

Cognitive & Behavioral Assessment

Long-term impact of intensive lifestyle intervention on cognitive function assessed with the National Institutes of Health Toolbox: The Look AHEAD study

Kathleen M. Hayden<sup>a,\*</sup>, Laura D. Baker<sup>a,b</sup>, George Bray<sup>c</sup>, Raymond Carvajal<sup>d</sup>, Kathryn Demos-McDermott<sup>e</sup>, Andrea L. Hergenroeder<sup>f,g</sup>, James O. Hill<sup>h</sup>, Edward Horton<sup>i</sup>, John M. Jakicic<sup>f,g</sup>, Karen C. Johnson<sup>j</sup>, Rebecca H. Neiberg<sup>k</sup>, Stephen R. Rapp<sup>a,l</sup>, Thomas A. Wadden<sup>d</sup>, Michael E. Miller<sup>k</sup>, for the Look AHEAD Movement and Memory and Look AHEAD Brain MRI Ancillary Study Groups

<sup>a</sup>Department of Social Sciences and Health Policy, Wake Forest School of Medicine, Winston-Salem, NC, USA

<sup>b</sup>Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA

<sup>c</sup>Pennington Biomedical Research Center, Louisiana State University, Baton Rouge, LA, USA

<sup>d</sup>Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA

<sup>e</sup>Department of Psychiatry and Human Behavior, The Miriam Hospital, Providence, RI, USA

<sup>f</sup>Department of Health and Physical Activity, University of Pittsburgh, Pittsburgh, PA, USA

<sup>g</sup>Department of Physical Therapy, University of Pittsburgh, Pittsburgh, PA, USA

<sup>h</sup>Anschutz Health and Wellness Center, Aurora, CO, USA

<sup>i</sup>Joslin Diabetes Center, Boston, MA, USA

<sup>j</sup>Department of Preventive Medicine, University of Tennessee Health Science Center, Memphis, TN, USA

<sup>k</sup>Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, NC, USA

<sup>l</sup>Department of Psychiatry and Behavioral Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA

Abstract

**Introduction:** This study sought to determine whether 10 years of assignment to intensive lifestyle intervention (ILI) relative to diabetes support and education leads to better cognition. We examine intervention effects overall and among clinical subgroups, and report correlations between computer-administered and interviewer-administered cognitive batteries.

**Methods:** The Action for Health in Diabetes (Look AHEAD) was a 16-site randomized controlled trial with overweight/obese individuals (aged 45–76) who had type 2 diabetes. The NIH Toolbox Cognition Battery tests developed to measure cognition across the lifespan were used to evaluate cognition. Results were compared with standard paper-and-pencil tests. The Toolbox and paper-and-pencil tests were administered an average of 10.9 years after randomization to 1002 participants.

**Results:** Toolbox measures significantly correlated with interviewer-administered measures, with the strongest correlations between the Toolbox Fluid Cognition Composite and Trails B ( $r = -0.64, P < .0001$ ) and Digit Symbol Coding ( $r = 0.63, P < .0001$ ), and between the Toolbox Dimensional Change Card Sort ( $r = 0.55, P < .0001$ ) and the Digit Symbol Coding test. Overall, ILI and diabetes support and education groups had similar adjusted mean cognitive outcomes ( $P > .05$  for all). Subgroup analyses identified different intervention effects within baseline body mass index groups for Picture Sequence Memory ( $P = .01$ ), within baseline cardiovascular disease groups for Picture Vocabulary ( $P = .01$ ) and Fluid Cognition Composite ( $P = .02$ ) measures, and within baseline age groups for Picture Vocabulary ( $P = .02$ ).

**Discussion:** Correlations between Toolbox and interviewer-administered outcomes provide a measure of internal validity. Findings suggest no overall effect of the intervention on cognition and

that an ILI resulting in weight loss may have negative implications for cognition in individuals aged  $\geq 60$ , with previous history of cardiovascular disease, and those with body mass index  $\geq 40$ .

© 2017 The Authors. Published by Elsevier Inc. on behalf of the Alzheimer's Association. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Keywords:** Diabetes mellitus; Obesity; Cognition; Body mass index; Randomized controlled trial; Weight loss; Neuropsychological tests; Aged

## 1. Introduction

Midlife obesity is associated with an increased risk for cognitive deficits in later life [1–3], whereas midlife physical activity is associated with less cognitive decline [4,5]. Behavioral interventions targeting weight loss and increased physical activity may reduce risk of cognitive impairment [6,7]. However, evidence that weight loss will prevent cognitive decline is lacking. Moreover, late life weight loss can be a sign of increased risk for dementia [1], and midlife weight change in either direction may be associated with greater risk for dementia later in life [8,9].

Type 2 diabetes mellitus is also associated with increased risk for cognitive deficits in later life [10–12]. Many pathologic processes may lead to this outcome, including reduced vascular function, increased inflammation, impaired glucose metabolism, and concomitant disorders, such as hypertension and depression [13]. Weight loss through reduced caloric intake and increased physical activity has the potential to improve each of these conditions [13,14]. Adults with type 2 diabetes may thus be particularly sensitive to any cognitive benefits of behavioral intervention for weight loss.

The Action for Health in Diabetes (Look AHEAD) study was a randomized controlled clinical trial that compared 10 years of intensive lifestyle intervention (ILI) targeting weight loss and increased physical activity to a diabetes support and education (DSE) control among overweight or obese adults with type 2 diabetes [14,15]. Although no cognitive assessment was conducted at baseline, earlier cross-sectional analyses of Look AHEAD participants at year 8 showed that there were no differences in overall cognitive function between the ILI and the DSE based on a standardized interviewer-administered battery of cognitive tests [16,17].

The computer-administered NIH Cognitive Toolbox has been designed to provide greater precision in cognitive assessments, particularly for domains related to processing speed and executive function [18,19]. There are three primary objectives of this report. We examine whether 10 years of random assignment to ILI relative to a control condition leads to better performance on cognitive measures. We examine the consistency of any intervention effects among previously examined baseline clinical subgroups in Look AHEAD including those with higher body mass index (BMI) and previous history of cardiovascular disease (CVD) [17]. We also report correlations between computer-administered and interviewer-administered cognitive test bat-

teries to provide support for the use of the NIH Toolbox in cognitive research.

## 2. Methods

The design and methods of Look AHEAD have been previously described [14]. Look AHEAD was a 16-site randomized controlled trial that recruited 5145 individuals (from 2001 to 2004) who were overweight or obese and had type 2 diabetes. Participants were aged 45 to 76 years and had a BMI of  $>25$  kg/m<sup>2</sup> (or  $>27$  kg/m<sup>2</sup> if on insulin), glycated hemoglobin (HbA1c)  $<11\%$ , systolic/diastolic blood pressure  $<160/<100$  mm Hg, and triglycerides  $<600$  mg/dL. All participants completed a maximal treadmill test. All Participants provided informed consent and were randomly assigned (1:1) to ILI or DSE. Local institutional review boards approved the protocol. The present study involved six Look AHEAD sites (including Baton Rouge, Denver, Memphis, Philadelphia, Pittsburgh, and Providence) that implemented two ancillary studies, the Look AHEAD Movement and Memory study and the Look AHEAD Brain study. All active participants at these sites who were willing to participate and free of contraindications (e.g., for magnetic resonance imaging [MRI] studies) were invited to participate. Although there were no baseline cognitive assessments in Look AHEAD, both of these ancillary studies administered paper-and-pencil tests in addition to the NIH Toolbox cognitive assessments an average of 10.9 years after randomization. The study sample comprised 1002 participants at those sites who completed both the NIH Toolbox and the face-to-face cognitive assessment (described subsequently). We attempted to administer both batteries on the same day whenever possible.

### 2.1. Interventions

ILI participants received dietary and physical activity goals for weight loss, and were seen weekly by study staff for 6 months and subsequently three times per month for 6 months, using a combination of group and individual contact. Thereafter, ILI participants were offered two contacts per month plus optional group meetings and refreshers, and national campaign materials [15]. DSE participants were invited to attend three group sessions per year for the first 4 years, and one session per year thereafter [14].

In September 2012, the study's sponsor (National Institute of Diabetes and Digestive and Kidney Diseases) terminated interventions based on recommendations from the

trial's Data and Safety Monitoring Board. The reason was statistical futility for the trial's primary end point, which was a composite score based on death from CVD, nonfatal myocardial infarction (MI), nonfatal stroke, and hospitalization for angina [20]. Observation of the cohort continues with semiannual clinic visits.

## 2.2. Cognitive assessments

The NIH Cognitive Toolbox is a computer-administered battery that was developed to harmonize cognitive assessments across studies and increase sensitivity to detect impairment and change over time [18]. Reliability and validity of the performance of the Toolbox composite measures in adult populations have been reported [21]. The following cognitive domains are assessed: executive function (inhibitory control/attention: Flanker Task; cognitive flexibility: Dimensional Change Card Sort), working memory (List Sorting), short-term memory (Picture Sequence), and an overall composite score that combines these outcomes (Fluid Cognition Composite) [21]. We also included the Picture Vocabulary test. Centrally trained, certified, and blinded staff assisted in the administration of the Toolbox at six participating sites; 1002 of these participants provided valid data.

The interviewer-administered battery assessed attention and concentration, with the Trail Making Test-Part A [22]; verbal learning and memory with the Rey Auditory Verbal Learning test [23,24]; inhibitory control and divided attention with the Trail Making Test-Part B [22] and the modified Stroop Color Word test [25,26]; processing speed with Digit Symbol Coding [27]; and global cognitive function with the Modified Mini-Mental Status Examination [28].

## 2.3. Other measures

Demographic characteristics and self-reported medical history were collected at baseline. Participants brought current prescription medications to update Look AHEAD medication records. Weight and height were measured in duplicate using a digital scale and stadiometer. Baseline hypertension was identified based on use of antihypertensive medications or measured blood pressure. Previous history of CVD at randomization was based on any patient self-report of MI, heart bypass surgery, coronary artery bypass graft, carotid endarterectomy, lower leg angioplasty, aortic aneurysm, congestive heart failure, or stroke. Physical fitness at baseline was estimated using the achieved metabolic equivalent level (based on treadmill speed and grade) at the time of termination of the graded exercise test. The Short-Form Health Survey 36 (SF-36) was administered to assess overall mental health [29].

## 2.4. Statistical methods

Potential differences by intervention group in baseline risk factors for cognitive deficits (e.g., age and education) were examined using *t*- and chi-squared tests. These compar-

isons were performed to assess balance across groups given that this substudy represented a *nonrandom* sample of those randomized at baseline. Partial Spearman correlations, adjusted for education (<13, ≥13 years) were used to evaluate the association between cognitive batteries regardless of intervention arm.

Rather than using the scale scores provided by the NIH Toolbox that are adjusted to national norms, we used the unadjusted scale scores that are provided for each test and then adjusted for gender, age, race/ethnicity, among other covariates to investigate the effect of the intervention on cognition. Group differences in adjusted means for each cognitive outcome were examined using a general linear model with additional adjustments for site, time from randomization, and the following baseline characteristics: age, sex, BMI, education, race/ethnicity (African American/Black–non-Hispanic, Hispanic, non-Hispanic White, Other/Mixed race), duration of diabetes, insulin use, hypertension, history of CVD, and SF-36 score [29]. From these models, adjusted means and 95% confidence intervals (CIs) were obtained within intervention group and for the difference between intervention groups.

Prespecified subgroup comparisons [30] that have been previously examined in Look AHEAD were based on characteristics at the time of randomization: age (<60, 60+ years), BMI (<30, 30–39, ≥40 kg/m<sup>2</sup>), duration of diabetes (<5, ≥5 years), and sex. For each outcome, additional comparisons were made across groups defined by the presence or absence of previous history of CVD consistent with other Look AHEAD reports [17]. Heterogeneity of intervention effects among baseline subgroup levels was investigated by adding subgroup by intervention interaction terms to each model. Because we present 30 tests of subgroup hypotheses each at the 0.05 level, there is a 78% chance (i.e.,  $1-[1-0.05]^{30}$ ) that at least one of these tests would be statistically significant at an alpha level of .05, assuming independence between tests [30]; however, as the outcomes are correlated, the calculation under independence would overestimate the type I error rate [31]. On the basis of the inspection of observed results, additional models were fitted that contained both the BMI and previous history of CVD by intervention interaction terms to the same model.

## 3. Results

The NIH Toolbox tests were administered during the Look AHEAD Brain [32] or Look AHEAD Memory and Movement [33] ancillary studies in a subset of 1002 participants enrolled at six sites (ILI, N = 527; DSE, N = 475; dates of administration, March 23, 2013 to September 23, 2014). Look AHEAD participants in this NIH Toolbox substudy tended to be more highly educated, have slightly lower HbA1c, and slightly higher baseline cardiorespiratory fitness compared with Look AHEAD participants who were not included. NIH Toolbox substudy participants also were less likely to be Hispanic (see [Supplemental Table 1](#)). A

Table 1  
Baseline Characteristics of  $n = 1002$  Participants by Intervention Group

Characteristic	Intensive lifestyle intervention (N = 527)	Diabetes support and education (N = 475)	P value
Age, mean (SD) (y)	58.9 (6.6)	58.7 (6.5)	.68
Female sex, no. (%)	310 (58.8%)	284 (59.8%)	.76
Race, no. (%)			.16
African American/Black (non-Hispanic)	111 (21.1%)	103 (21.7%)	
White	387 (73.4%)	333 (70.1%)	
Hispanic	12 (2.3%)	23 (4.8%)	
Other/Mixed	17 (3.2%)	16 (3.4%)	
Education, no. (%)			.13
<13 y	93 (17.6%)	74 (15.6%)	
13–16 y	186 (35.3%)	193 (40.6%)	
>16 y	241 (45.7%)	192 (40.4%)	
BMI, mean (SD) (kg/m <sup>2</sup> )	35.8 (6.0)	36.2 (5.8)	.30
BMI group, no. (%)			.07
<30 kg/m <sup>2</sup>	89 (16.9%)	57 (12.0%)	
30 to <40 kg/m <sup>2</sup>	330 (62.6%)	306 (64.4%)	
≥40 kg/m <sup>2</sup>	108 (20.5%)	112 (23.6%)	
HbA1c, mean (SD) (%)	7.2 (1.1)	7.2 (1.2)	.87
HbA1c (%) group, no. (%)			.98
<6.5%	136 (25.8%)	120 (25.3%)	
6.5% to ≤7.5%	237 (45.0%)	214 (45.0%)	
>7.5%	154 (29.2%)	141 (29.7%)	
Hypertension, no. (%)	452 (85.8%)	397 (83.6%)	.34
Previous history of CVD, no. (%)	80 (15.2%)	54 (11.4%)	.08
Only MI, bypass surgery or CABG	62 (11.8%)	41 (8.6%)	
Only stroke	11 (2.1%)	3 (0.6%)	
Both MI/bypass surgery/CABG and stroke	4 (0.8%)	4 (0.8%)	
Other CVD event	3 (0.6%)	6 (1.3%)	
Diabetes duration, mean (SD) (y)	6.4 (6.4)	6.8 (6.4)	.36
Diabetes duration, no. (%)			.25
<5 y	250 (47.4%)	212 (44.6%)	
≥5 y	276 (52.4%)	259 (54.5%)	
Insulin use, no. (%)			.55
No	435 (85.0%)	397 (86.3%)	
Yes	77 (15.0%)	63 (13.7%)	
Smoking status, no. (%)			.84
Never	265 (50.3%)	247 (52.1%)	
Past	241 (45.7%)	208 (43.9%)	
Present	21 (4.0%)	19 (4.0%)	
Cardiorespiratory fitness, mean (SD), METS	7.4 (1.9)	7.4 (2.1)	.90
Cardiorespiratory fitness category, No. (%)			.61
<7.5 METS	291 (55.2%)	270 (56.8%)	
≥7.5 METS	236 (44.8%)	205 (43.2%)	

Abbreviations: BMI, body mass index; CABG, coronary artery bypass graft; CVD, cardiovascular disease; METS, metabolic equivalents; MI, myocardial infarction; SD, standard deviation.

comparison of individuals from the substudy sites who participated in the substudy versus those who did not participate similarly had lower HbA1c, fewer had previous history of CVD, and had higher levels of baseline cardiorespiratory fitness (see [Supplemental Table 2](#)).

Baseline characteristics by intervention groups for participants with Toolbox assessments are presented in [Table 1](#).

The two groups were generally comparable, the ILI group had nominally, although not significantly, more participants with a history of CVD ( $P = .08$ ) and slightly lower BMI ( $P = .07$ ). As reported elsewhere, overall the intervention was successful, and before the termination of the trial participants randomized to the intervention achieved their cardiovascular fitness goals of reduced hypertension and diabetes control [[34,35](#)].

Time from randomization to the single administration of the NIH Toolbox Cognitive Assessment was 10.9 years (standard deviation [SD] = 0.77) in the ILI group and 10.9 years (SD = 0.79) in the DSE group, with a mean age of 70 years at the time of the cognitive assessment in both groups. Among those with a history of CVD, the most common events included MI and bypass surgery or coronary artery bypass graft.

Unadjusted and adjusted means (SD) for Toolbox test outcomes are presented by intervention group in [Table 2](#). The ILI and DSE groups had similar scores for all tests ( $P > .05$ ). As expected, Toolbox test outcomes were significantly correlated with the interviewer-administered measures ( $P < .001$ ; [Table 3](#)). Eighty-eight percent of the pairings of interviewer-administered and computer-administered tests were administered on the same day, and approximately 95% of tests were administered within 60 days of each other. When the time between tests was limited to those administered on the same day, the correlations changed little.

When further explored by baseline subgroup ([Fig. 1](#) and [Supplemental Table 3](#)), tests of interaction were only significant at the nominal .05 level for four of 30 tests. Subgroup analyses identified different intervention effects within baseline BMI groups for Picture Sequence Memory ( $P = .01$ ), within baseline CVD groups for Picture Vocabulary ( $P = .01$ ) and Fluid Cognition Composite ( $P = .02$ ) measures, and within baseline age groups for Picture Vocabulary ( $P = .02$ ). Briefly, within the highest BMI group (BMI ≥ 40), the ILI group had poorer short-term memory compared with the DSE group (Picture Sequence Memory mean difference = 3.54; 95% CI, 0.96–6.11), and in participants with previous history of CVD, language performance was poorer in the ILI group than in the DSE group (Picture Vocabulary mean difference = 3.81; 95% CI, 0.59–7.03), as was performance on the Fluid Cognition Composite (mean difference = 2.95; 95% CI, 0.36–5.54).

#### 4. Discussion

We evaluated cognitive performance using the NIH Toolbox and interviewer-administered tests in a subsample of 1002 Look AHEAD participants who participated in a follow-up cognitive evaluation 10 years after randomization. NIH Toolbox cognitive outcomes were significantly correlated with interviewer-administered cognitive measures providing further support for the validity of the NIH Toolbox, which could have important implications for where and when cognitive assessments are performed. Our findings also suggest that the effects of an ILI for adults with a previous history of CVD or with high BMI may have negative implications for cognition.

Table 2  
Unadjusted and Adjusted\* Mean Scores From the NIH Toolbox by Intervention Group

Outcome	Intensive lifestyle intervention			Diabetes support and education			Difference between adjusted means (95% CI)	P value for difference between adjusted means
	N	Unadjusted mean (SD)	Adjusted mean (95% CI)	N	Unadjusted mean (SD)	Adjusted mean (95% CI)		
Flanker–inhibitory and attention	519	98.7 (9.2)	98.9 (98.2, 99.5)	474	98.9 (9.4)	99.0 (98.2, 99.7)	0.11 (−0.90, 1.12)	.83
Dimensional Change Card Sort	518	98.5 (9.7)	98.6 (97.9, 99.3)	469	98.6 (10.0)	98.7 (97.9, 99.5)	0.14 (−0.94, 1.22)	.80
List Sorting	515	98.1 (9.9)	98.0 (97.2, 98.8)	470	97.9 (10.1)	97.7 (96.9, 98.6)	−0.27 (−1.44, 0.90)	.65
Picture Sequence Memory	503	91.4 (10.4)	91.4 (90.5, 92.2)	455	91.6 (9.9)	91.4 (90.5, 92.3)	0.06 (−1.17, 1.29)	.93
Picture Vocabulary	524	120.3 (11.3)	120.2 (119.4, 121.0)	472	120.0 (10.5)	120.2 (119.3, 121.0)	−0.03 (−1.20, 1.13)	.95
Fluid Cognition Composite	477	91.9 (8.7)	91.9 (91.3, 92.6)	437	92.1 (8.1)	92.1 (91.4, 92.7)	0.15 (−0.77, 1.07)	.75

Abbreviations: CI, confidence interval; SD, standard deviation.

\*Adjusted for clinic site, time from randomization, and the following baseline characteristics: age, sex, education, race/ethnicity, duration of diabetes, insulin use, hypertension, history of cardiovascular disease, and SF-36 mental score.

There were no intervention group differences in cognitive function measured by the Toolbox, which is consistent with previous reports regarding interviewer-administered tests completed 8 to 10 years after Look AHEAD randomization [16,17] and consistent with intervention tests performed on the interviewer-administered tests within the current subset of participants (results not shown). Among those with a previous history of CVD, the ILI group performed worse than the DSE group on a language task and on the Fluid Cognition Composite. These findings are consistent with earlier Look AHEAD reports [16,17] but are in contrast to results from a Look AHEAD MRI substudy that included a subsample ( $n = 319$ ) who completed cognitive testing and MRI [32]. Although total brain and hippocampal volume did not differ between groups, ILI participants had smaller white matter hyperintensity and ventricular volumes indicating better brain health roughly equivalent to 2 years younger in age [32].

Similarly, in the Action to Control Cardiovascular Risk in Diabetes Memory in Diabetes (ACCORDION MIND) study, although the lifestyle intervention group initially showed small but significant benefit in total brain volume [36], later follow-up showed no differences on MRI measures between intervention groups [37]. This study differed in that the follow-up period was shorter and the cohort did not consist of exclusively overweight or obese participants.

The results of a 12-study meta-analysis suggest that intentional weight loss can lead to improved cognitive function [38]. Clinical trials testing the effects of physical exercise have demonstrated cognitive benefits in older adults with subjective memory complaints [39] and improvements in glucose metabolism and cognition in adults with amnesic mild cognitive impairment and thus at high risk for progression to dementia [40]. Our overall findings that showed no benefits of ILI on cognition and a potentially negative impact in adults

Table 3  
Spearman Correlations Adjusted for Education Level (<13, ≥ 13 years) and Randomization Arm

Measure	NIH Cognition Toolbox computer administered measures					
	Flanker–Inhibitory and Attention	Dimensional Change Card Sort	List Sorting	Picture Sequence Memory	Picture Vocabulary	Fluid Cognition Composite
Interviewer-administered cognitive measures						
DSC*	0.48		0.36	0.38	0.29	0.63
Trails B*	−0.49	−0.52	−0.48	−0.40	−0.37	−0.64
3MSE*	0.38	0.44	0.39	0.39	0.50	0.52
Delayed RAVLT*	0.26	0.31	0.30	0.47	0.28	0.45
Stroop*	−0.31	−0.41	−0.31	−0.34	−0.30	−0.48
Interviewer-administered cognitive measures administered on the same day as computer-administered tests						
DSC*	0.51	0.56	0.36	0.39	0.29	0.64
Trails B*	−0.49	−0.53	−0.47	−0.41	−0.35	−0.65
3MSE*	0.37	0.44	0.39	0.39	0.51	0.52
Delayed RAVLT*	0.26	0.32	0.30	0.48	0.29	0.45
Stroop*	−0.32	−0.42	−0.30	−0.36	−0.28	−0.49
NIH Cognition Toolbox computer-administered measures						
Dimensional Change Card Sort*	0.62					
List Sorting*	0.28	0.33				
Picture Sequence Memory*	0.30	0.36	0.36			
Picture Vocabulary*	0.28	0.31	0.39	0.23		
Fluid Cognition Composite*	0.73	0.79	0.62	0.65	0.39	

Abbreviations: DSC, Digit Symbol Coding; RAVLT, Rey Auditory Verbal Learning test; 3MSE, Modified Mini–Mental Status Examination.

\*All correlations are significant at  $P < .0001$ .

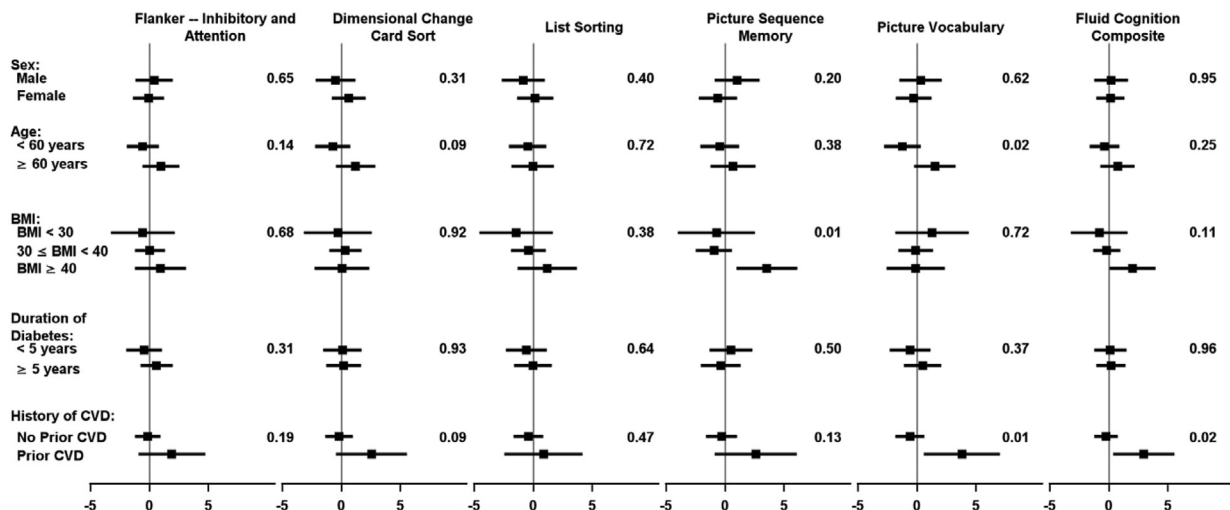


Fig. 1. Differences between intervention group means (95% CI) and tests of interaction among baseline subgroups. Estimated differences (DSE-ILI) and 95% CI obtained from least-squared means from general linear models adjusted for clinic site, time from randomization, and the following baseline characteristics: age, sex, education, race/ethnicity, duration of diabetes, insulin use, hypertension, history of cardiovascular disease, and SF-36 mental score. Interaction *P* values obtained by refitting the model with each intervention by subgroup effect added to separate models for each outcome and baseline subgroup. Abbreviations: CI, confidence interval; ILI, intensive lifestyle intervention.

with high BMI or with history of CVD are counter to expectation. It is possible that there were differential effects on participants by BMI or CVD status because of the complex associations between weight loss, cerebral blood flow, and vasodilation. Participants with a history of CVD and those who were morbidly obese may have already experienced some decrease in cognitive efficiency at baseline, possibly associated with diabetes [41]. Although speculative, the consequence of this accumulation of factors may be decreased cognitive function in members of the intervention group who were morbidly obese or had a previous history of CVD at baseline (for a more detailed discussion see [42]). However, as this finding also may be an artifact of multiple statistical tests, additional research is needed.

This study has both strengths and limitations. This is one of the few studies to compare performance on standardized cognitive tests with the NIH Toolbox in an older cohort of individuals with diabetes. In addition, the results support its validity and feasibility in large clinical trials with older adults. Because the cognitive assessments in Look AHEAD were administered years after the initial randomization to ILI or DSE, the study lacks a baseline evaluation of cognitive function. Although most computer-based and interviewer-based cognitive assessments were administered on the same day, for small portion of participants (5%) the two types of assessments took place up to 60 days apart. However, this variability did not alter the findings when examined in the statistical analysis. Although the analysis by baseline subgroups were determined a priori based on the number of tests performed, we cannot rule out the possibility that these results are because of chance alone [30]. Because the number of participants with previous history of CVD in both the ILI and DSE groups is relatively small, CIs on intervention effects are quite wide in this group. Overall, our findings provide evidence to support the internal validity of the

NIH Toolbox and utility of this computer test battery to assess cognition in older adults with type 2 diabetes.

**Acknowledgments**

*Clinical Sites*

*Pennington Biomedical Research Center:* George A. Bray, MD (Principal Investigator); Kristi Rau (Program Coordinator); Allison Strate, RN (Program Coordinator); Frank L. Greenway, MD (Coinvestigator); Donna H. Ryan, MD (Coinvestigator); Donald Williamson, PhD (Coinvestigator); Brandi Armand, LPN; Jennifer Arceneaux; Amy Bachand, MA; Michelle Begnaud, LDN, RD, CDE; Betsy Berhard; Elizabeth Caderette; Barbara Cerniauskas, LDN, RD, CDE; David Creel, MA; Diane Crow; Crystal Duncan; Helen Guay, LDN, LPC, RD; Carolyn Johnson, Lisa Jones; Nancy Kora; Kelly LaFleur; Kim Landry; Missy Lingle; Jennifer Perault; Cindy Puckett; Mandy Shipp, RD; Marisa Smith; Elizabeth Tucker.

*University of Colorado Health Sciences Center:* James O. Hill, PhD (Principal Investigator); Marsha Miller, MS, RD (Program Coordinator); Brent Van Dorsten, PhD (Coinvestigator); Judith Regensteiner, PhD (Coinvestigator); Ligia Coelho, BS; Paulette Cohrs, RN, BSN; Susan Green; April Hamilton, BS, CCRC; Jere Hamilton, BA; Eugene Leshchinskiy; Lindsey Munkwitz, BS; Loretta Rome, TRS; Terra Worley, BA; Kirstie Craul, RD, CDE; Sheila Smith, BS.

*The University of Tennessee Health Science Center: University of Tennessee East:* Karen C. Johnson, MD, MPH (Principal Investigator); Carolyn Gresham, RN (Program Coordinator); Stephanie Connelly, MD, MPH (Coinvestigator); Amy Brewer, RD, MS; Mace Coday, PhD; Lisa Jones, RN; Lynne Lichtermann, RN, BSN; Shirley Vosburg, RD, MPH; and J. Lee Taylor, MEd, MBA. *University of*

*Tennessee Downtown:* Abbas E. Kitabchi, PhD, MD (Principal Investigator); Ebenezer Nyenwe, MD (Coinvestigator); Helen Lambeth, RN, BSN (Program Coordinator); Amy Brewer, MS, RD, LDN; Debra Clark, LPN; Andrea Crisler, MT; Debra Force, MS, RD, LDN; Donna Green, RN; Robert Kores, PhD.

*University of Pittsburgh:* John M. Jakicic, PhD (Principal Investigator), David E. Kelley, MD (Principal Investigator); Jacqueline Wesche-Thobaben, RN, BSN, CDE (Program Coordinator); Lewis H. Kuller, MD, DrPH (Coinvestigator); Andrea Kriska, PhD (Coinvestigator); Amy D. Otto, PhD, RD, LDN (Coinvestigator), Lin Ewing, PhD, RN (Coinvestigator), Mary Korytkowski, MD (Coinvestigator), Daniel Edmundowicz, MD (Coinvestigator); Monica E. Yamamoto, DrPH, RD, FADA (Coinvestigator); Rebecca Danchenko, BS; Barbara Elnyczky; David O. Garcia, MS; George A. Grove, MS; Patricia H. Harper, MS, RD, LDN; Susan Harrier, BS; Nicole L. Helbling, MS, RN; Diane Ives, MPH; Juliet Mancino, MS, RD, CDE, LDN; Anne Mathews, PhD, RD, LDN; Tracey Y. Murray, BS; Joan R. Ritchea; Susan Urda, BS, CTR; Donna L. Wolf, PhD.

*Coordinating Center: Wake Forest School of Medicine.* Stephen B. Kritchevsky, PhD (Principal Investigator); Denise K. Houston, PhD (Coinvestigator); Jeff D. Williamson, MD (Coinvestigator); Stephen R. Rapp, PhD (Coinvestigator); Mark A. Espeland, PhD (Coinvestigator); Xiaoyina (Iris) Leng, MD, PhD (Coinvestigator); Gary Miller, PhD (Coinvestigator); Amelia Hodges, BS, CCRP (Program Coordinator); Michelle Gordon, MS; Jennifer Walker; Tara Becker; Jason Griffin, BS; Lea Harvin, BS; Kathy Lane, BS; Rebecca H. Neiberg, MS.

*Funding and Support:* This work was supported by the National Institute on Aging, National Institutes of Health, Department of Health and Human Services, R-01AG033087-01, and R-01AG033087-04S1. The Action for Health in Diabetes is supported through the following cooperative agreements from the National Institutes of Health: DK57136, DK57149, DK56990, DK57177, DK57171, DK57151, DK57182, DK57131, DK57002, DK57078, DK57154, DK57178, DK57219, DK57008, DK57135, and DK56992. The following federal agencies have contributed support: National Institute of Diabetes and Digestive and Kidney Diseases; National Heart, Lung, and Blood Institute; National Institute of Nursing Research; National Center on Minority Health and Health Disparities; Office of Research on Women's Health; the Centers for Disease Control and Prevention; and the Department of Veterans Affairs. This research was supported in part by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases. The Indian Health Service (IHS) provided personnel, medical oversight, and use of facilities. The opinions expressed in this article are those of the authors and do not necessarily reflect the views of the IHS or other funding sources.

Additional support was received from the University of Colorado Health Sciences Center General Clinical Research Center (M01RR00051) and Clinical Nutrition Research

Unit (P30 DK48520); the University of Tennessee at Memphis General Clinical Research Center (M01RR0021140); and the University of Pittsburgh General Clinical Research Center (GCRC) (M01RR000056), the Clinical Translational Research Center (CTRC) funded by the Clinical and Translational Science Award (UL1 RR 024153) and NIH grant (DK 046204); and the Frederic C. Bartter General Clinical Research Center (M01RR01346).

The following organizations have committed to make major contributions to Look AHEAD: FedEx Corporation; Health Management Resources; LifeScan, Inc, a Johnson & Johnson Company; OPTIFAST of Nestle HealthCare Nutrition, Inc; Hoffmann-La Roche Inc; Abbott Nutrition; and Slim-Fast Brand of Unilever North America.

### Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.dadm.2017.09.002>.

### RESEARCH IN CONTEXT

1. Systematic review: We searched PubMed for publications from studies that used the NIH Toolbox Cognition Battery. Few publications beyond the original validation studies were found.
2. Interpretation: In the Look AHEAD clinical trial, performance on the NIH Toolbox Cognition Battery was correlated with standard paper-and-pencil tests including Trail Making, the Rey Auditory Verbal Learning test, the Stroop Color Word test, Digit Symbol Coding, and the Modified Mini-Mental Status Examination. We found little difference in performance on the NIH Toolbox between intervention groups. Prespecified subgroup analysis found that those who were older at baseline, those with higher baseline body mass index, and those with a history of cardiovascular disease who were randomized to the intervention did not perform as well.
3. Future directions: Although modest, our results suggest negative outcomes in certain subgroups randomized to the intervention arm of the trial. These are consistent with earlier findings and suggest the need for further investigation.

### References

- [1] Fitzpatrick AL, Kuller LH, Lopez OL, Diehr P, O'Meara ES, Longstreth WT Jr, et al. Midlife and late-life obesity and the risk of dementia: cardiovascular health study. *Arch Neurol* 2009;66:336–342.

- [2] Gustafson D, Rothenberg E, Blennow K, Steen B, Skoog I. An 18-year follow-up of overweight and risk of Alzheimer disease. *Arch Intern Med* 2003;163:1524–1528.
- [3] Kivipelto M, Ngandu T, Fratiglioni L, Viitanen M, Kareholt I, Winblad B, et al. Obesity and vascular risk factors at midlife and the risk of dementia and Alzheimer disease. *Arch Neurol* 2005;62:1556–1560.
- [4] Rovio S, K areholt I, Helkala EL, Viitanen M, Winblad B, Tuomilehto J, et al. Leisure-time physical activity at midlife and the risk of dementia and Alzheimer's disease. *Lancet Neurol* 2005;4:705–711.
- [5] Chang M, Jonsson PV, Snaedal J, Bjornsson S, Saczynski JS, Aspelund T, et al. The effect of midlife physical activity on cognitive function among older adults: AGES–Reykjavik Study. *J Gerontol A Biol Sci Med Sci* 2010;65:1369–1374.
- [6] Barnes DE, Yaffe K. The projected effect of risk factor reduction on Alzheimer's disease prevalence. *Lancet Neurol* 2011;10:819–828.
- [7] Norton S, Matthews FE, Barnes DE, Yaffe K, Brayne C. Potential for primary prevention of Alzheimer's disease: an analysis of population-based data. *Lancet Neurol* 2014;13:788–794.
- [8] Beydoun MA, Beydoun H, Wang Y. Obesity and central obesity as risk factors for incident dementia and its subtypes: a systematic review and meta-analysis. *Obes Rev* 2008;9:204–218.
- [9] Buchman AS, Wilson RS, Bienias JL, Shah RC, Evans DA, Bennett DA. Change in body mass index and risk of incident Alzheimer disease. *Neurology* 2005;65:892–897.
- [10] Arvanitakis Z, Wilson RS, Bienias JL, Evans DA, Bennett DA. Diabetes mellitus and risk of Alzheimer disease and decline in cognitive function. *Arch Neurol* 2004;61:661–666.
- [11] Beeri MS, Goldbourt U, Silverman JM, Noy S, Schmeidler J, Ravona-Springer R, et al. Diabetes mellitus in midlife and the risk of dementia three decades later. *Neurology* 2004;63:1902–1907.
- [12] Ott A, Stolk RP, van Harskamp F, Pols HA, Hofman A, Breteler MM. Diabetes mellitus and the risk of dementia: the Rotterdam Study. *Neurology* 1999;53:1937–1942.
- [13] UK Prospective Diabetes Study (UKPDS) Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group. *Lancet* 1998;352:837–853.
- [14] Ryan DH, Espeland MA, Foster GD, Haffner SM, Hubbard VS, Johnson KC, et al. Look AHEAD (Action for Health in Diabetes): design and methods for a clinical trial of weight loss for the prevention of cardiovascular disease in type 2 diabetes. *Control Clin Trials* 2003;24:610–628.
- [15] Look AR, Wadden TA, West DS, Delahanty L, Jakicic J, Rejeski J, et al. The Look AHEAD study: a description of the lifestyle intervention and the evidence supporting it. *Obesity (Silver Spring)* 2006;14:737–752.
- [16] Espeland MA, Rapp SR, Bray GA, Houston DK, Johnson KC, Kitabchi AE, et al. Long-term impact of behavioral weight loss intervention on cognitive function. *J Gerontol A Biol Sci Med Sci* 2014;69:1101–1108.
- [17] Rapp SR, Luchsinger JA, Baker LD, Blackburn GL, Hazuda HP, Demos-McDermott KE, et al. Effect of a long-term intensive lifestyle intervention on cognitive function: Action for Health in Diabetes Study. *J Am Geriatr Soc* 2017;65:966–972.
- [18] Weintraub S, Dikmen SS, Heaton RK, Tulsky DS, Zelazo PD, Bauer PJ, et al. Cognitive assessment using the NIH Toolbox. *Neurology* 2013;80:S54–64.
- [19] Weintraub S, Dikmen SS, Heaton RK, Tulsky DS, Zelazo PD, Slotkin J, et al. The cognition battery of the NIH toolbox for assessment of neurological and behavioral function: validation in an adult sample. *J Int Neuropsychol Soc* 2014;20:567–578.
- [20] Look Ahead Research Group. Wing RR, Bolin P, Brancati FL, Bray GA, Clark JM, Coday M, et al. Cardiovascular effects of intensive lifestyle intervention in type 2 diabetes. *N Engl J Med* 2013;369:145–154.
- [21] Heaton RK, Akshoomoff N, Tulsky D, Mungas D, Weintraub S, Dikmen S, et al. Reliability and validity of composite scores from the NIH Toolbox Cognition Battery in adults. *J Int Neuropsychol Soc* 2014;20:588–598.
- [22] Reitan R. Validity of the Trail Making Test as an indicator of organic brain damage. *Percept Mot Skills* 1958;8:271–276.
- [23] Rey A. *L'examen Clinique en Psychologie*. Paris: Presses Universitaires de France; 1964.
- [24] Schmidt M. *Rey Auditory and Verbal Learning Test: a Handbook*. Los Angeles: Western Psychological Services; 1996.
- [25] Stroop JR. Studies of interference in serial verbal reactions. *Exp Psychol* 1935;18:643–662.
- [26] Houx PJ, Jolles J, Vreeling FW. Stroop interference: aging effects assessed with the Stroop Color-Word Test. *Exp Aging Res* 1993;19:209–224.
- [27] Wechsler D. *Wechsler Adult Intelligence Scale-III (WAIS-III)*. New York: Psychological Corporation/Harcourt, Inc; 1996.
- [28] Teng EL, Chui HC. The Modified Mini-Mental State (3MS) examination. *J Clin Psychiatry* 1987;48:314–318.
- [29] Ware JE, Snow KK, Kosinski M, Gandek B. *SF-36 Health survey manual and interpretation guide*. Boston, MA: The Health Institute, New England Medical Centre; 1993.
- [30] Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Statistics in medicine—reporting of subgroup analyses in clinical trials. *N Engl J Med* 2007;357:2189–2194.
- [31] Moy e L. *Multiple Analyses in Clinical Trials, Fundamentals for Investigators*. New York: Springer-Verlag; 2010.
- [32] Espeland MA, Erickson K, Neiberg RH, Jakicic JM, Wadden TA, Wing RR, et al. Brain and white matter hyperintensity volumes after 10 years of random assignment to lifestyle intervention. *Diabetes Care* 2016;39:764–771.
- [33] Houston DK, Leng X, Bray GA, Hergenroeder AL, Hill JO, Jakicic JM, et al. A long-term intensive lifestyle intervention and physical function: the look AHEAD Movement and Memory Study. *Obesity (Silver Spring)* 2015;23:77–84.
- [34] Gregg EW, Chen H, Wagenknecht LE, Clark JM, Delahanty LM, Bantle J, et al. Association of an intensive lifestyle intervention with remission of type 2 diabetes. *JAMA* 2012;308:2489–2496.
- [35] Jakicic JM, Egan CM, Fabricatore AN, Gaussoin SA, Glasser SP, Hesson LA, et al. Four-year change in cardiorespiratory fitness and influence on glycemic control in adults with type 2 diabetes in a randomized trial: the Look AHEAD Trial. *Diabetes Care* 2013;36:1297–1303.
- [36] Launer LJ, Miller ME, Williamson JD, Lazar RM, Gerstein HC, Murray AM, et al. Effects of intensive glucose lowering on brain structure and function in people with type 2 diabetes (ACCORD MIND): a randomised open-label substudy. *Lancet Neurol* 2011;10:969–977.
- [37] Murray AM, Hsu FC, Williamson JD, Bryan RN, Gerstein HC, Sullivan MD, et al. ACCORDION MIND: results of the observational extension of the ACCORD MIND randomised trial. *Diabetologia* 2017;60:69–80.
- [38] Siervo M, Arnold R, Wells JC, Tagliabue A, Colantuoni A, Albanese E, et al. Intentional weight loss in overweight and obese individuals and cognitive function: a systematic review and meta-analysis. *Obes Rev* 2011;12:968–983.
- [39] Lautenschlager NT, Cox KL, Flicker L, Foster JK, van Bockxmeer FM, Xiao J, et al. Effect of physical activity on cognitive function in older adults at risk for Alzheimer disease: a randomized trial. *JAMA* 2008;300:1027–1037.
- [40] Baker LD, Frank LL, Foster-Schubert K, Green PS, Wilkinson CW, McTiernan A, et al. Effects of aerobic exercise on mild cognitive impairment: a controlled trial. *Arch Neurol* 2010;67:71–79.
- [41] Hoogenboom WS, Marder TJ, Flores VL, Huisman S, Eaton HP, Schneiderman JS, et al. Cerebral white matter integrity and resting-state functional connectivity in middle-aged patients with type 2 diabetes. *Diabetes* 2014;63:728–738.
- [42] Espeland MA, Luchsinger JA, Baker LD, Neiberg R, Kahn SE, Arnold SE, et al. Effect of a long-term intensive lifestyle intervention on prevalence of cognitive impairment. *Neurology* 2017;88:2026–2035.