

Selecting Subjects for PNC Template

BBL, et al

October 10, 2015

Selecting Subjects for PNC Template

Goal here is to select a relatively matched sample (n=120) for the PNC template.

Data pulled from the July 30 PNC data release on Galton:

/home/analysis/redcap_data/201507//n1601_go1_datarel_073015.csv

Strategy is as follows:

1. Remove subjects who failed visual QA (could be debated) & health problems
2. Divide into 5 age bins (quantiles)
3. In each age bin, select 12 subjects who are healthy (no psychopathology/meds) and 12 who are not healthy (any psychopathology, meds ok)
4. For each of those 12 subjects , evenly matched on sex*race

=====

Now sanity checking to make sure this worked as plannedhere divided by age bins

female	male
12	12
12	12
12	12
12	12
12	12

caucasian	notCaucaisian
12	12
12	12
12	12
12	12
12	12

healthy	notHealthy
12	12
12	12
12	12
12	12
12	12

	healthy	notHealthy
female	30	30
male	30	30

	healthy	notHealthy
caucasian	30	30
notCaucasian	30	30

	female	male
caucasian	30	30
notCaucasian	30	30

=====

Check to see if age is similar across the different divisions. . . .

	ageMeanSex
female	15.34722
male	15.36111

	ageMeanHealth
healthy	15.30972
notHealthy	15.39861

	ageMeanRace
caucasian	15.36250
notCaucasian	15.34583

=====

Finally check that age is reasonably balanced over the range

