Dataset Integrity Check for Lifestyle Interventions for Expectant Moms (LIFE-Moms) Study

Prepared by NIDDK-CR February 16, 2023

Contents

1 Standard Disclaimer	. 2
2 Study Background	. 2
3 Archived Datasets	. 2
4 Statistical Methods	. 2
5 Results	. 2
6 Conclusions	. 3
Table A: Variables used to replicate results provided by the DCC	.4
Table B: Comparison of values computed in integrity check to DCC results	. 5
Attachment A: SAS Code	. 6

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 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

LIFE-Moms was a research consortium that consisted of seven independent but collaborative clinical trials and a research coordinating unit. The goal of the consortium was to identify effective behavioral and lifestyle interventions that would improve weight, glycemic control, and other pregnancy related outcomes among pregnant women that were overweight or obese. In addition, the consortium assessed if the lifestyle interventions reduced obesity and metabolic abnormalities in the children of LIFE-Moms participants.

3 Archived Datasets

All data files, as provided by the Data Coordinating Center (DCC), are located in the LIFE-Moms folder in the data package. A complete listing of datasets can be found in the Roadmap document. For this replication, variables were taken from the "primout.sas7bdat", "Im01niddkrelease.sas7bdat", "Im02niddkrelease.sas7bdat", "Im02cniddkrelease.sas7bdat", and "Im04niddkrelease.sas7bdat" datasets.

4 Statistical Methods

Analyses were performed to replicate results provided by the DCC. To verify the integrity of the data, only descriptive statistics were computed.

5 Results

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 provided by the DCC. The results of the replication are within expected variation to the DCC results.

Conclusions

The NIDDK Central Repository is confident that the LIFE-Moms data files to be distributed are a true copy of the study data.

Table A: Variables used to replicate results provided by the DCC

Table Variable	dataset.variable
Gestational age at randomization (weeks)	Primout.lmgroup
	Lm01niddkrelease.randgawk
Maternal age (years)	Primout.Imgroup
	Lm01niddkrelease.mage
Adjusted BMI at baseline (kg/m ²)	Primout.Imgroup
	Lm01niddkrelease.aprojbmi
Adjusted BMI at baseline category	Primout.Imgroup
	Lm01niddkrelease.aprojbmi
Race/Ethnicity	Primout.Imgroup
	Lm02niddkrelease.newrace
Nulliparous	Primout.Imgroup
	Lm02cniddkrelease.para
College education	Primout.Imgroup
	Lm02niddkrelease.college
Total family income	Primout.Imgroup
	Lm02niddkrelease.income
Married/living with significant other	Primout.Imgroup
	Lm02niddkrelease.bmarry
Neonatal sex	Primout.Imgroup
	Lm04niddkrelease.gsex

Table B: Comparison of values computed in integrity check to DCC results

Characteristic	DCC: Intervention	DSIC: Intervention	Diff.	DCC: Standard	DSIC: Standard	Diff.		
	(n=469)	(n=469)	(n=0)	of Care (n=464)	of Care (n=464)	(n=0)		
Gestational age at randomization (weeks)	14.3 [12.9-15.3]	14.3 [12.9-15.3]	0 [0-0]	14.3 [12.7-15.3]	14.3 [12.7-15.3]	0 [0-0]		
Maternal age (years)	30.5 ± 5.8	30.9 ± 4.5	0.4 ± 1.3	30.6 ± 5.7	30.8 ± 4.5	0.2 ± 1.2		
Adjusted BMI at baseline (kg/m ²)	30.8 [27.8-34.6]	30.8 [27.8-34.6]	0 [0-0]	30.6 [28.0-34.6]	30.6 [28.0-34.6]	0 [0-0]		
Adjusted BMI at baseline category								
Overweight	207 (44.1%)	207 (44.1%)	0 (0)	200 (43.1%)	200 (43.1%)	0 (0)		
Obese	262 (55.9%)	262 (55.9%)	0 (0)	264 (56.9%)	264 (56.9%)	0 (0)		
Race/Ethnicity								
Non-Hispanic Caucasian	157 (33.5%)	157 (33.5%)	0 (0)	177 (38.2%)	177 (38.2%)	0 (0)		
Non-Hispanic African American	177 (37.7%)	177 (37.7%)	0 (0)	160 (34.5%)	160 (34.5%)	0 (0)		
Hispanic	103 (21.9%)	103 (22.0%)	0 (0.1)	98 (21.1%)	98 (21.1%)	0 (0)		
Other, more than one race	32 (6.8%)	32 (6.8%)	0 (0)	29 (6.3%)	29 (6.3%)	0 (0)		
Nulliparous	207 (44.1%)	207 (44.1%)	0 (0)	170 (36.6%)	170 (36.6%)	0 (0)		
College education	233 (49.9%)	235 (50.1%)	2 (0.2)	229 (49.4%)	229 (49.4%)	0 (0)		
Total family income								
< \$25,000	161 (34.8%)	161 (34.8%)	0 (0)	167 (36.2%)	167 (36.2%)	0 (0)		
\$25,000 - \$74,999	119 (25.7%)	119 (25.7%)	0 (0)	119 (25.8%)	119 (25.8%)	0 (0)		
≥ \$75,000	183 (39.5%)	183 (39.5%)	0 (0)	176 (38.1%)	176 (38.1%)	0 (0)		
Married/living with significant other	350 (74.8%)	350 (74.8%)	0 (0)	363 (78.2%)	363 (78.2%)	0 (0)		
Neonatal sex								
Male	211 / 463 (45.6%)	211 / 463 (45.6%)	0 (0)	237 / 453 (52.3%)	237 / 453 (52.3%)	0 (0)		
Female	252 / 463 (54.4%)	252 / 463 (54.4%)	0 (0)	216 / 453 (47.7%)	216 / 453 (47.7%)	0 (0)		

Attachment A: SAS Code

libname life "X:\NIDDK\niddk-dr_studies2\LIFE-Moms\private_orig_data\NIDDK Data Repository LIFEMoms Upload 20210713\LIFE-Moms Release Datasets sas7bdat files";

*Participants in each treatment group: Standard of Care and Intervention; *Variable: LMGROUP, coding: X=Standard of Care, Y=Intervention; proc freq data=life.primout; tables Imgroup/missing; run;

*Gestational Age at Randomization; data work.lm01; set life.lm01niddkrelease; run;

data work.primout; set life.primout;
run;

proc sort data=lm01; by releaseid; run;

proc sort data=primout; by releaseid; run;

data ga; merge Im01 (in=a) primout (in=b); by releaseid; run;

*Ga at randomization; data ga_1; set ga; randgawk = (randgad/7); run;

proc means data=ga_1 median q1 q3; var randgawk; class Imgroup; run;

*Maternal Age at baseline;

```
data ga_1; set ga;
if mage = 0 then mage = 25;
if mage = 2 then mage = 37.78;
run;
proc means data=ga_1 mean std;
var mage;
class Imgroup;
run;
*BMI;
proc means data=ga median q1 q3;
var aprojbmi;
class Imgroup;
run;
*BMI Car;
data bmi; set ga;
bmi cat = 0;
/*if aprojbmi < 25 then bmi_cat = 0;*/</pre>
if /*aprojbmi >= 25 AND*/ aprojbmi < 30 then bmi_cat = 1;
if aprojbmi >=30 then bmi_cat = 2;
run;
proc freq data=bmi;
tables bmi_cat*Imgroup/norow nopercent;
run;
*Race/Ethnicity;
*variable: newrace, values: 1 = Hispanic, 2 = Non-Hispanic AA, 3 = Non-hispanic Caucasian, 4 = Other;
data Im02; set life.Im02niddkrelease;
run;
proc sort data=lm02;
by releaseid;
run;
data race; merge
primout (in=a)
Im02 (in=b);
by releaseid;
run;
proc freq data=race;
tables newrace*Imgroup/norow nopercent;
run;
*College Education;
```

proc freq data=race; tables college*Imgroup/norow nopercent; run;

*Total fam income; **proc freq** data=race; tables income*Imgroup/norow nopercent; **run**;

*married/living together; data marry; set race; mar_liv = 0; if bmarry = . then mar_liv = .; if bmarry = 1 OR bmarry = 2 then mar_live = 1; if bmarry = 3 or bmarry = 4 then mar_live = 2; run;

proc freq data=marry;

tables mar_live*Imgroup/norow nopercent; **run**;

*nulliparous;
data prevpreg; set life.lm02cniddkrelease;
run;

proc sort data=prevpreg; by releaseid; run;

data preg; merge primout (in=a) prevpreg (in=b); by releaseid; run;

proc freq data=preg; tables para*Imgroup/norow nopercent; run;

*neonatal sex; data lm04; set life.lm04niddkrelease; run;

proc sort data=lm04; by releaseid; run; data sex; merge Im04 (in=a) primout (in=b); by releaseid; run;

proc freq data=sex; tables gsex*Imgroup/norow nopercent; run;