Measuring retention within the adolescent brain cognitive development (ABCD)[SM] study

Sarah W. Feldstein Ewing [a,*], Genevieve F. Dash [d], Wesley K. Thompson [b,c], Chase Reuter [b,c], Vanessa G. Diaz [e], Andrey Anokhin [f], Linda Chang [g,h], Linda B. Cottler [i], Gayathri J. Dowling [j], Kimberly LeBlanc [j], Robert A. Zucker [k], Susan F. Tapert [e], Sandra A. Brown [e], Hugh Garavan [l]

[a] Department of Psychology and Interdisciplinary Neuroscience, University of Rhode Island, Chafee Hall, 130 Flagg Road, Kingston, RI 02881, USA
[b] Center for Human Development, University of California San Diego, 9500 Gilman Drive, MC 0115, La Jolla, CA 92093, USA
[c] Herbert Wertheim School of Public Health, University of California San Diego, 9500 Gilman Drive, MC 0115, La Jolla, CA 92093, USA
[d] Department of Psychological Sciences, University of Missouri, 210 McAlester Hall, Columbia, MO 65211, USA
[e] Departments of Psychology and Psychiatry, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA
[f] Department of Psychiatry, Washington University School of Medicine, 660 S. Euclid Ave, Box 8134, St. Louis, MO 63110, USA
[g] Departments of Diagnostic Radiology & Nuclear Medicine, and Neurology, University of Maryland School of Medicine, 655 West Baltimore Street, Baltimore, MD 21201, USA
[h] Department of Neurology Johns Hopkins School of Medicine, 655 West Baltimore Street, Baltimore, MD 21201, USA
[i] Department of Epidemiology, University of Florida, 2004 Mowry Road, Gainesville, FL 32610, USA
[j] Division of Extramural Research, National Institute on Drug Abuse, Bethesda, MD 20892, USA
[k] Department of Psychiatry, University of Michigan, 1301 Catherine Street, Ann Arbor, MI 48109, USA
[l] Department of Psychiatry, University of Vermont, 1 South Prospect Street, MS 446AR7, Burlington, VT 05401, USA

ARTICLE INFO

ABSTRACT

The Adolescent Brain Cognitive Development (ABCD)[SM] study aims to retain a demographically diverse sample of youth and one parent across 21 sites throughout its 10-year protocol while minimizing selective (systematic) attrition. To evaluate the effectiveness of these efforts, the ABCD Retention Workgroup (RW) has employed a data-driven approach to examine, track, and intervene via three key metrics: (1) which youth completed visits late; (2) which youth missed visits; and (3) which youth withdrew from the study. The RW actively examines demographic (race, education level, family income) and site factors (visit satisfaction, distance from site, and enrollment in ancillary studies) to strategize efforts that will minimize disengagement and loss of participating youth and parents. Data showed that the most robust primary correlates of late visits were distance from study site, race, and parental education level. Race, lower parental education level, parental employment status, and lower family income were associated with higher odds of missed visits, while being enrolled in one of the ancillary studies was associated with lower odds of missed visits. Additionally, parents who were primary Spanish speakers withdrew at slightly higher rates. These findings provide insight into future targets for pro-active retention efforts by the ABCD RW.

1. Introduction

The Adolescent Brain Cognitive Development (ABCD) study's recruitment started in Fall 2016 and is now fully into its tenure, in what could likely become one of the most impactful NIH-funded longitudinal studies of adolescent neurocognitive health and development (www.abcdstudy.org). Several factors enhance the potential reach of this study, including the large size and geographically-diverse sample enrolled across 21 sites within the United States (U.S.), and the active interactions and cross-national communication by a team of on-the-

ground study members and national experts. This team, the ABCD Retention Workgroup (ABCD RW), has been closely attending to and monitoring racial, socioeconomic (SES), and demographic factors throughout enrollment and retention to ensure that the sample maintains diversity throughout the longitudinal study (Feldstein Ewing et al., 2018).

Importantly, uneven attrition across sociodemographic groups or with respect to variables of interest may bias the results of longitudinal analyses. As such, carefully overseeing and minimizing such selective/ systematic attrition within this demographically-diverse sample of youth and their families is critical to mitigating potential threats to the study's validity. More specifically, the goal of carefully attending to retention is to prevent non-random attrition (Poulton et al., 2015), which can impact the future generalizability of study findings – particularly if data are systematically late, missing, and/or differentially absent in a certain region and/or within a particular demographic group.

Especially evident during the COVID-19 pandemic (Nooner et al., 2021) is how life events can disrupt children and parents from participating in research. In turn, staying connected with participating youth and their parents during difficult times is especially important and timely to ensure that the study is not missing precisely the over-burdened families whose representation is crucial to ensure. It is critical to make every possible effort to bolster and support the continued engagement and participation of our highest burdened families during the remaining 7–8 years of this study (Feldstein Ewing et al., 2018).

Further, with notable exceptions including by key members of our ABCD RW (Cottler et al., 2016, 2017; Montanaro et al., 2015; Nooner et al., 2021), retaining sociodemographically diverse participants over time in longitudinal research studies has proven difficult for numerous previous studies and many scientific teams. As a consequence, racially, SES-, and demographically-diverse participants remain comparatively underrepresented across many existing large-scale longitudinal research studies.

To protect against omission or differential retention of certain populations (e.g., youth and parents living in poverty; those in rural areas; those who may not trust researchers), which can bias research outcomes and restrict future generalizability (Western et al., 2016), additional efforts are necessary to ensure robust retention. During enrollment, ABCD sites made many targeted efforts to develop strong positive relationships across communities, including actively seeking input and feedback from all members within the consortium and from our community liaison boards (CLBs). These dialogs included, and continue to involve, an open awareness of prior historical research experiences, along with an understanding of how underrepresented communities may be reticent to participate in research. In terms of retention, the ABCD approach has been vigilant of these current and historical research experiences and related perceptions that might interfere with a family's interest in continued engagement (Feldstein Ewing et al., 2018).

1.1. Present study

Through these approaches, and by maintaining connections with youth and their parents, sites have undertaken extensive efforts to ensure that participants stay supported and engaged. While we have made numerous efforts to retain our highly diverse sample, we have also been actively evaluating the impact of our retention efforts on an empirical level via three key metrics found to be relevant in prior large-scale longitudinal studies (late visits, missed visits, and withdrawals) and how these metrics vary by site/study factors (distance to site, MRI review, ancillary study participation), demographics (gender, race, parental language preference, marital status) and socioeconomic indicators (parental education level and employment, family income) (Poulton et al., 2015).

While we have been evaluating and optimizing these metrics during

our monthly RW meetings throughout the course of the study, the RW uses this opportunity to conduct a comprehensive empirical analysis of the ABCD retention efforts thus far. This paper represents the first report of how well our target factors (site factors, demographics, socioeconomic status) have operated in terms of predicting late visits, missed visits, and withdrawals. In close collaboration with ABCD's Data Analysis, Informatics, and Resource Center (DAIRC), we report retention rates and to what degree sociodemographic variables (e.g., gender, race, education level, and family income) and site/study factors (e.g., visit satisfaction, distance from site, enrollment in two ancillary studies [i.e., additional, optional components of the ABCD protocol]) predict frequency of late and missed visits. At this time, rates of withdrawal were too low to model in this examination reliably; as such, only descriptive statistics of group differences between individuals who withdrew vs. remained in the study are presented. With respect to sociodemographic factors, we hypothesized socioeconomic status (e.g., lower parental education level, parental unemployment, lower family income) would be associated with more late visits and missed visits. With respect to study factors, we anticipated that lower visit satisfaction, greater distance from study site, and the added burden of being enrolled in one or more ancillary stud(ies) would be associated with more late visits and missed visits.

2. Materials and methods

For these analyses, we opted to focus on pre-COVID data; in turn, all statistical tests were conducted on data collected up until March 20th, 2020. This timeline was implemented for several reasons, including to accurately capture and represent the standing approach utilized by our team (outside of the pandemic), which will be our primary approach once the pandemic resolves, thereby rendering pre-pandemic statistics more informative of and generalizable to the larger body of retention approaches used for this study. Additionally, at the time that the manuscript was written, COVID was still very much underway, making it too soon to assess the direct and indirect impacts of the pandemic on retention efforts.

2.1. Data

All annual follow-up visits before March 20th, 2020 were included in analyses, including visits at the 1 year, 2 year, and 3 year time points. Annual visits (1 year, 2 year, 3 year) included in-person interviews (3–4 h) with longer interviews and additional imaging components (added 3–4 h) occurring biennially (2 year). Visits were either classified as completed or missing; completed visits done outside of the due date window for the respective visit were classified as late. Visits outside the due date window and not completed were labeled as missing. Due date windows varied by assessment type, with smaller ranges for non-imaging assessments, and larger ranges for imaging years.

Several baseline and time-varying sociodemographic and site/study related factors were included in the analyses. Specific baseline related variables such as participant age, participant gender, participant race, parental language preference, parental education level, parental employment status, parental marital status (married, widowed, divorced, separated, never married or living with a partner), family income, study site, and distance from site were all included in the analyses. In order to facilitate equitably powered analyses, the largest group (married) was used as the index, and compared against the other statuses for analysis. Parental language preference measured a preference for Spanish-speaking. Parental education level measured the highest level of schooling and parental employment status measured the employment status for the parent filling out the questionnaire. Family income measured the total combined family income.

Additionally, time-varying factors were also included, such as visit satisfaction, ABCD ancillary study enrollment (i.e., participation in "add-on" assessments that are not part of the standard assessment

protocol), and review of participant MRI with family. Visit satisfaction is a feedback questionnaire summary score measured at the annual visits (e.g., year 1, 2, and 3); the questionnaire contained parent items such as "I am ok with not knowing the answers to the questions my child answered" and youth items such as "I liked the drink choices," and "I felt comfortable in the scanner" (2 = yes, 1 = don't know, 0 = no). Ancillary study enrollment labels participants according to when they were first included in the particular ancillary study, as well as any subsequent time points after ancillary study enrollment (e.g., a participant first enrolled in an ancillary study at the 2 year visit would also have the 3 year time points labeled as being enrolled in an ancillary study). Review of participant MRI with family denotes whether the MRI report was reviewed with the family (which during this study occurred only in instances where there was the detection of a potential clinical finding during the MRI review), with possibility of referral recommendation or clinical follow-up, and was only measured at the 2 year visit.

Additional variable transformations were used at the aggregated participant-level analyses. Specifically, visit satisfaction was averaged across all time points for participants, and review of participant MRI, as well as ancillary study enrollment, were used as indicators for whether either of these occurred at any point for the participant in the study. These items were included in these analyses as they could each independently impact late and missed visits. Specifically, in the context of participant MRI review, additional stress could emerge from reviewing MRI incidental findings, which could cause families distress. For ancillary study enrollment, additional participation requirements within ancillary study participation could contribute to and/or exacerbate the sense of participant burden.

2.2. Statistical methods

Logistic regression was used to estimate how the predictors of interest influenced the probability of visits being late or missing, at a particular assessment timepoint. All completed visits were included as data in the late visit models, in which a late vs. non-late indicator was used as the dependent variable. For the missing visits models, all completed and missing visits were included, and an indicator of missing vs. completed was used as the dependent variable. Visit satisfaction was omitted in the missing visits models since this data would be empty for missed visits. Separate logistic regression models were fit for each of the outcomes (i.e., late visits and missed visits) and at each of the target time points.

Rates of withdrawal were too low to conduct logistic regressions reliably; as such, only descriptive statistics of group differences between individuals who withdrew vs. remained in the study are presented. Analyses for withdrawn participants were conducted at the aggregated participant-level. Youth were classified into two groups: withdrew at any point in the study or remained in the study. Differences between groups were assessed using t-tests and chi-squared tests on continuous and categorical variables mentioned above.

All analyses were conducted in R (version 3.5.3), including package *tableone*. Bonferroni correction was applied to the logistic regression models with alpha = 0.0042 and to the descriptive tests with alpha = 0.0036. Person-level analyses were two-tailed with alpha = 0.05.

3. Results

Our analytical sample contained 49,529 scheduled visits from the 11,878 youth. Out of these scheduled visits, 1940 (3.9%) were classified as missing. For the 47,589 visits that did occur, 1913 (4%) were classified as late. Out of the 11,878 participants in the study 127 (1.1%) withdrew.

3.1. Late visits

In terms of specific time-varying predictor z-scores and odds ratios

for models predicting late visits, farther distance to study sites was associated with a higher probability of late visits. Significant effects were observed at the 2-year time points, with odds ratio (OR) and 95% confidence interval (CI) model estimates of 1.13 (1.06, 1.21). Associated with a higher probability of late visits were parental education levels of HS/GED degree and some college at the 1-year visit as compared to post-graduation degree: OR (95% CI) were 2.48 (1.52, 4.04). Compared to white participants, Asian participants had a higher probability of a late visit at the 2-year visit: OR (95% CI) = 3.35 (1.79, 6.26). Youth with married parents tended to have a lower probability of late visits at the 2-year visits compared to non-married parents: OR (95% CI) 0.61 (0.44, 0.85). Compared with participants with parents who reported working full or part time, parents who refused to answer the employment question had a higher probability of late visits at the 1-year visits: OR (95% CI) = 10.4 (2.58, 41.89). Effects at the year 3 timepoint had wide confidence intervals, likely due to lower sample size for this timepoint (1098 for year 3 vs. 8246 for year 1 and 6607 for year 2). See Table 1.

3.2. Missed visits

Specific time-varying predictor z-scores and odds ratios for models predicting missed visits are presented in Table 2. Compared to parents with a post-graduate degree, parents with an education lower than HS diploma and HS/GED degree had a higher probability of missed visits at the 1 year visit, OR (95% CI) for: lower than HS = 3 (1.7, 5.29), 3.4 (2.04, 5.68), and 3.07 (1.84, 5.1). HS/GED degree = 2.32 (1.4, 3.83), 3.1 (1.98, 4.85), and 3.3 (2.17, 5.04). Similarly, parents who had attended some college had a higher probability of missed visits at the 1 year visit: OR (95% CI) 2.28 (1.55, 3.34). Compared to white participants, African American participants had a higher probability of missed visits at 1 year visit: OR (95% CI) = 1.82 (1.35, 2.46). Compared to parents working full or part time, stay home parents had a higher probability of missed visits at the 2 year visits: OR (95% CI): 1.91 (1.28, 2.86). As well, participants enrolled in Ancillary Study B had a lower probability of missed visits at the 1 year visit: OR (95% CI) = 0.02 (0, 0.08). Effects at the year 3 timepoint are reporter, but notably have slightly less power in the year 3 models.

3.3. Withdrawal

Withdrawal rates within the study were low (n = 127; 0.01%). Overall, participants who withdrew reported lower visit satisfaction [1.54 (0.40) vs 1.69 (0.35)] and more missed visits [0.48 (0.50) vs 0.06 (0.29)]. Additionally, there was a higher proportion of primaru Spanish speaking parents (11.8% vs 5.4%) and a lower proportion of ancillary study participants (0% vs 21.2%) among those who had withdrawn. Also, participants who withdrew had fewer parents working full or part time (63.8% vs 69.3%), more retired parents (2.4% vs 0.6%), temporarily laid off parents (3.1% vs 0.7%), and parents who were unemployed and not looking for work (2.4% vs 0.7%). Additionally, mean (SD) differences were observed. See Table 3.

4. Discussion

The present study aimed to evaluate factors related to the core metrics implemented by the ABCD Retention Workgroup (RW): late visits, missed visits, and withdrawal from the study. Identifying factors that predict these outcomes at each time point can assist in developing and implementing proactive measures to prevent loss of participating families as well as issues that could later contribute to potential differential attrition. The former is critical to ensuring equitable representation in research, and the latter is critical to preserving the validity of the study.

Consistent with initial hypotheses, lower socioeconomic status (lower parental educational level, parent not employed full/part time), and greater distance to study site were associated with late visits,

Table 1
Odds ratios (95% confidence intervals) for late visits across three years of the ABCD study.

	Year 1 (N = 8246)		Year 2 (N = 6607)		Year 3 (N = 1098)	
Variable (Reference)	OR	95% CI	OR	95% CI	OR	95% CI
Visit Satisfaction	0.79	(0.63,0.98)	0.71	(0.51,1)	0.42	(0.13,1.38)
Distance	1.08	(1,1.18)	**1.13**	**(1.06,1.21)**	1.51	(0.97,2.36)
Age	0.99	(0.97,1.01)	1.01	(1,1.03)	1.13	(1,1.27)
Male (vs. Female)	0.91	(0.68,1.22)	0.85	(0.66,1.1)	1.56	(0.31,7.8)
Married (vs. Unmarried)	0.59	(0.4,0.85)	**0.61**	**(0.44,0.85)**	1.59	(0.25,10.03)
Spanish Preferred (vs. English Preferred)	1.13	(0.55,2.32)	1.11	(0.53,2.32)	0.7	(0.03,16.63)
Race (vs. White)						
Asian	0.7	(0.17,2.92)	**3.35**	**(1.79,6.26)**	*	*
African American	1.18	(0.76,1.82)	0.83	(0.53,1.31)	0.28	(0.01,6.14)
Other/Mixed Race	0.82	(0.53,1.26)	1.02	(0.7,1.48)	4.29	(0.62,29.81)
Parental Education (vs. Post Graduate)						
Less Than HS Diploma	2.98	(1.38,6.42)	1.05	(0.48,2.31)	2.08	(0.06,74.95)
HS/GED	**3.55**	**(1.96,6.42)**	1.62	(0.94,2.81)	1.36	(0.07,24.86)
Some College	**2.48**	**(1.52,4.04)**	1.3	(0.88,1.91)	1	(0.08,12.74)
Bachelor's Degree	1.69	(1.06,2.69)	1	(0.7,1.44)	0.01	(0,28.63)
Parental Employment (vs. Full/Part Time)						
Maternity Leave	*	*	3.66	(0.67,19.87)	*	*
Sick Leave	*	*	*	*	*	*
Temporarily Laid Off	2.09	(0.62,7.11)	0.71	(0.09,5.41)	127.55	(2.27,7172.34)
Stay at Home Parent	1.36	(0.93,2.01)	0.76	(0.51,1.13)	1.33	(0.19,9.15)
Unemployed, Looking	1.41	(0.76,2.61)	0.56	(0.24,1.32)	*	*
Unemployed, Not Looking	0.71	(0.09,5.35)	0.62	(0.08,4.66)	*	*
Student	0.92	(0.28,2.99)	0.33	(0.08,1.38)	*	*
Disabled	0.9	(0.35,2.3)	0.37	(0.09,1.55)	5.69	(0.3,109.26)
Retired	1.39	(0.33,5.92)	0.44	(0.06,3.3)	*	*
Other	1.33	(0.53,3.35)	0.96	(0.38,2.43)	12.39	(0.54,284.43)
Refuse to Answer	**10.4**	**(2.58,41.89)**	8.61	(1.63,45.48)	*	*
Family Income (vs. ≥ $100k)						
< $50 K	0.64	(0.38,1.05)	0.71	(0.46,1.12)	1.8	(0.11,29.97)
≥ $50 K & < $100 K	0.76	(0.51,1.13)	0.84	(0.6,1.19)	0.42	(0.03,5.8)
MRI reviewed (vs. Not Reviewed)	*	*	0.81	(0.32,2.06)	*	*
Ancillary Study Participation (vs. Not Being in the Ancillary Study)						
Ancillary Study A	*	*	1.43	(0.61,3.38)	*	*
Ancillary Study B	0.81	(0.49,1.32)	0.54	(0.33,0.89)	*	*

Note. Odds ratios reflect estimates from logistic regression models predicting missed visits; *Odds ratio and 95% CI reflect either data not collected at that time point, or unstable estimates; Bold indicates significant effects.

although it is important to note that several of the associations were no longer significant at the more stringent Bonferroni-corrected *p*-value. By more conservative estimates, the primary correlates of late visits were distance from study site, child race, parental education level, and an absence of parents answer to the employment status query. Contrary to hypotheses, participants enrolled within one of the two ancillary studies were also less likely to have late visits at Year 2, though this association was no longer significant at the more stringent Bonferroni-corrected p-value threshold. Lower parental educational level and parental employment status were associated with more missed visits, while being enrolled in one of the ancillary studies was associated with fewer missed visits. Greater involvement with the study could serve to foster greater feelings of personal connection to the study team and personal investment in the study itself, subsequently leading to greater show rates. For some families, the additional financial compensation from the ancillary study may also be a meaningful incentive.

Several other notable patterns emerged that may help inform targeted retention strategies for the ABCD RW. First and foremost, participants whose parents declined to provide information about their employment status were more likely to have late visits at year 1. This status may be a signal that this subgroup of parents could benefit from additional relationship-building efforts by ABCD study sites.

Second, participants with stay-at-home parents were more likely to have missed visits at year 2. This may reflect childcare challenges, particularly for families who also have non-participating ABCD siblings at home. In turn, childcare assistance that can be provided by study sites may help minimize missed visits and improve retention among these families (Robinson et al., 2016). Notably, the year 2 visit is longer than the other visits, and thus could present added complications and

scheduling difficulties for families (e.g., in terms of organizing parent time off from work, transportation, and childcare coverage). Although not examined in the presented data examined prior to the pandemic, this situation is likely to be exacerbated during periods where parents are also maintaining additional childcare oversight efforts, such as during the pandemic, when many families have experienced significant shifts in childcare and day-to-day oversight of their children. These data underscore how important it is for sites to maintain close connections with families, in order to address potential life events that may be disrupting show rates and engagement.

Finally, the finding that primarily Spanish speaking parents were more likely to withdraw is an important flag for preventing potential selective attrition of this important ABCD population. Given that the prevalence of Spanish language use often declines with generational status (Portes and Schauffler, 1994; Umana-Taylor et al., 2009), it is possible that primarily Spanish speaking parents may have more recently immigrated wand subsequently may be navigating other additional life obstacles. Because Spanish-speaking communities are largely represented within a handful of study sites, the RW may benefit from developing a collective approach among primary Spanish-speaking sites to faciliate next steps towards enhancing engagement among primary Spanish-speaking families and develop best practices moving forward. Ensuring site staff who are culturally competent, bilingual, and able to connect with Spanish-speaking participants are efforts that many ABCD sites have already undertaken, and will be critical to continue in years forward.

In terms of participant withdrawal from the study, few conclusions can be drawn at this time due to the low rates of overall withdrawal. In general, participants who withdrew had more missed visits and lower

Table 2
Odds ratios (95% confidence intervals) for missed visits across three years of the ABCD study.

	Year 1 (N = 9940)		Year 2 (N = 6810)		Year 3 (N = 1353)	
Variable (Reference)	OR	95% CI	OR	95% CI	OR	95% CI
Distance	1.07	(1.01,1.15)	1.02	(0.91,1.16)	0	-146.03
Age	1	(0.99,1.02)	1	(0.98,1.03)	0.88	(0.75,1.03)
Male (vs. Female)	0.89	(0.72,1.1)	0.83	(0.61,1.13)	11.76	(0.63,220.06)
Married (vs. Unmarried)	0.75	(0.58,0.98)	0.68	(0.47,1)	0.44	(0.04,4.88)
Spanish Preferred (vs. English Preferred)	0.8	(0.5,1.3)	1.44	(0.74,2.81)	*	*
Race (vs. White)						
Asian	0.91	(0.36,2.3)	1.99	(0.8,4.97)	*	*
African American	**1.82**	**(1.35,2.46)**	1.12	(0.68,1.86)	0.47	(0.01,32.26)
Other/Mixed Race	1.3	(0.97,1.73)	0.99	(0.64,1.53)	1.58	(0.09,26.86)
Parental Education (vs. Post Graduate)						
Less Than HS Diploma	**3.4**	**(2.04,5.68)**	0.99	(0.46,2.15)	*	*
HS/GED	**3.1**	**(1.98,4.85)**	1.46	(0.77,2.77)	*	*
Some College	**2.28**	**(1.55,3.34)**	1.27	(0.78,2.06)	*	*
Bachelor's Degree	1.35	(0.92,1.99)	0.75	(0.47,1.2)	*	*
Parental Employment (vs. Full/Part Time)						
Maternity Leave	1.64	(0.44,6.08)	2.67	(0.27,26.48)	*	*
Sick Leave	3.21	(0.63,16.32)	*	*	*	*
Temporarily Laid Off	*	*	0.92	(0.12,7.18)	*	*
Stay at Home Parent	1.42	(1.06,1.9)	**1.91**	**(1.28,2.86)**	0.81	(0.05,12.97)
Unemployed, Looking	1.28	(0.83,1.98)	1.92	(0.95,3.88)	*	*
Unemployed, Not Looking	1.35	(0.45,4.03)	4.46	(1.43,13.96)	*	*
Student	1.97	(1.07,3.65)	1.33	(0.45,3.9)	*	*
Disabled	0.85	(0.44,1.63)	2.02	(0.8,5.14)	*	*
Retired	1.94	(0.73,5.18)	4.64	(1.4,15.39)	*	*
Other	1.37	(0.71,2.65)	1.06	(0.32,3.53)	*	*
Refuse to Answer	2.6	(0.71,9.51)	1.45	(0.15,13.57)	*	*
Income (vs. ≥ $100k)						
< $50 K	1.54	(1.05,2.24)	1.2	(0.7,2.06)	*	*
≥ $50 K & < $100 K	1.09	(0.78,1.52)	1.04	(0.67,1.6)	2.31	(0.18,29.41)
Ancillary Study Participation (vs. Not Being in the Ancillary Study)						
Ancillary Study A	0.68	(0.2,2.26)	0.54	(0.12,2.33)	*	*
Ancillary Study B	**0.02**	**(0,0.08)**	*	*	*	*

Note. Odds ratios reflect estimates from logistic regression models predicting missed visits; *Odds ratio and 95% CI reflect either data not collected at that time point, or unstable estimates; Bold indicates significant effects

mean visit satisfaction, were more likely to have primary Spanish-speaking parents, and were less likely to be enrolled in an ancillary study. However, as was observed in the late and missed visit findings, these are factors that merit continued monitoring and deeper investigation in order to prevent future loss of these families; these data provide the ABCD RW with a critical roadmap for tracking missed visits and other related withdrawal factors to proactively intervene with these families to prevent their loss from the study.

Consistent with past research on retention in studies with parents and children (Robinson et al., 2016), parental education level and employment status emerged as the most consistent and potentially useful indicators of participants at risk for late visits, missed visits, and study withdrawal. These factors are likely intertwined, given the clear connection between parental educational attainment and employment. Participants who had parents with less than a college education were more likely to not attend visits as scheduled, as were participants with stay at home parents. The latter status may have been less consistent and more variable schedules, which could potentially interfere with visit participation and successful attendance (Robinson et al., 2016). Such information can be identified to ensure that these participants are provided with the necessary support to remain engaged in the study.

Considering targets for retention strategies with respect to more burdensome in-person visits, distance from study site and parental employment status may be the most helpful markers for developing protocols to prevent late visits, missed visits, and withdrawal. Distance from study site presents a quantifiable metric by which to identify relative risk for late visits. Feasible approaches to mitigating this barrier may include providing additional appointment reminders prior to the visit, scheduling visits at low-traffic times of day/weekends, and providing assistance with transportation. With respect to parental employment, identifying parents whose schedules may be more variable

and therefore may have more barriers to attendance may be accommodated with implementing virtual visits as necessary, along with childcare for siblings of participants.

4.1. Future directions: examining the impact of the COVID-19 pandemic

The analyses presented in this manuscript targeted retention prior to the onset of pandemic. Arising from closures at testing facilities and guided by a primary concern for the health and well-being of participants, the ABCD study quickly transitioned to virtual assessment procedures during the pandemic. Non-imaging assessments were prioritized as virtual visits early in the pandemic to continue data collection and maintain connection with participating families to help facilitate retention. Imaging assessments in which the participant had MRI contraindications, the family lived far away from the site, or the family did not feel comfortable completing any assessments onsite were also prioritized as virtual visits. Sites then transitioned to hybrid assessments (e.g., questionnaires and cognitive testing completed at home with shorter visits to testing facilities for MRI, biospecimens and anthropometrics) both for imaging and non-imaging assessments. Currently, sites are transitioning back to on-site assessments with COVID safety modifications. The impact of these disruptions on retention is being carefully monitored despite significant challenges (e.g., late visits now largely reflect pandemic-related delays) and increased complexity (e.g., comparisons between on-site, virtual, and hybrid assessments). Fortunately, at this time, we have not detected appreciable increase in withdrawals arising from the pandemic.

However, based on an analysis of the Year 2 assessments (two-thirds of which preceded the pandemic and almost all of which are now completed), we have now begun to observe an increase in missed assessments, from 3% pre-pandemic to 14% post-pandemic. Some of the

Table 3
Predictor differences in participants who remained in the study vs. withdrew.

Variable	Remained in Study (n = 11,751)	Withdrew (n = 127)	
	Mean (SD)	Mean (SD)	*p*-value
# Missed Annual Visits	0.06 (0.29)	0.48 (0.50)	**< 0.001**
Average Visit Satisfaction (range [0,2])	1.69 (0.35)	1.54 (0.40)	**< 0.001**
Distance (kilometers)	24.60 (95.22)	18.22 (30.94)	0.463
Age (months)	118.93 (7.46)	119.96 (7.36)	0.122
	n (%)	n (%)	*p*-value
Male Child	6127 (52.2)	61 (48.0)	0.4
Parental Marital Status	7900 (67.8)	90 (70.9)	0.524
Parental Spanish Preference	636 (5.4)	15 (11.8)	**0.003**
Child's Race			0.077
White	7443 (64.3)	80 (65.0)	
Asian	271 (2.3)	5 (4.1)	
African American	1856 (16.0)	11 (8.9)	
Other/Mixed	2009 (17.4)	27 (22.0)	
Parental Education			0.759
Less Than HS Diploma	773 (6.6)	12 (9.4)	
HS/GED	1246 (10.6)	12 (9.4)	
Some College	3448 (29.4)	38 (29.9)	
Bachelor's Degree	3296 (28.1)	35 (27.6)	
Post Graduate	2965 (25.3)	30 (23.6)	
Parental Employment			**0.017**
Working Full/Part Time	8137 (69.3)	81 (63.8)	
Maternity Leave	30 (0.3)	0 (0.0)	
Sick Leave	21 (0.2)	0 (0.0)	
Temporarily Laid Off	85 (0.7)	4 (3.1)	
Stay at Home Parent	2047 (17.4)	24 (18.9)	
Unemployed, Looking	494 (4.2)	4 (3.1)	
Unemployed, Not Looking	84 (0.7)	3 (2.4)	
Student	242 (2.1)	3 (2.4)	
Disabled	256 (2.2)	3 (2.4)	
Retired	71 (0.6)	3 (2.4)	
Other	228 (1.9)	2 (1.6)	
Refuse to Answer	54 (0.5)	0 (0.0)	
Household Income			0.244
< $50 K	3185 (29.6)	37 (33.3)	
≥ $50 K & < $100 K	3034 (28.2)	36 (32.4)	
≥ $100k	4527 (42.1)	38 (34.2)	
MRI Reviewed	510 (4.3)	6 (4.7)	1
Ancillary Study Participation	2488 (21.2)	0 (0.0)	**< 0.001**

Note. Sociodemographic variables (i.e. Age, Distance, Male, Married, Spanish preferred, Race, Education, Employment, and Household Income), are values at baseline only. Mixed/other race includes American Indian, Alaskan Native, Guamanian, Samoan, Other Pacific Islander, and other races not specified. Average Visit satisfaction averages participants' visit satisfaction across all visits. MRI Reviewed and Ancillary study participant variables are indicators for whether these occurred at any point in the study. # Missed Visits variable is cumulative across the study.
Bold type indicates significance at Bonferroni-corrected p < .0036.

patterns of biased missingness reported above appear to have been exacerbated by the pandemic. While the transition to virtual assessments ensured that data collection and interactions with families has continued, we note that missed fully-virtual assessments appear higher in lower-income families. Challenges with internet connectivity and participant privacy may be relevant factors in this equation. The ABCD RW is continuing to monitor the impact of the COVID-19 pandemic on retention of participants, and this remains an important future direction for the efforts of the ABCD RW.

5. Conclusions

As evidenced by extremely low rates of withdrawal from the study, retention efforts made by the ABCD study thus far appear to have been successful. Fully detailed in recent publications by our team (Feldstein Ewing et al., 2018), to date, our retention efforts have largely revolved around facilitating connection with families via communication and support, including: (a) anticipating families' needs (e.g., offering snacks, childcare, assistance with transportation), (b) being positive and respectful of families (e.g., through prompt and timely compensation/payment after participating), (c) offering care for family members (e.g., ensuring that MRI scans and assessments are on days/times that are mutually convenient for families), and (d) making efforts to actively develop rapport with families (e.g., interacting with families in a positive way so that the families want to return). Our evaluation presented here is promising. To this end, we did not uncover evidence indicating pervasive patterns of selective attrition, particularly in the latter half of the study timepoints presented here (2–3 years), where differences between groups were relatively few in number. However, there remains room for improvement, particularly with respect to preventing selective attrition of primary Spanish-speaking participants, which will be critical to implement within the remaining years of the study.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

As part of the data use agreement, de-identified data is available for this study. Please see: https://nda.nih.gov/abcd.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.dcn.2022.101081.

References

Cottler, L.B., Hu, H., Smallwood, B.A., Anthony, J.C., Wu, L.T., Eaton, W.W., 2016. Nonmedical opioid pain relievers and all-cause mortality: a 27-year follow-up from the epidemiologic catchment area study. Am. J. Public Health 106, 509–516.

Cottler, L.B., Striley, C.W., Elliott, A.L., Zulich, A.E., Kwiatkowski, E., Nelson, D.R., 2017. Pragmatic trial of a study navigator model (NAU) vs. ambassador model (N+) to increase enrollment to health research among community members who use illicit drugs. Drug Alcohol Depend. 175, 146–150.

Feldstein Ewing, S.W., Chang, L., Cottler, L.B., Tapert, S.F., Dowling, G.J., Brown, S.A., 2018. Approaching retention within the ABCD study. Dev. Cogn. Neurosci. 32, 130–137.

Montanaro, E., Feldstein Ewing, S.W., Bryan, A.D., 2015. What works? An empirical perspective on how to retain youth in longitudinal HIV and substance risk reduction studies. Subst. Abus. 36, 493–499.

Nooner, K.B., Chung, T., Feldstein Ewing, S.W., Brumback, T., Arwood-Wenke, Z., Tapert, S.F., Cottler, L.B., 2021. Retaining adolescent and young adult participants in research during a pandemic: best practices from two large-scale developmental neuroimaging studies (NCANDA and ABCD). Front. Behav. Neurosci. 14, 587902.

Portes, A., Schauffler, R., 1994. Language and the second generation: bilingualism yesterday and today. Int. Migr. Rev. 28, 640–661.

Poulton, R., Moffit, T.E., Silva, P.A., 2015. The Dunedin Multidisciplinary Health and Development Study: overview of the first 40 years, with an eye to the future. Soc. Psychiatry Psychiatr. Epidemiol. 50, 679–693.

Robinson, L., Adair, P., Coffey, M., Harris, R., Burnside, G., 2016. Identifying the participant characteristics that predict recruitment and retention of participants to randomised controlled trials involving children: a systematic review. Trials 17, 1–17.

Umana-Taylor, A.J., Alfaro, E.C., Bamaca, M.Y., Guimond, A.B., 2009. The central role of familial ethnic socialization in Latino adolescents' cultural orientation. J. Marriage Fam. 71, 46–60.

Western, B., Braga, A., Hureau, D., Sirois, C., 2016. Study retention as bias reduction in a hard-to-reach population. Proc. Natl. Acad. Sci. USA 113 (20), 5477–5485. https://doi.org/10.1073/pnas.1604138113.