Longitudinal Analysis Manuscript: Working Draft

Samuel W. Hawes, PhD.[1*], Andrew K. Littlefield, PhD.[2], Daniel A. Lopez[3], Kenneth J. Sher[4], Wesley K. Thompson[5], Additional Co-authors1[1], Additional Co-authors2[1]

1 Center for Children and Families, Florida International University, Miami, FL, United States,
2 Psychological Sciences, Texas Tech University, Lubbock, TX, United States,
3 Department of Public Health, University of Rochester Medical Center, Rochester, NY, United States,
4 Psychological Sciences, University of Missouri, Columbia, MO, United States,
5 Population Neuroscience and Genetics Center, Laureate Institute for Brain Research, Tulsa, OK, United States,

☯These authors contributed equally to this work.

¤Current Address: Dept/Program/Center, Institution Name, City, State, Country

†Deceased

¶Biostatistics Working Group membership list can be found in the Acknowledgments sections

* shawes@fiu.edu

Abstract

The Adolescent Brain Cognitive Development (ABCD) Study presents a unique opportunity for researchers to investigate developmental processes in a large, diverse cohort of children and adolescents. Given the complex nature of the longitudinal data collected in the ABCD Study, researcher are likely to encounter a myriad of methodological and analytic considerations and concerns. This review provides a comprehensive examination of key issues and techniques related to longitudinal data analysis, specifically focusing on the ABCD Study. The text discusses model assumptions, common violations (e.g., independent and identically distributed residuals, *heterogeneous variability*) and their implications for valid inference. The importance of appropriately modeling covariance structures, understandings trade-offs between model fit and parsimony, and challenges related to sample size, attrition, missing data are highlighted. Consideration is given to the importance of selecting appropriate statistical models to account for correlations in repeated measurements and the assumptions underlying these models. The review also differentiates between linear and non-linear models in the context of continuous and discrete data, emphasizing various distributional assumptions and the necessity of choosing appropriate models and statistical methods. By addressing these complexities, the review seeks to equip researchers with the necessary knowledge and tools to make informed decisions as they navigate effectively analyzing and interpreting data available in the ABCD Study.

Author summary

The Adolescent Brain Cognitive Development (ABCD) Study offers valuable insights into adolescent brain development. Researchers working with this data face methodological and analytic concerns, including sample size, attrition, missing data, and measurement. This review addresses complexities in analyzing longitudinal data, with a focus on statistical models accounting for correlations in repeated measurements and their assumptions. We emphasize the importance of choosing appropriate models and analytic functions based on the data and the specific scientific question to be addressed. This concise resource aims to help researchers navigate longitudinal data analysis, make informed methodological decisions, and enhance the accuracy and reliability of their findings when working with data from the ABCD Study.

Introduction

The Adolescent Brain Cognitive Development (ABCD) Study® is the largest long-term investigation of neurodevelopment and child health in the United States. Conceived and initiated by the National Institutes of Health (NIH), this landmark prospective longitudinal study aims to transform our understanding of the genetic and environmental influences on brain development and their roles in behavioral and health outcomes in adolescents [1]. At its heart, the study is designed to chart the course of human development across multiple, interacting domains from late childhood to early adulthood and to identify factors that lead to both positive and negative developmental outcomes. Central to achieving these goals is the ABCD Study's® commitment to an open science framework designed to facilitate access to and sharing of scientific knowledge by espousing practices that increase openness, integrity, and reproducibility of scientific research (e.g., public data releases). In this sense, the ABCD Study® is a collaboration with the larger research community, with the rich longitudinal nature of the ABCD Study dataset allowing researchers to perform a variety of analyses of both methodological and substantive interest. Together, this presents a unique opportunity to significantly advance our understanding of how a multitude of biopsychosocial processes emerge and unfold across critical periods of development.

[section still be developed. . .]

The ABCD Study® Data

Participants enrolled in the ABCD Study include a large cohort of youth (n=11880) aged 9-10 years at baseline and their parents/guardians. The study sample was recruited from household populations in defined catchment areas for each of the 21 study sites across the United States (information regarding funding agencies, recruitment sites, investigators, and project organization can be obtained at the ABCD Study website). The ABCD Study is collecting longitudinal data on a rich variety of outcomes that will enable the construction of realistically-complex etiological models by incorporating factors from many domains simultaneously. Each new wave of data collection provides the building blocks for conducting probing longitudinal analyses that allow us to characterize normative development, identify variables that presage deviations from prototypic development, and assess a range of outcomes associated with variables of interest. This data includes a neurocognitive battery [2, 3], mental and physical health assessments [4], measures of culture and environment [5], substance use [add citation], biospecimens [6], structural and functional brain imaging [7, 8], geolocation-based environmental exposure data, wearables, and mobile technology [9], and whole genome genotyping [10]. Many of these measures are collected at in-person

annual visits, with brain imaging collected at baseline and every other year going forward. A limited number of assessments are collected in semi-annual telephone interviews between in-person visits. Data are publicly released on an annual basis through the NIMH Data Archive. By necessity, the study's earliest data releases were cross-sectional (i.e., the baseline data), however, the most recent public data release (NDA Release 4.0) contains data collected across three annual assessments, including two imaging assessments (baseline and year 2 follow-up visits).

Organization of current manuscript

The rich longitudinal nature of the ABCD Study dataset will allow researchers to perform analyses of both methodological and substantive interest. This report describes methods for longitudinal analyses of ABCD Study data that can address its fundamental scientific aims, as well as challenges inherent in a large population-based long-term study of adolescents. The manuscript is organized as follows:

[section still be developed...]

Part I: Longitudinal Research: Basic Concepts and Considerations

There are several important concepts to consider when conducting longitudinal analyses in a developmental context. These include different ways of thinking about developmental course, whether certain periods of development are relatively sensitive or insensitive to various types of insults or stressors, whether some time periods or situations inhibit the expression of individual differences due to extreme environmental pressures, and whether the same behavior manifested at different times represent the same phenomenon or different ones. Further, in the case of developmentally focused longitudinal research, each new measurement occasion not only provides a more extended portrait of the child's life course (and not just characterize growth during this period but also assesses the durability/chronicity of prior effects/consequences) but also brings with it greater methodological opportunities to exploit the statistical properties of longitudinal data in the furtherance of critical scientific questions. That is, we can ask more nuanced questions and make stronger inferences as our number of time-ordered observations grows, assuming we have assessed the "right" variables and the timings of our observations comport with the temporal dynamics of the mechanisms of interest. Appreciation of these and other issues can help to guide the analysis and interpretation of data and aid translation to clinical and public health applications.

Vulnerable periods. Development normatively progresses from less mature to more mature levels of functioning. However, unique epochs and experiences can alter the course of this idealized form of development. Consider research that shows cannabis use during adolescence is associated with later psychosis to a greater degree than cannabis use initiated later in development [add citation]; or, similarly, experimental research on rodents that shows rodent brains to be especially sensitive to the neurotoxic effects of alcohol on brain structure and learning early in development (corresponding to early adolescence in humans)[add citation]. These examples highlight the importance of considering the role of vulnerable periods – temporal windows of rapid brain development or remodeling during which the effects of environmental stimuli (e.g. cannabis exposure) on the developing brain may be particularly pronounced– when trying to establish an accurate understanding of the association between exposures and outcomes.

Developmental disturbances. Whereas vulnerable periods heighten 83
neurobiological susceptibility to environmental influences, at other times environmental 84
pressures will tend to suppress stability and disrupt the orderly stochastic process of 85
normative development (e.g., xxx-xxx). This situation reflects a developmental 86
disturbance in that the normal course of development is "disturbed" for some time by 87
some time-limited process. In such cases, we might find that prediction of behavior in 88
the period of the disturbance is reduced and/or, similarly, the behavior exhibited during 89
the disturbance might have less predictive power with respect to distal outcomes 90
compared to the behavior exhibited before and following the disrupted period. That is, 91
once the environmental stimuli are removed (or the individual is removed from the 92
environment), individual differences are again more freely expressed and the 93
autoregressive effects increase to levels similar to those before entering the environment. 94

Developmental snares and cascade effects. Normative development can also be 95
upended by experiences (e.g., drug use) that, through various mechanisms, disrupt the 96
normal flow of development wherein each stage establishes a platform for the next. For 97
instance, substance use could lead to association with deviant peers, precluding 98
opportunities for learning various adaptive skills and prosocial behaviors, in effect, 99
creating a "snare" that retards psychosocial development. Relatedly, the consequences 100
of these types of events can cascade (e.g., school dropout, involvement in the criminal 101
justice system) so that the effects of the snare are amplified. Although conceptually 102
distinct from vulnerable periods, both of these types of developmental considerations 103
highlight the importance of viewing behavior in the context of development and the 104
importance of attempting to determine how various developmental pathways unfold. 105

Distinguishing developmental change from experience effects. One can 106
often observe systematic changes over time in a variable of interest and assume this 107
change is attributable to development. To this point, cognitive abilities (e.g, verbal 108
ability, problem-solving) normatively grow earlier in development and often decline in 109
late life (e.g., memory, speed of processing). However, the observed patterns of growth 110
and decline often differ between cross-sectional vs. longitudinal effects [11] where 111
subjects gain increasing experience with the assessment with each successive 112
measurement occasion. Such experience effects on cognitive functioning have been 113
demonstrated in adolescent longitudinal samples similar to ABCD [12] and highlight the 114
need to consider these effects and address them analytically. In the case of 115
performance-based measures [e.g., matrix reasoning related to neurocognitive 116
functioning; see **(author?)** [11]], this can be due to "learning" the task from previous 117
test administrations (e.g., someone taking the test a second time performs better than 118
they did the first time simply as a function of having taken it before). Even in the case 119
of non-performance-based measures (e.g., levels of depression), where one cannot easily 120
make the argument that one has acquired some task-specific skill through learning, it 121
has been observed that respondents tend to endorse lower levels on subsequent 122
assessments [e.g., 13; see 14] and this phenomenon has been well documented in 123
research on structured diagnostic interviews [15]. While it is typically assumed that 124
individuals are rescinding or telling us less information on follow-up interviews, there is 125
reason to suspect that in some cases the initial assessment may be artefactually elevated 126
[see 16]. Some designs (specifically, accelerated longitudinal designs) are especially well 127
suited for discovering these effects and modeling them. While ABCD was not designed 128
as an accelerated longitudinal design, the variability in age at the time of baseline 129
recruitment (9 years, 0 months to 10 years, 11 months) allows some measures, collected 130
every year, to be conceptualized as an accelerated longitudinal design. Moreover, it is 131
possible that in later waves, patterns of longitudinal missing data will allow some 132
analyses to assess the confounded effects of age and the number of prior assessments. 133

However, ABCD is fundamentally a single-cohort, longitudinal design, where a number 134
of prior assessments and age are highly confounded, and for, perhaps, most analyses, 135
the possible influence of experience effects needs to be kept in mind. 136

Part II Longitudinal Data: Interpretation / Issues / Pitfalls & Assumption 137 138

Defining Features of Longitudinal Data Analysis. The hallmark characteristic of 139
longitudinal data analysis is its application to repeated assessments of the same 140
assessment targets (e.g., individuals, families) across time. While the primary reason for 141
collecting longitudinal data is in pursuit of addressing scientific questions, from a 142
methodological perspective, having multiple observations over time allows researchers to 143
identify potentially problematic observations when highly improbable longitudinal 144
patterns are observed. That is, we can ask more nuanced questions and make stronger 145
inferences as our number of time-ordered observations grows assuming we have assessed 146
the "right" variables and the timings of our observations comport with the temporal 147
dynamics of the mechanisms of interest . 148

Modeling Data Across Two Time Points versus Three or More Time Points. 149 150

Although the clear leap to the realm of longitudinal data involves going from one 151
assessment to two or more assessments, there are also notable distinctions in designs 152
based on two-assessment points versus three or more measurement occasions. Just as 153
cross-sectional data can be informative in some situations, two waves of data can be 154
beneficial in contexts such as when experimental manipulation is involved (e.g., 155
pre/post tests), or if the central goal is prediction (e.g., trying to predict scores on 156
Variable A at time T as a function of prior scores on Variable A and Variable B at time 157
T-1). At the same time, data based on two assessments are inherently limited on 158
multiple fronts. As [17] noted approximately forty years ago, "Two waves of data are 159
better than one, but maybe not much better". These sentiments are reflected in more 160
contemporary recommendations regarding best-practice guidelines for prospective data, 161
which increasingly emphasize the benefits of additional measurement occasions for 162
model identification and accurate parameter estimation. It is also consistent with 163
research recommending that developmental studies include three or more assessment 164
points, given it is impossible for data based on two-time points to determine the shape 165
of development (since linear, straight line change is the only possible form, given two 166
assessments; see [18]). Research designs that include three or more time points allow for 167
increasingly nuanced analyses that more adequately tease apart sources of variation and 168
covariation among the repeated assessments [19]– a key aspect of inferential research. 169
To illustrate, developmental theories are typically interested in understanding patterns 170
of within-individual change over time (discussed in further detail, below); however, two 171
data points provide meager information on change at the person level. This point is 172
further underscored in a recent review of statistical models commonly touted as 173
distinguishing within-individual vs between-individual sources of variance in which the 174
study authors concluded "... researchers are limited when attempting to differentiate 175
these sources of variation in psychological phenomenon when using two waves of data" 176
and perhaps more concerning, "...the models discussed here do not offer a feasible way 177
to overcome these inherent limitations" **(author?)** [20]. It is important to note, 178
however, that despite the current focus on two-wave designs versus three or more 179
assessment waves, garnering three assessment points is not a panacea for longitudinal 180

modeling. Indeed, several contemporary longitudinal models designed to isolate [181] within-individual variability [e.g., the Latent Curve Model with Structured Residuals; [182] **(author?)** [21]] require at least four assessments to parameterize fully and, more [183] generally, increasingly accurate parameter estimates are obtained as more assessment [184] occasions are used [18]. [185]

Types of stability and change [186]

If one were to try to sum up what development in a living organism is exactly, one [187] could plausibly argue it's the characterization of stability and change as the organism [188] traverses the life course. There are a few different ways to think of stability (and [189] change). Consider we measure the height of all youth in a 6th-grade class, once in the [190] fall at the beginning of the school year and once again in the spring at the end of the [191] school year. A common first step may be to compare the class's average height values [192] obtained at these two different measurement occasions. This comparison of the average [193] scores for the same group of individuals at multiple time points is referred to as [194] "mean-level" stability as it provides information about continuity and change in the [195] group level of an outcome of interest (e.g., height) over time. Another type of stability [196] involves calculating the correlation between the values obtained at different time points [197] (e.g., 'height in the fall' with 'height in the spring'). This type of "rank-order" stability [198] evaluates between-individual change by focusing on the degree to which individuals [199] retain their relative placement in a group across time. Consider, someone who is the [200] shortest person in their class in 6th grade may grow considerably over the school year [201] (i.e., exhibit mean level change), but remain the shortest person among their classmates. [202] That is, the individual is manifesting a type of rank-order stability. Both types of [203] stability and change are important. Mean-level change in certain traits might help to [204] explain why, in general, individuals are particularly vulnerable to social influences at [205] some ages more than others; rank order change might help to quantify the extent to [206] which certain characteristics of the individual are more trait-like. For example, in some [207] areas of development, there is considerable mean-level change that occurs over time [208] (e.g., changes in Big 5 personality traits), but relatively high rank-order stability. [209] Despite the useful information afforded by examining mean-level and rank-order change, [210] these approaches are limited in that they provide little information about patterns of [211] "within-individual" change and, in turn, can result in fundamental misinterpretations [212] about substantial or meaningful changes in an outcome of interest. [213]

There is growing recognition that statistical models commonly applied to [214] longitudinal data often fail to comport with the developmental theory they are being [215] used to assess (e.g., Curran, Lee, Howard, Lane, & MacCallum, 2012; Hoffman, 2015; [216] Littlefield et al., 2021. Specifically, developmental studies typically involve the use of [217] prospective data to inform theories that are concerned with clear within-person (i.e., [218] intraindividual) processes (e.g., how phenotypes change or remain stable within [219] individuals over time) [e.g., see 22]. Despite this, methods generally unsuited for [220] disaggregating between- and within-person effects (e.g., cross-lagged panel models [221] [CLPM]) remain common within various extant literatures. As a result, experts [222] increasingly caution about the need to xxxxxxxx [add citation]. Fortunately, there [223] exists a range of models that have been proposed to tease apart between- and [224] within-person sources of variance across time [see 20, 23]. Most of these contemporary [225] alternatives incorporate time-specific latent variables to capture between-person sources [226] of variance and model within-person deviations around an individual's mean (or trait) [227] level across time [e.g., RI-CLPM, 24; LCM-SR, 21]. It is important to note however [228] that these models require multiple assessments waves (e.g., four or more to fully specify [229] the LCM-SR), additional expertise to overcome issues with model convergence, and [230]

appreciation of modeling assumptions when attempting to adjudicate among potential   231
models in each research context [see 20, for further discussion].   232

Model Assumptions   233

Many statistical models assume certain characteristics about the data to which they are   234
being applied. As an example, common assumptions of parametric statistical models   235
include normality, linearity, and equality of variances. These assumptions must be   236
carefully considered before conducting analysis so that valid inferences can be made   237
from the data; that is, violation of a model's assumptions can substantively alter the   238
interpretation of results. Similarly, statistical models employed in the analyses of   239
longitudinal data often entail a range of assumptions that must be closely inspected.   240
One central issue for repeated measurements on an individual is how to account for the   241
correlated nature of the data; another common feature of longitudinal data is   242
heterogeneous variability; that is, the variance of the response changes over the duration   243
of the study. Traditional techniques, such as a standard regression or ANOVA model,   244
assume residuals are independent and thus are inappropriate for designs that assess (for   245
example) the same individuals across time. That is, given the residuals are no longer   246
independent, the standard errors from the models are biased and can produce   247
misleading inferential results. Although there are formal tests of independence for time   248
series data (e.g., the Durbin-Watson statistic; Durbin & Watson, 1950), more commonly   249
independence is assumed to be violated in study designs with repeated assessments.   250
Therefore, an initial question to be addressed by a researcher analyzing prospective data   251
is how to best model the covariance structure of said data.   252

Covariance Structures   253

Statistical models for longitudinal data include two main components to account for   254
assumptions that are commonly violated when working with repeated measures data: a   255
model for the covariance among repeated measures (both the correlations among pairs   256
of repeated measures on an individual and the variability of the responses on different   257
occasions), coupled with a model for the mean response and its dependence on   258
covariates (eg, treatment group in the context of clinical trials). This allows for the   259
specification of a range of so-called covariance structures, each with its own set of   260
tradeoffs between model fit and parsimony [e.g., see 25].   261

Accounting for Correlated Data   262

As an example, one alternative structure that attempts to handle the reality that   263
correlations between repeated assessments tend to diminish across time is the   264
autoregressive design. As the name implies, the structure assumes a subsequent   265
measurement occasion (e.g., assessment at Wave 2) is regressed onto (that is, is   266
predicted by) a prior measurement occasion (e.g., assessment at Wave 1). The most   267
common type of autoregressive design is the AR(1), where assessments at time T + 1   268
are regressed on assessments at Time T. Identical to compound symmetry, this model   269
assumes the variances are homogenous across time. Diverting from compound   270
symmetry, this model assumes the correlations between repeated assessments decline   271
exponentially across time rather than remaining constant. For example, per the AR(1)   272
structure, if the correlation between Time 1 and Time 2 data is thought to be .5, then   273
the correlation between Time 1 and Time 3 data would be assumed to be .5*.5 = .25,   274
and the correlation between Time 1 and Time 4 data would be assumed to be .5.5*.5 =*   275
.125. As with compound symmetry, the basic AR(1) model is parsimonious in that it   276
only requires two parameters (the variance of the assessments and the autoregressive   277

coefficient). Notably, the assumption of constant autoregressive relations between ₂₇₈ assessments is often relaxed in commonly employed designs that use autoregressive ₂₇₉ modeling (e.g., cross-lagged panel models [CLPM]). These designs still typically assume ₂₈₀ an AR(1) process (e.g., it is sufficient to regress the Time 3 assessment onto the Time 2 ₂₈₁ assessment and is not necessary to also regress the Time 3 assessment onto the Time 1 ₂₈₂ assessment, which would result in an AR(2) process). However, the magnitude of these ₂₈₃ relations is often allowed to differ across different AR(1) pairs of assessment (e.g., the ₂₈₄ relation between Time 1 and Time 2 can be different from the relation between Time 2 ₂₈₅ and Time 3). These more commonly employed models also often relax the assumption ₂₈₆ of equal variances of the repeated assessments. Although the AR(1) structure may ₂₈₇ involve a more realistic set of assumptions compared to compound symmetry, in that ₂₈₈ the AR(1) model allows for diminishing correlations across time, the basic AR(1) model, ₂₈₉ as well as autoregressive models more generally, can also suffer from several limitations ₂₉₀ in contexts that are common in prospective designs. In particular, recent work ₂₉₁ demonstrates that if a construct being assessed prospectively across time is trait-like in ₂₉₂ nature, then autoregressive relations fail to adequately account for this trait-like ₂₉₃ structure, with the downstream consequence that estimates derived from models based ₂₉₄ on AR structures (such as the CLPM) can be misleading and fail to adequately ₂₉₅ demarcate between- vs. within-person sources of variance [24]. ₂₉₆

Linear vs non-linear models ₂₉₇

Identification of optimal statistical models and appropriate mathematical functions ₂₉₈ requires an understanding of the type of data being used. Repeated assessments can be ₂₉₉ based on either continuous or discrete measures. Examples of discrete measures include ₃₀₀ repeated assessments of binary variables (e.g., past 12-month alcohol use disorder status ₃₀₁ measured across ten years), ordinal variables (e.g., a single item measuring the level of ₃₀₂ agreement to a statement on a three-point scale including the categories of "disagree", ₃₀₃ "neutral", and "agree" in an ecological momentary assessment study that involves ₃₀₄ multiple daily assessments), and count variables (e.g., number of cigarettes smoked per ₃₀₅ day across a daily diary study). In many ways, the distributional assumptions of ₃₀₆ indicators used in longitudinal designs mirror the decision points and considerations ₃₀₇ when delineating across different types of discrete outcome variables, a topic that spans ₃₀₈ entire textbooks [e.g., see 26]. For example, the Mplus manual [27] includes examples of ₃₀₉ a) censored and censored-inflated models, b) linear growth models for binary or ordinal ₃₁₀ variables, c) linear growth models for a count outcome assuming a Poisson model, d) ₃₁₁ linear growth models for a count outcome assuming a zero-inflated Poisson model and e) ₃₁₂ discrete- and continuous-time survival analysis for a binary outcome. Beyond these ₃₁₃ highlighted examples, other distributions (e.g., negative binomial) can be assumed for ₃₁₄ the indicators when modeling longitudinal data. These models can account for issues ₃₁₅ that can occur when working with discrete outcomes, including overdispersion (when ₃₁₆ the variance is higher than would be expected based on a given distribution) and ₃₁₇ zero-inflation [when more zeros occur than is expected based on a given distribution; see ₃₁₈ **(author?)** [26]]. Models involving zero-inflation parameters are referred to as two-part ₃₁₉ models, given one part of the model predicts the zero-inflation whereas the other part of ₃₂₀ the model predicts outcomes consistent with a given distribution [e.g., Poisson ₃₂₁ distribution; see **(author?)** [28], for a review of two-part models for longitudinal data]. ₃₂₂ Although there exist several alternative models for discrete indicators, some more recent ₃₂₃ models that have been proposed for prospective data are only feasible in cases where ₃₂₄ indicators are assumed to be continuous rather than discrete [e.g., LCM-SR; **(author?)** ₃₂₅ [21]]. Given the sheer breadth of issues relevant to determining better models for ₃₂₆ discrete outcomes, it is not uncommon for texts on longitudinal data analysis to only ₃₂₇ cover models and approaches that assume continuous indicators [e.g., 29]. However, ₃₂₈

some textbooks on categorical data analysis provide more detailed coverage of the myriad issues and modeling choices to consider when working with discrete outcomes [e.g., **(author?)** [26], Chapter 11 for matched pair/two-assessment designs; Chapter 12 for marginal and transitional models for repeated designs, such as generalized estimating equations, and Chapter 13 for random effects models for discrete outcomes].

Missing Data/Attrition

As recently reviewed by Littlefield (in press), investigators of prospective data are confronted with study attrition (i.e., participants may not provide data at a given wave of assessment) and thus approaches are needed to confront the issue of missing data. Three models of missingness are typically considered in the literature [see 30]. These three models are data: a) missing completely at random (MCAR), b) missing at random (MAR), and c) missing not at random (MNAR). Data that are MCAR means missing data is a random sample of all the types of participants (e.g., males) in a given dataset. MAR suggests conditionally missing at random [see 31]. That is, MAR implies missingness is completely random (i.e., does not hinge on some unmeasured variables) once missingness has been adjusted by all available variables in a dataset (e.g., biological sex). Data that are MNAR are missing as a function of unobserved variables. **(author?)** [31] provides an excellent and easy-to-digest overview of further details involving missing data considerations.

Multiple approaches have been posited to handle missing data. Before the advent of more contemporary approaches, common methods included several ad hoc procedures. These include eliminating the data of participants with missing data (e.g., listwise or pairwise deletion) or using mean imputation (i.e., replacing the missing value with the mean score of the sample that did participate). However, these methods are not recommended because they can contribute to biased parameter estimates and research conclusions [see 31]. More specifically, the last observation carried forward (LOCF) is a common approach to imputing missing data. LOCF replaces a participant's missing values after dropout with the last available measurement [32]. This approach assumes stability (i.e., a given participant's score is not anticipated to increase or decline after study dropout) and that the data are MCA R. However, as described by **(author?)** [32], it is common for treatment groups to show higher attrition compared to control groups in studies of dementia drugs. Given that dementia worsens over time, using LOCF biases the results in favor of the treatment group [see 32, for more details].

More modern approaches, such as using maximum likelihood or multiple imputation to estimate missing data, are thought to avoid some of the biases of older approaches [see 33, 31]. **(author?)** [31] noted several "myths" regarding missing data. For example, Graham notes many assume the data must be minimally MAR to permit estimating procedures (such as maximum likelihood or multiple imputation) compared to other, more traditional approaches (e.g., using only complete case data). Violations of MAR impact both traditional and more modern data estimation procedures, though as noted by Graham, violations of MAR tend to have a greater effect on older methods. Graham thus suggests that estimating missing data is a better approach compared to the older procedures in most circumstances, regardless of the model of missingness [i.e., MCAR, MAR, MNAR; see **(author?)** [31]].

Attrition from a longitudinal panel study such as ABCD is inevitable and represents a threat to the validity of longitudinal analyses and cross-sectional analyses conducted at later time points, especially since attrition can only be expected to grow over time. While, to date, attrition in ABCD has been minimal (some cite here), it remains an important focus for longitudinal analysis and its significance is likely to only grow as the

cohort ages. Ideally, one tries to minimize attrition through good retention practices 378 from the outset via strategies designed to maintain engagement in the project 379 [34, 35, 36]. However, even the best-executed studies need to anticipate growing 380 attrition over the length of the study and implement analytic strategies designed to 381 provide the most valid inferences. Perhaps the most key concern when dealing with 382 data that is missing due to attrition is determining the degree of bias in retained 383 variables that is a consequence of attrition. Assuming that the data are not missing 384 completely at random, attention to the nature of the missingness and employing 385 techniques designed to mitigate attrition-related biases need to be considered in all 386 longitudinal analyses. Several different approaches can be considered and employed 387 depending upon the nature of the intended analyses, the degree of missingness, and data 388 available to help estimate missing and unobserved values. 389

Quantifying effect sizes longitudinally 390

Given longitudinal data involve different sources of variance, quantifying effect sizes 391 longitudinally is a more difficult task compared to deriving such estimates from 392 cross-sectional data. Effect size can be defined as, "a population parameter (estimated 393 in a sample) encapsulating the practical or clinical importance of a phenomenon under 394 study." [37]. Common effect size metrics include r (i.e., the standardized covariance, or 395 correlation, between two variables) and Cohen's d [38]. Adjustments to common effect 396 size calculations, such as Cohen's d, are required even when only two time points are 397 considered [e.g., see 39]. **(author?)** [40] note there are multiple approaches to obtaining 398 standardized within-person effects, and that commonly suggested approaches (e.g., 399 global standardization) can be problematic [see 40, for more details]. Thus, obtaining 400 effect size metrics based on standardized estimates that are relatively simple in 401 cross-sectional data (such as r) becomes more complex in the context of prospective 402 data. **(author?)** [41] noted that equations for effects sizes used in studies involving 403 growth modeling analysis (e.g., latent growth curve modeling) were not mathematically 404 equivalent, and the effect sizes were not in the same metric as effect sizes from 405 traditional analysis [see 41, for more details]. Given this issue, there have been various 406 proposals for adjusting effect size measures in repeated assessments. **(author?)** [42] 407 reviews the approach for effect size metrics for analyses based on growth modeling, 408 including when considering linear and non-linear (i.e., quadratic) growth factors. 409 **(author?)** [39] review various equations for effect size calculations relevant to when 410 combining estimates in meta-analysis with repeated measures and independent-groups 411 designs. Other approaches to quantifying effect sizes longitudinally may be based on 412 standardized estimates from models that more optimally disentangle between- and 413 within-person sources of variance (as reviewed above). As an example, within a 414 RI-CLPM framework, standardized estimates between random intercepts (i.e., the 415 correlation between two random intercepts for two different constructs assessed 416 repeatedly) could be used to index the between-person relation, whereas standardized 417 estimates among the structured residuals could be used as informing the effect sizes of 418 within-person relations. 419

General guidelines for this Quarto template 420

This template shows how to use PLOS template from 421 https://plos.org/resources/writing-center/. Each journal have a submission guideline 422 page, please refer to it. 423

- PLOS Biology 424
- PLOS Climate 425

- PLOS Digital Health
- PLOS Computational Biology
- PLOS Genetics
- PLOS Global Public Health
- PLOS Medicine
- PLOS Neglected Tropical Diseases
- PLOS ONE
- PLOS Pathogens
- PLOS Sustainability and Transformation
- PLOS Water

This template file contains some guidelines and recommandation initially given in `plos_latex_template.tex` that can be found in https://github.com/quarto-journals/plos/blob/main/style-guide/plos_latex_template.tex

Metadata

About journal id field

This is an identifier for the target journal. It can be derived from https://plos.org/resources/writing-center/ following submission guidelines link, the identifier is the part of the URL after `https://journals.plos.org/<id>/s/submission-guidelines`

Journal	id
PLOS Biology	plosbiology
PLOS Climate	climate
PLOS Digital Health	digitalhealth
PLOS Computational Biology	ploscompbiol
PLOS Genetics	plosgenetics
PLOS Global Public Health	globalpublichealth
PLOS Medicine	plosmedicine
PLOS Neglected Tropical Diseases	plosntds
PLOS ONE	plosone
PLOS Pathogens	plospathogens
PLOS Sustainability and Transformation	sustainabilitytransformation
PLOS Water	water

Example :

```
format:
  plos-pdf:
    journal:
      id: water
```

Once your paper is accepted for publication,

Do not include track change in LaTeX file and leave only the final text of your manuscript. PLOS recommends the use of latexdiff to track changes during review, as this will help to maintain a clean tex file. Visit https://www.ctan.org/pkg/latexdiff?lang=en for info or contact us at latex@plos.org.

This should not be a problem using Quarto but still a recommandation from the journal

There are no restrictions on package use within the LaTeX files except that no packages listed in the template may be deleted.

Please do not include colors or graphics in the text. Color can be used to apply background shading to table cells only.

The manuscript LaTeX source should be contained within a single file (do not use\input, \externaldocument, or similar commands).

Please contact latex@plos.org with any questions submission guidelines. For anything Quarto related, please open an issue in https://github.com/quarto-journals/plos. If this is related to the LaTeX template, this could also be a good idea to contact PLOS directly.

Figures and Tables

Please include tables/figure captions directly after the paragraph where they are first cited in the text.

Figures

However, do not include graphics in your manuscript

- Figures should be uploaded separately from your manuscript file.
- Figures generated using LaTeX should be extracted and removed from the PDF before submission.
- Figures containing multiple panels/subfigures must be combined into one image file before submission.

This means that, depending on how you create your figure, a manual post processing will be required.

For figure citations, please use "Fig" instead of "Figure". This has been made the default in this Quarto format:

```
crossref:
  fig-title: Fig
```

Also, place figure captions after the first paragraph in which they are cited.

See PLOS figure guidelines at https://journals.plos.org/plosone/s/figures and in your specific journal guideline.

Tables

Tables should be cell-based and may not contain:

- spacing/line breaks within cells to alter layout or alignment
- do not nest tabular environments (no tabular environments within tabular environments)
- no graphics or colored text (cell background color/shading OK)

See PLOS table guidelines at http://journals.plos.org/plosone/s/tables and in your specific journal guideline.

For tables that exceed the width of the text column, use the adjustwidth environment as illustrated in the example table in text below. If you are in this case, you'll either need to manually post process the `.tex` file and recreate the PDF, or you need to include LaTeX tables directly.

Also, place tables after the first paragraph in which they are cited.

Equations, math symbols, subscripts, and superscripts

Below are a few tips to help format your equations and other special characters according to our specifications. For more tips to help reduce the possibility of formatting errors during conversion, please see our LaTeX guidelines at http://journals.plos.org/plosone/s/latex

- For inline equations, please be sure to include all portions of an equation in the math environment. For example, `x$^2$` is incorrect; this should be formatted as x^2 (or x² if the romanized font is desired).

- Do not include text that is not math in the math environment. For example, `CO2` should be written as `CO\textsubscript{2}` giving CO_2 instead of `CO$_2$`.

- Please add line breaks to long display equations when possible in order to fit size of the column.

- For inline equations, please do not include punctuation (commas, etc) within the math environment unless this is part of the equation.

- When adding superscript or subscripts outside of brackets/braces, please group using {}. For example, change `"[U(D,E,\gamma)]^2"` to `"{[U(D,E,\gamma)]}^2"`.

- Do not use `\cal` for caligraphic font. Instead, use `\mathcal{}`

Title and headings

Please use "sentence case" for title and headings (capitalize only the first word in a title (or heading), the first word in a subtitle (or subheading), and any proper nouns).

PLOS does not support heading levels beyond the 3rd, meaning no 4th level headings. Header 4 levels `####` is used for the *Supporting information* section

Abstract and author summary

Abstract must be kept below 300 words.

Author Summary must be kept between 150 and 200 words and first person must be used.

For PLOS ONE, author summary won't be included as it is not valid for submission.

Supplementary information syntax

Use this markdown syntax to create the supplementary information block with a custom block of class `.supp`

```
::: {.supp}
## SI TYPE {#id}

First paragraph is a title sentence that will be bold. (required)

Optionnaly, add descriptive text after the title of the
item. No third paragraph is allowed
:::
```

They need to be referenced in text using `nameref` by using this syntax
`[id](.nameref)` where `id` will be the id used on the header.

References

Within Quarto, `natbib` will be used with `plos2015.bst`, which expect numeric style
citation. Use brackets for references, e.g `[@ref]`.

Quarto features limitation

Some features are not working with this format in PDF:

- Callouts
- Code highlighting customization (border left, background color)

Following content of this document is from the LaTeX template
content to demo journal style.

Introduction

Lorem ipsum dolor sit [**?**] amet, consectetur adipiscing elit. Curabitur eget porta erat.
Morbi consectetur est vel gravida pretium. Suspendisse ut dui eu ante cursus gravida
non sed sem. Nullam Eq 1 sapien tellus, commodo id velit id, eleifend volutpat quam.
Phasellus mauris velit, dapibus finibus elementum vel, pulvinar non tellus. Nunc
pellentesque pretium diam, quis maximus dolor faucibus id. [**?**] Nunc convallis sodales
ante, ut ullamcorper est egestas vitae. Nam sit amet enim ultrices, ultrices elit pulvinar,
volutpat risus.

$$\mathrm{P_Y} = \underbrace{H(Y_n) - H(Y_n|\mathbf{V}_n^Y)}_{S_Y} + \underbrace{H(Y_n|\mathbf{V}_n^Y) - H(Y_n|\mathbf{V}_n^{X,Y})}_{T_{X \to Y}} \tag{1}$$

Materials and methods

Etiam eget sapien nibh

Nulla mi mi, Fig 1 venenatis sed ipsum varius, volutpat euismod diam. Proin rutrum
vel massa non gravida. Quisque tempor sem et dignissim rutrum. Lorem ipsum dolor sit
amet, consectetur adipiscing elit. Morbi at justo vitae nulla elementum commodo eu id
massa. In vitae diam ac augue semper tincidunt eu ut eros. Fusce fringilla erat porttitor
lectus cursus, vel sagittis arcu lobortis. Aliquam in enim semper, aliquam massa id,
cursus neque. Praesent faucibus semper libero.

Figure 1. Bold the figure title. Figure caption text here, please use this space for
the figure panel descriptions instead of using subfigure commands. A: Lorem ipsum
dolor sit amet. B: Consectetur adipiscing elit.

Results

Results and Discussion can be combined.

Nulla mi mi, venenatis sed ipsum varius, Table 2 volutpat euismod diam. Proin rutrum vel massa non gravida. Quisque tempor sem et dignissim rutrum. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Morbi at justo vitae nulla elementum commodo eu id massa. In vitae diam ac augue semper tincidunt eu ut eros. Fusce fringilla erat porttitor lectus cursus, S1 Video vel sagittis arcu lobortis. Aliquam in enim semper, aliquam massa id, cursus neque. Praesent faucibus semper libero.

Table 2. Table caption Nulla mi mi, venenatis sed ipsum varius, volutpat euismod diam.

Heading1				Heading2			
$cell1row1$	cell2 row 1	cell3 row 1	cell4 row 1	cell5 row 1	cell6 row 1	cell7 row 1	cell8 row 1
$cell1row2$	cell2 row 2	cell3 row 2	cell4 row 2	cell5 row 2	cell6 row 2	cell7 row 2	cell8 row 2
$cell1row3$	cell2 row 3	cell3 row 3	cell4 row 3	cell5 row 3	cell6 row 3	cell7 row 3	cell8 row 3

Table notes Phasellus venenatis, tortor nec vestibulum mattis, massa tortor interdum felis, nec pellentesque metus tortor nec nisl. Ut ornare mauris tellus, vel dapibus arcu suscipit sed.

LOREM and IPSUM nunc blandit a tortor

3rd level heading

Maecenas convallis mauris sit amet sem ultrices gravida. Etiam eget sapien nibh. Sed ac ipsum eget enim egestas ullamcorper nec euismod ligula. Curabitur fringilla pulvinar lectus consectetur pellentesque. Quisque augue sem, tincidunt sit amet feugiat eget, ullamcorper sed velit. Sed non aliquet felis. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Mauris commodo justo ac dui pretium imperdiet. Sed suscipit iaculis mi at feugiat.

1. react

2. diffuse free particles

3. increment time by dt and go to 1

Sed ac quam id nisi malesuada congue

Nulla mi mi, venenatis sed ipsum varius, volutpat euismod diam. Proin rutrum vel massa non gravida. Quisque tempor sem et dignissim rutrum. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Morbi at justo vitae nulla elementum commodo eu id massa. In vitae diam ac augue semper tincidunt eu ut eros. Fusce fringilla erat porttitor lectus cursus, vel sagittis arcu lobortis. Aliquam in enim semper, aliquam massa id, cursus neque. Praesent faucibus semper libero.

- First bulleted item.

- Second bulleted item.

- Third bulleted item.

Discussion

Nulla mi mi, venenatis sed ipsum varius,see Table 2 volutpat euismod diam. Proin rutrum vel massa non gravida. Quisque tempor sem et dignissim rutrum. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Morbi at justo vitae nulla elementum

commodo eu id massa. In vitae diam ac augue semper tincidunt eu ut eros. Fusce ⁵⁸⁶ fringilla erat porttitor lectus cursus, vel sagittis arcu lobortis. Aliquam in enim semper, ⁵⁸⁷ aliquam massa id, cursus neque. Praesent faucibus semper libero [**?**]. ⁵⁸⁸

Conclusion ⁵⁸⁹

CO_2 Maecenas convallis mauris sit amet sem ultrices gravida. Etiam eget sapien nibh. ⁵⁹⁰ Sed ac ipsum eget enim egestas ullamcorper nec euismod ligula. Curabitur fringilla ⁵⁹¹ pulvinar lectus consectetur pellentesque. Quisque augue sem, tincidunt sit amet feugiat ⁵⁹² eget, ullamcorper sed velit. ⁵⁹³

Sed non aliquet felis. Lorem ipsum dolor sit amet, consectetur adipiscing elit. ⁵⁹⁴ Mauris commodo justo ac dui pretium imperdiet. Sed suscipit iaculis mi at feugiat. Ut ⁵⁹⁵ neque ipsum, luctus id lacus ut, laoreet scelerisque urna. Phasellus venenatis, tortor nec ⁵⁹⁶ vestibulum mattis, massa tortor interdum felis, nec pellentesque metus tortor nec nisl. ⁵⁹⁷ Ut ornare mauris tellus, vel dapibus arcu suscipit sed. Nam condimentum sem eget ⁵⁹⁸ mollis euismod. Nullam dui urna, gravida venenatis dui et, tincidunt sodales ex. Nunc ⁵⁹⁹ est dui, sodales sed mauris nec, auctor sagittis leo. Aliquam tincidunt, ex in facilisis ⁶⁰⁰ elementum, libero lectus luctus est, non vulputate nisl augue at dolor. For more ⁶⁰¹ information, see S1 Appendix. ⁶⁰²

Supporting information ⁶⁰³

S1 Fig. Bold the title sentence. Add descriptive text after the title of the item ⁶⁰⁴ (optional). ⁶⁰⁵

S2 Fig. Lorem ipsum. Maecenas convallis mauris sit amet sem ultrices gravida. ⁶⁰⁶ Etiam eget sapien nibh. Sed ac ipsum eget enim egestas ullamcorper nec euismod ligula. ⁶⁰⁷ Curabitur fringilla pulvinar lectus consectetur pellentesque. ⁶⁰⁸

S1 File. Lorem ipsum. ⁶⁰⁹

S1 Video. Lorem ipsum. Maecenas convallis mauris sit amet sem ultrices gravida. ⁶¹⁰ Etiam eget sapien nibh. Sed ac ipsum eget enim egestas ullamcorper nec euismod ligula. ⁶¹¹ Curabitur fringilla pulvinar lectus consectetur pellentesque. ⁶¹²

S1 Appendix. Lorem ipsum. Maecenas convallis mauris sit amet sem ultrices ⁶¹³ gravida. Etiam eget sapien nibh. Sed ac ipsum eget enim egestas ullamcorper nec ⁶¹⁴ euismod ligula. Curabitur fringilla pulvinar lectus consectetur pellentesque. ⁶¹⁵

S1 Table. Lorem ipsum. Maecenas convallis mauris sit amet sem ultrices gravida. ⁶¹⁶ Etiam eget sapien nibh. Sed ac ipsum eget enim egestas ullamcorper nec euismod ligula. ⁶¹⁷ Curabitur fringilla pulvinar lectus consectetur pellentesque. ⁶¹⁸

Acknowledgments ⁶¹⁹

References

1. Volkow ND, Koob GF, Croyle RT, Bianchi DW, Gordon JA, Koroshetz WJ, et al. The conception of the ABCD study: From substance use to a broad NIH collaboration. Developmental Cognitive Neuroscience. 2018;32:4–7. doi:10.1016/j.dcn.2017.10.002.

2. Luciana M, Bjork JM, Nagel BJ, Barch DM, Gonzalez R, Nixon SJ, et al. Adolescent neurocognitive development and impacts of substance use: Overview of the adolescent brain cognitive development (ABCD) baseline neurocognition battery. Developmental cognitive neuroscience. 2018;32:67–79.

3. Thompson WK, Barch DM, Bjork JM, Gonzalez R, Nagel BJ, Nixon SJ, et al. The structure of cognition in 9 and 10 year-old children and associations with problem behaviors: Findings from the ABCD study's baseline neurocognitive battery. Developmental cognitive neuroscience. 2019;36:100606.

4. Barch DM, Albaugh MD, Avenevoli S, Chang L, Clark DB, Glantz MD, et al. Demographic, physical and mental health assessments in the adolescent brain and cognitive development study: Rationale and description. Developmental cognitive neuroscience. 2018;32:55–66.

5. Zucker RA, Gonzalez R, Ewing SWF, Paulus MP, Arroyo J, Fuligni A, et al. Assessment of culture and environment in the Adolescent Brain and Cognitive Development Study: Rationale, description of measures, and early data. Developmental cognitive neuroscience. 2018;32:107–120.

6. Uban KA, Horton MK, Jacobus J, Heyser C, Thompson WK, Tapert SF, et al. Biospecimens and the ABCD study: Rationale, methods of collection, measurement and early data. Developmental Cognitive Neuroscience. 2018;32:97–106. doi:10.1016/j.dcn.2018.03.005.

7. Casey BJ, Cannonier T, Conley MI, Cohen AO, Barch DM, Heitzeg MM, et al. The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. Developmental Cognitive Neuroscience. 2018;32:43–54. doi:10.1016/j.dcn.2018.03.001.

8. Hagler DJ, Hatton S, Cornejo MD, Makowski C, Fair DA, Dick AS, et al. Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. NeuroImage. 2019;202:116091. doi:10.1016/j.neuroimage.2019.116091.

9. Bagot KS, Matthews SA, Mason M, Squeglia LM, Fowler J, Gray K, et al. Current, future and potential use of mobile and wearable technologies and social media data in the ABCD study to increase understanding of contributors to child health. Developmental cognitive neuroscience. 2018;32:121–129.

10. Loughnan RJ, Palmer CE, Thompson WK, Dale AM, Jernigan TL, Fan CC. Polygenic score of intelligence is more predictive of crystallized than fluid performance among children. bioRxiv. 2020; p. 637512.

11. Salthouse TA. Why are there different age relations in cross-sectional and longitudinal comparisons of cognitive functioning? Current directions in psychological science. 2014;23(4):252–256.

12. Sullivan EV, Brumback T, Tapert SF, Prouty D, Fama R, Thompson WK, et al. Effects of prior testing lasting a full year in NCANDA adolescents: contributions from age, sex, socioeconomic status, ethnicity, site, family history of alcohol or

drug abuse, and baseline performance. Developmental cognitive neuroscience. 2017;24:72–83.

13. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Archives of general psychiatry. 1961;4(6):561–571.

14. French DP, Sutton S. Reactivity of measurement in health psychology: how much of a problem is it? What can be done about it? British journal of health psychology. 2010;15(3):453–468.

15. Robins L. Epidemiology: Reflections on Testing the Validity of Psychiatric Interviews | JAMA Psychiatry | JAMA Network; 1985. Available from: https://jamanetwork.com/journals/jamapsychiatry/article-abstract/493658.

16. Shrout PE, Stadler G, Lane SP, McClure MJ, Jackson GL, Clavél FD, et al. Initial elevation bias in subjective reports. PNAS. 2018;115(1):E15–E23. doi:10.1073/pnas.1712277115.

17. Rogosa D, Brandt D, Zimowski M. A growth curve approach to the measurement of change. Psychological bulletin. 1982;92(3):726.

18. Duncan TE, Duncan SC. The ABC's of LGM: An Introductory Guide to Latent Variable Growth Curve Modeling. Social and Personality Psychology Compass. 2009;3(6):979–991. doi:10.1111/j.1751-9004.2009.00224.x.

19. King KM, Littlefield AK, McCabe CJ, Mills KL, Flournoy J, Chassin L. Longitudinal modeling in developmental neuroimaging research: Common challenges, and solutions from developmental psychology. Developmental Cognitive Neuroscience. 2018;33:54–72. doi:10.1016/j.dcn.2017.11.009.

20. Littlefield AK, King KM, Acuff SF, Foster KT, Murphy JG, Witkiewitz K. Limitations of cross-lagged panel models in addiction research and alternative models: An empirical example using project MATCH. Psychology of Addictive Behaviors;doi:10.1037/adb0000750.

21. Curran PJ, Howard AL, Bainter S, Lane ST, McGinley JS. The Separation of Between-person and Within-person Components of Individual Change Over Time: A Latent Curve Model with Structured Residuals. J Consult Clin Psychol. 2014;82(5):879–894. doi:10.1037/a0035297.

22. Curran PJ, Bauer DJ. The disaggregation of within-person and between-person effects in longitudinal models of change. Annual review of psychology. 2011;62:583–619.

23. Orth U, Clark DA, Donnellan MB, Robins RW. Testing prospective effects in longitudinal research: Comparing seven competing cross-lagged models. Journal of Personality and Social Psychology. 2021;120(4):1013.

24. Hamaker EL, Kuiper RM, Grasman RPPP. A critique of the cross-lagged panel model. Psychological Methods. 2015;20(1):102–116. doi:10.1037/a0038889.

25. Kincaid C. Guidelines for selecting the covariance structure in mixed model analysis, paper 198-30 in Proceedings of the Thirtieth Annual SAS Users Group Conference. Inc, Cary, North Carolina. 2005;.

26. Lenz ST. Alan Agresti (2013): Categorical data analysis. Statistical Papers. 2016;57(3):849.

27. Muthén LK. Mplus user's guide. Los Angeles: Muthén & Muthén; 1998; 2017.

28. Farewell VT, Long DL, Tom BDM, Yiu S, Su L. Two-Part and Related Regression Models for Longitudinal Data. Annual Review of Statistics and Its Application. 2017;4(1):283–315. doi:10.1146/annurev-statistics-060116-054131.

29. Little TD. The Oxford Handbook of Quantitative Methods, Vol. 2: Statistical Analysis. Oxford University Press; 2013.

30. LITTLE RJA, RUBIN DB. The Analysis of Social Science Data with Missing Values. Sociological Methods & Research. 1989;18(2-3):292–326. doi:10.1177/0049124189018002004.

31. Graham JW. Missing Data Analysis: Making It Work in the Real World. Annual Review of Psychology. 2009;60(1):549–576. doi:10.1146/annurev.psych.58.110405.085530.

32. Molnar FJ, Hutton B, Fergusson D. Does analysis using "last observation carried forward" introduce bias in dementia research? Cmaj. 2008;179(8):751–753.

33. Enders CK. Applied Missing Data Analysis. Guilford Press; 2010.

34. Cotter RB, Burke JD, Stouthamer-Loeber M, Loeber R. Contacting participants for follow-up: how much effort is required to retain participants in longitudinal studies? Evaluation and Program Planning. 2005;28(1):15–21.

35. Hill KG, Woodward D, Woelfel T, Hawkins JD, Green S. Planning for long-term follow-up: Strategies learned from longitudinal studies. Prevention Science. 2016;17(7):806–818.

36. Watson N, Leissou E, Guyer H, Wooden M. Best Practices for Panel Maintenance and Retention. In: Advances in Comparative Survey Methods. John Wiley & Sons, Ltd; 2018. p. 597–622. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118884997.ch29.

37. Kraemer HC. Effect size. The encyclopedia of clinical psychology. 2014; p. 1–3.

38. Cohen J. Statistical power. Analysis for the behavioral sciences. 1988; p. 273–406.

39. Morris SB, DeShon RP. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. Psychological methods. 2002;7(1):105.

40. Wang L, Zhang Q, Maxwell SE, Bergeman CS. On standardizing within-person effects: Potential problems of global standardization. Multivariate behavioral research. 2019;54(3):382–403.

41. Feingold A. Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. Psychological methods. 2009;14(1):43.

42. Feingold A. Time-varying effect sizes for quadratic growth models in multilevel and latent growth modeling. Structural equation modeling: a multidisciplinary journal. 2019;26(3):418–429.