Prevalence of Biochemically-Verified Substance Use in Healthy Adolescents Across the United States: Hair Toxicology Results in the ABCD Study

Natasha E. Wade[1], Yajuan Si[2], Susan F. Tapert[1], Janosch Linkersdörfer[3], Krista M. Lisdahl[4], Hailley R. Moore, Laila Tally, Biplabendu Das[3], Marilyn A. Huestis[5], Alexander L. Wallace[1], Ryan M. Sullivan[1], Veronica Szpak[1], Le Zhang[3], Laura R. Ziemer[3], Wesley K. Thompson[6]

[1]Department of Psychiatry, University of California, San Diego, USA
[2]Survey Research Center, University of Michigan, Ann Arbor, MI, USA
[3]Center for Multimodal Imaging & Genetics, J. Craig Venter Institute, La Jolla, USA
[4]Univeristy of Wisconsin Milwaukee
[5]Huestis & Smith Toxicology, LLC, Fredericksburg, VA 22405
[6]Center for Population Neuroscience and Genetics, Laureate Institute for Brain Research, Tulsa, OK 74136

*Correspondence and reprint requests to:
Natasha E. Wade, Ph.D.
Assistant Professor, Department of Psychiatry
University of California, San Diego
9500 Gilman Drive, MC 0405
La Jolla, CA 92093
email: nwade@ucsd.edu

Acknowledgements: N.E. Wade was supported by National Institute on Drug Abuse (DA050779, PI: Wade; DA054980, MPI: Doran/Jacobus). This work was also supported by T32 AA013525 (PI: Riley/Spadoni to Wallace, Sullivan, and Szpak), DA062011 (PI: Wallace), and DA064409 (PI: Sullivan). Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development™ (ABCD) Study, held in the NIH Brain Development Cohorts Data Sharing Platform. This is a multisite, longitudinal study designed to recruit more than 10,000 children aged 9–10 and follow them over 10 years into early adulthood.
The ABCD Study® is supported by the **National Institutes of Health** and additional federal partners under award numbers:
U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041

134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, U24DA041147.
A full list of supporters is available at Federal Partners – ABCD Study.
ABCD Consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD Consortium investigators. The ABCD data repository grows and changes over time. The ABCD data used in this report came from https://doi.org/10.82525/jy7n-g441. Additional support for this work was made possible from NIEHS R01-ES032295 and R01-ES031074.

Abstract

Objective: To determine correspondence between self-reported substance use and biochemical verification through hair samples and to estimate U.S. prevalence of adolescent substance use.

Methods: Data came from the nationwide Adolescent Brain Cognitive Development Study (n=11,868; age 9-10 at Baseline, age 15-16 at Wave 6). Past-3-month substance use was measured annually from 2016 to 2024. Hair samples objectively detected moderate+ substance use in a subsample of participants ($n_{samples}$=11,865; n=6,133 unique participants). Multi-step weighting methods estimated national prevalence trends of cannabis, alcohol, and nicotine use over time, adjusting for discrepancies in sample representation due to recruitment demography, missed visits, and hair samples testing.

Results: Correspondence between self-report and toxicological data improved with age (ages 11-12=<1%; ages 15-16=45%). Weighted estimates of biochemically verified substance use indicated 7.1% of 15-16 year olds engaged in moderate-to-heavy cannabis use, 0.3% heavily used alcohol, and 4.7% heavily used nicotine.

Conclusions: Youth reported substance use patterns demonstrated improved biochemical verification with age. Biochemical verification reflects substantive cannabis and nicotine use by ages 15-16, supporting combining toxicological and self-report data to improve identification of substance use in youth.

Introduction

Recent scientific efforts across the globe enlist broad recruitment of large, diverse cohorts to investigate physical and mental health risk and resilience factors, across the US[1-3], UK[4-6], Europe[7], Netherlands[8], and Switzerland[9]. In studies of youth and young adults, a common aim is to investigate predictors and consequences of substance use. As substance use represents sensitive information, self-report could be inaccurate in youth recruited into normative studies, especially those with parental involvement[9-11].

Biochemical verification through toxicological samples can inform self-reported substance use, particularly hair analysis, which provides up to 3 months of detection. Initial data from several large, epidemiologically based studies suggest that many adolescents and young adults may underreport substance use. In one Zurich study, hair testing of young adults reveals substance underreporting by 30-60%[9]. In the Adolescent Brain Cognitive Development Study (ABCD), 1,390 samples collected at ages 9-13 out of 11,868 enrolled participants indicated that hair testing could double identification of youth using substances[10]. These data highlight the utility of hair analysis to improve substance use estimate accuracy, which is key as the primary study aims focus on identifying the sequelae of substance use[12]. Due to the costs of hair toxicology, ABCD samples selected for analyses prioritized participants who reported risk characteristics for early substance use[11], with a small portion of typical risk youth randomly selected for hair analyses.

Prioritization of participant characteristics and attrition in longitudinal studies can limit the generalizability of toxicology-based estimates. Probability sampling is the gold standard for representative prevalence estimation. Cohort studies, a crucial type of nonprobability sample, require adjustment of sample representation to match the target population[13,14]. Population

weighting adjusts the influence of individual responses and is often recommended to better represent population norms when implemented using an appropriate external benchmark dataset[15]; accordingly, ABCD population weights are included in annual data releases[15]. While the ABCD sample was recruited through epidemiologically guided processes, it is not directly representative of the U.S., so population weights (computed in conjunction with the American Community Survey [ACS]) can help mitigate selection bias[16-18]. Further, attrition or fluctuating participation in longitudinal cohorts is expected, which may further impact generalizability. ABCD's retention rate is 95%[19], yet not all participants complete each wave of data collection, so accounting for attrition and correcting potential biases through weighting are crucial to generalizability.

Here, we report correspondence between hair testing results and self-reported substance use patterns, and assess the prevalence of positive test results after adjusting sample representation with the ABCD cohort through inverse propensity weighting. To estimate the biochemically verified three-month prevalence of substance use across the ABCD follow-up waves, we compute weighted hair testing results, where we combine the ABCD population weights at Baseline[15] with two sets of newly created weights to sequentially adjust for (1) whether a participant completed or missed a given wave and (2) whether hair samples were successfully assayed by drug test for that wave. We apply the final weight to positive hair results for cannabis, alcohol, and nicotine by wave of data collection, estimating the population prevalence of these three most used substances in youth ages 9-16 across the U.S.

Methods

Data were drawn from ABCD Data Release 6.0 (https://doi.org/10.82525/jy7n-g441), including all annual visits from Baseline (ages 9-10) through half of Wave 6 follow-ups (ages

15-16) in 11,868 youth. Hair testing was conducted on a subset of participants ($n_{samples}$=11,865 from 6,133 unique participants). ABCD participants were enrolled from their local catchment centers through epidemiologically guided school recruitment, with four sites mainly recruiting participants from twin registries[20]. Since one of ABCD's primary aims is to study how substance use influences development[12], the cohort was recruited at an age when they were expected to be substance naïve[21]. Inclusion criteria at Baseline included living within the catchment area; being 9-10 years old and fluent in English; having no history of substantial substance use, severe mental illness, moderate-to-severe intellectual or developmental disability, or history of major neurological disorder including severe traumatic brain injury; birthweight >2 lbs; and no magnetic resonance imaging contraindication. Participants needed an English or Spanish speaking parent/guardian to consent and participate. Eligible participants attend data collection sessions at their local site annually and complete brief mid-year interviews remotely. The present analyses include all participants with data from any visit (n=11,868).

Measures

Hair Sample Collection and Toxicological Analysis. ABCD research assistants collected ~100mg of hair closest to the root around the crown of the head from all participants who had hair styles that would not be negatively impacted by collection and were willing. Hair samples were securely stored in envelopes inside plastic sleeves and placed in locked cabinets until selected for shipping. Upon receipt by the laboratory, hair was measured and trimmed to 3.9 cm, if necessary, and washed to mitigate against environmental exposure contamination, as recommended by the Society of Hair Testing[22]. The hair sample then underwent screening by immunoassays for most analytes and tested directly on mass spectrometry for cannabinoids and ethyl glucuronide (EtG). Presumptive positives were confirmed via liquid chromatography

tandem mass spectrometry (LC-MS/MS) or gas chromatography tandem mass spectrometry (GC-MS/MS). Collection and testing details are in the Supplement.

A full drug panel was assessed; here we report on: (1) cannabinoids (THCCOOH, the definitive indicator of personal ingestion[23]); (2) alcohol (EtG); and (3) nicotine (cotinine). The two testing laboratories employed by ABCD used different cutoffs to determine positive results across drug classes, so to harmonize results across labs, positive results were identified in two ways: for a more conservative estimate ("high threshold"), quantitative results with the higher cutoff between the two labs were used; for a more sensitive estimate ("low threshold"), samples above the limit of detection (LOD) at each laboratory were identified. While the high threshold partially reduces sensitivity as described in prior reports[10], it allows for more direct comparison across labs. LODs and cutoffs are detailed in the Supplement.

Self-Reported Substance Use. Trained research staff annually administered the Timeline Followback (TLFB) interview[24,25] assessing use pattern, frequency, and quantity of all major substance categories in standard units including cannabis (flower, concentrates, edibles, tinctures, synthetics), alcohol, nicotine (tobacco products, electronic nicotine delivery systems or vapes, and nicotine replacement products), and other drug use. Through reviewing a calendar of lifetime use (from annual visit to annual visit) reported in standardized units of use,[21,26] past-3-month substance use can be calculated (e.g., through the ABCDscores package[27]).

Predictors. We fit logistic regression models to predict wave-specific participation and hair testing selection and use the inverse propensities as weighting adjustment factors. Predictors included in logistic regressions for creating the weights were selected based on substantive knowledge and their correlations to ensure no potential confounding factors were missed. Time-varying measures included age, parent- and youth-reported youth substance use and attitudes,

youth mental health, and visit type (in-person, hybrid, or remote). Several time-invariant predictors collected at Baseline were also included: sex at birth, race, ethnicity, household income, highest level of parental education, family history of mental health disorders, the Childhood Opportunity Index, and state laws regarding recreational and medicinal cannabis[28] (see the Supplement for a detailed list).

Statistical Analyses

Correspondence Between Self-Report and Hair Toxicology. Descriptive statistics were computed for participants' self-report and hair toxicology results. As hair testing is not expected to detect single use episodes[29-31] but typically requires at least several uses or more, a threshold of self-reporting at least two use days a month on average was used (i.e., ≥6 use days in past three months; consistent with prior findings for positive cannabinoid hair testing[32]). Self-reported and hair testing rates of cannabis, alcohol, and nicotine use were compared, and analyses of the most reported substance within participants with tested hair samples (cannabis) assessed concordance of self-report and biochemical verification over time. Correspondence rates were unweighted.

Multi-step Weighting. To obtain unbiased estimates of substance use prevalence across time, we adjusted for sample representation to reduce selection bias through population weighting: Baseline weights, created by Heeringa & Berglund[15]; attrition/missing an annual visit (Wave Participation weights); and whether a hair sample was tested (Hair Test weights). Specifically, starting from the ABCD Study Baseline weights, we (1) fit logistic regression models to predict whether individuals participate in the follow-up wave and use the inverse propensity of participation as the weighting adjustment factor for attrition; (2) fit logistic regression models to predict whether hair samples were successfully tested by drug class,

including as covariates selection criteria for testing and other relevant factors (e.g., sex, study site), and use the inverse propensity of being tested as the weighting adjustment factor of testing; and (3) multiplied the three weights, corresponding to Baseline selection, attrition, and hair testing, trimmed weights as necessary, and used these products as the final weights. Note that the final weight for cases with hair tested at Baseline is a product of the Baseline and Hair Test weights. All ABCD individual records with potential relevance for selection for visit completion, hair testing, and substance use risk were considered for inclusion in models for Wave Participation and Hair Test weights, respectively. The Baseline weights were included as a predictor in the logistic regression for computing Wave Participation weights, and both Baseline and Wave Participation weights were included as predictors in the logistic regression for computing Hair Test weights. The weighting algorithms are described in more detail below, with specific variables included in the Supplement.

Baseline Weights. The procedure for computing the ABCD Study Baseline weights has been described previously[15], with details in the Supplement. Their validity depends on whether the Baseline weights can substantially balance Baseline sample enrollment on factors related to prevalence of pertinent study outcomes. Baseline sociodemographic and recruitment information was used to generate propensity weights within the ABCD cohort to reflect national representation for 9-10 year-olds[15]. These weights, available in the ABCD data release and generated leveraging a high-quality external source (i.e., the ACS) that benchmarks the US population, can reduce the effect of selection bias and improve generalizability of ABCD analyses. Here, we assume that the Baseline weights have substantially accounted for selection factors related to both Baseline sample enrollment and this study's outcomes.

Wave Participation Weights. Participation rate at each wave is: Wave 1 = 94.5%; Wave 2 = 92.5%; Wave 3 = 82.1%; Wave 4 = 82.1%; Wave 5 = 74.9%; Wave 6 = 42.6% (not all participants' data are included in Release 6.0 from Waves 5 and 6). We created wave-specific weights using Baseline variables in logistic regressions to predict whether a participant participated in a given wave. Variables included were potential barriers to study visit participation (e.g., youth spending a lot of time in multiple homes; see Supplement). Wave-specific weights for each participant were computed by taking the inverse of the probability of wave participation estimated from the logistic regression.

Hair Test Weights by Substance. We assigned a value of 1 to a successfully assayed hair sample for the examined substance and 0 to those participating at this wave but without hair testing and fit a logistic regression to predict the testing propensity of the substance at each wave. Samples sent in for testing that had insufficient quantity to test were also given a 0 outcome. Predictors included those used to determine which hair samples to prioritize for testing[11] and other variables that may influence substance use, created separately for each substance, given different rates of successful testing and prediction characteristics. Wave- and substance-specific hair testing weights for each participant were computed by taking the inverse of the probability of successful assaying estimated from the logistic regression. Hair Test weights were trimmed at the 95th percentile to prevent the influence of extreme weights.

Calculating Substance Use Prevalence Estimates. For prevalence estimates of the entire cohort over time for any positive hair results, data were restricted to participants who had hair testing results by substance (number of hair samples: $n_{cannabis}$ = 11,607 ; $n_{alcohol}$ = 11,373; $n_{nicotine}$ = 11,720). We used the *survey* package in R[33] to account for the final weights and family clustering effects in the prevalence estimation of any positive hair results by substance-specific

hair testing results (cannabis [THCCOOH], alcohol, and nicotine) by wave. Confidence intervals of weighted estimates were obtained using the logit (Wald-type) method. Sensitivity analyses assessed trimming effects, where results including only the trimmed Hair Test weights were presented, given minimal difference between cannabis models (see **Tables S4, S5**). Finally, in the Supplement, we estimated prevalence of *all* cannabis positives (i.e., THCCOOH, CBD, THC, Δ8-THC) in **Table S6** and prevalence of all THCCOOH positives but *not* accounting for Baseline weights (**Table S7**; i.e., adjusting to the cohort instead of the population), illustrating the crucial roles of weighting.

Missing Predictor Values. Missing predictor values were included as another level for each variable included in the prediction models. Continuous predictors were transformed into categorical variables (e.g., Externalizing T-score ≥65[34]), except for age and weights. Where the missingness category perfect predicted whether or not hair samples were tested, multiple imputation via the *mice* package were employed[35]: cannabis state-level policies ($n_{missing}$=10; 0.01%), number of peers who drink alcohol ($n_{missing}$=986; 1.4%) and get drunk ($n_{missing}$=987; 1.4%), and using drugs for non-medical purposes ($n_{missing}$=901; 1.3%), Externalizing T-score ($n_{missing}$=902; 1.3%), parental education ($n_{missing}$=14; 0.1%), and household income ($n_{missing}$=1,017; 8.6%).

Results

Assayed Hair Sample Characteristics. 11,865 hair samples were assayed from 6,133 participants across all 21 sites, with 1 to 6 samples per participant (1 sample=2,588; 2=1,925; 3=1,103; 4=466; 5 or 6=50); 66.8% of samples were assayed by US Drug Testing Laboratory (USDTL) and 33.2% by Psychemedics. Most assayed samples were from Waves 2 (n=3,394) and 4 (n=2,900; see **Table S8** for details). Insufficient hair quantity was available in 1,046 of samples

to confirm testing of at least one drug analyte following a presumptive positive (see Supplement). For cannabinoid testing, 11,607 were successfully assayed; for alcohol, $n_{samples}$=11,373; for nicotine, $n_{samples}$=11,720 (see Supplement**)**. Hair results across drug classes are available in **Table S9**. Sociodemographics of the tested samples and ABCD cohort, relative to ACS at the same age as Baseline, are presented in **Table 1** (breakdown by Wave in **Table S10**).

Table 1. Sample Characteristics of Tested Hair Relative to the ABCD Cohort and ACS Benchmark Distribution

Characteristic	**Category**	**Hair Tested** N=6,133	**ABCD Cohort** N=11,868	**ACS** 2011-2015
Sex				
	Male	52.4%	52.2%	51.2%
	Female	47.6%	47.8%	48.8%
Race				
	White	63.3%	66.9%	52.4%
	Black	5.6%	3.9%	13.4%
	Asian	1.6%	1.4%	5.9%
	Hispanic	19.1%	17.5%	24%
	Other	10.3%	10.2%	4.2%
Family Income				
	<$25K	10.5%	13.8%	21.5%
	$25K-$49K	11.3%	13.4%	21.7%
	$50K-$74K	11.1%	12.6%	17.0%
	$75K-$99K	12.2%	13.2%	12.5%
	$100K-$199K	31.1%	27.9%	20.5%
	$200K +	13.5%	10.5%	6.8%
	Other/Decline	9.2%	8.5%	--
Region				
	Northeast	19.2%	16.3%	16.3%
	Midwest	15.1%	14.7%	21.6%
	South	29.7%	28.1%	38.0%
	West	36.0%	40.9%	24.1%

Notes: Hair tested reflect Baseline sociodemographic characteristics of participants whose hair sample was tested, regardless of how many waves of hair samples were tested; ABCD Cohort reflects the Baseline characteristics of the full ABCD cohort; ACS reflects the American Community Survey sociodemographics used to derive the ABCD Baseline weights.

Self-Report and Hair Toxicology Comparison. For cannabis products (see **Figure 1a**), 333 (2.8%, unweighted) reported use at least 2 times per month in the past 3 months. Products and amount of use varied, with some participants reporting up to 91 days of smoked cannabis (including vaped flower, smoked flower, concentrate use, and vaped oil), up to 90 days of edible use, and up to 33 days of cannabidiol use. Correspondence between hair results and self-report for cannabinoids ranged from 92.7-95.1%, demonstrating some under-reporting of cannabis use.

Figure 1. Correspondence between hair results and self-report, by drug class.

a)

Cannabis

	+ Hair	**- Hair**
+ Self	2.0-2.0%	0.8-0.8%
- Self	4.1-6.5%	90.7-93.1%

b)

Alcohol

	+ Hair	**- Hair**
+ Self	<0.1%	0.4-1.6%
- Self	0.7-0.7%	98.1-99.3%

c)

Nicotine/Tobacco

	+ Hair	**- Hair**
+ Self	0.6-0.7%	1.7-1.9%
- Self	1.8-1.8%	95.6-95.9%

Notes: All concordance rates are unweighted. + *Hair* indicates: positive hair sample; - *Hair*: negative hair sample; + *Self*: self-reported substance use of 6 use days in the past 3 months; - *Self*: less than 6 use days self-reported. Percentages reflect the 11,865 samples tested across all waves, with percentile range reflecting whether using high and low threshold cut-offs for identifying positive hair results.

For alcohol use (see **Figure 1b**), 83 (0.7%) participants self-reported 6+ alcohol use days in the past 3 months. Correspondence between hair results and self-reported alcohol use ranged from 98.1-99.3%; <10 participants were positive on hair testing and self-reported ≥6 alcohol use days in the past 3 months.

For use of nicotine or tobacco products (see **Figure 1c**), 283 (2.4%, unweighted) self-reported 6+ use days in the past 3 months. Participants reported up to daily nicotine use via cigarettes and e-cigarette/vaping products. Correspondence between hair results and self-reported nicotine or tobacco use ranged from 96.2-96.6%.

Cannabis was most frequently confirmed on hair testing and most frequently reported amongst those with hair samples tested. At Baseline no participants self-reported ≥2 cannabis use days per month in the past 3 months and had hair positive for cannabinoids (see **Table 2)**; by Wave 6, 45% (unweighted) of participants who were positive for cannabinoids self-reported ≥2x/month cannabis use days. Overall concordance (matching hair results and self-report) varies by year, with lower overall concordance in the years with the highest number of hair cannabinoid positive results.

Table 2. Concordance between Hair Cannabinoids and Self-Report by Wave ($n_{samples}$=11,865)

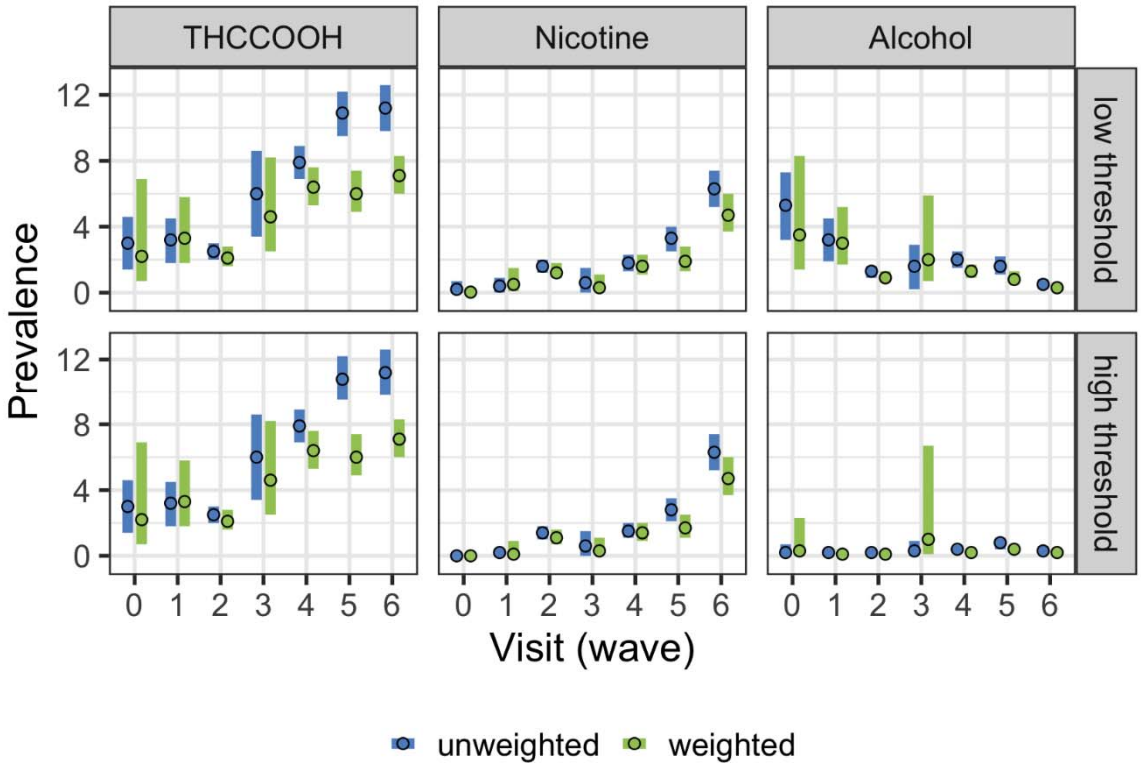
	% Self-Reported Cannabis	**% + Hair Cannabinoids**	**% Concordance**	**% Positive + Concordant/ Cannabinoids**
Baseline	0%	4.3%	95.6%	0%
Wave 1	<1%	6.6%	93.4%	2.0%
Wave 2	<1%	4%	96%	<1%
Wave 3	<1%	13.8%	85.6%	0%
Wave 4	1.9%	9.3%	90.5%	17.8%
Wave 5	5.4%	12.4%	86.2%	32.8%
Wave 6	8.1%	11.5%	85.6%	45.0%

Notes: All concordance rates are unweighted. % Self-Report is the percentage of participants who self-reported cannabis use ≥2x/month by wave; % + Hair Cannabinoids is the percentage of samples tested that were positive for cannabinoids on confirmation testing; % Concordant/All Samples represent percentage of youth whose hair results matched their self-report (either positive hair results and self-reported, or negative hair results and denied use); % Positive + Concordant/Positive Cannabinoids shows that same percentage, but is restricted to those who are concordant who tested positive for cannabinoids. The lower threshold for positive hair testing identification was used here.

Hair Toxicology Results. As can be seen in **Figure 2** (full data in **Table S11**), there is an increasing prevalence rate with age. Weighting largely attenuates estimates, particularly in later waves when participants are older. Consistent with expectations, more strict thresholds for determining positive results ("High Threshold") result in a lower estimated prevalence than "Low Threshold" cutoffs which result in higher prevalence estimates. Variation in confidence intervals likely reflects sample size differences (smaller confidence intervals are observed with larger numbers of tested hair samples) and methodological testing differences, such as the methods used within and between testing labs and the cutoff used to establish positives.

Relative to large, important epidemiological studies, weighted prevalence is more consistent regarding cannabis use but less clearly matched for alcohol and nicotine. Specifically, the wave of ABCD data that most closely corresponds to 8th grade (Wave 4), 6.4% had confirmed cannabis use, 0.2% with alcohol use, and 1.4% with nicotine use, whereas the Monitoring the Future Study[36] found 8th graders reported a past-month rate of 4.3% for cannabis, 1.5% for being drunk, and 4.9% for nicotine. Rates of hair positive for alcohol more closely fit with data indicating that 0.5% of 12-17 year-olds report binge drinking on 5 or more days in the past month in the National Survey on Drug Use and Health[37]. Similarly, in the National Youth Tobacco Survey, around 1.5% of middle schoolers report using e-cigarette products ≥6 days/month[38].

Figure 2. Weighted Prevalence Estimates by Substance and Wave



Notes: Shaded regions represent 95% confidence interval. High threshold refers to a higher cutoff to determine a positive hair test and a more conservative estimate; low threshold refers to a cutoff set at the level of detection (LOD) of the lab to identify a positive hair test, resulting in a more sensitive estimate.

Discussion

Based on current raw and weighted prevalence estimates of biochemically verified substance use, substantive substance use is present in the ABCD cohort, perhaps indicating more frequent use than is captured within large, epidemiological studies that do not have biochemical confirmation. Consistent with prior reports[9,10], our data suggests under-reporting of substance use across adolescents. However, concordance improves with age: while at Wave 2 (ages ~11-12), <1% of participants were positive for hair cannabinoids also self-reported cannabis use at a level which could be detectable on hair testing, by Wave 6 (ages ~15-16) that rate increased to nearly half (45%) of participants. As expected, in our weighting analyses, whether using more

sensitive or conservative thresholds, hair toxicology results indicate increased substance use with age, with estimated prevalence of moderate to heavy cannabis[31] use in 7.1%, heavy alcohol use in 0.3%, and heavy nicotine use in 4.7% of 15-16 year-olds.

Determining the consequences of adolescent substance use requires accurate categorization of substance use onset, and under-reporting of substance use is a significant concern. While studies recruiting specifically for substance use appear to show less underreporting (e.g., Wade et al.[32]), community-based studies of youth (including ABCD) appear to have misreporting[9-11,39]. Results indicate that, when comparing self-report to unweighted hair testing results, concordance increases with adolescent age, perhaps due to increased comfort in disclosure with age, increased knowledge of one's substance use with age, or increased trust with study staff[40-42]. Use of both self-report and objective, biochemical verification through toxicological data are recommended to provide the most accurate estimate of adolescent substance use[43].

Results here are helpful for establishing that non-randomized selection of biochemical verification for substance use is informative for research, though not directly generalizable. Without weighting, unweighted rates underestimate positives in early years (Waves 0-3) and may overestimate prevalence in later years (Waves 4-6). Considering primary (e.g., Figure 2) and follow-up (Table S7 for assessing within-cohort generalizability) weighting models, unweighted hair data are not representative of the broader population or the full ABCD cohort, respectively. Mechanistic studies of specific substances or analytes (e.g., THCCOOH, Δ8-THC) can reliably use hair data to confirm use, while other analyses may choose to incorporate hair data as one data point within a range of measures to identify substance use.

Weighted results for hair cannabinoids map well to other large epidemiological studies, while alcohol and nicotine in hair may be less readily detected. Notably, hair captures a larger window of use (~3 months) than is captured in other large projects like Monitoring the Future, and only detects more intense use across substances[29-32]. Alcohol particularly requires substantial use for detection on hair testing[44,45], explaining the low detection seen here. Therefore, results suggest consistency with epidemiological studies, but also that the low level of use in some self-reported and epidemiological studies (e.g., ≥1 use) may underestimate the severity and frequency of substance use across adolescence.

Some factors may increase noise within these estimates. For instance, the low number of samples tested on some years, use of different labs, or within lab method changes (e.g., using GC-MS/MS for a brief period for EtG) increase noise, demonstrated through increased confidence intervals in weighted estimates (e.g., at Wave 3). Yet the totality of the pattern indicate hair toxicology data are a robust means of enhancing identification of adolescent substance use. Applying weights in this way was previously shown to reveal differential impact on neurocognition and functional imaging[17], and may be used similarly with hair data.

Strengths of the present analyses are use of a benchmark dataset (ACS) for weighting[15], broad consideration of factors potentially influencing hair testing results, and thorough measurement of self-reported and biochemically verified substance use. Limitations include that hair samples cannot be collected from all participants, and methods are key for comfortable collection across participants[46]. Basic drugs are deposited in higher concentrations in hair than acidic or neutral drugs due to strong melanin binding, although cannabinoids do not[47,48]. Two different laboratories analyzed hair samples, with one lab changing EtG methods for a brief

period of time. Considering different cutoffs, LOD, and level of quantitation (LOQ) can help inform interpretation of these results across laboratories.

In conclusion, biochemical verification of substance use data from a diverse, large longitudinal study indicates substantive use of cannabis (7.1%) and nicotine (4.7%) by age 15-16. This is consistent with epidemiological self-report data, and provides novel insight into higher intake levels of use than previously measured in some self-report studies. Combining toxicological data with self-report provides clear indications of more accurate self-report as youth age. Results robustly support use of biochemical verification of substance use is to ensure accurate determination of substance use consequences in youth.

404 1. Lyles CR, Lunn MR, Obedin-Maliver J, Bibbins-Domingo K. The new era of precision population
405 health: insights for the All of Us Research Program and beyond. *J Transl Med.* 2018;16(1):211.
406 2. Jernigan TL, Brown SA, Coordinators AC. Introduction. *Dev Cogn Neurosci.* 2018;32:1-3.
407 3. Jordan CJ, Weiss SRB, Howlett KD, Freund MP. Introduction to the Special Issue on "Informing
408 Longitudinal Studies on the Effects of Maternal Stress and Substance Use on Child Development:
409 Planning for the HEALthy Brain and Child Development (HBCD) Study". *Advers Resil Sci.*
410 2020;1(4):217-221.
411 4. Downey P, Peakman TC. Design and implementation of a high-throughput biological sample
412 processing facility using modern manufacturing principles. *Int J Epidemiol.* 2008;37 Suppl 1:i46-
413 50.
414 5. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the
415 causes of a wide range of complex diseases of middle and old age. *PLoS Med.*
416 2015;12(3):e1001779.
417 6. Golding J, Pembrey M, Jones R, Team. AS. ALSPAC--the Avon Longitudinal Study of Parents and
418 Children. I. Study methodology. *Paediatr Perinat Epidemiol.* 2001;15(1):74-87.
419 7. Mascarell Maricic L, Walter H, Rosenthal A, et al. The IMAGEN study: a decade of imaging
420 genetics in adolescents. *Mol Psychiatry.* 2020;25(11):2648-2671.
421 8. Kooijman MN, Kruithof CJ, van Duijn CM, et al. The Generation R Study: design and cohort
422 update 2017. *Eur J Epidemiol.* 2016;31(12):1243-1264.
423 9. Steinhoff A, Shanahan L, Bechtiger L, et al. When Substance Use Is Underreported: Comparing
424 Self-Reports and Hair Toxicology in an Urban Cohort of Young Adults. *J Am Acad Child Adolesc*
425 *Psychiatry.* 2023;62(7):791-804.
426 10. Wade NE, Sullivan RM, Tapert SF, et al. Concordance between substance use self-report and hair
427 analysis in community-based adolescents. *Am J Drug Alcohol Abuse.* 2023;49(1):76-84.
428 11. Wade NE, Tapert SF, Lisdahl KM, Huestis MA, Haist F. Substance use onset in high-risk 9-13 year-
429 olds in the ABCD study. *Neurotoxicol Teratol.* 2022;91:107090.
430 12. Volkow ND, Koob GF, Croyle RT, et al. The conception of the ABCD study: From substance use to
431 a broad NIH collaboration. *Dev Cogn Neurosci.* 2018;32:4-7.
432 13. Wu C. Statistical inference with non-probability survey samples. *Surv Methodol.* 2022;48:283-
433 311.
434 14. Elliott MR, Valliant R. Inference for Nonprobability Samples. *Statistical Science.* 2017;32(2).
435 15. Heeringa SG, Berglund PA. A Guide for Population-based Analysis of the Adolescent Brain
436 Cognitive Development (ABCD) Study Baseline Data. *bioRxiv.* 2020.
437 16. Gard AM, Hyde LW, Heeringa SG, West BT, Mitchell C. Why weight? Analytic approaches for
438 large-scale population neuroscience data. *Dev Cogn Neurosci.* 2023;59:101196.
439 17. Lin Z, Molloy MF, Sripada C, Kang J, Y. S. Population-weighted Image-on-scalar Regression
440 Analyses of Large Scale Neuroimaging Data. *medRxiv [Preprint].* 2025.
441 18. Si Y, Lee S, Heeringa SG. Population Weighting in Statistical Analysis. *JAMA Intern Med.*
442 2024;184(1):98-99.
443 19. Feldstein Ewing SW, Dash GF, Thompson WK, et al. Measuring retention within the adolescent
444 brain cognitive development (ABCD)(SM) study. *Dev Cogn Neurosci.* 2022;54:101081.
445 20. Garavan H, Bartsch H, Conway K, et al. Recruiting the ABCD sample: Design considerations and
446 procedures. *Dev Cogn Neurosci.* 2018;32:16-22.
447 21. Lisdahl KM, Sher KJ, Conway KP, et al. Adolescent brain cognitive development (ABCD) study:
448 Overview of substance use assessment methods. *Dev Cogn Neurosci.* 2018;32:80-96.
449 22. Cooper GA, Kronstrand R, Kintz P, Society of Hair T. Society of Hair Testing guidelines for drug
450 testing in hair. *Forensic Sci Int.* 2012;218(1-3):20-24.

23. Hill VA, Schaffer MI, Stowe GN. Carboxy-THC in Washed Hair: Still the Reliable Indicator of Marijuana Ingestion. *J Anal Toxicol.* 2016;40(5):345-349.
24. Robinson SM, Sobell LC, Sobell MB, Leo GI. Reliability of the Timeline Followback for cocaine, cannabis, and cigarette use. *Psychol Addict Behav.* 2014;28(1):154-162.
25. Sobell LC, Sobell MB. *Timeline Follow-Back.* Humana Press; 1992.
26. Lisdahl KM, Tapert S, Sher KJ, et al. Substance use patterns in 9-10 year olds: Baseline findings from the adolescent brain cognitive development (ABCD) study. *Drug Alcohol Depend.* 2021;227:108946.
27. Zhang L, Celhay O, Das B, et al. ABCDscores: An R package for computing summary scores in the ABCD Study®. *bioRxiv* 2025:2025.2009.2004.674066.
28. Fan CC, Marshall A, Smolker H, et al. Adolescent Brain Cognitive Development (ABCD) study Linked External Data (LED): Protocol and practices for geocoding and assignment of environmental data. *Dev Cogn Neurosci.* 2021;52:101030.
29. Doran N, Correa JB, Myers MG, Tully L. Associations Between Self-Reported and Biological Measures of Nicotine Consumption Among Young Adult Nondaily Cigarette Smokers. *Am J Addict.* 2020;29(6):471-475.
30. Bertol E, Vaiano F, Boscolo-Berto R, et al. Alcohol, caffeine, and nicotine consumption in adolescents: hair analysis versus self-report. *Am J Drug Alcohol Abuse.* 2017;43(3):341-349.
31. Taylor M, Lees R, Henderson G, et al. Comparison of cannabinoids in hair with self-reported cannabis consumption in heavy, light and non-cannabis users. *Drug Alcohol Rev.* 2017;36(2):220-226.
32. Wade NE, Wallace AL, Baca R, et al. Cannabinoid Profiling Across Toxicology Samples in Adolescents and Young Adults by Route of Administration and in Relation to Depression Symptoms. *Journal of Analytical Toxicology.* 2025:bkaf051.
33. Lumley T. survey: analysis of complex survey samples. *R package version 44.* 2024.
34. Achenbach TM. The Achenbach System of Empirically Based Assessemnt (ASEBA): Development, Findings, Theory, and Applications.2009, Burlington, VT.
35. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software.* 2011;45:1-67.
36. Miech RA, Johnston LD, Patrick ME, O'Malley PM. Monitoring the Future national survey results on drug use, 1975–2023: Overview and detailed results for secondary school students.2024. Located at: Monitoring the Future Monograph Series, Ann Arbor, MI.
37. SAMHSA. Key substance use and mental health indicators in the United States: Results from the 2023 National Survey on Drug Use and Health (HHS Publication No. PEP24□07□021, NSDUH Series H□59).2024.
38. Park-Lee E, Jamal A, Cowan H, et al. Notes from the Field: E-Cigarette and Nicotine Pouch Use Among Middle and High School Students — United States, 2024. *MMWR Morb Mortal Wkly Rep* 2024;73:774-778.
39. Folk JB, Hirschtritt ME, McCrary QD, Kalapatapu RK. Agreement between Youth Self-Report and Biospecimen-Confirmed Substance Use: A Systematic Review. *Subst Use Misuse.* 2022;57(4):531-538.
40. Mensch BS, Kandel DB. Underreporting of substance use in a national longitudinal youth cohort: Individual and interviewer effects. *Public Opinion Quarterly.* 1988;52:100-124.
41. Latkin CA, Edwards C, Davey-Rothwell MA, Tobin KE. The relationship between social desirability bias and self-reports of health, substance use, and social network factors among urban substance users in Baltimore, Maryland. *Addict Behav.* 2017;73:133-136.
42. Johnson TP. Sources of Error in Substance Use Prevalence Surveys. *Int Sch Res Notices.* 2014;2014:923290.

499 43. Fortuna LR, Bagot KS. Editorial: Inconclusive (or "Elusive"): Concordance Between Substance Use
500 Self-report and Toxicology for Diverse Youth Populations. *J Am Acad Child Adolesc Psychiatry.*
501 2023;62(7):721-723.
502 44. Biondi A, Freni F, Carelli C, Moretti M, Morini L. Ethyl glucuronide hair testing: A review. *Forensic*
503 *Sci Int.* 2019;300:106-119.
504 45. Boscolo-Berto R, Favretto D, Cecchetto G, et al. Sensitivity and specificity of EtG in hair as a
505 marker of chronic excessive drinking: pooled analysis of raw data and meta-analysis of
506 diagnostic accuracy studies. *Ther Drug Monit.* 2014;36(5):560-575.
507 46. Wright KD, Ford JL, Perazzo J, et al. Collecting Hair Samples for Hair Cortisol Analysis in African
508 Americans. *J Vis Exp.* 2018(136).
509 47. Huestis MA, Gustafson RA, Moolchan ET, et al. Cannabinoid concentrations in hair from
510 documented cannabis users. *Forensic Sci Int.* 2007;169(2-3):129-136.
511 48. Mieczkowski T. Assessing the potential of a "color effect" for hair analysis of 11-nor-9-carboxy-
512 delta(9)-tetrahydrocannabinol: analysis of a large sample of hair specimens. *Life Sci.*
513 2003;74(4):463-469.

514