# ABCD Transparency and Reproducibility

This is the form that will be filled out for each manuscript. Please view the reference guide if you want to look at some examples: Reference Guide.

**Video tutorial** can be found here: https://drive.google.com/file/d/1IzeKJnA414T3aBHXC-UO0WDcaOTPObZh/view?usp=sharing

Email me (lopdanie@ohsu.edu) or Rebekah Huber (Rebekah.Huber@utah.edu) if you have any questions.

* Indicates required question

1. Publication Title *

   Please refer to the spreadsheet and copy+paste the publication citation. For example, " Associations between 24 hour movement behaviours and global cognition in US children: a cross-sectional observational study"

2. Paper Number *

   Please refer to the spreadsheet and enter the corresponding "paper number." The paper number for the example above is 1001.

3. Name *

   Write your name so we know who completed the form.

4. Did the paper conduct any statistical analyses? *

Here we want to know if the paper analyzed ABCD data. Some ABCD papers are purely methods papers (e.g., how the MRI data was processed etc). These types of papers will not be included in the current project. Make sure to still submit the form.

*Mark only one oval.*

- Yes *Skip to question 5*
- No

## Analysis-level reproducibility

5. 1. Was the analysis code/scripts shared in the manuscript? For example, is there a digital object identifier (DOI) or URL link that takes you directly to the code on Github/ Open Science Framework, or some other site? *

This can be found in the methods section. Authors will typically state something along the lines of "analysis script used for this study can be found at _____." The URL will typically be a Github, Open Science Foundation, Zenodo, or Figshare link. Some authors might specify that it is in the supplementary materials. Supplementary materials are typically linked within the journal article link. **Their code/scripts should be showing their analysis using ABCD data.**

*Mark only one oval.*

- Yes
- No *Skip to question 7*

## 1b

6. 1b. For manuscripts with shared code, is the url still active and working (i.e., is it archived in a long-term accessible location on the web)?

If a URL was provided, check to see if it still works.

*Mark only one oval.*

- Yes
- No

Question 2

7. 2. Is the software used for the analysis included in the text (e.g., R Studio, Matlab)? *

This one can be found in the methods section. They will state the software used to do the statistical analysis. Most common ones are R, R Studio, Matlab, Mplus, Python, SAS, SPSS, STATA, JASP.

*Mark only one oval.*

- ◯ Yes
- ◯ No *Skip to question 9*

2b

8. 2b. Is the software version included in the text (e.g., R Studio version 3.1)?

Related to question #2. The version number is typically stated right after the software used.

*Mark only one oval.*

- ◯ Yes
- ◯ No

3-5

9. 3. Was there mention of which ABCD version number was used in the analyses? * (e.g., ABCD version 4.0).

The currently available ABCD versions are 1.0, 1.1. 2.0, 2.0.1, 3.0, 4.0, 5.0. Earlier papers stated the release was in the first annual curated dataset. This would qualify as version 1.0.

*Mark only one oval.*

- ◯ Yes
- ◯ No

10. 4. Did the authors include the ABCD Study data release Data Object Identifier (DOI)? *

This is typically included towards the end of the text.

*Mark only one oval.*

- Yes
- No

11. 5. Did the authors state the ABCD Study visits included in their analyses? (e.g., baseline, only ABCD year-2 visit) *

Authors should state the visits included in their analysis.

*Mark only one oval.*

- Yes
- No

## Analytic Sample

12. 6. Was there a table showing descriptive statistics of the sample included in the analysis (e.g., age, sex, race & ethnicity)? *

The descriptive statistics are typically included in Table 1 or Table 2. Some authors include them in their supplementary materials. Some authors do not put them in table format, and instead describe them in the methods or results section. Here, we're looking more for a detailed description of the sample included in the analyses (e.g., descriptive statistics for age, race/ethnicity, marital status, household income etc.), and not just a brief written description (e.g., the mean age of our sample was 10.8 years old).

*Mark only one oval.*

- Yes
- No

13. 7. Did the paper specify the number of subjects included in the data analysis after exclusions? (i.e., final sample size of their analytic model). *

This one relates to the final number of subjects included in the analysis. Did they mention the sample size included in their analysis? Even though ABCD had 11,000+ participants, it's very rare to see an analysis that uses more than ~11,000. The sample size varies a lot depending on the measures used and the study visit. There are lots of reasons for this varying sample size. First, many participants (>10%) failed the imaging data quality control. Second, many researchers use listwise deletion when running statistical models. Listwise deletion removes a participant if they have ANY missing data. Third, attrition due to withdrawal or loss to follow-up will lead to smaller sample sizes.

*Mark only one oval.*

- ◯ Yes
- ◯ No

14. 8. Did the paper specify how many participants were excluded due to exclusion criteria (e.g., excluded due to missing covariate data)? *

Was there any mention of approximately how many were removed from the final analysis due to missing data or other exclusion criteria? We don't need the exact number removed. **We're just trying to determine whether they acknowledged that their final sample size was reduced.** If no participants were excluded (e.g., they only report the imputed results) then select the final option.

*Mark only one oval.*

- ◯ Yes
- ◯ No
- ◯ No participants were excluded

15. 9. Was there a description of the sample excluded from the analysis? *

This one is asking for something like Table 1 or a description in the text of those that were excluded from the analysis. Papers that describe the excluded sample typically put the table in the supplementary materials or the description in the Results section. We just want to know if there was any potential for bias due to the missing data (e.g., bias due to differential patterns of missingness). If no participants were excluded (e.g., they only report the imputed results) then select the final option.

*Mark only one oval.*

- ◯ Yes
- ◯ No
- ◯ No participants were excluded

16. 10. Did the authors report the percentage/proportion of missing data? *

We just want to know if there was some numerical discussion in the manuscript about missing data for individual variables. The most common variable with missing data in the ABCD Study is household income. Select “yes” if this was reported for any variable in the article even if not reported for every variable. Select "no" if the article did not explicitly state and did not include a way of determining the level of missing data for individual variables (e.g., in the text or in a table).

*Mark only one oval.*

- ◯ Yes
- ◯ No

17. 11. Was there any mention of the missing data mechanism (e.g., missing at random, missing not at random)? *

Some authors will examine the missing data mechanism to determine the nature of missingness. This can dictate the appropriate method to use for dealing with missing data (e.g., listwise deletion, multiple imputation).

*Mark only one oval.*

- ◯ Yes
- ◯ No

18. 12. Did the authors use data imputation methods to account for missing data (e.g., multiple imputation, full information maximum likelihood)? *

Typically included in the Methods section. We just want to know if they used any imputation methods. There are a few, but the most common one is multiple imputation by chained equations (MICE).

*Mark only one oval.*

- ◯ Yes
- ◯ No

## Methods and Results

19. 13. Was there inclusion of race and/or ethnicity in the statistical model? *

Discussion of variables included in the statistical model is typically in the Methods section. Race and/or ethnicity could be used in many different ways (e.g., as a control variable, as a primary exposure of interest).

*Mark only one oval.*

- ◯ Yes
- ◯ No *Skip to question 21*

## 13b

20. 13b. Was there an explanation of why race and/or ethnicity was included in the statistical model? *

We're looking for some sort of explanation that cites relevant literature or uses other criteria for including race and/or ethnicity (e.g., race and/or ethnicity was included for adjustment after review of the relevant literature, after using causal diagrams to examine potential confounding variables in the association between x and y, as the primary exposure of interest). We want to make sure there was meaningful inclusion of race and/or ethnicity.

*Mark only one oval.*

- ◯ Yes
- ◯ No

14

21. 14. Was there inclusion of household income in the statistical model? *

*Mark only one oval.*

- ◯ Yes
- ◯ No *Skip to question 23*

14b

22. 14b. Was there an explanation of why household income was included in the statistical model? *

We're looking for some sort of explanation that cites relevant literature or uses other criteria for including household income (e.g., household income was included for adjustment after review of the relevant literature, after using causal diagrams to examine potential confounding variables in the association between x and y). We want to make sure there was meaningful inclusion of household income.

*Mark only one oval.*

- ◯ Yes
- ◯ No

15

23. 15. Was there inclusion of participant sex in the statistical model? *

*Mark only one oval.*

- Yes
- No *Skip to question 25*

15b

24. 15b. Was there an explanation of why sex was included in the statistical model? *

We're looking for some sort of explanation that cites relevant literature or uses other criteria for including sex (e.g., sex was included for adjustment after review of the relevant literature, after using causal diagrams to examine potential confounding variables in the association between x and y). We want to make sure there was meaningful inclusion of sex.

*Mark only one oval.*

- Yes
- No

16

25. 16. Was there inclusion of participant age in the statistical model? *

*Mark only one oval.*

- Yes
- No *Skip to question 27*

16b

26. 16b. Was there an explanation of why age was included in the statistical model? *

We're looking for some sort of explanation that cites relevant literature or uses other criteria for including age (e.g., age was included for adjustment after review of the relevant literature, after using causal diagrams to examine potential confounding variables in the association between x and y). We want to make sure there was meaningful inclusion of age.

*Mark only one oval.*

- Yes
- No

17-18

27. 17. Was there an explanation of why other variables (e.g., parental monitoring) were included in the statistical model? *

Related to question #13-#16. We're looking for some sort of explanation why other variables were included in the statistical model. This question does not include race/ethnicity, sex, age, and household income.

*Mark only one oval.*

- Yes
- No
- No other variables were included in the statistical model

28. 18. Was there a description of any data manipulation (e.g., changing a variable from continuous to categorical)? *

This one should be in the methods section. Sometimes researchers manipulate a variable prior to using it in a statistical model. For example, some might take a continuous variable and change it into a categorical variable. This is commonly done with clinical thresholds (e.g., continuous BMI categorized as low/medium/high etc. or blood pressure categorized as Normal/At Risk/High Blood Pressure). Sometimes it's done arbitrarily (e.g., taking a continuous variable and labeling those greater than the 80th percentile as the high group). **Other variable manipulation methods include z-transformation and mean centering.**

*Mark only one oval.*

- ◯ Yes
- ◯ No *Skip to question 30*

18b

29. 18b. Was there a reason given for any data manipulation? *

Did the paper state why the variable was manipulated?

*Mark only one oval.*

- ◯ Yes
- ◯ No

19

30. 19. Was there an effect estimate reported (e.g., point estimate, beta coefficient, odds ratio, risk ratio, Cohen's d)? *

Typically, in the methods and results section. The results of a regression model (or any kind of statistical model) are typically effect estimates/point estimates. It can take the form of a beta coefficient (β), odds ratio (OR), risk ratio (RR), mean difference, Cohen's d, etc.

*Mark only one oval.*

- ◯ Yes
- ◯ No *Skip to question 32*

19b

31. 19b. Was there a quantification of uncertainty of the estimated parameters (e.g., confidence intervals or credible intervals, standard errors)? *

Related to question #19. Authors typically state this in Methods or Results. There is usually a measure of uncertainty reported alongside an effect estimate. Most common ones are confidence intervals and standard errors. For Bayesian statistics, this would be the credible interval. Confidence intervals/credible intervals are sometimes presented in the Results table. They are often enclosed in brackets.

*Mark only one oval.*

- ◯ Yes
- ◯ No

20

32. 20. Was a p-value reported? *

Typically reported in the Methods and Results section. Did the study report p-values? P-values are commonly reported in papers that do any sort of hypothesis testing. They are not typically reported in papers that use Bayesian statistics.

*Mark only one oval.*

- ◯ Yes
- ◯ No

33. 21. Did the authors mention issues with multiple comparisons? *

Multiple comparison issues are common in studies using MRI and genetics data. But they can also be a problem in studies using only behavioral measures. We want to know whether the paper acknowledged the issue of multiple comparisons in any way.

*Mark only one oval.*

- ◯ Yes
- ◯ No *Skip to question 35*

21b

34. 21b. Was there any correction for multiple testing? (This only applies to analyses that had many different tests of significance) *

Related to question #21. Did they use some type of correction for multiple comparisons (e.g., Bonferroni correction, False Discovery Rate)? Did they justify not doing so (e.g., pre-registered hypothesis, risk of a false negative)?

*Mark only one oval.*

- Yes
- No

22

35. 22. Was there a sensitivity analysis conducted? *

A sensitivity analysis is usually described in the Methods section. We just want to know if they did any type of sensitivity analysis (no need to describe what they did).

*Mark only one oval.*

- Yes
- No

36. 23. Did the authors use a goodness of fit test to select their final model? (e.g., Akaike Information Criterion, Bayesian Information Criterion). *

We want to know whether the authors used any method to determine the fit of the model. This can help find a statistical model that is parsimonious to minimize issues with overfitting. Some popular 'goodness of fit' tests include Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), and Leave-One-Out Information Criterion (LOOIC).

*Mark only one oval.*

- Yes
- No

37. 24. Did the paper use the Data Exploration and Analysis Portal (DEAP) to run their analysis? *

Typically included in the Methods section. We want to know if they used the ABCD sponsored Data Exploration and Analysis Portal (DEAP) to conduct their statistical analysis. This is sometimes used in place of more traditional software (e.g., R, Stata, Matlab, SPSS, SAS).

*Mark only one oval.*

- Yes
- No

38. 25. Were potential outliers discussed in the manuscript? *

Mention of outliers will typically be in the Methods section. Outlying data points can be problematic in certain analyses (e.g., small samples).

*Mark only one oval.*

- Yes
- No

39. 26. Was skewed data discussed in the manuscript? *

Skewed data can be problematic for obtaining accurate effect estimates. We want to know if there was any mention of skewed data.

*Mark only one oval.*

- ◯ Yes
- ◯ No

40. 27. Did the study use machine-learning methods? *

We just want to know if machine learning was used to investigate the associations of interest.

*Mark only one oval.*

- ◯ Yes
- ◯ No

## Miscellaneous

41. 28. Was imaging data used in the analysis? *

The different options for the imaging data are task-related functional magnetic resonance imaging (fMRI), structural MRI (sMRI), diffusion MRI(dMRI), resting-state fMRI.

*Mark only one oval.*

- ◯ Yes
- ◯ No *Skip to question 43*

## 28b

42. 28b. Was there a rationale and method for selecting the particular regions of interest (e.g., did they specify why they included the nucleus accumbens in their analysis)? *

For studies that used imaging data, was there a reason given for selecting the regions of interest? Reasons can range from a priori selection after review of the literature to exploratory reasons (i.e., no pre-specified hypothesis). If the study used all brain regions in their analysis (e.g., for a factor analysis) then please select the third choice. If the authors used all regions then select “Conducted whole brain analysis.”

*Mark only one oval.*

- ◯ Yes
- ◯ No
- ◯ Conducted whole brain analyses
- ◯ Other: _____

29

43. 29. Was genetics data used in the analysis? *

We just want to know if any of the ABCD genetics data was used in any way. For example, to determine polygenic risk scores.

*Mark only one oval.*

- ◯ Yes
- ◯ No

44. 30. Was the study pre-registered? (e.g, pre-registered on Open Science Framework). *

This might be included at the end of the Introduction section or in the methods section. Typically there will be a link to the Open Science Framework pre-registration page.

*Mark only one oval.*

- ◯ Yes
- ◯ No

45. 31. Was there mention of any impact to generalizability due to missing data/attrition? *

This one will typically be in the Discussion section – included with the other limitations of the study. If the study had no missing data (e.g., due to using only imputed data) and no attrition (e.g., if it only included visit 1) then select the final option.

*Mark only one oval.*

- Yes
- No
- No missing data/attrition

46. 32. Did the authors indicate author contributions? *

We just want to know if the authors indicated the role of each co-author. This may be journal specific.

*Mark only one oval.*

- Yes
- No

Other

47. Flag this article as a special case (i.e., requires whole group discussion). *

*Mark only one oval.*

- Yes
- No

48. Open-ended notes about important/noteworthy things.

Skip if there is nothing to note.