Childhood trauma scores showed no meaningful link to amygdala-prefrontal coupling in 187 adults

A preregistered equivalence test of childhood trauma against resting amygdala-vmPFC connectivity on open data from the Transdiagnostic Connectome Project

Zain Dana Harper

Version 1.0, 9 November 2026. Finding F50 in the findings ledger. Honest null (tier D).

Abstract

A widely discussed idea holds that childhood maltreatment weakens the coupling between the amygdala, which responds to threat, and the ventromedial prefrontal cortex, which helps regulate that response. We tested one resting-state form of the idea with the analysis fixed before any questionnaire value or image was seen. The data are the Transdiagnostic Connectome Project on OpenNeuro (CC0): 241 adults with and without psychiatric diagnoses, preprocessed in full, of whom 187 met the motion and questionnaire rules. Childhood trauma scores were unrelated to amygdala-vmPFC resting connectivity in 187 adults: partial r -0.006, 95% interval -0.152 to 0.140. The 90% interval, -0.129 to 0.117, lies inside -0.20 to 0.20, so effects of 0.20 or larger in size are excluded at the preregistered bound. A secondary region gave r -0.025. Shuffling the trauma scores 10,000 times placed the observed value at the 47th percentile. Fifteen exploratory sensitivity analyses moved the primary r between -0.081 and 0.081, and none reached an association. The emotional-neglect and physical-abuse subscale analyses, r 0.081 and -0.081, are inconclusive because their 90% intervals touch 0.20. Power to detect a correlation of 0.20 was 0.77, so true effects between 0.10 and 0.20 remain open.

1 What this means for a reader

Small studies have reported that people who were maltreated as children show weaker resting coupling between the amygdala and the prefrontal cortex. In 187 adults, with the test fixed in advance, we found no association of a size that would matter, and a much larger study found the same in 1,325 adults. Treat a strong resting-state version of the claim with caution. A weak effect is still possible, two trauma subscales gave inconclusive results, and other kinds of brain measurement were not tested.

2 Introduction

The amygdala and the ventromedial prefrontal cortex (vmPFC) form a circuit often linked to threat and its regulation. A common hypothesis is that maltreatment in childhood changes this circuit, and that the change shows up as weaker resting-state functional connectivity between the two.

Birn and colleagues [2] studied 27 male combat veterans, 17 of them with PTSD, and modelled amygdala and hippocampus connectivity from childhood maltreatment, combat exposure and PTSD symptoms together. Their abstract reports that scores on the Childhood Trauma Questionnaire (CTQ) predicted lower amygdala connectivity to the vmPFC.

Goltermann and colleagues [3] tested the same family of claims in 1,325 adults (624 with major depression, 701 healthy controls), with the CTQ and regions chosen from earlier reports, among them the amygdala. Maltreatment was not associated with the connectivity of any region of interest in the full sample. Within the depressed group, maltreatment went with greater, not weaker, connectivity between the amygdala and dorsolateral frontal areas. The authors conclude that most earlier connectivity correlates of maltreatment did not

replicate.

Our study adds a smaller, independent test on open data with one feature the larger study did not state in its abstract: a preregistered smallest effect of interest and an equivalence test (two one-sided tests) in the form Lakens [4] describes for correlations. An equivalence test can support "no effect of a size that matters", which a non-significant result alone cannot.

3 Preregistration and design

The preregistration was frozen on 9 October 2026, before any phenotype value or image from the dataset was downloaded or viewed. Only file names, the dataset description and the CTQ definitions file had been read. Three choices were made at that point: an anatomical vmPFC, the full sample with no diagnostic filter, and a smallest effect of interest of 0.20 in size, about 4 percent shared variance.

The decision rule had three outcomes, fixed in advance:

- Association: the 95 percent interval excludes 0 and, by a deviation logged before any data were processed, the Holm correction-adjusted p across the two registered tests is below 0.05.
- No meaningful association: the 90 percent interval lies wholly inside -0.20 to 0.20.
- Inconclusive: neither.

The preregistration also stated that effects between 0.10 and 0.20 could be missed at the expected sample size, and that any write-up must say so.

Six deviations are logged in the preregistration, each dated with the point at which it was decided:

- 1 to 3, on 9 October, before any participant was processed: keep aCompCor components up to five per mask when fewer are available; require the Holm-adjusted p for an association call; use each session's acquired field maps for distortion correction, with a field-map-free fallback where they do not match.
- 4, on 10 October: a fallback preprocessing path if a software licence was not in place. It was not used; every participant went through one pipeline.
- 5 and its correction 5a, on 10 October, after one participant's motion check and before any trauma or connectivity value was opened: the data are multiband with a 0.8 second repetition time, and breathing showed up as apparent motion. The motion rule became a 0.1 Hz low-pass filter on the motion parameters, with volumes censored above 0.1 mm, following Gratton and colleagues [5]. That rule was validated on single-band data, so here it runs outside its validated setting. Deviation 5 first specified a band-stop filter that did not remove the breathing signal; 5a replaced it.
- 6, on 29 October, after preprocessing and before any trauma-connectivity value was opened, decided from counts of the questionnaire coding only: the dataset codes missing CTQ answers as 88, which the registered code did not list, so 88 became a missing code.

4 Data and provenance

The data are the Transdiagnostic Connectome Project, OpenNeuro dataset ds005237, snapshot 1.1.3 [1], released under CC0, a public-domain dedication. It is a transdiagnostic adult sample with up to four resting-state runs per participant. This paper reports group statistics only; per-participant tables were not published and no individual record appears here.

All 241 participants were preprocessed with fMRIPrep 25.2.6, with no processing errors. Of these, 230 had four runs, 6 had three and 5 had two. Distortion correction used acquired field maps for 198 participants and the field-map-free method for 43.

The analysed sample of 187 had a mean age of 33.4 years (SD 13.3). It included 112 women, 73 men and 2 people who reported another sex; 122 were scanned at one site and 65 at the other; 114 were in the patient

group and 73 in the general-population group. The mean CTQ total was 43.0 (SD 17.3) on a scale of 25 to 125.

5 Methods

Regions

Regions came from the Harvard-Oxford atlas at the 50 percent threshold in standard space: left and right amygdala combined; "Frontal Medial Cortex" for the primary vmPFC target; and "Subcallosal Cortex", the label closest to Brodmann area 25, as the pre-declared secondary target.

Denoising and motion

Each run was cleaned by regression of 24 motion parameters, aCompCor (up to five white-matter and five cerebrospinal-fluid components) and cosine drift terms, with no global signal regression in the primary analysis. Framewise displacement was computed from low-pass filtered motion parameters, and censoring removed volumes above 0.1 mm together with uncensored segments shorter than five volumes. A run was unusable with more than 20 percent of volumes censored, and a participant needed at least two usable runs.

Connectivity

For each usable run j, r_j is the Pearson correlation of the mean denoised time series of the two regions. It is converted with the Fisher z transform,

$$z_j = \frac{1}{2} \ln \frac{1 + r_j}{1 - r_j} \tag{1}$$

and averaged over the participant's R usable runs:

$$Z = \frac{1}{R} \sum_{j} z_j \tag{2}$$

The test

The exposure is the CTQ total, the sum of the five subscales. The statistic is the partial correlation r_p between CTQ total and Z, adjusting for age, sex, site, mean filtered framewise displacement and the number of retained volumes, which the result file records as $k = 6$ model terms. Intervals use the Fisher transform of r_p with standard error

$$SE = \frac{1}{\sqrt{n - 3 - k}} \tag{3}$$

and run from $z_p - c \cdot SE$ to $z_p + c \cdot SE$, transformed back to the correlation scale, with c 1.96 for the 95 percent interval and 1.645 for the 90 percent interval. With $n = 187$ this gives the intervals in Table 1. Statistical power uses the same normal approximation.

6 Controls and planted checks

The control shuffled CTQ totals across participants 10,000 times (seed 20261029) and recomputed the partial correlation each time, giving a null distribution for both the primary and the secondary test. It shows where the observed value sits among values that carry no link by construction.

There is no positive control. No known effect was planted in the data, and no planted check shows that this pipeline would find a connectivity effect of size 0.20 when one is present. The equivalence call rests on the interval width, which depends on the sample size and not on the pipeline's sensitivity. A pipeline that destroyed real signal would also produce a narrow interval around zero. This is the main weakness of the design.

7 Results

Of 241 participants, 17 were excluded for having fewer than two usable runs after motion censoring, 16 had no CTQ record and 21 were missing at least one subscale, leaving 187.

Childhood trauma scores were unrelated to amygdala-vmPFC resting connectivity in 187 adults: partial r -0.006, 95% interval -0.152 to 0.140. The 90% interval, -0.129 to 0.117, lies inside -0.20 to 0.20, so effects of 0.20 or larger in size are excluded at the preregistered bound. The decision is "no meaningful association". The secondary region gave the same decision (Table 1).

Table 1. The two registered tests. Holm-adjusted p is across the two tests.

Test	n	Partial r	95% interval	90% interval	Holm p
Primary, frontal medial cortex	187	-0.006	-0.152 to 0.140	-0.129 to 0.117	1.0
Secondary, subcallosal cortex	187	-0.025	-0.171 to 0.121	-0.147 to 0.098	1.0

The unadjusted p values were 0.94 and 0.74. Under the shuffle control the primary value sat at the 47th percentile of 10,000 shuffles (two-sided permutation p 0.93) and the secondary at the 37th (p 0.73), as Figure 1 shows. The shuffled primary values ran from -0.119 to 0.124 between the 5th and 95th percentiles.

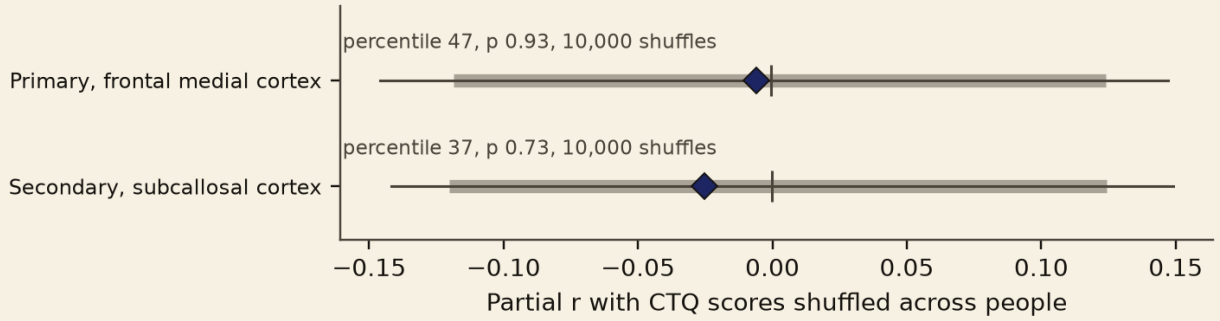


Figure 1. Observed partial r (diamonds) against the 5th to 95th percentile (bar), 2.5th to 97.5th percentile (line) and median (tick) of 10,000 shuffles of CTQ scores across participants. Drawn from the study's group-level result file.

Power to detect a correlation of 0.20 was 0.77, so true effects between 0.10 and 0.20 remain open. At the Holm-adjusted threshold the power was 0.68. The smallest correlation detectable with 80 percent power was 0.207. The chance that the equivalence test succeeds when the true correlation is zero was 0.71.

Fifteen exploratory sensitivity analyses moved the primary r between -0.081 and 0.081, and none reached an association. Figure 2 draws all of them with the two registered tests; Table 2 in Appendix A lists them. Thirteen of the fifteen met the equivalence rule on the primary region. The emotional-neglect and physical-abuse subscale analyses, r 0.081 and -0.081, are inconclusive because their 90% intervals touch 0.20. Their 90% intervals run from -0.041 to 0.201 and from -0.202 to 0.042. Each of these single subscales neither shows an association nor excludes one of 0.20.

The same sensitivity analyses were also run on the secondary region. Two of them were inconclusive there: the band-stop motion rule (n 131, r -0.110, 90% interval -0.253 to 0.039) and the sexual-abuse subscale (r -0.098, 90% interval -0.218 to 0.025). None reached an association.

8 Discussion

The registered result is an honest null with a stated bound. Within this sample and this measure, the correlation between CTQ total and amygdala-vmPFC resting connectivity is smaller than 0.20 in size. That is consistent with

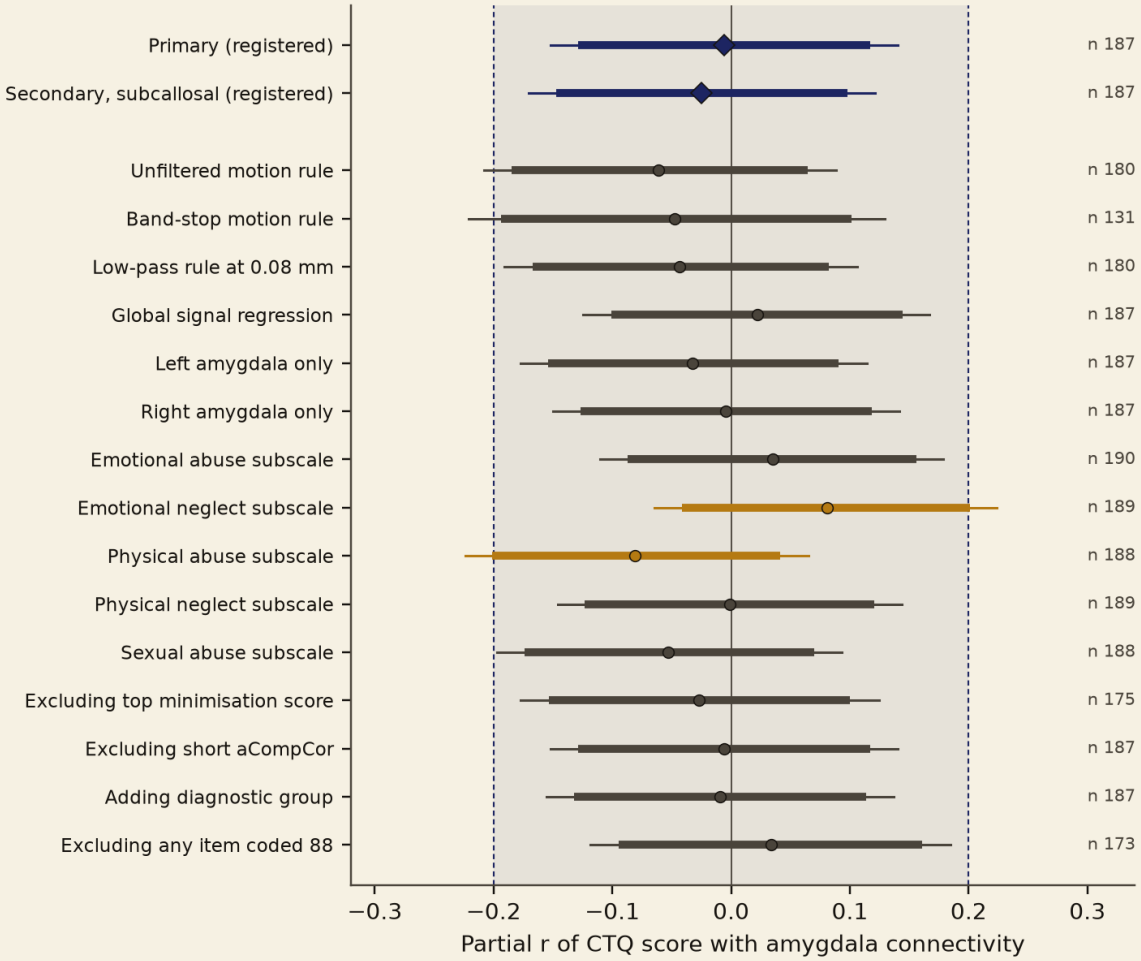


Figure 2. Partial r with 90% intervals (thick) and 95% intervals (thin) for the two registered tests (top) and the fifteen exploratory sensitivity analyses on the primary region. The shaded band is the preregistered equivalence region, -0.20 to 0.20. Ochre marks an inconclusive result. Drawn from the study's group-level result files.

Goltermann and colleagues [3], whose abstract reports no maltreatment association with the connectivity of its regions of interest in 1,325 adults.

It does not contradict Birn and colleagues [2] directly. Their sample was 27 male veterans, most with PTSD, and their model held combat exposure and PTSD symptoms fixed alongside maltreatment; ours is a mixed transdiagnostic sample, mostly women and men, with demographic and motion covariates. An effect confined to a population like theirs could exist and be diluted here. The difference in design is the main reason the two results can both stand.

The inconclusive subscale results deserve plain reporting. Emotional neglect and physical abuse gave correlations of opposite sign, each with a 90 percent interval that crosses 0.20 by a small margin. They come from fifteen exploratory analyses with no correction, so one or two such edges are expected by chance. They do not support an association, and they also do not let us extend the equivalence claim to each subscale.

The deeper limit is the missing positive control. A null from a pipeline never shown to detect a known effect is weaker than a null from one that has. Adding one, for example a well-replicated connectivity difference with age or between homologous regions, is the next step before this pipeline is used for further trauma questions.

9 Limitations

- One resting-state measure, with anatomical masks, in one transdiagnostic adult sample. Task-evoked regulation and other connectivity methods were not tested.
- The CTQ is retrospective self-report.
- The primary motion rule uses a low-pass filter on framewise displacement outside the single-band data it was validated on.
- There is no positive control, so the pipeline's ability to find a known effect is not shown.
- A true correlation between 0.10 and 0.20 in size could give this result; power at 0.20 was 0.77.
- Two subscale analyses on the primary region and two analyses on the secondary region were inconclusive.
- The analysis was run by our own code and was not blind.

10 Verified

Verified: on 9 November 2026 the flow counts (241, 17, 16, 21, 187), the primary and secondary correlations with their 95% and 90% intervals, the unadjusted and Holm p values, the permutation percentiles, p values and quantiles, the power values, all fifteen sensitivity analyses on both regions, and the group descriptions of the sample were re-read from the study's two group-level result files. They match the findings ledger, with the dated correction of 9 November for the two subscale analyses. The interval formula was read from the analysis script and agrees with the reported primary interval. The frozen preregistration hash recorded in both result files matches the hash in the study's hash register. The current preregistration file has a different SHA-256 because the deviation log was appended after freezing; the frozen body was not re-hashed for this paper. Both figures were drawn by a script that reads only the two result files. The study was not re-run. The literature was found by searches run on 9 November 2026 and read at the level of abstracts and publisher or repository pages; the full texts were not read. For Gratton and colleagues [5], the preregistration records that its methods section was read on 10 October 2026.

11 Corrections

- 10 October 2026: deviation 5 contained an error; its supporting numbers came from a low-pass check while its rule named a band-stop filter. Deviation 5a corrected it before any trauma or connectivity value was opened.
- 9 November 2026: the findings ledger said of the fifteen sensitivity analyses only that none reached an association. The emotional-neglect and physical-abuse subscale analyses (r 0.081 and -0.081) are inconclusive, their 90% intervals touching 0.20. This paper reports them that way, and also reports the two inconclusive secondary-region analyses.
- 9 November 2026: first published version.

12 Reproduction

- Data: download OpenNeuro ds005237, snapshot 1.1.3 (CC0).
- Preprocess with fMRIPrep 25.2.6 into standard space. Denoise with 24 motion parameters, aCompCor and cosine terms. Censor volumes with low-pass filtered framewise displacement above 0.1 mm and segments shorter than five volumes; drop runs with more than 20 percent censored; keep participants with at least two usable runs.
- Extract mean time series from the Harvard-Oxford amygdala and Frontal Medial Cortex masks at the 50 percent threshold, correlate, apply eq. 1 and average with eq. 2.
- Sum the five CTQ subscales, treating 88 as missing. Compute the partial correlation with the six covariate

terms and its intervals with eq. 3, then apply the decision rule at 0.20.

- Shuffle CTQ totals 10,000 times with seed 20261029 for the null.
- Preregistration, frozen body SHA-256: 4d4681b9f44188b37e60fd9c53d94a85e01cd1529f4d2be7228a42a3ff36f753. Current file with the deviation log: f185c72a449a8649dda7177a3c93a4107da54bc243b4001edd439ad8cceadb82. This study ran before our study runner was used for it, so it has no plan hash or run id; the result files carry the frozen hash and the SHA-256 of every phenotype input.
- Cost: the analysis session used 50.4 kJ (14.0 Wh), measured at the machine. The 19-day preprocessing batch was not metered.

13 Appendix A. Sensitivity analyses on the primary region

All fifteen were declared as exploratory, most of them in the preregistration body and the rest in the deviation log. The decision column applies the registered rule to each.

Table 2. Exploratory sensitivity analyses, primary region (amygdala to Frontal Medial Cortex).

Analysis	n	Partial r	90% interval	Decision
Unfiltered motion rule	180	-0.061	-0.185 to 0.065	no meaningful
Band-stop motion rule	131	-0.047	-0.194 to 0.101	no meaningful
Low-pass rule at 0.08 mm	180	-0.043	-0.167 to 0.082	no meaningful
Global signal regression	187	0.022	-0.101 to 0.144	no meaningful
Left amygdala only	187	-0.032	-0.154 to 0.091	no meaningful
Right amygdala only	187	-0.004	-0.127 to 0.118	no meaningful
Emotional abuse subscale	190	0.035	-0.087 to 0.156	no meaningful
Emotional neglect subscale	189	0.081	-0.041 to 0.201	inconclusive
Physical abuse subscale	188	-0.081	-0.202 to 0.042	inconclusive
Physical neglect subscale	189	-0.001	-0.123 to 0.121	no meaningful
Sexual abuse subscale	188	-0.053	-0.174 to 0.070	no meaningful
Excluding top minimisation score	175	-0.027	-0.153 to 0.100	no meaningful
Excluding short aCompCor	187	-0.006	-0.129 to 0.117	no meaningful
Adding diagnostic group	187	-0.009	-0.132 to 0.114	no meaningful
Excluding any clinical item coded 88	173	0.034	-0.095 to 0.161	no meaningful

14 Glossary

Each term is defined where it is first used in the text. The same definitions, with the field each comes from and what each excludes, are collected here.

- **amygdala**. A pair of almond-shaped structures deep in the temporal lobes, one per hemisphere, involved in detecting threat and in emotional learning. Field: neuroanatomy. Excludes its separate subnuclei, which the anatomical masks here do not distinguish.
- **ventromedial prefrontal cortex (vmPFC)**. The lower middle part of the frontal lobe, linked to the regulation of emotional responses; here defined by the Harvard-Oxford label Frontal Medial Cortex. Field: neuroanatomy. Excludes functional definitions from a task, and the dorsal and lateral prefrontal cortex.
- **resting-state functional connectivity**. The correlation over time between the blood-oxygen signals of two brain regions while a person lies in the scanner with no task. Field: neuroimaging. Excludes anatomical connections, activity evoked by a task, and the direction of any influence between regions.
- **Childhood Trauma Questionnaire (CTQ)**. A 28-item adult self-report questionnaire about childhood experiences, scored as five subscales (emotional abuse, physical abuse, sexual abuse, emotional neglect, physical neglect) of 5 to 25 each and a total of 25 to 125, plus minimisation items that flag denial. Field: clinical psychology. Excludes records of maltreatment from the time it happened, and clinical diagnosis.
- **smallest effect of interest**. The size of effect, fixed before the data are seen, below which an effect is treated

as too small to matter for the decision at hand; here a correlation of 0.20 in size. Field: statistics (equivalence testing). Excludes a claim that smaller effects are zero.

- **equivalence test (two one-sided tests)**. A test that supports the absence of a meaningful effect when the 90 percent interval of the estimate lies wholly inside the band set by the smallest effect of interest, which is the same as rejecting both edges of the band at the 5 percent level. Field: statistics. Excludes a test of whether the effect is exactly zero.
- **preregistration**. A dated written plan for a study, fixed and hash-locked before any result is seen: the question, the data, the statistic, the controls, the forecasts and the kill conditions. Field: research method (metascience). Excludes a plan revised after looking at the outcome, and a protocol written only for the record after the run.
- **deviation**. A dated, reasoned change to a preregistered plan, written into the plan's log, with the point in the study at which it was decided. Field: research method (metascience). Excludes a silent change to the analysis.
- **Holm correction**. A step-down adjustment of p values across a family of tests that keeps the chance of any false positive in the family at the stated level. Field: statistics (multiple testing). Excludes adjustment for analyses outside the declared family, such as the exploratory sensitivity checks.
- **fMRIPrep**. An open preprocessing pipeline for functional MRI that corrects head motion and distortion, aligns images to a standard brain space and writes nuisance signals for later cleaning. Field: neuroimaging software. Excludes the denoising and statistics that follow, which were run by our own code.
- **Harvard-Oxford atlas**. A probabilistic atlas of cortical and subcortical regions in standard brain space; at a 50 percent threshold a voxel belongs to a region when at least half of the reference brains placed it there. Field: neuroimaging. Excludes regions drawn from each participant's own anatomy or from functional data.
- **aCompCor**. Anatomical component correction: the main components of the signal in white matter and cerebrospinal fluid, where little neural signal is expected, are regressed out as noise. Field: neuroimaging (denoising). Excludes global signal regression, which removes the mean of the whole brain.
- **global signal regression**. Removing the average signal of the whole brain from every region before connectivity is computed. Field: neuroimaging (denoising). Excludes aCompCor and motion regression, which remove other nuisance signals.
- **framewise displacement**. A summary of how far the head moved between two consecutive brain volumes, in millimetres, from the six rigid-body motion parameters. Field: neuroimaging (motion quality control). Excludes motion within a single volume.
- **censoring**. Dropping single brain volumes whose head motion exceeds a threshold before connectivity is computed. Field: neuroimaging (motion quality control). Excludes dropping whole runs or whole participants, which follow separate rules.
- **Fisher z**. The transform z = artanh(r) of a correlation, which makes its sampling spread close to normal so that correlations can be averaged and given intervals. Field: statistics. Excludes any change to the sign or ranking of correlations.
- **partial correlation**. The correlation between two quantities after the linear effect of a set of covariates has been removed from both. Field: statistics. Excludes adjustment for covariates that were not measured, and any statement about cause.
- **statistical power**. The chance that a test detects an effect of a stated size, given the sample size and the threshold for significance. Field: statistics. Excludes the chance that a detected effect is real.
- **control**. A version of the data with the structure under test removed (for example shuffled or pooled), scored the same way as the real data, so the real score has something to beat. Field: statistics and experimental design. Excludes a comparison with a different dataset or with a published number.
- **null distribution**. The spread of scores a control gives over many random draws; its 95th percentile is the

bar a real score must clear. Field: statistics. Excludes a theoretical distribution assumed rather than drawn from the data.

- **positive control**. A test of a known effect run through the same pipeline, to show that the pipeline can find an effect of the expected kind. Field: experimental design. Excludes shuffled controls, which show what no effect looks like.
- **planted check**. Data built so that the right answer is known, run through the same analysis before the real data, to show that the analysis can see the effect when it is there. Field: research method; called a positive control in experimental science. Excludes any check made on the real data, which cannot show what the analysis misses.
- **honest null**. A working test that found no effect beyond a stated bound, a forecast that failed, or an error of our own found and fixed, reported with the same weight as a positive result. Field: research method (our usage). Excludes a void run, and a result left unpublished because it was negative.
- **findings ledger**. The append-only, dated list of every result we have recorded, numbered F1, F2 and on; corrections are added as new dated lines and old entries are never rewritten. Field: our records. Excludes a curated list of successes.

15 References

[1] Transdiagnostic Connectome Project. OpenNeuro dataset ds005237, snapshot 1.1.3. CC0. doi:10.18112/openneuro.ds005237.v1.1.3

[2] Birn, R. M., Patriat, R., Phillips, M. L., Germain, A. and Herringa, R. J. (2014). Childhood maltreatment and combat post-traumatic stress differentially predict fear-related fronto-subcortical connectivity. Depression and Anxiety 31, 880-892. doi:10.1002/da.22291

[3] Goltermann, J., Winter, N. R., Meinert, S., Sindermann, L., Lemke, H., Leehr, E. J. and others (2022). Resting-state functional connectivity patterns associated with childhood maltreatment in a large bicentric cohort of adults with and without major depression. Psychological Medicine, online 27 June 2022. doi:10.1017/S0033291722001623

[4] Lakens, D. (2017). Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. Social Psychological and Personality Science 8, 355-362. doi:10.1177/1948550617697177

[5] Gratton, C., Dworetsky, A., Coalson, R. S., Adeyemo, B., Laumann, T. O., Wig, G. S. and others (2020). Removal of high frequency contamination from motion estimates in single-band fMRI saves data without biasing functional connectivity. NeuroImage 217, 116866. doi:10.1016/j.neuroimage.2020.116866