Dataset Integrity Check for Rare and Atypical Diabetes Network (RADIANT) M001 Parikh

Prepared by NIDDK-CR
August 8, 2024

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 Rare and Atypical DIAbetes NeTwork (RADIANT) is a multicenter prospective cohort study that is dedicated to characterizing (discovering and defining) rare and atypical forms of diabetes. Individuals will be screened for evaluation of suspected atypical diabetes of unknown origin. Among the pool of evaluated individuals, those found to have a known form of diabetes will be excluded from further study. The remaining participants will be prioritized for genetic/genomic analysis and further testing.

The M001 study sought to develop an analytical approach to identify and cluster phenotypes of atypical diabetes (AD).

3 Archived Datasets

A full listing of the archived datasets included in the package can be found in the Roadmap document. All data files, as provided by the Data Coordinating Center (DCC), are located in the RADIANT folder in the data package. For this replication, variables were taken from the "m_001_hparikh_niddk_04july2022_1.sas7bdat" and "m_001_hparikh_niddk_04july2022_2.sas7bdat" datasets.

4 Statistical Methods

Analyses were performed to replicate results for the data in the publication by Parikh et al. [1]. To verify the integrity of the data, only descriptive statistics were computed.

5 Results

For Table 1 and Table 2 in the publication [1], Clinical parameters for GAD65Ab-positive and -negative participants with diabetes from the Cameron County Hispanic Cohort (CCHC) (Study A) and Clinical parameters for autoantibody-positive and -negative participants with diabetes from the Texas Children's Hospital Registry for New-Onset Type 1 Diabetes (TCHRNO-1) (Study B), respectively, Tables A1 and A2 list the variables that were used in the replication, and Tables B1 and B2 compare the results calculated from the archived data files to the results in Tables 1 and 2. The results of the replication are within expected variation to the published results.

6 Conclusions

The NIDDK Central Repository is confident that the RADIANT M001 data files to be distributed are a true copy of the study data.

7 References

[1] Parikh HM, Remedios CL, Hampe CS, Balasubramanyam A, Fisher-Hoch SP, Choi YJ, Patel S, McCormick JB, Redondo MJ, Krischer JP. Data Mining Framework for Discovering and Clustering Phenotypes of Atypical Diabetes. The Journal of Clinical Endocrinology and Metabolism, 108(4), 834-846, April 2023. doi: https://doi.org/10.1210/clinem/dgac632

Table A1: Variables used to replicate Table 1 – Clinical parameters for GAD65Ab-positive and -negative participants with diabetes from the Cameron County Hispanic Cohort (CCHC) (Study A)

Table Variable	**dataset.variable**
N (male/female)	m_001_hparikh_niddk_04july2022_1.gender
GAD65Ab levels (U/mL)	m_001_hparikh_niddk_04july2022_1.gad65ab_index
Age at onset (years)	m_001_hparikh_niddk_04july2022_1.diab_age
Disease duration (years)	m_001_hparikh_niddk_04july2022_1.disease_duration
Body mass index (kg/m^2)	m_001_hparikh_niddk_04july2022_1.bmi1
Waist circumference (cm)	m_001_hparikh_niddk_04july2022_1.waist
Blood pressure systolic (mmHg)	m_001_hparikh_niddk_04july2022_1.corrsys1
Blood pressure diastolic (mmHg)	m_001_hparikh_niddk_04july2022_1.corrdia1
Triglycerides (mg/dL)	m_001_hparikh_niddk_04july2022_1.trig
High-density lipoprotein cholesterol (mg/dL)	m_001_hparikh_niddk_04july2022_1.hdlc
Fasting blood glucose (mg/dL)	m_001_hparikh_niddk_04july2022_1.mfbg
Homeostasis Model Assessment – Insulin Resistance Index	m_001_hparikh_niddk_04july2022_1.homa_ir
Homeostasis Model Assessment – Beta-cell Function (%)	m_001_hparikh_niddk_04july2022_1.homa_beta

Table A2: Variables used to replicate Table 2 – Clinical parameters for autoantibody-positive and -negative participants with diabetes from the Texas Children's Hospital Registry for New-Onset Type 1 Diabetes (TCHRNO-1) (Study B)

Table Variable	**dataset.variable**
N (male/female)	m_001_hparikh_niddk_04july2022_2.female
Age at diagnosis (years)	m_001_hparikh_niddk_04july2022_2.dxage
Body mass index percentile (%)	m_001_hparikh_niddk_04july2022_2.v1bmicentile
pH at diagnosis	m_001_hparikh_niddk_04july2022_2.ph
Beta-hydroxybutyrate (mmol/L) at diagnosis	m_001_hparikh_niddk_04july2022_2.bohb
Bicarbonate (mmol/L) at diagnosis	m_001_hparikh_niddk_04july2022_2.bicarb
Glucose (mg/dL) at diagnosis	m_001_hparikh_niddk_04july2022_2.glucose
Hemoglobin A1c (%) at diagnosis	m_001_hparikh_niddk_04july2022_2.dxa1c
C-peptide levels (ng/mL) at diagnosis	m_001_hparikh_niddk_04july2022_2.dxcpep
Insulin autoantibodies levels (U/mL) at diagnosis	m_001_hparikh_niddk_04july2022_2.dxiaa
Islet cell antigen 512 autoantibodies levels (U/mL) at diagnosis	m_001_hparikh_niddk_04july2022_2.dxica512
Glutamic acid decarboxylase autoantibodies levels (U/mL) at diagnosis	m_001_hparikh_niddk_04july2022_2.dxgad65

Table B1: Comparison of values computed in integrity check to reference article Table 1

Characteristic	Pub: GAD65Ab negative (n=185)	DSIC: GAD65Ab negative (n=185)	Diff. (n=0)	Pub: GAD65Ab positive (n=93)	DSIC: GAD65Ab positive (n=93)	Diff. (n=0)
N (male/female)	185 (66/119)	185 (66/119)	0 (0)	93 (30/63)	93 (30/63)	0 (0)
GAD65Ab levels (U/mL)	0.03 (0-0.05)	0.03 (0-0.05)	0 (0)	0.08 (0.05-1.35)	0.08 (0.05-1.35)	0 (0)
Age at onset (years)	50 (11-91)	50 (11-91)	0 (0)	44 (4-81)	45 (4-81)	1 (0)
Disease duration (years)	6 (0-37)	6 (0-37)	0 (0)	10 (0-36)	10 (0-36)	0 (0)
Body mass index (kg/m^2)	33.6 (17-65.6)	33.6 (17-65.6)	0 (0)	30.2 (19.5-62.6)	30.2 (19.5-62.6)	0 (0)
Waist circumference (cm)	109.5 (78-162.2)	109.5 (78-162.2)	0 (0)	104 (73.5-166)	104 (73.5-166)	0 (0)
Blood pressure systolic (mmHg)	126 (85-191)	126 (85-191)	0 (0)	122 (89-179)	122 (89-179)	0 (0)
Blood pressure diastolic (mmHg)	74 (44-98)	74 (44-98)	0 (0)	73 (42-92)	73 (42-92)	0 (0)
Triglycerides (mg/dL)	148 (49-2728)	148 (49-2728)	0 (0)	155 (49-1227)	155 (49-1227)	0 (0)
High-density lipoprotein cholesterol (mg/dL)	44 (6.2-106)	44 (6.2-106)	0 (0)	43 (20-150)	43 (20-150)	0 (0)
Fasting blood glucose (mg/dL)	128 (64-402)	128 (64-402)	0 (0)	144 (42-356)	144 (42-356)	0 (0)
Homeostasis Model Assessment – Insulin Resistance Index	3.8 (0.4-18.6)	3.8 (0.4-18.6)	0 (0)	3.4 (0.1-24.2)	3.4 (0.1-24.2)	0 (0)
Homeostasis Model Assessment – Beta-cell Function (%)	65.4 (3.5-100)	65.4 (3.5-100)	0 (0)	41.6 (0-100)	41.6 (0-100)	0 (0)

Table B2: Comparison of values computed in integrity check to reference article Table 2

Characteristic	Pub: Autoantibody negative (n=46)	DSIC: Autoantibody negative (n=46)	Diff. (n=0)	Pub: Autoantibody positive (n=712)	DSIC: Autoantibody positive (n=712)	Diff. (n=0)
N (male/female)	46 (28/18)	46 (28/18)	0 (0)	712 (368/344)	712 (368/344)	0 (0)
Age at diagnosis (years)	12.3 (1.1-17.9)	12.3 (1.1-17.9)	0 (0)	9.8 (0.3-48.5)	9.8 (0.3-48.5)	0 (0)
Body mass index percentile (%)	86 (12-100)	86 (12-100)	0 (0)	77 (3-100)	77 (3-100)	0 (0)
pH at diagnosis	7.3 (7.0-7.4)	7.3 (7.0-7.4)	0 (0)	7.3 (6.7-12)	7.3 (6.7-12)	0 (0)
Beta-hydroxybutyrate (mmol/L) at diagnosis	2.7 (0.03-18)	2.7 (0.03-18)	0 (0)	4.2 (0.01-22)	4.2 (0.01-22)	0 (0)
Bicarbonate (mmol/L) at diagnosis	23 (4-31)	23 (4-31)	0 (0)	19 (2-91)	19 (2-91)	0 (0)
Glucose (mg/dL) at diagnosis	367 (118-716)	367 (118-716)	0 (0)	370 (63-1967)	370 (63-1967)	0 (0)
Hemoglobin A1c (%) at diagnosis	12.7 (6.4-17.4)	12.7 (6.4-17.4)	0 (0)	11.8 (1.6-20)	11.8 (1.6-20)	0 (0)
C-peptide levels (ng/mL) at diagnosis	0.7 (0.09-9)	0.7 (0.09-9)	0 (0)	0.4 (0.04-18)	0.4 (0.04-18)	0 (0)
Insulin autoantibodies levels (U/mL) at diagnosis	0.3 (0.3-0.3)	0.3 (0.3-0.3)	0 (0)	0.3 (0.3-151)	0.3 (0.3-151)	0 (0)
Islet cell antigen 512 autoantibodies levels (U/mL) at diagnosis	0.7 (0.7-0.7)	0.7 (0.7-0.7)	0 (0)	9 (0.7-51)	9 (0.7-51)	0 (0)
Glutamic acid decarboxylase autoantibodies levels (U/mL) at diagnosis	0.1 (0.1-0.1)	0.1 (0.1-0.1)	0 (0)	5.8 (0.1-109.5)	5.8 (0.1-109.5)	0 (0)

Attachment A: SAS Code

```
libname m1 "X:\NIDDK\niddk-
dr_studies6\RADIANT\private_created_data\DEID\m_001_hparikh_niddk_submission_v2";

proc contents data=m1.m_001_hparikh_niddk_04july2022_1;
run;

*CCHC Study A;

*Sex;
proc freq data=m1.m_001_hparikh_niddk_04july2022_1;
tables gender GAD65Ab_index;
run;

data test; set m1.m_001_hparikh_niddk_04july2022_1;
gad = (gad65AB_index * 1000);
run;

proc freq data=test;
tables gad;
run;

data test1; set test;
gad_cat = . ;
if gad < 50 then gad_cat = 0;
if gad > 50 then gad_cat = 1;
run;

proc freq data=test1;
tables gad_cat*gender/norow nopercent;
run;

*GAD65Ab levels;
proc means data=test1 n median min max;
var GAD65Ab_index;
class gad_cat;
run;

*Age at onset;
proc means data=test1 n median min max;
var diab_age;
class gad_cat;
run;

*Disease duration years;
proc means data=test1 n median min max;
```

```
var disease_duration;
class gad_cat;
run;

*BMI, waist, sys bp, dia bp, trigs, hdl, blood glucose, IR, Beta-cell;
proc means data=test1 n median min max;
var bmi1 waist corrsys1 corrdia1 trig hdlc mfbg homa_ir HOMA_beta;
class gad_cat;
run;

*TCHRNO-1 Study B;
proc freq data=m1.m_001_hparikh_niddk_04july2022_2;
tables NegativeAb;
run;

data studyb; set m1.m_001_hparikh_niddk_04july2022_2;
if negativeab = 1 OR negativeab = . then neg_cat = 1;
if negativeab = 0 then neg_cat = 0;
run;

*sex;
proc freq data=studyb;
tables Female*Neg_cat/missing norow;
run;

*Age at diagnosis, BMI, pH at diagnosis, beta-hydroxybutyrate, bicarb, glucose, hemoglobin a1c, c-
peptide, insulin autoantibodies, islet cell antigen 512, glutamic acid;
proc means data=studyb n median min max;
var dxage v1bmicentile ph bohb bicarb glucose dxa1c dxcpep dxiaa dxICA512 dxGAD65;
class neg_cat;
run;
```