

Dataset Integrity Check for the Look AHEAD Continuation Study - Cognitive Function

Prepared by NIDDK-CR
August 27, 2021

Contents

1 Standard Disclaimer	2
2 Study Background	2
3 Archived Datasets	2
4 Statistical Methods	2
5 Results	2
6 Conclusions	3
7 References	3
Table A: Variables used to replicate Table 1 – Baseline characteristics of Look AHEAD participants who had cognitive assessments, by treatment group assignment	4
Table B: Comparison of values computed in integrity check to reference article Table 1 values ¹	5
Attachment A: SAS Code	7

1 Standard Disclaimer

The intent of this DSIC is to provide confidence that the data distributed by the NIDDK repository is a true copy of the study data. Our intent is not to assess the integrity of the statistical analyses reported by study investigators. As with all statistical analyses of complex datasets, complete replication of a set of statistical results should not be expected in secondary analysis. This occurs for a number of reasons including differences in the handling of missing data, restrictions on cases included in samples for a particular analysis, software coding used to define complex variables, etc. Experience suggests that most discrepancies can ordinarily be resolved by consultation with the study data coordinating center (DCC), however this process is labor-intensive for both DCC and Repository staff. It is thus not our policy to resolve every discrepancy that is observed in an integrity check. Specifically, we do not attempt to resolve minor or inconsequential discrepancies with published results or discrepancies that involve complex analyses, unless NIDDK Repository staff suspect that the observed discrepancy suggests that the dataset may have been corrupted in storage, transmission, or processing by repository staff. We do, however, document in footnotes to the integrity check those instances in which our secondary analyses produced results that were not fully consistent with those reported in the target publication.

2 Study Background

The Look AHEAD Continuation Study (Look AHEAD-C) builds on the Look AHEAD study to determine the long-term impact of an intensive lifestyle intervention on 1) physical function and mobility disability, and 2) cognitive function and cognitive impairment. Collection of cognitive function measures began in Year 8 of the study and continued through Year 13.

3 Archived Datasets

All SAS data files, as provided by the Data Coordinating Center (DCC), are located in the Look AHEAD folder in the data package. For this replication, variables were taken from the “batterysummary.sas7bdat”, “baseline_combined.sas7bdat”, and “cf_apoe_v3.sas7bdat” datasets.

4 Statistical Methods

Analyses were performed to replicate results for the data published by Rapp et al. [1] for Effect of a Long-Term Intensive Lifestyle Intervention on Cognitive Function: Action for Health in Diabetes Study. To verify the integrity of the datasets, descriptive statistics were computed.

5 Results

For Table 1 in the publication [1], Baseline characteristics of Look AHEAD participants who had cognitive assessments, by treatment group assignment, Table A lists the variables that were used in the replication, and Table B compares the results calculated from the archived data files to the results published in Table 1. The results of the replication are within expected variation to the published results given that only consented and randomized participant data were shared with the NIDDK Central Repository.

6 Conclusions

Since only data for participants who consented to share data with the NIDDK repository are available, the results of the replication are within expected variation to the published results.

7 References

[1] Rapp SR, Luchsinger JA, Baker LD, Blackburn GL, Hazuda HP, Demos-McDermott KE, Jeffery RW, Keller JN, McCaffery JM, Pajewski NM, Evans M, Wadden TA, Arnold SE, Espeland MA. Effect of a Long-Term Intensive Lifestyle Intervention on Cognitive Function: Action for Health in Diabetes Study. *Journal of the American Geriatrics Society*, 65(5), 966-972, May 2017. doi: <https://doi.org/10.1111/jgs.14692>

Table A: Variables used to replicate Table 1 – Baseline characteristics of Look AHEAD participants who had cognitive assessments, by treatment group assignment

Table Variable	dataset.variable
Age, years	baseline_combined.psage
BMI (kg/m ²)	baseline_combined.bmi
Gender	baseline_combined.psgender
History of cardiovascular disease	baseline_combined.cvdhis
Race/ethnicity	baseline_combined.racevar
Diabetes duration	baseline_combined.diab_dur
Education	baseline_combined.sdeduc
Hypertension	baseline_combined.hypertension
Prescription medications, Antihypertensive	baseline_combined.htndrug
Prescription medications, Any oral diabetes meds	baseline_combined.diabdrug
Prescription medications, Biguanides	baseline_combined.biguanides
Prescription medications, Insulin	baseline_combined.insulins
Prescription medications, Lipid lowering	baseline_combined.lipidrug
Short Form-36 General	baseline_combined.genhlth
Short Form-36 Physical Component	baseline_combined.pcs
Short Form-36 Mental Component	baseline_combined.mcs
Physical fitness, METs	baseline_combined.maxexmets
HbA1c	baseline_combined.hba1cpct
Apolipoprotein E4 alleles	cf_apoe_v3.ApoE4

Table B: Comparison of values computed in integrity check to reference article Table 1 values¹

Variable	Manuscript: Diabetes Support and Education (DSE) (n=1,861)	DSIC: DSE (n=1,846)	Diff. (n=15)	Manuscript: Intensive Lifestyle Intervention (ILI) (n=1,890)	DSIC: ILI (n=1,871)	Diff. (n=19)	Manuscript p-value	DSIC p-value	Diff.
Age, %							0.52	0.26	0.26
45-54	26	24	2	27	25	2			
55-64	57	58	1	57	59	2			
65-76	18	19	1	16	17	1			
Mean (SD)	58.2 (6.6)	58.6 (6.6)	0.4 (0)	58.0 (6.5)	58.3 (6.5)	0.3 (0)	0.34	0.14	0.20
Body mass index (kg/m ²), %							0.24	0.44	0.20
25-29.9	15	15	0	17	16	1			
30-39.9	63	63	0	62	62	0			
≥ 40	22	22	0	22	22	0			
Mean (SD)	36.0 (5.8)	36.0 (5.8)	0 (0)	35.6 (5.9)	35.7 (5.9)	0.1 (0)	0.08	0.19	0.11
Gender, %							0.46	0.46	0
Women	62	61	1	61	59	2			
Men	38	39	1	39	41	2			
History of cardiovascular disease, %							0.19	0.29	0.10
No	89	88	1	88	87	1			
Yes	11	12	1	12	13	1			
Race/Ethnicity, %							0.96	0.72	0.24
African-American	17	18	1	16	17	1			
American Indian	6	0	6	6	0	6			
Hispanic/Latino	13	13	0	14	14	0			
Non-Hispanic white	61	65	4	61	65	4			
Other/multiple	3	4	1	3	4	1			
Diabetes duration, %							0.43	0.69	0.26
< 5 years	46	47	1	48	48	0			
≥ 5 years	54	53	1	52	52	0			

Education, %							0.11	0.21	0.10
High school graduate	51	50	1	48	48	0			
College graduate	21	21	0	24	24	0			
Post college degree	19	20	1	19	20	1			
Other	9	9	0	9	8	1			
Hypertension, %							0.41	0.62	0.21
No	18	17	1	17	17	0			
Yes	82	83	1	83	83	0			
Prescription medications, %									
Antihypertensive	71	71	0	72	72	0	0.31	0.49	0.18
Any oral diabetes meds	86	86	0	86	86	0	0.94	0.60	0.34
Biguanides	61	61	0	62	63	1	0.63	0.29	0.34
Insulin	15	15	0	15	15	0	0.81	0.72	0.09
Lipid lowering	48	49	1	51	51	0	0.34	0.24	0.10
Short Form-36, Mean (SD)									
General	47.3 (8.7)	47.5 (8.6)	0.2 (0.1)	47.3 (9.1)	47.3 (9.1)	0 (0)	0.83	0.59	0.24
Physical Component	48.3 (7.8)	48.3 (7.8)	0 (0)	48.1 (7.9)	48.2 (7.9)	0.1 (0)	0.57	0.78	0.21
Mental Component	53.7 (7.8)	54.4 (7.7)	0.7 (0.1)	53.4 (8.2)	54.0 (8.0)	0.6 (0.2)	0.26	0.15	0.11
Physical Fitness, METS, Mean (SD)	7.3 (2.0)	7.3 (2.0)	0 (0)	7.4 (2.0)	7.3 (1.9)	0.1 (0.1)	0.60	0.73	0.13
HbA1c, %	7.25 (1.16)	7.25 (1.17)	0 (0.01)	7.22 (1.13)	7.22 (1.13)	0 (0)	0.44	0.42	0.02
Apolipoprotein E4 alleles, %							0.97	0.84	0.13
0	77	78	1	77	77	0			
1	21	21	0	21	21	0			
2	2	2	0	2	2	0			

¹ Documentation included in the package from the DCC states that some participants (e.g., from Native American sites) were excluded per consent limitations, and only the consented and randomized participants were sent to the NIDDK Central Repository.

Attachment A: SAS Code

```
* Look AHEAD Cognitive Function;
* DSIC;

*****;
* Read in data      *;
*****;

libname baseline "Z:\NIDDK\niddk-
dr_data_curation2\Look_AHEAD_V5\Data\Baseline\Analysis_Data";
libname cogfunc "Z:\NIDDK\niddk-
dr_studies2\Look_AHEAD\private_orig_data\LA.Cognitive
Function.1_v3\Cognitive Function - NIDDK Release\3-Cognitive Battery
Data\3a-Cognitive Battery Data Sets";

*****;
* Making temp datasets *;
*****;

data work.base; set baseline.baseline_combined;
run;

data work.apoe; set cogfunc.cf_apoe_v3;
run;

data work.batterysummary; set cogfunc.CF_BATTERYSUMMARY_V3;
run;

*****;
* Check data      *;
*****;

proc contents data=work.batterysummary; run;
* 5123 obs and 34 vars (equivalent to 5123 assessments);

proc contents data=work.base; run;
* 4906 obs and 294 vars;
* intervention: treatmentArm;
* age: psage;
* body mass index: bmi;
* gender: psgender;
* history of cardiovascular disease: cvdhis;
* race: racevar;
* diabetes duration: diab_dur;
* education: sdeduc;
* hypertension: hypertension;
* prescription medications, antihypertensive: htndrug;
* prescription medications, diabetes meds: diabdrug;
```



```

* prescription medications, biguanides: biguanides;
* prescription medications, insulin: insulins;
* prescription medications, lipid lowering: lipidrug;
* SF-36 general: genhlth;
* SF-36 physical component: pcs;
* SF-36 mental component: mcs;
* Physical fitness, METs: maxexmets;
* HbA1c: hbalcpct;

proc contents data=work.apoe; run;
* 4005 obs and 2 vars;
* Apolipoprotein E4 alleles: ApoE4;

*****;
* Match population *;
*****;

* Restrict to first visit per participant to match total N of
participants;
proc sort data=work.batterysummary out=work.batterysummary2 nodupkey;
    by MaskID;
run; *3717 observations/participants;

*****;
* Data management *;
* Sorting and merging datasets, *;
* keeping and formatting appropriate variables *;
*****;

data work.base2; set work.base;
    maskid = input(P_id, 8.); *rename ID to facilitate merge;
run;

proc sort data=work.base2;
    by maskid;
run;

proc sort data=work.apoe;
    by maskID;
run;

data work.combined1;
    merge
        work.batterysummary2 (IN=A)
        work.base2
        work.apoe;
    by maskid;
    if A;
run;

```

```

proc contents data=work.combined1; run;

*****;
* Recreating Table 1          *;
*****;

proc freq data=work.combined1;
  tables treatmentArm;
run;

* create a categorical variable for age;
data work.combined2;
  set work.combined1;
  if psage >= 45 AND psage <= 54 then psage_c = 1;
  if psage >=55 AND psage <=64 then psage_c = 2;
  if psage >= 65 AND psage <= 76 then psage_c = 3;
run;

proc freq data=work.combined2;
  tables psage_c*treatmentArm / chisq;
run;

proc means data=work.combined2;
  var psage;
  class treatmentArm;
run;

proc ttest data=work.combined2;
  class treatmentArm;
  var psage;
run;

* create a categorical variable for bmi;
data work.combined3;
  set work.combined2;
  if bmi <30 then bmi_c = 1;
  if bmi >=30 AND bmi <40 then bmi_c = 2;
  if bmi >=40 then bmi_c = 3;
run;

proc freq data=work.combined3;
  tables bmi_c*treatmentArm / chisq;
run;

proc means data=work.combined3;
  var bmi;
  class treatmentArm;
run;

proc ttest data=work.combined3;
  class treatmentArm;

```

```

    var bmi;
run;

proc freq data=work.combined3;
    tables psgender*treatmentArm / chisq;
run;

proc freq data=work.combined3;
    tables cvdhis*treatmentArm / chisq;
run;

proc freq data=work.combined3;
    tables racevar*treatmentArm / chisq;
run;

* create a categorical variable for diabetes duration;
data work.combined4;
    set work.combined3;
    if diab_dur <5 then diab_dur2 = 1;
    if diab_dur >=5 then diab_dur2 = 2;
run;

proc freq data=work.combined4;
    tables diab_dur2*treatmentArm / chisq;
run;

* create a categorical variable for education;
data work.combined5;
    set work.combined4;
    if sdeduc = "Associate degree (junior college)" OR sdeduc = "High
school diploma or equivalency (GED)" OR sdeduc = "Some vocational
school" OR sdeduc = "Some college" then sdeduc2=1;
    if sdeduc = "Bachelor's degree" OR sdeduc = "Some graduate
school" then sdeduc2 = 2;
    if sdeduc = "Doctorate" OR sdeduc = "Master's degree" OR sdeduc =
"Professional (MD, JD, DDS, etc.)" then sdeduc2=3;
    if sdeduc = "Less than high school" OR sdeduc = "Missing" OR
sdeduc = "Other" then sdeduc2=4;
run;

* confirm variable creation;
proc freq data=work.combined5;
    tables sdeduc*sdeduc2;
run;

proc freq data=work.combined5;
    table sdeduc2* treatmentArm / chisq;
run;

proc freq data=work.combined5;
    tables hypertension*treatmentArm / chisq;
run;

```

```

* recode htndrug to set "Missing" to .;
data work.combined6;
    set work.combined5;
    if htndrug = "Yes" then htndrug2 = 1;
    if htndrug = "No" then htndrug2 = 2;
run;

proc freq data=work.combined6;
    tables htndrug2*treatmentArm / chisq;
run;

* recode diabdrug to set "Missing" to .;
data work.combined7;
    set work.combined6;
    if diabdrug = "Yes" then diabdrug2 = 1;
    if diabdrug = "No" then diabdrug2 = 2;
run;

proc freq data=work.combined7;
    tables diabdrug2*treatmentArm / chisq;
run;

* recode biguanides to set "Missing" to .;
data work.combined8;
    set work.combined7;
    if biguanides = "Yes" then biguanides2 = 1;
    if biguanides = "No" then biguanides2 = 2;
run;

proc freq data=work.combined8;
    tables biguanides2*treatmentArm / chisq;
run;

* recode insulins to set "Missing" to .;
data work.combined9;
    set work.combined8;
    if insulins = "Yes" then insulins2 = 1;
    if insulins = "No" then insulins2 = 2;
run;

proc freq data=work.combined9;
    tables insulins2*treatmentArm / chisq;
run;

* recode lipidrug to set "Missing" to .;
data work.combined10;
    set work.combined9;
    if lipidrug = "Yes" then lipidrug2 = 1;
    if lipidrug = "No" then lipidrug2 = 2;
run;

```

```
proc freq data=work.combined10;  
    tables lipiddrug2*treatmentArm / chisq;  
run;  
  
proc means data=work.combined10;  
    var genhlth;  
    class treatmentArm;  
run;  
  
proc ttest data=work.combined10;  
    class treatmentArm;  
    var genhlth;  
run;  
  
proc means data=work.combined10;  
    var pcs;  
    class treatmentArm;  
run;  
  
proc ttest data=work.combined10;  
    class treatmentArm;  
    var pcs;  
run;  
  
proc means data=work.combined10;  
    var mcs;  
    class treatmentArm;  
run;  
  
proc ttest data=work.combined10;  
    class treatmentArm;  
    var mcs;  
run;  
  
proc means data=work.combined10;  
    var maxexmets;  
    class treatmentArm;  
run;  
  
proc ttest data=work.combined10;  
    class treatmentArm;  
    var maxexmets;  
run;  
  
proc means data=work.combined10;  
    var hbalcpct;  
    class treatmentArm;  
run;  
  
proc ttest data=work.combined10;  
    class treatmentArm;  
    var hbalcpct;
```

```
run;  
  
proc freq data=work.combined10;  
  tables ApoE4*treatmentArm / chisq;  
run;
```