03)	--	--

Abbreviations: AD = Alzheimer's disease; Aβ = Amyloid beta; APOE = Apolipoprotein E; BMI = Body Mass Index; Dem = Dementia

[a] study population data is for entire cohort

[b] only continuous RR reported for AD outcome

[c] number of events estimated from published data

Note: confidence intervals reflect published results, before random-effects weighting

Table 1.2 – Baseline Characteristics of Additional Studies

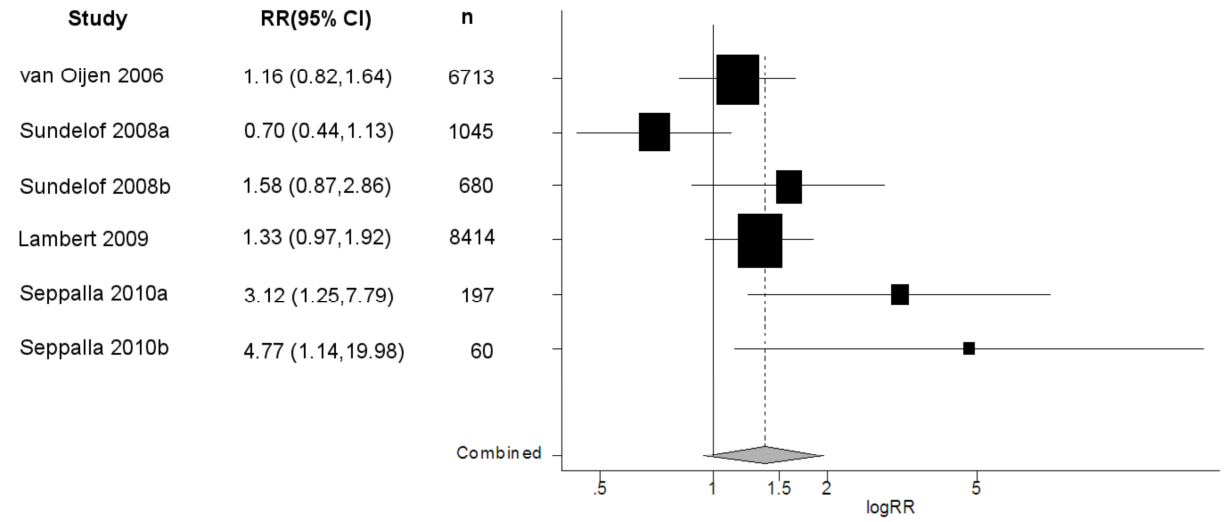
Study	Design	Outcome	No. at risk	No. of Events		Mean Age	Sex (female)	Follow-up time (yrs)	Adjusted For	Reason not included
				Dem.	AD					
Blasko 2010	cohort	AD	406	--	33	75.8	56.5%	5	sex, education, creatinine, smoking, stroke/infarction in MRI, sGDS score, interaction between Aβ42 and APOE	only continuous RR reported
Cosentino 2010	cohort	cognitive change	880	--	--	76.1	68.0%	4.5	age, sex, race, BMI, APOE, recruitment wave	cognitive change as outcome
Lopez 2008	cohort	AD	274	--	88	79.3	60.9%	4.5	age	only continuous RR reported
Okereke 2009	cohort	cognitive change	481	--	--	63.6	100.0%	10	age, education, BMI, hypertension, dyslipidemia,heart disease, smoking, HT,physical activity, alcohol, depression	cognitive change as outcome
Yaffe 2011	cohort	cognitive change	997	72	--	74.0	55.1%	10	age, race, education, diabetes, smoking, APOE	cognitive change as outcome

Abbreviations: APOE = Apolipoprotein E; BMI = Body Mass Index; Dem = Dementia; HT = hormone therapy; MRI = Magnetic Resonance Imaging; sGDS = short form of the Geriatric Depression Scale

Note: confidence intervals reflect published results, before random-effects weighting

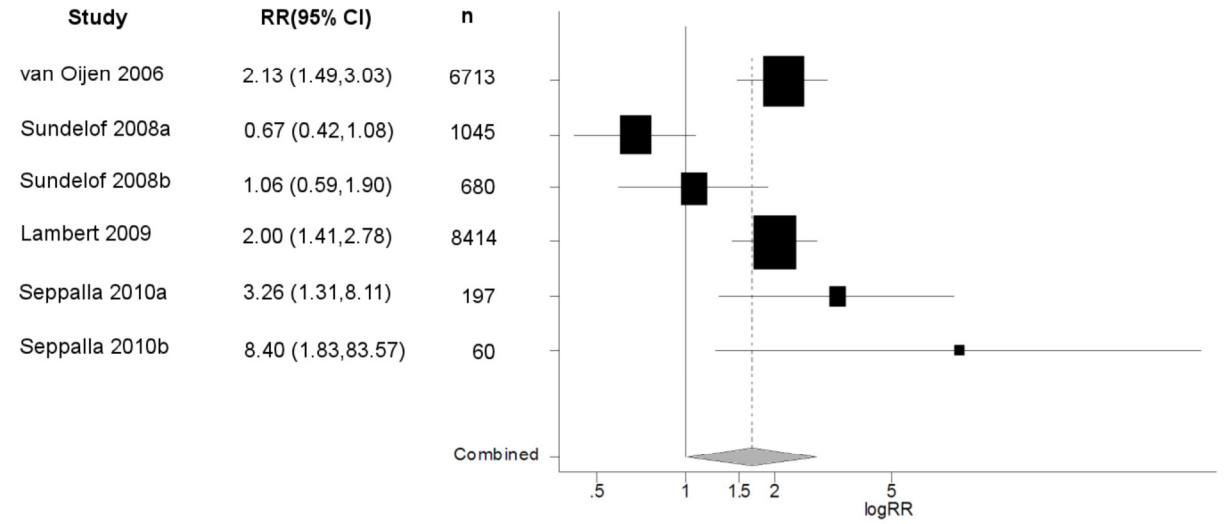
Figure 1.3 – Meta-Analysis of Plasma Aβ and Incident Dementia

Plasma $A\beta_{42}$ and Incident Dementia



Pooled RR = 1.37 (95% CI: 0.95,1.98), p = 0.10
I^2 = 64%, p = 0.02

Plasma $A\beta_{42}$:$A\beta_{40}$ Ratio and Incident Dementia



Pooled RR = 1.67 (95% CI 1.02,2.75), p = 0.04
I^2 = 80%, p < 0.001

Pooled relative risks (RR's) of incident dementia for $A\beta_{42}$ and $A\beta_{42}$:$A\beta_{40}$ ratio. I^2 indicates degree of heterogeneity.

Four studies reported effect estimates for the association between $A\beta_{40}$ levels and risk of dementia, and all used categorical exposures[7, 15, 16]. A pooled analysis indicated no relation between $A\beta_{40}$ and dementia development (RR=1.01; 95% CI: 0.60,1.71; p=0.97).

1.4.4 Plasma Amyloid ß-Protein and Cognitive Decline

Four studies reported effect estimates for the association between plasma Aβ levels and cognitive decline[6, 11, 23, 24]. No association between $A\beta_{42}$ levels and cognitive decline was found in two[6, 11] of the studies. One study reported a statistically significant association between decreased baseline $A\beta_{42}$ levels and subsequent cognitive decline[24], while the remaining study also reported a significant association but for increased baseline $A\beta_{42}$ levels[23]. Thus, findings for plasma $A\beta_{42}$ as a predictor of cognitive decline were inconsistent. However, similar to the meta-analysis of dementia, three of four studies reported that lower $A\beta_{42}$:$A\beta_{40}$ ratio at baseline significantly predicted greater cognitive decline[6]. In addition, one of these studies further measured change over ten years of the $A\beta_{42}$:$A\beta_{40}$ ratio, reporting that a decrease in this ratio over time predicted greater subsequent cognitive decline[11]. The most recent study[24] reported an interaction between plasma $A\beta_{42}$:$A\beta_{40}$ and cognitive reserve, such that the relation between $A\beta_{42}$:$A\beta_{40}$ and cognition was strongest in those with the least education. No other studies have examined such an interaction, although the Nurses' Health Study[11] found that $A\beta_{42}$:$A\beta_{40}$ predicted cognitive decline in a well-educated population of women.

1.4.5 Plasma Amyloid ß-Protein and Alzheimer's Disease

Nine studies reported risk ratios for plasma $A\beta_{42}$ as a predictor of the development of clinically diagnosed AD. Six of these studies employed categorical exposures, and most of the studies were smaller in size than those of total dementia (Table 1.1). Results for $A\beta_{42}$ were inconsistent: three studies

reported risk ratios above 1.0 for lower baseline levels of $A\beta_{42}$ (with 2 achieving statistical significance), while three reported risk ratios below 1.0 (with 2 achieving statistical significance). Reflecting these results, a pooled analysis (Figure 1.3) showed no relation between levels of $A\beta_{42}$ and risk of developing AD (RR=1.01, 95% CI=0.48,2.11; p=0.99). Of the three studies reporting continuous levels of $A\beta_{42}$, which were thus not included in the meta-analysis, one benefited from a fairly large sample (n=1756; 289 cases), but still observed a null result[16]. The two remaining studies showed a statistically significant decreased risk for lower levels of $A\beta_{42}$[25], although one was no longer significant in the maximally adjusted model[8].

Among eight studies that reported risk ratios for plasma $A\beta_{42}$:$A\beta_{40}$ ratio and risk of developing AD, six employed ordered categories of $A\beta_{42}$ (Table 1.1). These data were more consistent than for $A\beta_{42}$ alone. Five of the six studies reported an increase in the likelihood of developing AD for the lowest compared to highest quantile of $A\beta_{42}$:$A\beta_{40}$ ratio, and three were statistically significant[4, 6, 7]. In our pooled analysis, there was a significant increase in risk of AD comparing the bottom versus top quantiles of $A\beta_{42}$:$A\beta_{40}$ ratios (RR=1.60, 95% CI=1.04,2.46; p=0.03) (Figure 1.4). Of the two studies not included in the meta-analysis due to the absence of categorical data on Aβ, one published null results[8] while the other study reported that a lower ratio of $A\beta_{42}$:$A\beta_{40}$ was a significant predictor of a higher rate of AD development[16], consistent with the findings of our meta-analysis.

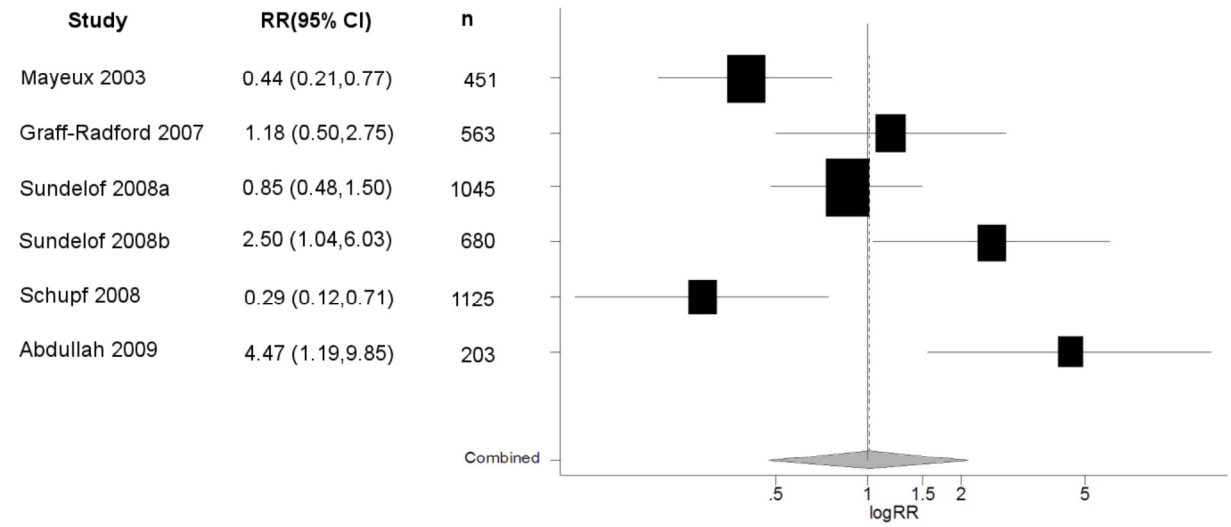
In all meta-analyses of plasma Aβ species and AD, the pooled estimate did not significantly change when using results from the most adjusted models or when excluding the all-male cohort. Heterogeneity was found in the above meta-analyses with an I^2 statistic of 64% for $A\beta_{42}$ and 80% for $A\beta_{42}$:$A\beta_{40}$ ratio; thus, the summary risk ratios must be interpreted cautiously.

Seven studies reported effect estimates for plasma $A\beta_{40}$ as a predictor of AD, of which four presented Aβ levels in ordered categories. Our meta-analysis found an elevated risk of AD with lower $A\beta_{40}$, but the confidence interval was fairly wide and the summary estimate was not statistically

significant (RR=1.66; 95% CI=0.98,2.83; p=0.06). Of the three studies not included in the meta-analysis, two reported null results[4, 8], and one showed a statistically significant decreased risk of AD with lower levels of $A\beta_{40}$[16]. Heterogeneity was not significant with an I^2 statistic of 45% for $A\beta_{40}$. Results remained consistent when $A\beta_{40}$ was analyzed as a continuous exposure.

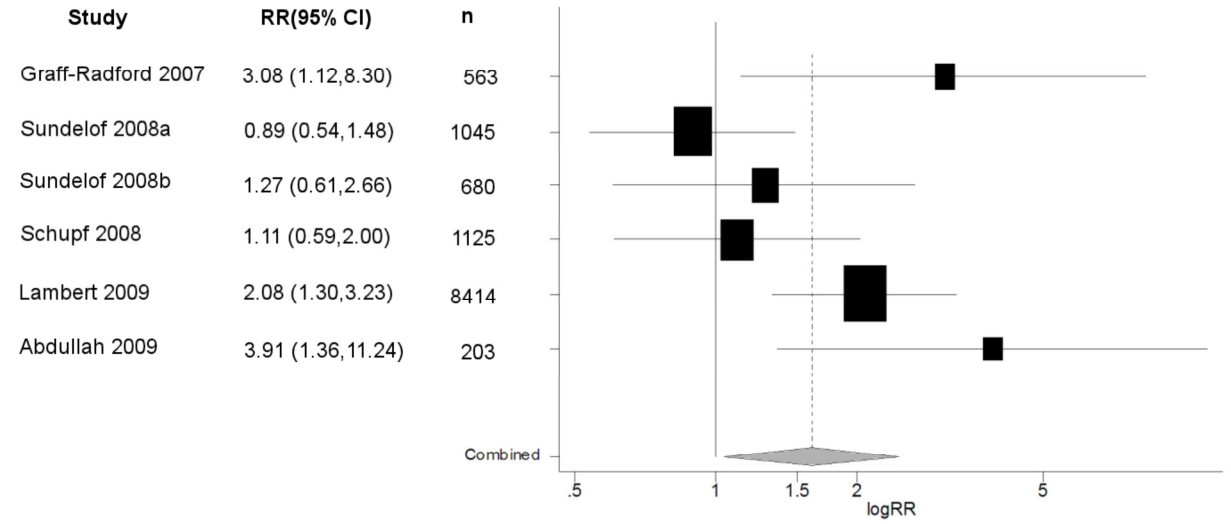
Figure 1.4 – Meta-Analysis of Plasma Aβ and Incident Alzheimer's Disease

Plasma $A\beta_{42}$ and Incident Alzheimer's Disease



Pooled RR = 1.01(95% CI 0.48,2.11), p = 0.99
I^2 = 81%, p < 0.001

Plasma $A\beta_{42}$:$A\beta_{40}$ Ratio and Incident Alzheimer's Disease



Pooled RR = 1.60(95% CI 1.04,2.46), p = 0.03
I^2 = 59%, p = 0.03

Pooled relative risks (RR's) of incident Alzheimer's disease for $A\beta_{42}$ and $A\beta_{42}$:$A\beta_{40}$ ratio. I^2 indicates degree of heterogeneity.

1.5 Discussion

In this systematic review, we examined the literature regarding plasma $A\beta_{42,}$ as well as the ratio of $A\beta_{42}$:$A\beta_{40,}$ as predictors of dementia and AD. We found that plasma levels of $A\beta_{42}$ alone were not strong predictors of dementia or AD risk, with non-significant risk ratios across studies of both these outcomes. In contrast, the data across studies of $A\beta_{42}$:$A\beta_{40}$ were more promising; we found a significant elevated risk for developing dementia or AD in subjects with lower $A\beta_{42}$:$A\beta_{40}$ ratios, with most studies reporting fairly similar findings. These results were robust to sensitivity analyses when using data from the most adjusted models reported, when analyzed as continuous levels rather than as ordered categories, and when the single study comprised of an all male cohort was excluded. No evidence for publication bias was found. Moreover, four studies reported results for cognitive decline as the outcome, and results from three of these were consistent with our meta-analysis in observing that a lower ratio of $A\beta_{42}$:$A\beta_{40}$ predicted worse cognitive decline. Collectively, the existing research offers cautious support of the hypothesis that lower levels of the plasma $A\beta_{42}$:$A\beta_{40}$ ratio reflect a process of selective deposition of $A\beta_{42}$ in the brain as insoluble amyloid plaques, thus predictive of dementia development.

While we calculated summary risk ratios across studies to produce a quantitative estimate of effect from the existing data, a key limitation of our findings was the significant heterogeneity in each pooled estimate, necessitating caution in our interpretation of findings. We could not formally identify the source of heterogeneity, however we can hypothesize as to its causes. First, we suspect that much of the heterogeneity was likely the consequence of known measurement issues for plasma Aβ. The studies used in this meta-analysis employed varying ELISA assays and multiplex platforms, resulting in a wide distribution of median Aβ levels between studies. The development of a standardized assay is therefore highly important to achieve more comparable results in further research on plasma Aβ. Second, Aβ levels likely have differing implications at different stages in the pathogenesis of dementia,

and the follow-up time varied considerably among studies; although meta-regression did not show this to significantly contribute to heterogeneity. Yet, variation of baseline Aβ levels and the subjects' degree of underlying, preclinical dementia at baseline may have been important contributors to heterogeneity, regardless of follow-up time. Most studies in the meta-analysis did not assess baseline levels of mild cognitive impairment in participants, and criteria for pre-clinical AD have only recently been promoted[3]. Thus there was no clear means of evaluating any influence of varying levels of cognitive health at baseline. In two small studies which reported MCI prevalence, estimates ranged from 9.6%[9] to 19.3%[14], indicating there is likely wide variation in the level of early or underlying dementia across studies. Future research will be informed by more standard assessment of preclinical dementia both at baseline and during follow-up These issues can be addressed, in part, by measuring the relative change of Aβ levels at multiple time points, as opposed to a single baseline measurement. Some studies have already employed this temporal design and have been relatively consistent in showing significant associations between decreasing levels of $A\beta_{42}$ or $A\beta_{42:}A\beta_{40}$ ratio over time and cognitive status[9, 11, 12, 14, 23]. Overall, our results emphasize the need for further research to better understand all of the issues pertaining to heterogeneity before plasma Aβ can be of broader predictive utility as a biomarker of impending dementia.

Other limitations of our meta-analysis should be considered. Results from individual studies are subject to potential unmeasured confounding and bias, and a meta-analysis cannot eliminate these issues (although we found similar results when using both minimally and maximally adjusted RR's). Missing data could introduce bias if missingness is related to both the exposure and outcome, which is often likely, although the majority of studies reported reasonable follow-up rates. Additionally, some studies may be inappropriate to pool together. For example, one study[15] included an all male cohort, potentially limiting the generalizability of this study, as women formed the majority of the other

cohorts. However, excluding this study did not alter findings of our meta-analysis. Lastly, data extraction was not blinded, which may also be a source of bias, although this issue is debatable[26].

There are numerous reasons why plasma Aβ is a particularly appealing biomarker: (1) most interventions currently under investigation for AD focus on manipulating Aβ levels, and thus an Aβ-based biomarker may be especially relevant for identifying those who will benefit if such treatments become available; (2) Aβ accumulation appears to be the initial step in AD pathogenesis [27] and thus an Aβ-based biomarker should be especially suitable for identifying patients at the earliest stages of the disease process, when intervention will likely be most effective; and (3) a plasma-based biomarker is simple, inexpensive and non-invasive, all of which are important qualities for population-based screening tools.

In conclusion, despite the limitations of existing research and heterogeneity across the studies considered, this systematic review and meta-analysis suggests that the ratio of plasma $A\beta_{42}$:$A\beta_{40}$ may have value in predicting the risk of later development of dementia or AD and merits further investigation.

Endogenous Sex Hormones and Cognitive Function in Older Women

Alain K. Koyama, Shelley S. Tworoger, A. Heather Eliassen, Olivia I. Okereke, Marc G. Weisskopf,

Bernard Rosner, Kristine Yaffe, Francine Grodstein

2.1 Abstract

Introduction: We examined the association between endogenous sex hormones and both objective and subjective measures of cognitive function.

Methods: We followed 3,044 women up to 23 years in a prospective cohort study. We measured plasma levels of estrone, estrone sulfate, estradiol, androstenedione, testosterone, dehydroepiandrosterone (DHEA), and dehydroepiandrosterone sulfate (DHEA-S) in 1989-1990, conducted neuropsychologic testing in 1999-2008, and inquired about subjective cognition in 2012.

Results: Overall, we observed little relation between plasma levels of hormones and either neuropsychologic test performance or subjective cognition. However, after adjustment for age and education, we observed a borderline significant association of higher levels of plasma estrone with higher scores for both overall cognition (p trend=0.10) and verbal memory (p trend=0.08).

Conclusion: There were no clear associations of endogenous hormone levels at mid-life and cognition in later life, although a suggested finding of higher levels of plasma estrone associated with better cognitive function merits further research.

2.2 Introduction

Despite the public health burden of cognitive impairment on an aging population, the etiology of cognitive decline is still not well understood. Much biological evidence suggests sex hormones may play a role in the development of cognitive decline. For example, estrogen receptors are expressed in many key regions of the brain involved in cognitive function, including the hippocampus and other limbic structures, cingulate and the frontal cortex[28]. Laboratory studies also suggest both direct and indirect neuroprotective effects of estrogens including promotion of hippocampal synaptic plasticity and protection against apoptosis and oxidative stress[29]. Less research exists on the cognitive effects of androgens in women. As with estrogens, androgens can bind to receptors in the brain and may exert neuroprotective effects such as protection against beta-amyloid induced apoptosis and the hyperphosphorylation of tau protein[30, 31]. Additionally, androgen receptors are particularly concentrated in the hippocampus[32], a critical region for learning and memory and one of the earliest regions impacted in the pathogenesis of Alzheimer disease.

Conflicting with the biologic evidence, the pivotal Women's Health Initiative Memory Study randomized controlled trial demonstrated a detrimental effect of combination estrogen and progestin therapy on cognitive function when administered to older women[33]. Observational studies of endogenous hormones (in the absence of exogenous hormone use) may help to reconcile some of the differences in findings with the biological evidence, and could reduce some biases inherent in observational research on hormone therapy[34]. Furthermore, the limited use of androgen therapy in women prohibits large-scale research of exogenous androgens and cognition. While some existing research has indeed addressed the role of endogenous sex hormones in late-life cognitive decline, results have been inconsistent and many studies are limited by cross-sectional analyses or short follow-up times [35, 36].

Finally, there is increasing interest in the use of subjective cognitive concerns (SCC) as an indicator of cognitive function. Existing studies suggest SCC are associated with grey matter atrophy[37], white matter tract degeneration[38], amyloid burden[39], as well as cognitive function[40, 41]. Thus, SCC may provide a complementary outcome in cognitive aging research. We therefore conducted a study to prospectively investigate if plasma levels of sex hormones and their prohormones were associated with objective and subjective measures of cognitive function in a population of older women who provided blood samples at mid-life.

2.3 Methods

2.3.1 Study Population

The Nurses' Health Study (NHS) is an ongoing prospective study of registered nurses in the United States[42]. The study began in 1976, when 121,700 female nurses aged 30 to 55 years completed and returned a mailed questionnaire. Follow-up questionnaires are mailed biennially and a follow-up rate of approximately 90% has been maintained. Baseline for the present analyses occurred from 1989 to 1990, when 32,826 women provided blood samples by overnight mail and completed a short questionnaire. For the present analyses, measures of sex hormones were utilized from previous studies in NHS, including nested case-control studies of breast cancer, ovarian cancer, colon cancer, rheumatoid arthritis, inflammatory bowel disease, stroke and myocardial infarction. Among the 32,826 women with blood samples, 25,964 did not have any sex hormones measured, 1,713 did not have cognitive data (cognitive assessments were only administered to the oldest segment of the cohort), 2,043 were cases from the nested case-control studies, and 62 were missing data on age or age at menopause, resulting in an analytic cohort of 3,044 women with at least one sex hormone measured.

2.3.2 Biomarker Assessment

Upon receipt, blood samples were aliquotted into plasma, white blood cell, and red blood cell components, and stored in liquid nitrogen freezers at -130° C. Further details on the collection and storage procedures have been reported previously[43]. Measured hormones included bound levels of plasma estrone, estrone sulfate, estradiol, androstenedione, testosterone, dehydroepiandrosterone (DHEA), and dehydroepiandrosterone sulfate (DHEA-S).

Estrone, estrone sulfate, estradiol, androstenedione and testosterone were measured by radioimmunoassay at the Quest Diagnostics Nichols Institute (San Juan Capistrano, CA) or by liquid chromatography-tandem mass spectrometry (ThermoFisher Scientific, Franklin, MA and Applied Biosystems-MDS Sciex, Foster City, CA) at the Mayo Medical Laboratories (Rochester, MN). DHEA was measured by radioimmunoassay (Diagnostic Systems Laboratories, Webster, TX) at Quest Diagnostics or by the quantitative sandwich enzyme immunoassay technique at Dr. Nader Rifai's laboratory at the Department of Laboratory Medicine, Children's Hospital Boston (Boston, MA). DHEA-S was measured by the Immulite 2000 a solid-phase, chemiluminescent immunoassay (Siemens Medical Solutions, Los Angeles, CA) at Quest Diagnostics and Mayo Medical Laboratories, or by a coated-tube radioimmunoassay at Dr. Rifai's laboratory. In a prior study in the NHS cohort, levels of sex hormones measured using different assays were highly correlated (R=0.87 for estrone to 0.98 for testosterone)[44]. The assay detection limits were for 10 pg/ml estrone, 40 pg/ml for estrone sulfate, 2 pg/ml for estradiol, 5 ng/dL for androstenedione, 0.5-2 ng/dL for testosterone, 10 ng/dL for DHEA, and 5-15 ug/dL for DHEA-S. Values below the detection limit were set to half the limit.

Average overall coefficients of variation from the measured batches were within acceptable ranges (estrone: 11.3%, estrone-S: 12.6%, estradiol: 13.5%, androstenedione: 9.3%, testosterone: 13.3%,

DHEA: 10.9%, DHEA-S: 6.6%). We adjusted for inter-batch variation using the average-batch calibration method, described by Rosner et. al.[45]. In brief, we assumed the combined batches represented an average batch, and calibrated all hormone levels to have a comparable distribution to the average batch. This was done by regressing hormone levels on their strongest predictors (age and BMI) and indicator variables for each batch. Hormone levels were calibrated by subtracting the difference of the value of the coefficient for the batch and the average of all batch coefficients, effectively adjusting for inter-batch variability independent of differences in age and BMI distribution between batches.

2.3.3 Cognitive Assessment

From 1995 to 2001, a cognitive substudy was initiated in which 19,415 women aged 70 years and older without a history of stroke were administered cognitive testing via telephone. The battery included six cognitive tests. We administered the Telephone Interview of Cognitive Status (TICS)[46], a telephone version of the Mini Mental State Examination (MMSE)[47]; verbal memory was measured using the immediate and delayed recall of the TICS 10-word list, and immediate and delayed recalls of the East Boston Memory Test[48]. The category fluency test, a measure of semantic memory, required participants to recite as many names of animals as possible in one minute[49]. Backward Digit Span, a test of working memory and information processing, required participants to repeat a series of numbers in the reverse order they were given[50]. After the baseline cognitive interview, up to three follow-up assessments were conducted approximately every two years. In addition, in 2012 all women in the parent cohort were asked a series of questions regarding SCC on the mailed questionnaire, which included difficulties in memory, remembering a short list, remembering recent events, understanding or following spoken instructions, understanding a group conversation or the plot of a television program, and finding one's way on familiar streets.

2.3.4 Measurement of Covariates

In primary analyses, we controlled for covariates near blood draw, which were chosen a priori from factors plausibly associated with both endogenous hormone levels and cognitive function based on existing literature. Demographic variables included age and education (registered nurse/associate's degree, bachelor's degree, graduate degree). Because education was only collected from women in the cognitive substudy, in the analysis of SCC we used data on occupational status in 2012 (not working, working full- or part-time). Body mass index (BMI, kg/m^2) was calculated from self-reported height and weight (<22, 22-24.9, 25-29.9, ≥30). Lifestyle factors included smoking status (current, former, never), alcohol consumption (non-drinker, 1-14 g/day, ≥15 g/day), and physical activity, which was measured using a validated physical activity questionnaire (quintiles of metabolic equivalents per week). Comorbidities included a history of self-reported physician diagnosis of diabetes, hypertension, and myocardial infarction. Because depression can be highly correlated with cognitive function, we used measures of depression near cognitive assessment instead of at blood draw. For women in the cognitive substudy, the SF-36 Mental Health Index (MHI) was used to measure depressive symptoms (quintiles). Scores for the MHI range from 0 to 100, with higher scores indicating fewer depressive symptoms[51]. For the analysis of SCC, the 15-item Geriatric Depression Scale (GDS) was used because this scale was used on the 2012 questionnaire in the parent study. Scores on the GDS range from 0 to 15, with scores above 5 suggestive of depression[52]. In both analyses, we also controlled for current antidepressant use.

2.3.5 Statistical Analysis

To test the association between quartiles of plasma hormone levels and the composite outcomes of overall cognition and verbal memory on the neuropsychologic test battery, we used multivariate linear regression models. To reduce measurement error, overall cognition was estimated by

creating a composite score at each time point, averaging the z-scores of each of the six individual cognitive tests (using the variation at baseline to calculate z-scores). Because the NHS cohort is a relatively young and well-educated population, there was not substantial change overall in cognitive scores over the follow-up period. Moreover, the follow-up time from blood draw to cognitive assessment was longer than the follow-up time from the first to last cognitive assessment. Therefore, we conducted analyses with the primary outcome of cognitive status by averaging the composite scores from each time point to create a single measure of cognitive status in older age. Verbal memory was estimated similarly, using the average of the z-scores for the each of the four tests of verbal memory.

In addition, we used multivariate logistic regression to test the association between plasma hormone levels and reporting one or more SCC (versus none) on the 2012 parent questionnaire. For all analyses, we used two models: a basic model adjusted for age and education/occupation, and a full model further adjusted for other potential confounders (BMI, alcohol use, physical activity, age at menopause, depression, antidepressant use). Tests of trend were conducted by modeling the median value of each quartile of hormone level as a continuous variable. In all analyses examining estrone, estrone sulfate, and estradiol, women reporting postmenopausal hormone use at blood draw were excluded to minimize misclassification or confounding from factors associated with hormone therapy use.

To assess possible sources of bias, we conducted several secondary analyses. First, while we did not adjust for cardiovascular disease or diabetes in our primary analyses because they may be causal intermediates, we adjusted for these factors in a secondary analysis. In another analysis, we adjusted for covariates measured near the initial cognitive assessment rather than at blood draw, due to the extended period of time from blood draw until the cognitive interviews. Because the role of any risk

factors can differ among individuals already in the early stages of cognitive disease, we conducted another analysis excluding women with a TICS score of 30 or lower[53] at baseline. Additionally, we were particularly concerned about the influence of depressive symptoms on subjective measures of cognition. Therefore, for a secondary analysis of the association between plasma hormones and SCC, we additionally excluded women who had GDS scores over 5, which is suggestive of depression. Because some women who were not current hormone therapy users at blood draw later reported hormone therapy use over the follow-up period, we conducted another analysis excluding these women to ensure that hormone levels at blood draw best reflected long-term hormone levels. Lastly, in another secondary analysis, we used a repeated measures model with an autoregressive covariance pattern instead of averaging scores over all time points, in order to assess the trajectories of cognitive scores over time. SAS 9.3 (SAS Institute Inc., Cary, NC) was used for all statistical analyses.

2.4 Results

2.4.1 Population Characteristics

Among the 3,044 women, 1,261 had measures for estrone, 912 for estrone sulfate, 1,338 for estradiol, 855 for androstenedione, 2,569 for testosterone, 1,248 for DHEA, and 2,265 for DHEA-S. At blood draw, women ranged in age from 43 to 69 years (mean = 30.1). The mean follow-up time was 9.5 ± 1.4 years from blood draw to the first cognitive interview and 22.6 ± 0.4 years to the SCC measured in 2012. All women were postmenopausal at blood draw and 98.9% were Caucasian. Additional characteristics of the population, by quartiles of estradiol and testosterone, are shown in **Table 2.1** and **Table 2.2**, respectively. On average, women with higher levels of plasma estrone had higher BMI, lower alcohol intake, and a greater likelihood of reporting a history of diabetes. Women with higher levels of plasma DHEA were on average younger, but did not otherwise substantially differ.

Table 2.1 – Characteristics of Participants, by Quartile of Estradiol Level (n=1,338)

	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Age (mean ± SD)	61.7 ± 4.7	61.9 ± 4.6	61.6 ± 5.0	61.0 ± 5.2
Age at menopause (mean ± SD)	47.9 ± 5.4	48.2 ± 5.8	48.1 ± 5.8	48.9 ± 5.1
Education (n, %)[1]				
RN	138 (73.4%)	150 (73.2%)	142 (77.6%)	130 (76.9%)
Bachelor's	36 (19.2%)	40 (19.5%)	25 (13.7%)	30 (17.8%)
Graduate	14 (7.5%)	15 (7.3%)	16 (8.7%)	9 (5.3%)
Smoking (n, %)				
Never	152 (45.5%)	161 (48.4%)	164 (48.8%)	153 (45.7%)
Former	137 (41.0%)	148 (44.4%)	135 (40.2%)	147 (43.9%)
Current	45 (13.5%)	24 (7.2%)	37 (11.0%)	35 (10.5%)
Body mass index (n, %)	23.4 ± 3.3	24.4 ± 3.1	26.3 ± 4.2	29.2 ± 5.2
Alcohol (n, %)				
Non-drinker	123 (36.8%)	123 (36.9%)	141 (42.0%)	173 (51.6%)
1-14 g/day	182 (54.5%)	179 (53.8%)	159 (47.3%)	131 (39.1%)
≥15 g/day	29 (8.7%)	31 (9.3%)	36 (10.7%)	31 (9.3%)
Physical activity, MET-hr/week (mean ± SD)	18.2 ± 21.1	19.1 ± 21.5	17.5 ± 21.8	14.1 ± 17.6
Diabetes (n, %)	5 (1.5%)	8 (2.4%)	16 (4.8%)	21 (6.3%)
Hypertension (n, %)	91 (27.3%)	97 (29.1%)	104 (31.0%)	133 (39.7%)
Myocardial infarction (n, %)	10 (3.0%)	8 (2.4%)	11 (3.3%)	9 (2.7%)
SF-36 Mental Health Index (mean ± SD)[2,3]	82.2 ± 12.5	83.9 ± 10.9	80.9 ± 11.8	82.3 ± 10.5
Current antidepressant use (n, %)[2]	10 (5.3%)	4 (2.0%)	13 (7.1%)	10 (5.9%)

[1]only available in the cognitive substudy

[2]assessed at the most recent measurement prior to the first cognitive interview

[3]range: 0-100 (lower scores indicate more depressive symptoms)

Table 2.2 – Characteristics of Participants, by Quartile of Testosterone Level (n=2,569)

	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Age (mean ± SD)	60.6 ± 5.3	60.0 ± 5.2	60.3 ± 5.5	60.4 ± 5.8
Age at menopause (mean ± SD)	46.6 ± 6.4	47.3 ± 6.2	47.8 ± 5.7	47.9 ± 5.8
Education (n, %)[1]				
RN	231 (77.5%)	196 (72.9%)	216 (75.0%)	224 (75.4%)
Bachelor's	46 (15.4%)	53 (19.7%)	51 (17.7%)	52 (17.5%)
Graduate	21 (7.1%)	20 (7.4%)	21 (7.3%)	21 (7.1%)
Smoking (n, %)				
Never	303 (47.3%)	311 (48.2%)	312 (48.8%)	295 (45.9%)
Former	288 (44.9%)	275 (42.6%)	262 (40.9%)	267 (41.5%)
Current	50 (7.8%)	59 (9.2%)	66 (10.3%)	81 (12.6%)
Body mass index (n, %)	25.3 ± 4.4	25.4 ± 4.3	25.4 ± 4.3	25.2 ± 4.6
Alcohol (n, %)				
Non-drinker	237 (37.0%)	267 (41.4%)	268 (41.9%)	264 (41.1%)
1-14 g/day	348 (54.3%)	324 (50.2%)	309 (48.3%)	311 (48.4%)
≥15 g/day	56 (8.7%)	54 (8.4%)	63 (9.8%)	68 (10.6%)
Physical activity, MET-hr/week (mean ± SD)	17.1 ± 20.0	18.0 ± 30.9	18.1 ± 20.6	17.9 ± 30.5
Diabetes (n, %)	21 (3.3%)	17 (2.6%)	26 (4.1%)	21 (3.3%)
Hypertension (n, %)	206 (32.1%)	173 (26.8%)	189 (29.5%)	192 (29.9%)
Myocardial infarction (n, %)	22 (3.4%)	19 (3.0%)	19 (3.0%)	11 (1.7%)
SF-36 Mental Health Index (mean ± SD)[2,3]	81.2 ± 12.3	82.1 ± 11.8	83.8 ± 11.5	81.0 ± 11.5
Current antidepressant use (n, %)[2]	18 (6.0%)	15 (5.6%)	14 (4.9%)	15 (5.1%)

[1]only available in the cognitive substudy

[2]assessed at the most recent measurement prior to the first cognitive interview

2.4.2 Hormone Levels and Composite Cognitive Scores

Table 2.3 displays the mean differences in overall cognition, by quartile of plasma hormone level. In the model adjusted for age and education, plasma hormone levels were not significantly associated with overall cognition. However, women with higher levels of plasma estrone had higher mean scores for overall cognition, which was borderline statistically significant (p trend=0.10). In the full model, further adjusted for BMI, alcohol use, physical activity, age at menopause, depression status and antidepressant use, results remained largely unchanged. In secondary analyses, results remained similar after further adjustment for hypertension and diabetes, after using the most recent covariates, after exclusion of women with low TICS scores, or exclusion of women who reported hormone therapy use between blood draw and cognitive assessment (data not shown).

Plasma hormone levels were not significantly associated with verbal memory in age and education-adjusted models (**Table 2.4**), similar to results for overall cognition. Women with higher levels of plasma estrone had higher mean verbal memory scores, with borderline statistical significance (p trend=0.08). In the full model, results were similar with no significant associations between plasma hormone levels and verbal memory. Results did not appreciably change after secondary analyses.

2.4.3 Plasma Hormone Levels and SCC

Table 2.5 shows odds ratios for the association between quartile of plasma hormone level and reporting one or more SCC. In the age and occupation-adjusted model, women with higher levels of plasma estrone sulfate in earlier life had a lower odds of SCC (Q4 vs Q1: OR=0.65 [95% CI: 0.43, 1.01]; p trend=0.03). In the full model, these odds ratios were somewhat attenuated (Q4 vs Q1: OR=0.75 [95% CI: 0.47, 1.19]; p trend=0.13). In the basic model, women with higher levels of DHEA had a borderline

significant increased odds of reporting one or more SCC (Q4 vs Q1: OR=1.46 [95% CI: 1.00, 2.12]; p trend=0.07). In the full model, this association was similar (Q4 vs Q1: OR=1.51 [95% CI: 1.03, 2.22]; p trend=0.05). Similar results were seen for DHEA-S (basic model: Q4 vs Q1: OR=1.30 [95% CI: 0.98, 1.71]; p trend=0.09; full model: Q4 vs Q1: OR=1.437 [95% CI: 1.03, 1.83]; p trend=0.04). In secondary analyses, results did not appreciably change. Plasma levels of other hormones were not significantly associated with SCC.

Table 2.3 – Mean Differences in Overall Cognition, by Quartile of Plasma Hormone Level*

	Q1 (ref.)	Q2 vs Q1	Q3 vs Q1	Q4 vs Q1	p trend
Estrone (n=655)					
Model 1	0.00	-0.01 (-0.15, 0.13)	0.00 (-0.14, 0.14)	0.10 (-0.04, 0.24)	0.10
Model 2	0.00	-0.01 (-0.15, 0.13)	-0.04 (-0.18, 0.10)	0.09 (-0.05, 0.24)	0.18
Estrone Sulfate (n=446)					
Model 1	0.00	0.13 (-0.04, 0.30)	0.15 (-0.02, 0.32)	0.08 (-0.08, 0.25)	0.57
Model 2	0.00	0.13 (-0.04, 0.29)	0.12 (-0.05, 0.29)	0.09 (-0.09, 0.26)	0.56
Estradiol (n=688)					
Model 1	0.00	-0.05 (-0.19, 0.09)	-0.04 (-0.18, 0.10)	0.01 (-0.13, 0.15)	0.70
Model 2	0.00	-0.07 (-0.21, 0.07)	-0.03 (-0.17, 0.12)	-0.01 (-0.16, 0.14)	0.85
Androstenedione (n=400)					
Model 1	0.00	0.01 (-0.17, 0.19)	-0.03 (-0.21, 0.15)	-0.09 (-0.27, 0.09)	0.26
Model 2	0.00	-0.04 (-0.23, 0.14)	-0.07 (-0.25, 0.11)	-0.14 (-0.32, 0.05)	0.12
Testosterone (n=1,063)					
Model 1	0.00	0.00 (-0.12, 0.11)	-0.02 (-0.13, 0.10)	0.04 (-0.07, 0.16)	0.46
Model 2	0.00	-0.01 (-0.12, 0.11)	-0.04 (-0.16, 0.07)	0.02 (-0.09, 0.14)	0.68
DHEA (n=522)					
Model 1	0.00	0.02 (-0.14, 0.18)	0.04 (-0.12, 0.20)	-0.04 (-0.20, 0.13)	0.63
Model 2	0.00	0.02 (-0.14, 0.18)	0.03 (-0.13, 0.19)	-0.06 (-0.22, 0.10)	0.44
DHEA Sulfate (n=900)					
Model 1	0.00	0.13 (0.01-0.26)	0.06 (-0.07, 0.18)	0.03 (-0.10, 0.15)	0.83
Model 2	0.00	0.13 (0.01-0.25)	0.05 (-0.07, 0.18)	0.00 (-0.13, 0.13)	0.51

*values indicate standard units of averaged z-scores from each cognitive test

Model 1: adjusted for age, education

Model 2: adjusted for age, education, BMI, alcohol use, physical activity, age at menopause, depression status, antidepressant use

Table 2.4 – Mean Differences in Verbal Memory, by Quartile of Plasma Hormone Level*

	Q1 (ref.)	Q2 vs Q1	Q3 vs Q1	Q4 vs Q1	p trend
Estrone (n=655)					
Model 1	0.00	-0.01 (-0.17, 0.15)	0.00 (-0.16, 0.16)	0.12 (-0.03, 0.28)	0.08
Model 2	0.00	0.00 (-0.16, 0.16)	-0.03 (-0.19, 0.13)	0.11 (-0.05, 0.28)	0.16
Estrone Sulfate (n=446)					
Model 1	0.00	0.19 (0.00, 0.38)	0.14 (-0.05, 0.32)	0.09 (-0.10, 0.27)	0.77
Model 2	0.00	0.19 (0.00, 0.37)	0.10 (-0.09, 0.29)	0.08 (-0.11, 0.28)	0.78
Estradiol (n=688)					
Model 1	0.00	-0.05 (-0.21, 0.10)	-0.03 (-0.19, 0.12)	0.00 (-0.16, 0.15)	0.83
Model 2	0.00	-0.07 (-0.22, 0.09)	-0.02 (-0.18, 0.14)	-0.03 (-0.20, 0.14)	0.94
Androstenedione (n=400)					
Model 1	0.00	-0.03 (-0.23, 0.17)	-0.07 (-0.26, 0.13)	-0.08 (-0.28, 0.12)	0.40
Model 2	0.00	-0.07 (-0.27, 0.14)	-0.09 (-0.29, 0.10)	-0.11 (-0.32, 0.09)	0.30
Testosterone (n=1,063)					
Model 1	0.00	0.01 (-0.11, 0.14)	0.01 (-0.12, 0.13)	0.05 (-0.08, 0.18)	0.43
Model 2	0.00	0.01 (-0.11, 0.14)	-0.01 (-0.14, 0.12)	0.04 (-0.09, 0.17)	0.59
DHEA (n=522)					
Model 1	0.00	-0.07 (-0.25, 0.11)	-0.02 (-0.19, 0.16)	-0.06 (-0.24, 0.12)	0.67
Model 2	0.00	-0.06 (-0.24, 0.12)	-0.02 (-0.20, 0.16)	-0.07 (-0.25, 0.11)	0.53
DHEA Sulfate (n=900)					
Model 1	0.00	0.14 (0.01, 0.28)	0.06 (-0.08, 0.20)	0.05 (-0.08, 0.19)	0.84
Model 2	0.00	0.14 (0.00, 0.27)	0.06 (-0.08, 0.19)	0.03 (-0.11, 0.17)	0.90

*values indicate standard units of averaged z-scores from each cognitive test measuring verbal memory

Model 1: adjusted for age, education

Model 2: adjusted for age, education, BMI, alcohol use, physical activity, age at menopause, depression status, antidepressant use

Table 2.5 – Odds of Subjective Cognitive Concerns, by Quartile of Plasma Hormone Level*

	Q1 (ref.)	Q2 vs Q1	Q3 vs Q1	Q4 vs Q1	*p* trend
		OR (95% CI)	OR (95% CI)	OR (95% CI)	
Estrone (n=949)					
Model 1	1.00	0.88 (0.60-1.28)	1.03 (0.70-1.50)	0.78 (0.54-1.14)	0.26
Model 2	1.00	0.86 (0.58-1.27)	1.09 (0.73-1.61)	0.80 (0.53-1.20)	0.39
Estrone Sulfate (n=701)					
Model 1	1.00	1.05 (0.67-1.62)	1.00 (0.65-1.54)	0.65 (0.43-1.01)	0.03
Model 2	1.00	1.12 (0.71-1.77)	1.06 (0.67-1.67)	0.75 (0.47-1.19)	0.13
Estradiol (n=1,008)					
Model 1	1.00	0.90 (0.63-1.30)	1.04 (0.72-1.49)	0.84 (0.58-1.20)	0.38
Model 2	1.00	0.90 (0.62-1.32)	0.98 (0.66-1.45)	0.84 (0.55-1.28)	0.47
Androstenedione (n=655)					
Model 1	1.00	0.63 (0.40-0.99)	0.71 (0.45-1.13)	0.86 (0.54-1.37)	0.83
Model 2	1.00	0.70 (0.43-1.12)	0.73 (0.45-1.17)	0.96 (0.59-1.55)	0.94
Testosterone (n=2,053)					
Model 1	1.00	1.07 (0.83-1.38)	0.92 (0.72-1.19)	1.08 (0.83-1.39)	0.73
Model 2	1.00	1.06 (0.81-1.38)	0.90 (0.69-1.17)	1.07 (0.82-1.39)	0.78
DHEA (n=966)					
Model 1	1.00	1.32 (0.91-1.90)	1.50 (1.04-2.17)	1.46 (1.00-2.12)	0.07
Model 2	1.00	1.31 (0.90-1.92)	1.55 (1.06-2.27)	1.51 (1.03-2.22)	0.05
DHEA Sulfate (n=1,795)					
Model 1	1.00	1.12 (0.85-1.46)	1.04 (0.79-1.36)	1.30 (0.98-1.71)	0.09
Model 2	1.00	1.12 (0.84-1.48)	1.03 (0.77-1.37)	1.37 (1.03-1.83)	0.04

Model 1: adjusted for age, occupation

Model 2: adjusted for age, occupation, BMI, alcohol use, physical activity, age at menopause, depression status, antidepressant use

*Odds of any versus no SCC across 6 questionnaire items regarding self-perceived cognitive status

2.5 Discussion

We examined whether plasma levels of sex hormones and their prohormones in mid-life were associated with objective and subjective measures of cognitive function in a population of older women. Overall, levels of endogenous hormones were not significantly associated with cognitive function. However, we found a suggestion of modest associations between higher levels of plasma estrone and both overall cognition and verbal memory, and similar findings for the association between higher levels of plasma estrone sulfate and a decreased odds of SCC. There also was a suggestive positive association for DHEA and its primary circulating metabolite, DHEA-S, with SCC.

Despite biological evidence supporting a role of estradiol in cognitive function[29], we did not observe an association between levels of plasma estradiol and either objective or subjective measures of cognitive function in postmenopausal women. Prior epidemiologic studies have reported very mixed results, suggesting a protective[54], harmful[55], null[56], or J-shaped association[57]. The measurement of total estradiol may partly explain the divergent findings, since levels of free and bioavailable estradiol may better represent their potential biologic activity. Moreover, an increasing number of studies demonstrate that estradiol can be produced in the hippocampus[58, 59]. Therefore, it is possible that locally synthesized estradiol has greater potential capacity to affect neurodegenerative processes than circulating estradiol.

To our knowledge, three prior prospective studies have investigated the association between endogenous estrone or estrone sulfate and cognitive function. One study did not report a significant association, although may have been limited by a small sample size (n=148) and short follow-up (2 years)[54]. Two studies showed higher levels of estrone were associated with worse cognitive

outcomes[55, 56]. However, these studies included older populations than our study, and it is possible that the association between estrogen levels and cognitive function can differ with respect to age or time since menopause[60]. The majority of studies examining the association between endogenous estrogens and cognitive function have investigated estradiol, due to its biologic potency relative to estrone or estrone sulfate. However, estrone sulfate occurs in much higher circulating levels in postmenopausal women, and growing biological evidence suggests that neuroprotective effects of estrogens are not limited to estradiol but may also be attributed to estrone or estrone sulfate[61]. Moreover, as results for both SCC and objective measures of cognitive function were qualitatively similar in our study, further investigation on estrone is merited.

Testosterone and androstenedione were not associated with either objective or subjective measures of cognitive function. To our knowledge, no prospective studies have previously investigated the association between androstenedione and cognition in older adults. In contrast to studies of estrogens, epidemiologic studies investigating the role of testosterone have been more consistent with generally null findings[57, 62], in line with the current study. The association between higher levels of DHEA and DHEA-S and SCC was unexpected, as the collective findings from biologic and epidemiologic studies suggest either a null or protective association[36]. It is possible that this is a chance finding, because this association was only significant when using subjective measure of cognitive function, which can be highly variable.

Strengths of this study include the prospective design, long follow-up period including hormone data from mid-life, large sample size, and multiple methods of measuring cognitive status. Because this is an observational study, we cannot discount the possible effects of residual confounding. Another limitation is use of peripheral levels of hormones in the blood which may not correlate with levels in the

brain, possibly explaining null findings. Lastly, a single measurement at baseline may not represent long-term levels of plasma hormones, which may bias results towards the null if long-term levels are most important to cognitive status. However, prior studies in NHS suggest that a single measurement of plasma hormones can reliably represent average levels over up to a 10-year period[63, 64], and our secondary analyses excluding women who used hormone therapy subsequent to blood draw should also help to focus findings on long-term endogenous levels.

In conclusion, we found suggestive evidence that higher plasma levels of estrone and estrone sulfate are positively associated with both objective and subjective measure of cognitive function in older women. Plasma levels of other sex hormones were not clearly associated with cognitive function. Further large prospective studies of the wide range of endogenous hormones measured earlier in life may be particularly useful in consolidating inconsistencies in the collective findings to date and to help understand whether hormones may be important to cognition.

Evaluation of a Self-Administered Computerized Cognitive Battery in an Older Population

Alain K. Koyama, Kaitlin A. Hagan, Olivia I. Okereke, Marc G. Weisskopf, Bernard Rosner,

Francine Grodstein

3.1 Abstract

Objectives: Computerized cognitive testing offers several advantages over traditional methods of measuring cognitive function, yet is still novel in epidemiologic studies. We therefore aimed to assess the utility of the Cogstate, self-administered computerized neuropsychologic battery in a large population of older men.

Methods: We invited 7,167 men, aged 67-94 years, from the Health Professionals Follow-up Study, a prospective cohort of male health professionals. We considered individual Cogstate scores and composite scores measuring psychomotor speed and attention, learning and working memory, and overall cognition. Risk factor data was collected from questionnaires administered four years prior to cognitive testing, and at mid-life, 28 years prior to testing. Multivariate linear regression was used to assess the association between risk factors and each outcome.

Results: The 1,866 men who agreed to complete Cogstate testing were similar to the 5,300 non-responders. Many expected risk factors were associated with Cogstate scores. For example, in multivariate-adjusted models, increasing age was significantly associated with worse performance on all outcomes ($p < 0.001$). For overall cognition, a history of hypertension was significantly associated with worse performance (mean difference=-0.08 standard units [95% CI -0.15, -0.01]) and higher nut consumption was significantly associated with better performance (>2 servings/week vs. never or <1 serving/month: 0.16 [0.04, 0.28]).

Conclusions: The self-administered Cogstate battery showed significant associations with several risk factors known to be associated with cognitive function. Future epidemiologic studies of cognitive aging may benefit from the numerous advantages of self-administered computerized testing.

3.2 Introduction

With an aging global population, the public health burden of dementia is expected to rise rapidly in the near future. Increasing attention must be placed on dementia research to identify new risk factors and interventions. While technological advances in medicine, particularly in neuroimaging and genetics, have already made valuable contributions to our understanding of the disease[65-67], similar advances in the effective measurement of cognitive outcomes have not progressed as quickly.

Epidemiologic studies of cognitive aging typically rely on neuropsychologic tests, which can provide a breadth of data on cognitive function, but require trained interviewers (introducing both inter- and intra-interviewer variability) as well as substantial time and cost on the part of both investigators and study participants. In contrast, computerized cognitive testing offers numerous advantages over traditional neuropsychological testing such as substantially increased cost-efficiency and convenience, accurate response time measurement, and decreased susceptibility to sources of human error such as interviewer bias[68].

The Cogstate brief battery, a computerized series of neuropsychological tests[69], has demonstrated good validity and high test-retest reliability in cognitively normal older adults as well as those with mild cognitive impairment (MCI) or dementia[69-72]. The Cogstate battery is also sensitive enough to detect subtle cognitive decline over 12 months in a population of older adults with MCI[73] and to differentiate between normal cognitive function, MCI, and Alzheimer's disease[70]. However,

prior studies using the Cogstate battery in older populations are in predominately small samples and often used only in supervised clinical or research settings. While unsupervised self-administration (e.g. from the participant's home) can maximize efficiency and convenience, to our knowledge, no prior studies have involved unsupervised self-administration in a large population. As a valuable supplement or alternative to traditional methods of neuropsychological testing, it remains of interest to examine the feasibility of self-administered computerized testing in a large-scale setting, particularly in older populations. We therefore aimed to evaluate the usability and distribution of scores of the Cogstate brief battery in a population of older adults.

3.3 Methods

3.3.1 Study Population

The Health Professionals Follow-up Study (HPFS) is an ongoing longitudinal study which began in 1986, when 51,529 men aged 40-75 years in allied health professions were recruited. Participants were originally recruited via mailed questionnaires, with follow-up data collected using biennial questionnaires. Health and lifestyle data for the present study was collected using the 2010 questionnaire, allowing a slight lag between risk factor evaluation and cognitive assessment to reduce the possibility of reverse causation. Computerized cognitive testing was conducted in 2014, when email invitations to complete testing were sent out to the 7,167 men who had completed the 2014 mailed questionnaire and had email addresses available.

3.3.2 Measurements

Covariates were chosen a priori based on risk factors known to be associated with cognitive function in prior literature. Age was calculated from self-reported date of birth. Body mass index (BMI, kg/m^2) was calculated from self-reported height and weight (<22, 22-24.9, 25-29.9, ≥30). Lifestyle

factors included smoking status (never, former or current) and physical activity measured as estimated mean energy expended per week (quartiles of metabolic equivalents per week) using a validated physical activity questionnaire. Dietary factors, recorded using a validated semi-quantitative food frequency questionnaire[74, 75], included current multivitamin use, alcohol intake (none, 1-2 servings per day, >2 per day), nut intake (<1 serving/month, 1-3 servings/month, 1-2 servings/week, >2 servings/week), fish intake (<1 serving/month, 1-3 servings/month, 1-2 servings/week, >2 servings/week) and total energy intake (kcal/day). Because extensive dietary data was not available from the 2010 questionnaire, alcohol and nut intake were recorded using data from the 2006 questionnaire. Comorbidities included a history of self-reported physician diagnosis of diabetes, hypertension, and myocardial infarction. Because of the potential importance of mid-life factors, we also collected information on all covariates from the 1986 questionnaire.

3.3.3 Cognitive Assessment

Cognitive function was measured using the self-administered Cogstate computerized battery[69]. The Cogstate battery comprises four tasks taking approximately 15-20 minutes to complete in total. At the beginning of each task, participants view instructions for each task, perform a practice trial for that task, and then are given the actual task to complete. All tasks involve images of playing cards, due to their familiarity to most ages and cultures. Each task requires participants to respond to the playing cards, using the “K” and “D” keys on their computer keyboard which correspond to a “Yes” or “No” response, respectively Descriptions of each of the four tasks are presented below, with participants performing the tasks in the order presented.

The Detection Task (DET) measures psychomotor function and information processing speed. The participant views a series of joker playing cards on the screen turn over. When a card turns

over, the participant must then press the “Yes” key as quickly as possible.

The Identification Task (IDN) measures visual attention and vigilance. The screen shows red or black joker cards flipping over, and participants press the “Yes” and “No” keys as quickly as possible to note the red cards (i.e., “yes” if red, “no” if black).

The One Card Learning Task (OCL) measures visual learning and short-term memory. A series of playing cards is flipped over on the screen one at a time. Each time a card is revealed, the participant must then respond “Yes” or “No” to note whether that card has been previously shown at any time during the task.

The One Back Task (ONB) is designed to measure attention and working memory. A series of playing cards is flipped over on the screen one at a time. When each card is revealed, the participants responds “yes” or “no” to note whether the card is the same as the previous card.

For the DET, IDN, and ONB tasks, scores are the $\log_{10}$ transformed mean response times of correct trials, while for the OCL task, scores are the arcsine of the square root of the proportion of correct responses (transformations are applied to normalize the distribution). In addition to assessing each task individually, we also created composite scores since composite measurements may increase power through increased precision and sensitivity[76, 77]. Composite scores were created by averaging the z-scores of scores from individual tasks. Three composite scores were created: 1) DET and IDN, to measure psychomotor speed and attention; 2) OCL and ONB, to measure learning and working memory; 3) all four tests, as a measure of overall cognition.

3.3.4 Statistical Analysis

We created summary statistics to describe the mean, variability and range of scores on each Cogstate task. Univariate analyses were done to assess associations between response or non-response and the risk factors associated with cognitive function. Chi-square tests were used for categorical variables. We used Kruskal-Wallis tests for continuous variables because all continuous variables were non-normally distributed. To evaluate the association between risk factors and scores on each task or composite score, we conducted a separate multivariate linear regression model for each cognitive outcome. Covariates for all models included age, smoking status, BMI, physical activity, alcohol intake, nut intake, total energy intake, diabetes, hypertension, and myocardial infraction. Because mid-life factors may be important predictors for late life cognitive function, we conducted an additional analysis using risk factors measured at mid-life in 1986. In addition, we used integrity criteria to exclude any Cogstate scores below established thresholds[69], which were (in percent of trials correct) 80% for the DET task, 80% for IDN, 50% for OCL and 70% for ONB. In a secondary analysis, we used more inclusive integrity criteria, excluding only those scores for which participants scored 0% correct on a task. Linear tests of trend for ordinal variables (physical activity, nut intake, fish intake) were conducted by modeling the median value of each category as a continuous variable. All analyses were performed using SAS 9.3 (SAS Institute Inc., Cary, NC).

3.4 Results

3.4.1 Population Characteristics

Baseline characteristics of all participants invited to complete Cogstate testing are shown in **Table 3.1.** Men ranged in age from 63 to 95 years (mean = 71.0). Among the 7,167 men who were invited to participate, 1,866 (26%) conducted Cogstate testing. Overall, differences were small between men who responded and those who did not respond. On average, men who did not respond were

Table 3.1 – Baseline Characteristics of HPFS Participants, by Response Status (n=7,166)

		Responders (n=1,866)*	
	Non-Responders (n=5,300)	Complete battery (n=1,767)	Incomplete battery (n=279)
Age, mean ± SD	71.3 ± 6.2	69.7 ± 5.4	73.0 ± 5.9
Smoking, n (%)			
Never	2,132 (48.9%)	651 (47.8%)	104 (42.5%)
Former/Current	2,225 (51.1%)	712 (52.2%)	141 (57.6%)
Body mass index, kg/m^2, n (%)			
<22	438 (8.8%)	128 (8.4%)	24 (9.0%)
22-24.9	1,534 (31.0%)	467 (30.7%)	80 (30.0%)
25-29.9	2,248 (45.4%)	707 (46.5%)	121 (45.3%)
30+	733 (14.8%)	220 (14.5%)	42 (15.7%)
Physical activity, MET-hr/week, n (%)	40.7 ± 38.6	42.1 ± 38.5	38.6 ± 36.4
Alcohol, servings/day, n (%)			
None	2,078 (43.4%)	619 (41.8%)	107 (41.2%)
1-2	1,973 (41.2%)	630 (42.6%)	110 (42.3%)
>2	734 (15.3%)	231 (15.6%)	43 (16.5%)
Nut Intake, servings/day, n (%)			
<1/month	529 (10.9%)	156 (10.4%)	41 (15.4%)
1-3/month	542 (11.1%)	156 (10.4%)	32 (12.0%)
1-2/week	1,459 (29.9%)	468 (31.2%)	82 (30.7%)
>2/week	2,343 (48.1%)	722 (48.1%)	112 (42.0%)
Fish Intake, servings/day, n (%)			
<1/month	340 (7.0%)	109 (7.3%)	13 (4.9%)
1-3/month	565 (11.6%)	178 (11.9%)	32 (12.1%)
1-2/week	2,488 (51.2%)	778 (51.9%)	140 (53.0%)
>2/week	1,466 (30.2%)	433 (28.9%)	79 (29.9%)
Diabetes, n (%)	548 (10.3%)	138 (8.7%)	35 (12.5%)
Hypertension, n (%)	3,110 (58.7%)	843 (53.1%)	173 (62.0%)
Myocardial infarction, n (%)	393 (7.4%)	110 (6.9%)	25 (9.0%)
Stroke, n (%)	138 (2.6%)	29 (1.8%)	9 (3.2%)
Current multivitamin use, n (%)	3,402 (69.4%)	1,027 (68.4%)	176 (65.9%)

*Complete battery: participants who passed integrity criteria for all four Cogstate tasks; Incomplete battery: participants who did not pass integrity criteria for at least one Cogstate task

significantly more likely to be older than those who did respond (mean=71.3 and 70.2 years, respectively) and report a history of hypertension (prevalence=58.7% and 54.5%, respectively). In addition, among those who responded, 15% (n=279 men) did not pass integrity criteria for at least one task (99 did not pass integrity criteria on all four tasks, and were excluded from analysis). The number of integrity failures were generally greater for the more difficult tasks (DET: n=123, IDN: n=154, OCL: n=197, ONB: n=168). When using more inclusive integrity criteria (flagging participants who scored 0% correct on a task), this pattern reversed (DET: n=94, IDN: n=28, OCL: n=0, ONB: n=8). Compared to men who passed integrity criteria on all tasks, men who scored below integrity criteria on at least one task were on average, older, reported lower nut consumption, and had a history of hypertension and diabetes.

The distribution of scores for each task are shown in **Figure 3.1**. Mean scores for each task were 2.59 ± 0.09 for DET, 2.73 ± 0.07 for IDN, 1.00 ± 0.11 for OCL, and 2.92 ± 0.09 for ONB. Ranges for each score were 2.34 to 3.17 for DET, 2.55 to 3.18 for IDN, 0.79 to 1.38 for OCL, and 2.70 to 3.36 for ONB. Scores for the reaction time-based scores (DET, IDN, ONB) demonstrated a slight positive skew, while scores for OCL (based on proportion correct) showed a slight negative skew.

3.4.2 Association Between Participant Characteristics and Cogstate Scores

Results showing the associations between risk factors and scores on individual tasks are shown in **Table 3.2.** As expected, older men had significantly worse mean scores on all cognitive outcomes ($p < 0.001$). Men with higher BMI had significantly worse mean scores on the OCL task (30+ kg/m^2 vs. 22-24.9 kg/m^2: mean difference=-0.021 points [95% CI -0.040, -0.002]; 25-29.9 kg/m^2 vs. 22-24.9 kg/m^2: -0.016 [-0.029, -0.002]) and better mean scores on the ONB task (30+ kg/m^2 vs. 22-24.9 kg/m^2: mean difference =-0.023 [-0.037, -0.008]; 25-29.9 kg/m^2 vs. 22-24.9 kg/m^2: mean difference -0.012 [-0.022, -0.002]). Men

Figure 3.1 - Distribution of Scores on Cogstate Tasks. For the Detection, Identification, and One Back tasks, scores are the $\log_{10}$ transformed mean response times of correct trials. For the One Card Learning task, scores are the arcsine of the square root of the proportion of correct responses.

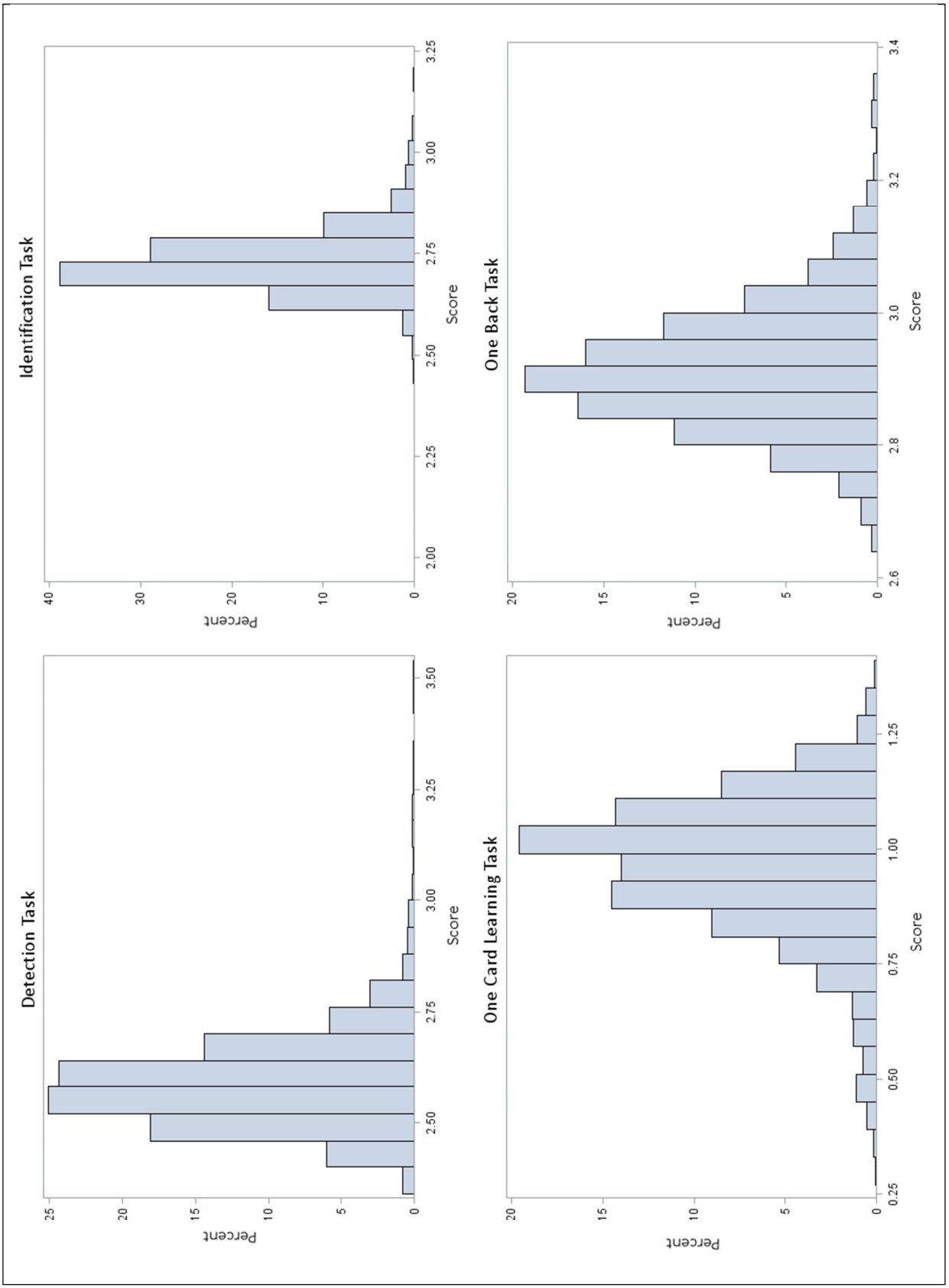

Identification Task
Percent
Score
Detection Task
Percent
Score
One Back Task
Percent
Score
One Card Learning Task
Percent
Score

Table 3.2 – Associations Between Risk Factors and Cogstate Scores

	DET (n=1,620)		IDN (n=1,591)		OCL (n=1,546)		ONB (n=1,576)	
	β	*p*	β	*p*	β	*p*	β	*p*
Age (years)	0.002 (0.002, 0.003)	<0.001	0.002 (0.001, 0.002)	<0.001	-0.004 (-0.005, -0.003)	<0.001	0.004 (0.003, 0.004)	<0.001
Body mass index								
<22 kg/m^2	0.007 (-0.011, 0.025)	0.46	-0.002 (-0.014, 0.011)	0.76	-0.008 (-0.030, 0.013)	0.44	0.010 (-0.007, 0.027)	0.23
22-24.9 kg/m^2 (ref.)	-	-	-	-	-	-	-	-
25-29.9 kg/m^2	-0.001 (-0.012, 0.010)	0.87	-0.003 (-0.010, 0.005)	0.49	-0.016 (-0.029, -0.002)	0.02	-0.012 (-0.022, -0.002)	0.02
30+ kg/m^2	-0.007 (-0.022, 0.009)	0.40	-0.008 (-0.019, 0.003)	0.14	-0.021 (-0.040, -0.002)	0.03	-0.023 (-0.037, -0.008)	0.002
Physical activity, MET-hr/wk								
First quartile (ref.)	-	-	-	-	-	-	-	-
Second quartile	-0.001 (-0.014, 0.012)	0.88	-0.004 (-0.013, 0.005)	0.38	0.014 (-0.003, 0.030)	0.10	0.002 (-0.010, 0.015)	0.74
Third quartile	0.001 (-0.013, 0.015)	0.88	-0.002 (-0.011, 0.008)	0.76	0.005 (-0.011, 0.022)	0.53	0.003 (-0.010, 0.016)	0.61
Fourth quartile	-0.007 (-0.021, 0.007)	0.32	-0.005 (-0.015, 0.004)	0.28	0.007 (-0.010, 0.023)	0.44	0.001 (-0.012, 0.014)	0.92
p-trend		0.29		0.42		0.89		0.99
Alcohol, servings/day								
None (ref.)	-	-	-	-	-	-	-	-
1-2	-0.012 (-0.022, -0.002)	0.02	-0.003 (-0.010, 0.004)	0.47	0.000 (-0.012, 0.013)	0.95	-0.008 (-0.017, 0.002)	0.10
>2	0.005 (-0.009, 0.019)	0.51	0.001 (-0.009, 0.011)	0.87	0.000 (-0.016, 0.017)	0.96	0.001 (-0.012, 0.014)	0.88
Nut Intake, servings								
<1/month (ref.)	-	-	-	-	-	-	-	-
1-3/month	0.001 (-0.018, 0.021)	0.89	0.002 (-0.012, 0.016)	0.81	0.017 (-0.007, 0.041)	0.15	-0.004 (-0.023, 0.014)	0.65
1-2/week	-0.012 (-0.028, 0.004)	0.15	-0.003 (-0.014, 0.009)	0.66	0.022 (0.003, 0.042)	0.03	-0.010 (-0.026, 0.005)	0.18
>2/week	-0.013 (-0.029, 0.003)	0.10	-0.004 (-0.016, 0.007)	0.46	0.018 (-0.001, 0.038)	0.06	-0.014 (-0.029, 0.001)	0.07
p-trend		0.12		0.33		0.59		0.10

Notes: Lower scores indicate better performance for DET, IDN, and ONB. Higher scores indicate better performance for OCL. All variables were placed into the same model.

Table 3.2 (Continued) – Associations Between Risk Factors and Cogstate Scores

	DET (n=1,620)		IDN (n=1,591)		OCL (n=1,546)		ONB (n=1,576)	
	β	*p*	β	*p*	β	*p*	β	*p*
Fish Intake, servings								
<1/month (ref.)	-	-	-	-	-	-	-	-
1-3/month	0.001 (-0.021, 0.023)	0.91	-0.001 (-0.016, 0.015)	0.95	-0.004 (-0.030, 0.022)	0.77	0.018 (-0.002, 0.038)	0.08
1-2/week	-0.006 (-0.025, 0.013)	0.53	-0.007 (-0.020, 0.006)	0.28	-0.005 (-0.028, 0.017)	0.64	0.004 (-0.014, 0.021)	0.69
>2/week	-0.007 (-0.026, 0.013)	0.49	-0.002 (-0.015, 0.012)	0.81	0.000 (-0.023, 0.024)	0.98	0.006 (-0.012, 0.024)	0.49
p-trend		0.38		0.74		0.59		0.84
Diabetes	0.013 (-0.003, 0.030)	0.11	0.009 (-0.003, 0.020)	0.13	0.016 (-0.004, 0.036)	0.12	0.000 (-0.015, 0.015)	0.99
Hypertension	0.004 (-0.006, 0.013)	0.45	0.007 (0.000, 0.013)	0.05	-0.009 (-0.020, 0.003)	0.13	0.006 (-0.002, 0.015)	0.15
Myocardial infarction	-0.001 (-0.019, 0.017)	0.92	-0.009 (-0.022, 0.003)	0.14	0.000 (-0.022, 0.022)	0.99	-0.005 (-0.021, 0.012)	0.60

Notes: Lower scores indicate better performance for DET, IDN, and ONB. Higher scores indicate better performance for OCL. All variables were placed into the same model.

Table 3.3 – Associations Between Risk Factors and Cogstate Composite Scores

	DET + IDN (n=1,580) (psychomotor speed, attention)		**OCL + ONB (n=1,505) (learning, working memory)**		**Overall Cognition (n=1,469)**	
	β	*p*	β	*P*	B	*P*
Age (years)	-0.03 (-0.04, -0.02)	<0.001	-0.04 (-0.04, -0.03)	<0.001	-0.03 (-0.04, -0.03)	<0.001
Body mass index						
<22 kg/m^2	-0.01 (-0.18, 0.17)	0.95	-0.08 (-0.22, 0.07)	0.29	-0.03 (-0.16, 0.10)	0.64
22-24.9 kg/m^2 (ref.)	-	-	-	-	-	-
25-29.9 kg/m^2	0.03 (-0.08, 0.13)	0.62	0.00 (-0.09, 0.09)	0.97	0.02 (-0.07, 0.10)	0.71
30+ kg/m^2	0.11 (-0.04, 0.26)	0.15	0.04 (-0.09, 0.16)	0.57	0.06 (-0.05, 0.18)	0.28
Physical activity, MET-hr/wk						
First quartile (ref.)	-	-	-	-	-	-
Second quartile	0.04 (-0.08, 0.17)	0.51	0.04 (-0.07, 0.15)	0.51	0.01 (-0.09, 0.11)	0.79
Third quartile	0.00 (-0.13, 0.13)	0.97	0.00 (-0.12, 0.11)	0.94	-0.03 (-0.13, 0.08)	0.61
Fourth quartile	0.07 (-0.06, 0.20)	0.29	0.01 (-0.11, 0.12)	0.92	0.00 (-0.10, 0.10)	0.99
p-trend		0.38		0.82		0.85
Alcohol, servings/day						
None (ref.)	-	-	-	-	-	-
1-2	0.08 (-0.02, 0.18)	0.10	0.03 (-0.05, 0.11)	0.51	0.04 (-0.03, 0.12)	0.27
>2	-0.04 (-0.17, 0.10)	0.58	-0.02 (-0.13, 0.10)	0.76	-0.03 (-0.14, 0.07)	0.53
Nut Intake, servings						
<1/month (ref.)	-	-	-	-	-	-
1-3/month	-0.03 (-0.22, 0.16)	0.76	0.10 (-0.07, 0.26)	0.25	0.04 (-0.11, 0.18)	0.64
1-2/week	0.09 (-0.07, 0.24)	0.29	0.16 (0.02, 0.29)	0.02	0.12 (0.00, 0.25)	0.05
>2/week	0.10 (-0.05, 0.26)	0.20	0.17 (0.04, 0.30)	0.01	0.15 (0.03, 0.27)	0.01
p-trend		0.16		0.08		0.02
Fish Intake, servings						
<1/month (ref.)	-	-	-	-	-	-
1-3/month	-0.01 (-0.22, 0.20)	0.91	-0.13 (-0.30, 0.05)	0.15	-0.08 (-0.24, 0.08)	0.33
1-2/week	0.08 (-0.09, 0.26)	0.36	-0.03 (-0.18, 0.11)	0.65	0.02 (-0.11, 0.16)	0.72
>2/week	0.05 (-0.14, 0.23)	0.61	-0.04 (-0.20, 0.12)	0.61	0.01 (-0.13, 0.15)	0.93
p-trend		0.72		0.75		0.60
Diabetes	-0.16 (-0.32, 0.00)	0.04	0.10 (-0.03, 0.24)	0.14	-0.05 (-0.17, 0.08)	0.47
Hypertension	-0.07 (-0.16, 0.02)	0.11	-0.08 (-0.16, 0.00)	0.04	-0.08 (-0.15, -0.01)	0.03
Myocardial infarction	0.09 (-0.08, 0.26)	0.31	0.02 (-0.13, 0.16)	0.82	0.04 (-0.09, 0.18)	0.51

Notes: Higher scores indicate better performance. All variables were placed into the same model

who reported 1-2 drinks per day had significantly better mean scores on the DET task compared to non-drinkers (mean difference =-0.012 [-0.022, -0.002]). On average, men with a history of hypertension had worse scores on the IDN task compared to those without a history of hypertension (mean difference = 0.007 [0.000, 0.013]). Men who reported more frequent nut intake generally had better mean scores on the OCL task (1-3 servings/month vs. <1 serving/month: mean difference 0.017 [-0.007, 0.041]; 1-2 servings/week vs. <1 serving/month: mean difference 0.022 [0.003, 0.042]; >2 servings/week vs. <1 serving/month: mean difference 0.018 [-0.001, 0.038]), although a linear test for trend was not significant (p=0.59). Results remained similar when using more inclusive integrity criteria.

Associations between risk factors and composite scores are shown in **Table 3.3**. Increased nut intake was associated with higher mean scores on overall cognition (>2 servings/week vs. <1 serving/month: 0.15 standard units [0.03, 0.27]; 1-2 servings/week vs. <1 serving/month: 0.12 standard units [0.00, 0.25]; 1-3 servings/month vs. <1 serving/month: 0.04 standard units [-0.11, 0.18]; p-trend = 0.02). On average, men with a history of diabetes had significantly worse scores for the composite outcome of psychomotor speed and attention (-0.16 standard units [-0.32, 0.00]). Lastly, men reporting a history of hypertension had on average, significantly worse scores on the composite outcomes of learning and working memory (-0.08 standard units [-0.16, 0.00]) and overall cognition (-0.08 standard units [-0.15, -0.01]).

When assessing the association between mid-life risk factors measured in 1986 and composite Cogstate scores, associations for some risk factors changed (**Table 3.4**). In contrast to the main analysis, men with higher levels of physical activity had better mean scores on learning and working memory (p-trend = 0.02) and overall cognition (p trend = 0.049). In addition, men who consumed fish more

Table 3.4 – Associations Between Mid-life Risk Factors and Cogstate Composite Scores

	DET + IDN (n=1,672) (psychomotor speed, attention)		**OCL + ONB (n=1,597) (learning, working memory)**		**Overall Cognition (n=1,559)**	
	β	*p*	B	*p*	β	*P*
Age (years)	-0.03 (-0.04, -0.02)	<0.001	-0.04 (-0.04, -0.03)	<0.001	-0.03 (-0.04, -0.03)	<0.001
Body mass index						
<22 kg/m^2	0.03 (-0.11, 0.18)	0.65	-0.07 (-0.19, 0.05)	0.28	-0.03 (-0.14, 0.08)	0.62
22-24.9 kg/m^2 (ref.)	-	-	-	-	-	-
25-29.9 kg/m^2	0.00 (-0.09, 0.10)	0.93	-0.02 (-0.10, 0.06)	0.63	-0.01 (-0.09, 0.06)	0.69
30+ kg/m^2	0.19 (-0.01, 0.40)	0.06	0.12 (-0.05, 0.29)	0.17	0.17 (0.01, 0.32)	0.04
Physical activity, MET-hr/wk						
First quartile (ref.)	-	-	-	-	-	-
Second quartile	-0.04 (-0.15, 0.08)	0.54	0.01 (-0.09, 0.11)	0.85	-0.04 (-0.14, 0.05)	0.34
Third quartile	0.06 (-0.06, 0.18)	0.31	0.13 (0.03, 0.24)	0.01	0.08 (-0.01, 0.18)	0.08
Fourth quartile	0.10 (-0.02, 0.22)	0.09	0.03 (-0.07, 0.13)	0.58	0.06 (-0.03, 0.15)	0.20
p-trend		0.02		0.56		0.049
Alcohol, servings/day						
None (ref.)	-	-	-	-	-	-
1-2	-0.02 (-0.11, 0.07)	0.67	0.00 (-0.08, 0.08)	0.96	-0.02 (-0.09, 0.05)	0.63
>2	-0.07 (-0.21, 0.07)	0.32	0.02 (-0.10, 0.13)	0.76	-0.04 (-0.15, 0.06)	0.45
Nut Intake, servings						
<1/month (ref.)	-	-	-	-	-	-
1-3/month	0.01 (-0.12, 0.14)	0.91	-0.01 (-0.12, 0.10)	0.85	0.02 (-0.08, 0.12)	0.73
1-2/week	-0.02 (-0.14, 0.09)	0.67	0.05 (-0.05, 0.14)	0.36	0.02 (-0.06, 0.11)	0.60
>2/week	0.05 (-0.09, 0.18)	0.51	0.08 (-0.03, 0.20)	0.16	0.08 (-0.02, 0.19)	0.13
p-trend		0.40		0.12		0.11
Fish Intake, servings						
<1/month (ref.)	-	-	-	-	-	-
1-3/month	0.29 (0.08, 0.49)	0.006	0.01 (-0.17, 0.19)	0.91	0.13 (-0.03, 0.29)	0.10
1-2/week	0.31 (0.13, 0.49)	<0.001	0.03 (-0.12, 0.19)	0.67	0.15 (0.01, 0.29)	0.03
>2/week	0.28 (0.09, 0.46)	0.004	0.04 (-0.12, 0.20)	0.64	0.15 (0.00, 0.29)	0.04
p-trend		0.42		0.65		0.38
Hypertension	-0.05 (-0.18, 0.08)	0.47	0.03 (-0.08, 0.14)	0.57	0.01 (-0.10, 0.11)	0.91
Myocardial infarction	-0.09 (-0.67, 0.49)	0.75	0.14 (-0.31, 0.60)	0.54	-0.01 (-0.44, 0.42)	0.97

Notes: Higher scores indicate better performance. All variables were placed into the same model.

frequently had higher scores on the composite outcomes of psychomotor speed and attention (>2 servings/week vs. <1 serving/month: 0.28 standard units [0.09, 0.46]; 1-2 servings/week vs. <1 serving/month: 0.31 standard units [0.13, 0.49]; 1-3 servings/month vs. <1 serving/month: 0.29 standard units [0.08, 0.49]; p-trend = 0.42) and overall cognition (>2 servings/week vs. <1 serving/month: 0.15 standard units [0.00, 0.29]; 1-2 servings/week vs. <1 serving/month: 0.15 standard units [0.01, 0.29]; 1-3 servings/month vs. <1 serving/month: 0.13 standard units [-0.03, 0.29]; p-trend = 0.38). Similar to the main analysis, men who reported the most frequent nut intake (>2 servings/week) had better mean scores for overall cognition compared to men who reported the least frequent nut intake (<1 serving/month), but this association only reached borderline significance (0.08 standard units [-0.02, 0.19]). When using more inclusive integrity criteria, results were similar.

3.5 Discussion

To our knowledge, this is the first large, population-based study to conduct unsupervised, self-administered, computerized cognitive testing in older adults. Although the participation rate was low in these older men, characteristics of participants who responded were generally similar to those who did not, suggesting that low participation is not differentially attributed to risk factors for cognitive decline; this is important since it suggests that non-participation would reduce sample size but would not introduce meaningful bias into research findings. In addition, several factors known to be associated with cognitive function were significantly associated with Cogstate scores, supporting the validity of the battery in measuring several cognitive domains.

In addition to the low participation, the proportion of participants who did not complete a Cogstate task above integrity criteria (15%), as well as participant feedback during data collection (e.g. confusion regarding task instructions), suggest the need in self-administered testing for clear and

unambiguous instructions and a user interface properly optimized for an older population. As prior studies in older adults suggest a willingness or even preference for digital interfaces in primary data collection[78-80], an appropriate interface and instructions can be vital to maximize task completion for a population who may be limited by sensory impairments and/or a low level of computer literacy. This should be a clear priority in future research involving self-administered cognitive evaluation in older adults.

Nonetheless, the distribution of scores on Cogstate tasks was generally similar to those reported in prior studies using trained, cognitively normal older participants. Mean scores on the IDN task in our study (2.73 ± 0.07) were very similar compared to prior studies (Hammers et al.: n=23, mean age=68.4 ± 9.5, mean score=2.73 ± 0.08; Fredrickson et al.: n=301, mean age=61.9 ± 7.2, mean score=2.72 ± 0.07; and Lim et al.: n=15, mean age 73.6 ± 6.9, mean score=2.73 ± 0.06)[69, 81, 82]. The OCL and ONB tasks, despite being the most difficult and thus likely to have greater variability, also had very similar mean scores compared to prior studies[69, 81, 82]. In contrast, mean scores for the DET task in our study (2.59 ± 0.09), were slightly worse than those reported in other studies of cognitively normal older adults (Hammers et al.: mean score=2.50 ± 0.11, Fredrickson et al.: mean score=2.52 ± 0.11[69], Lim et al.: mean score=2.56 ± 0.10) and more similar to scores among participants with mild cognitive impairment (Hammers et al.: n=20, mean age=73.5 ± 5.9, mean score=2.52 ± 0.08), Lim et al. (n=47, mean age 78.9 ± 6.9, mean score=2.59 ± 0.12). Although such differences may be attributed to random chance, it is also possible that worse performance on DET, the simplest task, reflected difficulties with task comprehension, given that this task was both administered first and had the highest proportion of participants scoring 0% correct.

A possible limitation of this study is the relative homogeneity and high education level of the population. Thus, participation and performance on the Cogstate battery and participants' level of task comprehension may differ in other older populations. It is also possible that the low participation and somewhat low task comprehension may not be generalizable to somewhat younger populations. However, task comprehension will likely improve in future cohorts as computer literacy increases in older adults[83]. Additionally, although prior studies of the Cogstate battery in older adults demonstrate good correlation with performance on other neuropsychological test instruments[72, 84], the Cogstate battery may not adequately measure some cognitive domains such as executive functions or semantic and verbal fluency.

In conclusion, the Cogstate self-administered test showed promising results, with performance on Cogstate tasks significantly associated with several known risk factors for cognitive decline. Further studies to establish psychometric standards and normative data in different populations would be helpful to promote more widespread application in clinical and research settings.

References

1. Alzheimer's Association, *2010 Alzheimer's disease facts and figures.* Alzheimers Dement, 2010. **6**(2): p. 158-94.
2. Eschweiler, G.W., et al., *New developments in the diagnosis of dementia.* Dtsch Arztebl Int, 2010. **107**(39): p. 677-83.
3. Alzheimer's Association, N.I.o.A. *Recommendations to Update Diagnostic Criteria*. 2010; Available from: http://www.alz.org/research/diagnostic_criteria/.
4. Abdullah, L., et al., *Serum Abeta levels as predictors of conversion to mild cognitive impairment/Alzheimer disease in an ADAPT subcohort.* Mol Med, 2009. **15**(11-12): p. 432-7.
5. Blasko, I., et al., *Conversion from cognitive health to mild cognitive impairment and Alzheimer's disease: prediction by plasma amyloid beta 42, medial temporal lobe atrophy and homocysteine.* Neurobiol Aging, 2008. **29**(1): p. 1-11.
6. Graff-Radford, N.R., et al., *Association of low plasma Abeta42/Abeta40 ratios with increased imminent risk for mild cognitive impairment and Alzheimer disease.* Arch Neurol, 2007. **64**(3): p. 354-62.
7. Lambert, J.C., et al., *Association of plasma amyloid beta with risk of dementia: the prospective Three-City Study.* Neurology, 2009. **73**(11): p. 847-53.
8. Lopez, O.L., et al., *Plasma amyloid levels and the risk of AD in normal subjects in the Cardiovascular Health Study.* Neurology, 2008. **70**(19): p. 1664-71.
9. Mayeux, R., et al., *Plasma A[beta]40 and A[beta]42 and Alzheimer's disease: relation to age, mortality, and risk.* Neurology, 2003. **61**(9): p. 1185-90.
10. Mayeux, R., et al., *Plasma amyloid beta-peptide 1-42 and incipient Alzheimer's disease.* Ann Neurol, 1999. **46**(3): p. 412-6.

11. Okereke, O.I., et al., *Ten-year change in plasma amyloid beta levels and late-life cognitive decline.* Arch Neurol, 2009. **66**(10): p. 1247-53.

12. Pomara, N., et al., *Selective reductions in plasma Abeta 1-42 in healthy elderly subjects during longitudinal follow-up: a preliminary report.* Am J Geriatr Psychiatry, 2005. **13**(10): p. 914-7.

13. Schupf, N., et al., *Peripheral Abeta subspecies as risk biomarkers of Alzheimer's disease.* Proc Natl Acad Sci U S A, 2008. **105**(37): p. 14052-7.

14. Seppala, T.T., et al., *Plasma A{beta}42 and A{beta}40 as markers of cognitive change in follow-up: a prospective, longitudinal, population-based cohort study.* J Neurol Neurosurg Psychiatry, 2010.

15. Sundelof, J., et al., *Plasma beta amyloid and the risk of Alzheimer disease and dementia in elderly men: a prospective, population-based cohort study.* Arch Neurol, 2008. **65**(2): p. 256-63.

16. van Oijen, M., et al., *Plasma Abeta(1-40) and Abeta(1-42) and the risk of dementia: a prospective case-cohort study.* Lancet Neurol, 2006. **5**(8): p. 655-60.

17. Viswanathan, A., et al., *Plasma Abeta, homocysteine, and cognition: the Vitamin Intervention for Stroke Prevention (VISP) trial.* Neurology, 2009. **72**(3): p. 268-72.

18. Greenland, S. and M.P. Longnecker, *Methods for trend estimation from summarized dose-response data, with applications to meta-analysis.* Am J Epidemiol, 1992. **135**(11): p. 1301-9.

19. Okereke, O.I., et al., *Performance characteristics of plasma amyloid-beta 40 and 42 assays.* J Alzheimers Dis, 2009. **16**(2): p. 277-85.

20. DerSimonian, R. and N. Laird, *Meta-analysis in clinical trials.* Control Clin Trials, 1986. **7**(3): p. 177-88.

21. M, E., S. GD, and A. DG, *Systematic reviews in health care : meta-analysis in context.* 2001, London: BMJ.

22. Egger, M., et al., *Bias in meta-analysis detected by a simple, graphical test.* BMJ, 1997. **315**(7109): p. 629-34.

23. Cosentino, S.A., et al., *Plasma {beta}-Amyloid and Cognitive Decline.* Arch Neurol, 2010.

24. Yaffe, K., et al., *Association of Plasma β-Amyloid Level and Cognitive Reserve With Subsequent Cognitive Decline* JAMA, 2011. **305**(3): p. 261-266.

25. Blasko, I., et al., *Plasma Amyloid Beta-42 Independently Predicts Both Late-Onset Depression and Alzheimer Disease.* Am J Geriatr Psychiatry, 2010.

26. Berlin, J.A., *Does blinding of readers affect the results of meta-analyses? University of Pennsylvania Meta-analysis Blinding Study Group.* Lancet, 1997. **350**(9072): p. 185-6.

27. Hardy, J. and D.J. Selkoe, *The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics.* Science, 2002. **297**(5580): p. 353-6.

28. Hogervorst, E., *Hormones, cognition and dementia : state of the art and emergent therapeutic strategies*. 2009, Cambridge, UK ; New York: Cambridge University Press. xii, 280 p., 2 p. of plates. .

29. Grossman, M., *Hormone Therapy and the Brain: A Clinical Perspective on the Role of Estrogen.* Ann Intern Med, 2000. **133**(8): p. 660.

30. Rosario, E.R. and C.J. Pike, *Androgen regulation of beta-amyloid protein and the risk of Alzheimer's disease.* Brain Res Rev, 2008. **57**(2): p. 444-53.

31. Papasozomenos, S. and A. Shanavas, *Testosterone prevents the heat shock-induced overactivation of glycogen synthase kinase-3 beta but not of cyclin-dependent kinase 5 and c-Jun NH2-terminal kinase and concomitantly abolishes hyperphosphorylation of tau: implications for Alzheimer's disease.* Proc Natl Acad Sci U S A, 2002. **99**(3): p. 1140-5.

32. Beyenburg, S., et al., *Androgen receptor mRNA expression in the human hippocampus.* Neurosci Lett, 2000. **294**(1): p. 25-8.

33. Rapp, S.R., et al., *Effect of estrogen plus progestin on global cognitive function in postmenopausal women: the Women's Health Initiative Memory Study: a randomized controlled trial.* JAMA, 2003. **289**(20): p. 2663-72.

34. O'Brien, J., et al., *Postmenopausal hormone therapy is not associated with risk of all-cause dementia and Alzheimer's disease.* Epidemiol Rev, 2014. **36**(1): p. 83-103.

35. Boss, L., et al., *Endogenous sex hormones and cognitive function in older adults: a systematic review.* West J Nurs Res, 2014. **36**(3): p. 388-426.

36. Maggio, M., et al., *DHEA and cognitive function in the elderly.* J Steroid Biochem Mol Biol, 2014.

37. Peter, J., et al., *Gray matter atrophy pattern in elderly with subjective memory impairment.* Alzheimers Dement, 2014. **10**(1): p. 99-108.

38. Selnes, P., et al., *White matter imaging changes in subjective and mild cognitive impairment.* Alzheimers Dement, 2012. **8**(5 Suppl): p. S112-21.

39. Amariglio, R.E., et al., *Subjective cognitive complaints and amyloid burden in cognitively normal older individuals.* Neuropsychologia, 2012. **50**(12): p. 2880-6.

40. Jacinto, A.F., et al., *Subjective memory complaints in the elderly: a sign of cognitive impairment?* Clinics (Sao Paulo), 2014. **69**(3): p. 194-7.

41. Amariglio, R.E., et al., *Specific subjective memory complaints in older persons may indicate poor cognitive function.* J Am Geriatr Soc, 2011. **59**(9): p. 1612-7.

42. Colditz, G.A., J.E. Manson, and S.E. Hankinson, *The Nurses' Health Study: 20-year contribution to the understanding of health among women.* J Womens Health, 1997. **6**(1): p. 49-62.

43. Hankinson, S.E., et al., *Plasma sex steroid hormone levels and risk of breast cancer in postmenopausal women.* J Natl Cancer Inst, 1998. **90**(17): p. 1292-9.

44. Tworoger, S.S., et al., *Inclusion of endogenous hormone levels in risk prediction models of postmenopausal breast cancer.* J Clin Oncol, 2014. **32**(28): p. 3111-7.

45. Rosner, B., et al., *Determination of blood pressure percentiles in normal-weight children: some methodological issues.* Am J Epidemiol, 2008. **167**(6): p. 653-66.

46. Brandt, J., M. Spencer, and M. Folstein, *The Telephone Interview for Cognitive Status.* Neuropsychiatry Neuropsychol Behav Neurol., 1988. **1**: p. 111-117.

47. Folstein, M.F., S.E. Folstein, and P.R. McHugh, *"Mini-mental state". A practical method for grading the cognitive state of patients for the clinician.* J Psychiatr Res, 1975. **12**(3): p. 189-98.

48. Albert, M., et al., *Use of brief cognitive tests to identify individuals in the community with clinically diagnosed Alzheimer's disease.* Int J Neurosci, 1991. **57**(3-4): p. 167-78.

49. Spreen, O. and E. Strauss, *A Compendium of Neuropsychological Tests: Administration, Norms and Commentary*. 1991, New York, NY: Oxford University Press.

50. The Psychological Corporation, *WAIS-III/WMS-III technical manual: Updated*. 2002, San Antonio, TX.

51. Ware, J., et al., *SF-36 Health Survey: Manual & Interpretation Guide*. 1993, Boston, MA: Nimrod Press.

52. Lesher, E.L. and J.S. Berryhill, *Validation of the Geriatric Depression Scale--Short Form among inpatients.* J Clin Psychol, 1994. **50**(2): p. 256-60.

53. Stampfer, M.J., et al., *Effects of moderate alcohol consumption on cognitive function in women.* N Engl J Med, 2005. **352**(3): p. 245-53.

54. Ryan, J., et al., *Hormone levels and cognitive function in postmenopausal midlife women.* Neurobiol Aging, 2012. **33**(3): p. 617.e11-22.

55. Laughlin, G.A., D. Kritz-Silverstein, and E. Barrett-Connor, *Endogenous oestrogens predict 4-year decline in verbal fluency in postmenopausal women: the Rancho Bernardo Study.* Clin Endocrinol (Oxf), 2010. **72**(1): p. 99-106.

56. Yaffe, K., et al., *Serum estrogen levels, cognitive performance, and risk of cognitive decline in older community women.* J Am Geriatr Soc, 1998. **46**(7): p. 816-21.

57. Carcaillon, L., et al., *High plasma estradiol interacts with diabetes on risk of dementia in older postmenopausal women.* Neurology, 2014. **82**(6): p. 504-11.

58. Yue, X., et al., *Brain estrogen deficiency accelerates Abeta plaque formation in an Alzheimer's disease animal model.* Proc Natl Acad Sci U S A, 2005. **102**(52): p. 19198-203.

59. Bian, C., et al., *Intriguing roles of hippocampus-synthesized 17β-estradiol in the modulation of hippocampal synaptic plasticity.* J Mol Neurosci, 2014. **54**(2): p. 271-81.

60. Maki, P.M., *Critical window hypothesis of hormone therapy and cognition: a scientific update on clinical studies.* Menopause, 2013. **20**(6): p. 695-709.

61. Barha, C.K. and L.A. Galea, *Influence of different estrogens on neuroplasticity and cognition in the hippocampus.* Biochim Biophys Acta, 2010. **1800**(10): p. 1056-67.

62. Yaffe, K., et al., *Endogenous sex hormone levels and risk of cognitive decline in an older biracial cohort.* Neurobiol Aging, 2007. **28**(2): p. 171-8.

63. Hankinson, S.E., et al., *Reproducibility of plasma hormone levels in postmenopausal women over a 2-3-year period.* Cancer Epidemiol Biomarkers Prev, 1995. **4**(6): p. 649-54.

64. Zhang, X., et al., *Postmenopausal plasma sex hormone levels and breast cancer risk over 20 years of follow-up.* Breast Cancer Res Treat, 2013. **137**(3): p. 883-92.

65. Karch, C.M., C. Cruchaga, and A.M. Goate, *Alzheimer's disease genetics: from the bench to the clinic.* Neuron, 2014. **83**(1): p. 11-26.

66. Dennis, E.L. and P.M. Thompson, *Functional brain connectivity using fMRI in aging and Alzheimer's disease.* Neuropsychol Rev, 2014. **24**(1): p. 49-62.

67. Cohen, A.D. and W.E. Klunk, *Early detection of Alzheimer's disease using PiB and FDG PET.* Neurobiol Dis, 2014.

68. Zygouris, S. and M. Tsolaki, *Computerized Cognitive Testing for Older Adults: A review.* Am J Alzheimers Dis Other Demen, 2014.

69. Fredrickson, J., et al., *Evaluation of the usability of a brief computerized cognitive screening test in older people for epidemiological studies.* Neuroepidemiology, 2010. **34**(2): p. 65-75.

70. Lim, Y.Y., et al., *Three-month stability of the CogState brief battery in healthy older adults, mild cognitive impairment, and Alzheimer's disease: results from the Australian Imaging, Biomarkers, and Lifestyle-rate of change substudy (AIBL-ROCS).* Arch Clin Neuropsychol, 2013. **28**(4): p. 320-30.

71. Harel, B.T., et al., *Examining the nature of impairment in visual paired associate learning in amnestic mild cognitive impairment.* Neuropsychology, 2011. **25**(6): p. 752-62.

72. de Jager, C.A., et al., *Detection of MCI in the clinic: evaluation of the sensitivity and specificity of a computerised test battery, the Hopkins Verbal Learning Test and the MMSE.* Age Ageing, 2009. **38**(4): p. 455-60.

73. Maruff, P., et al., *Subtle memory decline over 12 months in mild cognitive impairment.* Dement Geriatr Cogn Disord, 2004. **18**(3-4): p. 342-8.

74. Feskanich, D., et al., *Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire.* J Am Diet Assoc, 1993. **93**(7): p. 790-6.

75. Rimm, E.B., et al., *Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals.* Am J Epidemiol, 1992. **135**(10): p. 1114-26; discussion 1127-36.

76. Crane, P.K., et al., *Composite scores for executive function items: demographic heterogeneity and relationships with quantitative magnetic resonance imaging.* J Int Neuropsychol Soc, 2008. **14**(5): p. 746-59.

77. Gibbons, L.E., et al., *A composite score for executive functioning, validated in Alzheimer's Disease Neuroimaging Initiative (ADNI) participants with baseline mild cognitive impairment.* Brain Imaging Behav, 2012. **6**(4): p. 517-27.

78. Fanning, J. and E. McAuley, *A comparison of tablet computer and paper-based questionnaires in healthy aging research.* JMIR Res Protoc, 2014. **3**(3): p. e38.

79. Hohwü, L., et al., *Web-based versus traditional paper questionnaires: a mixed-mode survey with a Nordic perspective.* J Med Internet Res, 2013. **15**(8): p. e173.

80. Collerton, J., et al., *A comparison of computerized and pencil-and-paper tasks in assessing cognitive function in community-dwelling older people in the Newcastle 85+ Pilot Study.* J Am Geriatr Soc, 2007. **55**(10): p. 1630-5.

81. Lim, Y.Y., et al., *Three-month stability of the CogState brief battery in healthy older adults, mild cognitive impairment, and Alzheimer's disease: results from the Australian Imaging, Biomarkers, and Lifestyle-rate of change substudy (AIBL-ROCS).* Arch Clin Neuropsychol, 2013. **28**(4): p. 320-30.

82. Hammers, D., et al., *Reliability of repeated cognitive assessment of dementia using a brief computerized battery.* Am J Alzheimers Dis Other Demen, 2011. **26**(4): p. 326-33.

83. A., H.T., C.B. S., and H.C. G., *Evaluating websites for older adults: adherence to 'senior-friendly' guidelines and end-user performance.* Behav. Inf. Technol., 2008. **27**: p. 191-199.

84. Maruff, P., et al., *Validity of the CogState brief battery: relationship to standardized tests and sensitivity to cognitive impairment in mild traumatic brain injury, schizophrenia, and AIDS dementia complex.* Arch Clin Neuropsychol, 2009. **24**(2): p. 165-78.