Dataset Integrity Check for the Look AHEAD Brain Magnetic Resonance Imaging Study (Look AHEAD Brain)

Prepared by NIDDK-CR
July 23, 2021

Contents

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 ancillary study, Look AHEAD Brain, was conducted to assess whether participation in the 10-year lifestyle intervention, as part of Look AHEAD, had an impact on white matter hyperintensity and loss of brain tissue among individuals with type 2 diabetes. A subset of Look AHEAD study participants underwent standardized brain magnetic resonance imaging in conjunction with tests assessing cognitive function 10-12 years post-randomization.

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 "healthbehaviors_diet.sas7bdat", "baseline_combined.sas7bdat", "sociodemographicsbaseline.sas7bdat", "laboratorymeasures.sas7bdat", "healthoutcomes_meduse.sas7bdat", and "cf_brainvolume.sas7bdat" datasets.

4 Statistical Methods

Analyses were performed to replicate results for the data published by Espeland et al. [1] for Brain and White Matter Hyperintensity Volumes After 10 Years of Random Assignments to Lifestyle Intervention. To verify the integrity of the dataset, descriptive statistics were computed.

5 Results

For Table 1 in the publication [1], Characteristics at the time of enrollment into the Look AHEAD trial of participants who had successful MRI scans in the Look AHEAD Brain study, by intervention 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.

Please note that the category, “ApoE4 alleles”, in Table 1 was not able to be replicated as the underlying variable(s) in the datasets could not be identified. All other variables in Table 1 were identified and replicated successfully.

6 Conclusions

The results of the replication are within expected variation to the published results.

7 References

[1] Espeland MA, Erickson K, Neiberg RH, Jakicic JM, Wadden TA, Wing RR, Desiderio L, Erus G, Hsieh MK, Davatzikos C, Maschak-Carey BJ, Laurienti PJ, Demos-McDermott K, Bryan RN. Brain and White Matter Hyperintensity Volumes After 10 Years of Random Assignment to Lifestyle Intervention. Diabetes Care, 39(5), 764-771, May 2016. PMCID: http://www.ncbi.nlm.nih.gov/pmc/articles/pmc4839171/ doi: https://doi.org/10.2337/dc15-2230

Table A: **Variables used to replicate Table 1 – Characteristics at the time of enrollment into the Look AHEAD trial of participants who had successful MRI scans in the Look AHEAD Brain study, by intervention assignment**

Table Variable	**dataset.variable**
Age, years	cf_brainvolume.psage
Female sex	baseline_combined.psgender
Race/ethnicity	baseline_combined.racevar
Education, years	sociodemographicbaseline.sdeduc
BMI, kg/m^2	cf_brainvolume.bmi
HbA_{1C}, %	laboratorymeasures.hba1cpct
Diabetes duration, years	baseline_combined.diab_dur
Insulin use	healthoutcomes_meduse.insulins
Hypertension	baseline_combined.hypertension
Prior cardiovascular disease	baseline_combined.cvdhis
Depressive symptoms	baseline_combined.beck_score_nowgt
Antidepressant use	healthoutcomes_meduse.antidep
Alcohol intake, drinks/day	healthbehaviors_diet.FG27
Baseline smoking status	baseline_combined.smoking
Fitness, METs	baseline_combined.maxexmets
Paffenbarger activity questionnaire score	baseline_combined.paffenbarger

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

Variable	Manuscript: Diabetes Support and Education (DSE) (n=155)	DSIC: DSE (n=154)	Diff. (n=1)	Manuscript: Intensive Lifestyle Intervention (ILI) (n=164)	DSIC: ILI (n=164)	Diff. (n=0)	Manuscript *p*-value	DSIC *p*-value	Diff.
Age, years							0.50	0.53	0.03
45-54	44 (28.4)	43 (27.9)	1 (0.5)	40 (24.4)	40 (24.5)	0 (0.1)			
55-64	91 (58.7)	91 (59.1)	0 (0.4)	96 (58.5)	95 (58.3)	1 (0.2)			
65-76	20 (12.9)	20 (13.0)	0 (0.1)	28 (17.1)	28 (17.2)	0 (0.1)			
Female sex	114 (73.5)	113 (73.4)	1 (0.1)	110 (67.1)	110 (67.1)	0 (0)	0.21	0.22	0.01
Race/ethnicity							0.70	0.68	0.02
African American	37 (23.9)	37 (24.0)	0 (0.1)	33 (20.1)	33 (20.1)	0 (0)			
Non-Hispanic white	110 (71.0)	109 (70.8)	1 (0.2)	123 (75.0)	123 (75.0)	0 (0)			
Other/multiple	8 (5.2)	8 (5.2)	0 (0)	8 (4.9)	8 (4.9)	0 (0)			
Education, years							0.61	0.56	0.05
< 13	24 (16.1)	24 (16.2)	0 (0.1)	30 (18.8)	32 (19.8)	2 (1.0)			
13-16	58 (38.9)	57 (38.5)	1 (0.4)	54 (33.8)	54 (33.3)	0 (0.5)			
> 16	67 (45.0)	67 (45.2)	0 (0.2)	76 (47.5)	76 (46.9)	0 (0.6)			
BMI, kg/m^2							0.046	0.03	0.016
25-29	18 (11.6)	18 (11.7)	0 (0.1)	32 (19.5)	32 (19.5)	0 (0)			
30-39	99 (63.9)	98 (63.6)	1 (0.3)	106 (64.6)	107 (65.2)	1 (0.6)			
≥ 40	38 (24.5)	38 (24.7)	0 (0.2)	26 (15.9)	25 (15.2)	1 (0.7)			
HbA_{1c}, %							0.58	0.48	0.10
< 7.0	65 (42.5)	65 (42.2)	0 (0.3)	73 (44.8)	73 (44.5)	0 (0.3)			
7.0-8.9	75 (49.0)	74 (48.1)	1 (0.9)	81 (49.7)	81 (49.4)	0 (0.3)			
≥ 9.0	13 (8.5)	15 (9.7)	2 (1.2)	9 (5.5)	10 (6.1)	1 (0.6)			
Diabetes duration, years							0.97	0.90	0.07
< 5	73 (47.7)	74 (48.0)	1 (0.3)	76 (47.5)	80 (48.8)	4 (1.3)			
≥ 5	80 (52.3)	80 (52.0)	0 (0.3)	84 (52.5)	84 (51.2)	0 (1.3)			
Insulin use	20 (13.6)	20 (13.7)	0 (0.1)	18 (11.5)	18 (11.5)	0 (0)	0.57	0.56	0.01
Hypertension	129 (83.2)	128 (83.1)	1 (0.1)	135 (82.3)	135 (82.3)	0 (0)	0.83	0.85	0.02

Prior cardiovascular disease	13 (8.4)	13 (8.4)	0 (0)	13 (7.9)	13 (7.9)	0 (0)	0.88	0.87	0.01
Depressive symptoms							0.77	0.77	0
BDI score < 11	135 (87.7)	135 (87.7)	0 (0)	142 (86.6)	142 (86.6)	0 (0)			
BDI score ≥ 11	19 (12.3)	19 (12.3)	0 (0)	22 (13.4)	22 (13.4)	0 (0)			
Antidepressant use	28 (19.0)	28 (19.2)	0 (0.2)	32 (20.4)	32 (20.4)	0 (0)	0.77	0.79	0.02
Alcohol intake, drinks/day							0.56	0.80	0.24
< 1	143 (92.3)	150 (97.4)	7 (5.1)	154 (93.9)	159 (97.0)	5 (3.1)			
≥ 1	12 (7.7)	4 (2.6)	8 (5.1)	10 (6.1)	5 (3.0)	5 (3.1)			
Baseline smoking status							0.97	0.96	0.01
Never	77 (49.7)	76 (49.4)	1 (0.3)	83 (50.9)	83 (50.9)	0 (0)			
Past	72 (46.5)	72 (46.7)	0 (0.2)	74 (45.4)	74 (45.4)	0 (0)			
Current	6 (3.9)	6 (3.9)	0 (0)	6 (3.7)	6 (3.7)	0 (0)			
Fitness, METs							0.07	0.06	0.01
< 7.1	84 (54.2)	84 (54.5)	0 (0.3)	72 (43.9)	72 (43.9)	0 (0)			
≥ 7.1	71 (45.8)	70 (45.5)	1 (0.3)	92 (56.1)	92 (56.1)	0 (0)			
Paffenbarger activity questionnaire score							0.72	0.70	0.02
< 1,060	73 (47.1)	73 (47.4)	0 (0.3)	70 (42.7)	70 (42.7)	0 (0)			
≥ 1,060	35 (22.6)	35 (22.7)	0 (0.1)	41 (25.0)	41 (25.0)	0 (0)			
Not collected	47 (30.3)	46 (29.9)	1 (0.4)	53 (32.3)	53 (32.3)	0 (0)			

Attachment A: SAS Code

```
libname brain "X:\NIDDK\niddk-
dr_studies2\Look_AHEAD\private_orig_data\la_cogfunc_brainvolume\brainv
olume4";
libname baseline "X:\NIDDK\niddk-
dr_data_curation2\Look_AHEAD_V5\Data\Baseline\Analysis_Data";
libname socio "X:\NIDDK\niddk-
dr_data_curation2\Look_AHEAD_V5\Data\Baseline\Form_Data";
libname lab "X:\NIDDK\niddk-
dr_data_curation2\Look_AHEAD_V5\Data\End_of_Intervention\Measures_Data
";
libname health "X:\NIDDK\niddk-
dr_data_curation2\Look_AHEAD_V5\Data\End_of_Intervention\Forms_Data";

******************;
*Check Data      *;
******************;

proc contents data=health.healthbehaviors_diet;
run;
*FG27 - alc servings per day;

proc contents data=baseline.baseline_combined;
run;
*beck_score_wgt;
*beck_score_nowgt;
*paffenbarger;
*maximal MET value: maxexmets
*sex: psgender;
*age: age;
*intervention: treatmentarm;
*race/eth: racevar;
*diabetes duration: diab_dur;
*hypertension: hypertension;
*CVD: cvdhis;
*smoking status: smoking;

proc contents data=socio.sociodemographicbaseline;
run;

*education: sdeduc;

proc contents data=lab.laboratorymeasures;
run;

*HbA1C %: HbA1cpct;

proc contents data=health.healthoutcomes_meduse;
run;
```

```
*insulin: insulins;
*anitdepressants: antidep;

****************************************;
*Making Temp Datasets                   ;
****************************************;

data healthdiet; set health.healthbehaviors_diet;
if visit = "Baseline";
run;

data brain; set brain.cf_brainvolume;
run;

data base; set baseline.baseline_combined;
run;

data socio; set socio.sociodemographicbaseline;
run;

data labs; set lab.laboratorymeasures;
run;

data alcohol; set socio.alcohol;
run;

data healthout; set health.healthoutcomes_meduse;
run;

proc freq data=base;
tables p_id;
run;

proc freq data=brain;
tables MaskID;
run;

************************************************;
*Merging Datasets and keeping appropriate variables;
************************************************;

data base2; set base;
     maskid = input(P_id, 8.);
     CVDhistory = input(cvdhis, 8.);
run;

data socio2; set socio;
     maskid = input(p_id, 8.);
run;

data labs2; set labs;
if visit = "Baseline";
run;

data healthout2; set healthout;
```

```
if visit = "Baseline";
run;

data alcohol2; set alcohol;
maskid = input(p_id, 8.);
run;

proc sort data=healthdiet;
by maskid;
run;

proc sort data=alcohol2;
by maskid;
run;

proc sort data=brain;
by maskid;
run;

proc sort data=base2;
by maskid;
run;

proc sort data=socio2;
by maskid;
run;

proc sort data=labs;
by MaskID;
run;

proc sort data=healthbeh;
by MaskID;
run;

proc sort data=healthout;
by MaskID;
run;

data combined;
     merge
          brain      (in=a)
          base2      (in=b keep=maskid beck_score_wgt
beck_score_nowgt paffenbarger maxexmets psgender treatmentarm racevar
diab_dur hypertension cvdhistory smoking age total_alcohol)
          socio2     (in=c keep=maskid sdeduc)
          labs2      (in=d keep=maskid hba1cpct)
          healthout2 (in=e keep=maskid insulins antidep)
          healthdiet (in=f keep=maskid FG27);
     by maskid;
     if a=1;
run;

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

*creating an age categorical variable;
```

```
data combined2; set combined;
if psage <= 54 then age_cat = 1;
if psage >=55 AND age <=64 then age_cat = 2;
if psage >= 65 then age_cat = 3;
run;

proc freq data=combined2;
tables age_cat*treatmentarm/chisq;
run;

proc freq data=combined2;
tables psgender*treatmentarm/chisq;
run;

*need to alter the race variable to match publication;
data combined3; set combined2;
if racevar = "African American / Black (not Hispanic)" then race = 1;
if racevar = "White" then race = 2;
if racevar = "Hispanic" OR racevar = "Other/Mixed" then race = 3;
run;

proc freq data=combined3;
tables race*treatmentarm/chisq;
run;

*need to alter the education variable to match the publication;
data combined4; set combined3;
if sdeduc <=2 then educ = 1;
if sdeduc = 3 or sdeduc = 4 or sdeduc =5 then educ = 2;
if sdeduc >=6 and sdeduc <=10 then educ = 3;
run;

proc freq data=combined4;
tables sdeduc educ*treatmentarm/chisq;
run;

*need to create categorical BMI var;
data combined5; set combined4;
if bmi >=25 AND bmi <30 then bmi2 = 1;
if bmi >=30 AND bmi <40 then bmi2 = 2;
if bmi >=40 then bmi2 = 3;
run;

proc freq data=combined5;
tables bmi2*treatmentarm/chisq;
run;

*need to create a categorical HbA1c variable;
data combined6; set combined5;
if hba1cpct < 7 then hba1 = 1;
if hba1cpct >= 7 and hba1cpct < 9 then hba1 = 2;
if hba1cpct >= 9 then hba1 = 3;
run;

proc freq data=combined6;
tables HbA1*treatmentarm/chisq;
run;
```

```
*categorical variable for diabetes duration;
data combined7; set combined6;
if diab_dur <5 then diadur = 1;
if diab_dur >=5 then diadur = 2;
run;

proc freq data=combined7;
tables diadur*treatmentarm/chisq;
run;

proc freq data=combined7;
tables insulins*treatmentarm/chisq;
run;

proc freq data=combined7;
tables hypertension*treatmentarm/chisq;
run;

proc freq data=combined7;
tables cvdhis*treatmentarm/chisq;
run;

proc freq data=combined7;
tables antidep*treatmentarm/chisq;
run;

*categorical alcohol intake var;
data combined8; set combined7;
if fg27 < 1 then alc = 1;
if fg27 >= 1 then alc = 2;
run;

proc freq data=combined8;
tables alc*treatmentarm/chisq;
run;

data combined9; set combined8;
if smoking = "Missing" then smoking2 = .;
if smoking = "Never" then smoking2 = 1;
if smoking = "Past" then smoking2 = 2;
if smoking = "Present" then smoking2 = 3;
run;

proc freq data=combined9;
tables smoking2*treatmentArm/chisq;
run;

*categorical METs var;
data combined10; set combined9;
if maxexmets <7.1 then mets = 1;
if maxexmets >=7.1 then mets = 2;

proc freq data=combined10;
tables mets*treatmentarm/chisq;
run;

*paffenbarger categorical var;
data combined11; set combined10;
```

```
if paffenbarger <1060 then paff = 1;
if paffenbarger >=1060 then paff = 2;
if paffenbarger = . then paff = 3;
run;

proc freq data=combined11;
tables paff*treatmentarm/chisq;
run;

data combined12; set combined11;
if beck_score_nowgt <11 then beck_nowgt = 1;
if beck_score_nowgt >=11 then beck_nowgt = 2;
run;

proc freq data=combined12;
tables beck_nowgt*treatmentarm/chisq;
run;
```