

Dataset Integrity Check for Multidisciplinary Approach to the Study of Pelvic Pain – Twin Study (MAPP Twin)

Prepared by NIDDK-CR
November 7, 2024

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 Table 1 – Characteristics of individual twins on all key study variables	4
Table B: Comparison of values computed in integrity check to reference Table 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 Multidisciplinary Approach to the Study of Pelvic Pain – Twin Study (MAPP Twin) sought to understand how pain was felt in people with Interstitial Cystitis/Painful Bladder Syndrome. Twin pairs were followed for a year to collect additional data on the genetic and environmental influences on chronic pelvic pain.

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 MAPP Twin folder in the data package. For this replication, variables were taken from the “rev1_twin_baseline.sas7bdat” dataset.

4 Statistical Methods

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

5 Results

For the DCC provided Table 1, Characteristics of individual twins on all key study variables, 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 in Table 1. The results of the replication are within expected variation to the results in the DCC provided table.

6 Conclusions

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

Table A: Variables used to replicate Table 1 – Characteristics of individual twins on all key study variables

Table Variable	dataset.variable
Age	rev1_twin_baseline.age
Race	rev1_twin_baseline.race_cau rev1_twin_baseline.race_aa rev1_twin_baseline.race_asian rev1_twin_baseline.race_nativ_am rev1_twin_baseline.race_nativ_haw rev1_twin_baseline.race_other
Partner/Spouse	rev1_twin_baseline.partner_spouse
Education	rev1_twin_baseline.education
Employment	rev1_twin_baseline.emp_status
Income	rev1_twin_baseline.income
Childhood Death of Friend or Family	rev1_twin_baseline.ctes_q1
Childhood Parents Divorce/Separation	rev1_twin_baseline.ctes_q2
Childhood Traumatic Sexual Experience	rev1_twin_baseline.ctes_q3
Childhood Victim of Violence	rev1_twin_baseline.ctes_q4
Childhood Extremely Ill or Injured	rev1_twin_baseline.ctes_q5
Recent Death of Friend or Family	rev1_twin_baseline.rtes_q1
Recent Divorce or Separation	rev1_twin_baseline.rtes_q2
Recent Traumatic Sexual Experience	rev1_twin_baseline.rtes_q3
Recent Victim of Violence	rev1_twin_baseline.rtes_q4
Recent Extremely Ill or Injured	rev1_twin_baseline.rtes_q5
Recent Major Change in Work	rev1_twin_baseline.rtes_q6
HADS Total	rev1_twin_baselinehadsanxtot rev1_twin_baseline.hadsdeptot
SEARS Total	rev1_twin_baseline.sear_total_score
FSFI Total	rev1_twin_baseline.fsfi_total

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

Variable	Table 1 (n=98)	DSIC (n=98)	Diff. (n=0)
Age, Mean (SD)	36.7 (15.6)	36.7 (15.6)	0 (0)
Race			
Caucasian	40 (40.8)	74 (75.5)	34 (34.7)
Other	58 (59.2)	24 (24.5)	34 (34.7)
Partner/Spouse			
No	51 (52.0)	51 (52.0)	0 (0)
Yes	47 (48.0)	47 (48.0)	0 (0)
Education			
HS or GED	9 (9.2)	9 (9.2)	0 (0)
Some college	39 (39.8)	39 (39.8)	0 (0)
College/university	38 (38.8)	38 (38.8)	0 (0)
Post college/university	12 (12.2)	12 (12.2)	0 (0)
Employment			
Employed	72 (73.5)	72 (73.5)	0 (0)
Unemployed	13 (13.3)	13 (13.3)	0 (0)
Retired	9 (9.2)	9 (9.2)	0 (0)
Full-time homemaker	4 (4.1)	4 (4.1)	0 (0)
Income			
\$10,000 or less	14 (14.4)	14 (14.4)	0 (0)
\$10,001 to \$25,000	13 (13.4)	13 (13.4)	0 (0)
\$25,001 to \$50,000	20 (20.6)	20 (20.6)	0 (0)
\$50,001 to \$100,000	26 (26.8)	26 (26.8)	0 (0)
More than \$100,000	17 (17.5)	17 (17.5)	0 (0)
Prefer not to say	7 (7.2)	7 (7.2)	0 (0)
Childhood Death of Friend or Family			
No	55 (56.1)	55 (56.1)	0 (0)
Yes	43 (43.9)	43 (43.9)	0 (0)
Childhood Parents Divorce/Separation			
No	72 (73.5)	72 (73.5)	0 (0)
Yes	26 (26.5)	26 (26.5)	0 (0)
Childhood Traumatic Sexual Experience			
No	68 (70.8)	68 (70.8)	0 (0)
Yes	28 (29.2)	28 (29.2)	0 (0)
Childhood Victim of Violence			
No	78 (81.3)	78 (81.3)	0 (0)
Yes	18 (18.8)	18 (18.8)	0 (0)
Childhood Extremely Ill or Injured			
No	83 (85.6)	83 (85.6)	0 (0)
Yes	14 (14.4)	14 (14.4)	0 (0)
Recent Death of Friend or Family			
No	47 (48.0)	47 (48.0)	0 (0)
Yes	51 (52.0)	51 (52.0)	0 (0)

Variable	Table 1 (n=98)	DSIC (n=98)	Diff. (n=0)
Recent Divorce or Separation			
No	86 (89.6)	86 (89.6)	0 (0)
Yes	10 (10.4)	10 (10.4)	0 (0)
Recent Traumatic Sexual Experience			
No	94 (95.9)	94 (95.9)	0 (0)
Yes	4 (4.1)	4 (4.1)	0 (0)
Recent Victim of Violence			
No	93 (94.9)	93 (94.9)	0 (0)
Yes	5 (5.1)	5 (5.1)	0 (0)
Recent Extremely Ill or Injured			
No	87 (88.8)	87 (88.8)	0 (0)
Yes	11 (11.2)	11 (11.2)	0 (0)
Recent Major Change in Work			
No	51 (52.0)	51 (52.0)	0 (0)
Yes	47 (48.0)	47 (48.0)	0 (0)
HADS Total, Mean (Min-Max)	8.4 (0-29)	8.4 (0-29)	0 (0)
SEARS Total, Mean (SD)	70.2 (20.7)	70.2 (20.7)	0 (0)
FSFI Total, Mean (SD)	27.5 (5.4)	27.5 (5.4)	0 (0)

Attachment A: SAS Code

```
libname twin "Z:\NIDDK\niddk-  
dr_studies6\MAPP_I_EPS_Twin_Study\private_created_data\DEID\NIDDK Repository Deposit  
22nov2023";  
  
/*****/  
/* MAPP TWIN DSIC */  
/*****/  
proc contents data=twin.final_twin_longitudinal;  
run;  
  
proc contents data=twin.rev1_twin_baseline;  
run;  
  
*replicating table provided by DCC;  
  
*age;  
proc means data=twin.rev1_twin_baseline mean std;  
var age;  
run;  
  
*race;  
proc freq data=twin.rev1_twin_baseline;  
tables racecat race_cau race_3lev race_aa race_asian race_nativ_am race_nativ_haw race_other;  
run;  
  
data one; set twin.rev1_twin_baseline;  
race_rep = race_cau;  
if race_aa = 1 OR race_asian = 1 OR race_nativ_am = 1 OR race_nativ_haw = 1 OR race_other = 1 then  
race_rep = 0;  
run;  
  
proc freq data=one;  
tables race_rep;  
run;  
  
*partner/spouse, education, employment status, income, death of friend/family, parents divorce,  
traumatic sexual experience, victim of violence, extremely ill, recent death of friend/family, recent  
divorce, recent traumatic sexual experience, recent victim of violence, recent extremely ill or injured,  
recent major change in work;  
  
proc freq data=twin.rev1_twin_baseline;  
tables partner_spouse education emp_status income ctes_q1 ctes_q2 ctes_q3 ctes_q4 ctes_q5 rtes_q1  
rtes_q2 rtes_q3 rtes_q4 rtes_q5 rtes_q6;  
run;
```

```
*HADS total, SEARS Total, FSFI Total;  
data two; set one;  
hads_tot = hadsanxtot + hadsdeptot;  
run;  
  
proc means data=two mean std min max;  
var hads_tot sear_total_score fsfi_total;  
run;
```