Supplementary Appendix

Supplement to: Chen D, Berona J, Chan Y-M, et al. Psychosocial functioning in transgender youth after 2 years of hormones. N Engl J Med 2023;388:240-50. DOI: 10.1056/NEJMoa2206297

This appendix has been provided by the authors to give readers additional information about the work.

Table of Contents

METHODS

Measures

Demographic and Clinical Characteristics

Participants self-reported age, race/ethnicity, gender identity, and designated sex at birth. For age, participants were asked "How old are you?" For race/ethnicity, between the start of the study and May 2018, participants were asked "With which racial or ethnic group do you most closely identify? (Choose one) and provided with the following options: (a) American Indian or Alaska Native; (b) Asian; (c) Black or African American; (d) Hispanic or Latino; (e) Native Hawaiian or Other Pacific Islander; (f) White; (g) Other. After May 2018, participants were asked "What race or ethnicity are you? Check all that apply" and provided with the following options: (a) American Indian or Alaska Native; (b) Asian; (c) Black or African American; (d) Hispanic or Latino; (e) Native Hawaiian or other Pacific Islander; (f) White; (g) other. Those selecting "other" were asked to specify race or ethnicity in free text form. Participant responses were recoded into the following: (a) non-Latinx/Latine White; (b) Latinx/Latine, non-White; (c) Latinx/Latine, White; (d) Black/African American; (e) Asian/Pacific Islander; (f) Multiracial; (g) other; and (h) Unknown.

For gender identity, youth either selected from eight response options [male, female, transgender female (male-to-female), transgender male (female-to-male), gender fluid, gender queer, bigender, or nonbinary] or indicated "other" and specified. Responses were recoded into three categories: transmasculine, transfeminine, and nonbinary. For designated sex at birth, participants were asked "What was your assigned sex at birth?" with male and female as response options.

Longitudinal Outcomes

Appearance Congruence. Appearance congruence was captured through the 9-item appearance congruence subscale of the Transgender Congruence Scale.[1] Each item was rated on a 5-point scale from "strongly disagree" to "strongly agree" and averaged. Example items include: "My outward appearance represents my gender identity" and "I am happy with the way my appearance expresses my gender identity". Higher scores reflect greater appearance congruence.

Depression Symptoms. Depression symptoms were assessed using the 21-item Beck Depression Inventory-II (BDI-II).[2] Each item was rated on a 4-point scale, summed and compared to standardized cutoffs reflecting minimal (0-13), mild (14-19), moderate (20-28), or severe depression symptoms (29-63).

Anxiety Symptoms. Anxiety symptoms were assessed by the Revised Children's Manifest Anxiety Scale, Second Edition (RCMAS2).[3] Forty-nine items were rated "yes"/ "no". "Yes" responses were tallied and transformed into a *T* score; for this scale *T* scores >60 are considered clinically significant.

Positive Affect. Positive affect was assessed using the 10-item Positive Affect measure from the National Institutes of Health (NIH) Toolbox—Emotion Battery.[4] Participants were asked to rate how frequently they experienced a variety of positive feelings over the past seven days. Example items include "I felt joyful" and "I felt content". Each item was rated on a 5-point scale from 1 = "not at all" to 5 = "very much". Raw scores were summed and converted to *T* scores; higher scores indicate greater positive affect.

Life Satisfaction. Life satisfaction was assessed using the 10-item General Life Satisfaction measure from the NIH Toolbox—Emotion Battery.[4] Participants were asked to rate how much they agree or disagree with statements about their personal well-being. Example items

include “If I could live my life over, I would change almost nothing,” “I have what I want in life,” and “My life is going well.” Each item was rated on a 5-point scale from “strongly disagree” to “strongly agree”. Raw scores were summed and converted to *T* scores; higher scores indicate greater life satisfaction.

Rationale for Selecting Primary Mental Health Outcome Measures

The Trans Youth Care—United States (TYCUS) study used various measures to assess different domains of mental health and psychosocial functioning,[1] including the Youth Self-Report (YSR),[2] a widely used child-report measure that assesses problem behaviors along two “broadband scales” (Internalizing, Externalizing) and eight empirically-based syndrome and DSM-oriented scales and provides a Total Problems score, and the age-appropriate version of the MINI International Neuropsychiatric Interview (MINI)[3] or the MINI International Neuropsychiatric Interview for Children and Adolescents (MINI-KID).[4] We chose to use the BDI-II and RCMAS2 as our primary mental health outcome measures in this paper as they are more granular than the YSR and have clinical thresholds that aid in interpretation of findings. Furthermore, the YSR and MINI/MINI-KID were administered annually (baseline, 12-month, and 24-month) versus the BDI-II and RCMAS2 which were administered every 6 months. Having more datapoints to model change across time allowed us to explore whether change in these outcomes were non-linear in nature. Future work using the YSR and MINI/MINI-KID data will allow for comparison across samples, as these measures are widely used among other study teams.[5,6]

Statistical Analysis Plan

Missing Data

At least four out of five total time points were available for 75% of participants (Table S1). As a result, there was high covariance coverage with data available for the majority of the sample for each variable of interest at all time points (range of data present: 0.66-0.99; Table S2). Within our sample, data exhibited skew and were determined to be missing at random (Little's MCAR test: χ^2 [751] = 803.25, p = 0.09).[5,6] This type of missing data can be appropriately handled using maximum likelihood estimation methods (described below).

Longitudinal Modeling Approach

Analyses were conducted in a latent growth curve modeling (LGCM) framework using Mplus 8.8.[7] This approach provides a unified modeling framework with several pertinent computational techniques including specification of hierarchical data structure, accommodation of missing data, and integration of both maximum likelihood and Bayesian estimation techniques. Consistent with NEJM recommendations, we handled missing data using model-based methods.[8] More specifically, LGCM was conducted with a two-stage estimation process in which starting values were generated for parameter estimates using full-information maximum likelihood estimation (FIML) followed by optimization using the Bayes estimator. The Bayes estimator was used in the second stage optimization as it is recommended for use when variables of interest exhibit non-normal distributions.[9,10] Bayesian estimation uses Markov chain Monte Carlo (MCMC) resampling algorithms and do not require large sample sizes.[11,12] These methods accommodate multilevel models that would otherwise be computationally intractable due to small sample sizes, modest effect sizes, and skewed response distributions.[13]

Model Specifications

Latent growth curves were generated for each variable of interest. Linear and quadratic effects of time were explored for inclusion. In all cases, quadratic effects were either non-significant (i.e., confidence intervals included 0) or had small parameter estimates that did not alter interpretation of results. For parsimony, all growth curves included intercepts and linear slopes. Intercept priors were estimated based on median values from observed data. Models employed MCMC algorithms to generate a series of 50,000 random draws from 4 stationary Markov chains to approximate the multivariate posterior distribution of our sample, with a burn-in period of 2,500 iterations. Model convergence was determined by the Gelman-Rubin potential scale reduction factor (PSR) values, with values close to 1 indicating convergence.[14] Trace plots were also inspected to evaluate model fit. All PSR values (range: 1.01-1.03) and trace plots indicated that the models converged and fit the data well.

Table S1. Count of Visits Completed

Visits	n	Proportion present
1	12	0.04
2	27	0.09
3	38	0.11
4	76	0.24
5	162	0.51

Proportion present is out of N=315 eligible participants.

Table S2. Data Coverage for Key Variables

	Baseline		Month 6		Month 12		Month 18		Month 24	
Variable	n	present*	n	present	n	present	n	present	n	present
AC	310	0.98	283	0.90	249	0.79	212	0.67	221	0.70
BDI	307	0.97	281	0.89	248	0.79	210	0.67	219	0.70
RCMAS	308	0.98	282	0.90	248	0.79	209	0.66	216	0.69
NPA	311	0.99	284	0.90	250	0.79	211	0.67	223	0.71
NLS	312	0.99	282	0.90	250	0.79	210	0.67	224	0.71

Note. Proportion present is out of N=315 eligible participants. AC = appearance congruence. BDI = Beck Depressive Inventory. RCMAS = Revised Children's Manifest Anxiety Scale. NPA = NIH Toolbox Positive Affect. NLS = NIH Toolbox Life Satisfaction
*present= proportion present.

Table S3. Comparison of Analytic Sample (n=291) and Participants Excluded from Longitudinal Analysis (n=24)

	t	df	*p*	Cohen's *d*
Baseline Age	0.28	26.27	0.78	0.06
Appearance Congruence	-0.63	25.58	0.54	-0.13
Depression	1.99	22.17	0.06	0.48
Anxiety	1.02	21.42	0.32	0.24
Positive Affect	-0.09	23.07	0.93	-0.02
Life Satisfaction	-1.56	24.03	0.13	-0.35
	c^2	df	*p*	f
Designated sex	0.47	1	0.49	0.04
Early gender-affirming care	0.44	1	0.51	0.04
Racial/ethnic identity	0.002	1	0.97	0.002

Note. For continuous variables, negative *t*-scores and Cohen's *d* indicate higher scores among participants excluded from longitudinal analysis.

Table S4. Representativeness of Study Participants

Category	Example
Disease, problem, or condition under investigation	People who identify as transgender in the U.S.
Special considerations related to:	
Sex and gender	Of the estimated 1.3 million transgender adults, 38.5% are transgender women, 35.9% are transgender men, and 25.6% are nonbinary.
Age	Youth ages 13 to 17 comprise 7.6% of the U.S. population and represent 18% of the transgender population in the U.S. Youth ages 18 to 24 comprise 11% of the U.S. population and represent 24.4% of the transgender population in the U.S. Approximately 1.4% of youth ages 13 to 17 and 1.3% of youth ages 18 to 24 identify as transgender.
Race or ethnic group	The racial/ethnic distribution of youth and adults who identify as transgender appears generally similar to the U.S. population, though transgender youth and adults are more likely to report being Latinx and less likely to report being White compared to the U.S. population. Among youth ages 13 to 17, white youth represent 51.3% of the U.S. population and 46.3% of transgender youth are white. Black youth represent 13.4% of the U.S. population and 13.2% of transgender youth are Black. Asian youth represent 5% of the U.S. population and 3.6% of transgender youth are Asian. American Indian or Alaska Native (AIAN) youth represent 0.8% of the U.S. population and 1% of transgender youth are AIAN. Latinx youth represent 24.8% of the U.S. population and 31% of transgender youth are Latinx. Multiracial youth represent 4.7% of the U.S. population and 5% of transgender youth are multiracial.
Geography	Percentage of residents in U.S. regions who identify as transgender range from 1.8% in the Northeast to 1.2% in the Midwest for youth ages 13 to 17. At the state level, estimates range from 3% of youth ages 13-17 identifying as transgender in New York to 0.6% in Wyoming.
Other considerations	In the last decade, the number of youth presenting for gender-affirming medical care has increased exponentially. In addition, the number of youth reporting a nonbinary identity also has increased significantly in recent years.
Overall representativeness of this trial	Transmasculine participants are over-represented in our study and non-binary participants are under-represented. Non-Latinx white and multiracial participants are over-represented in our sample, whereas Black participants are vastly under-represented in our sample. The proportion of Latinx and Asian participants are

	comparable to population estimates. Because study recruitment occurred at 4 study sites in the Northeast, Midwest, and California, youth in the Southeastern and Southwestern United States are not represented in the sample.

Note. Numbers are predominately pulled from the most recent Williams Institute Executive Summary "How many adults and youth identify as transgender in the United States" published in June 2022 by Jody L. Herman, Andrew R. Flores, and Kathryn K. O'Neill.

Table S5. Paired Samples *t*-tests Comparing Scores at Baseline and 24 Months

	n	baseline	24 Months	*p*-value	effect size
Appearance congruence	213	2.86 (0.74)	3.86 (0.76)	<0.001	-1.12
Depression	211	16.39 (11.88)	13.95 (12.76)	<0.001	0.20
Anxiety	208	60.25 (11.18)	57.38 12.00)	<0.001	0.25
Positive affect	215	42.90 (10.05)	43.72 (12.03)	0.37	-0.05
Life satisfaction	217	39.92 (10.55)	44.61 (12.29)	<0.001	-0.39

Note. Variables are presented as mean (SD). Results are based on *t*-tests (baseline minus 24-months). Negative *t*-test values indicate increases in appearance congruence, positive affect, and life satisfaction. Effect sizes are Cohen's *d* (ranges: 0.20, small; 0.50, medium; 0.80, large).

Table S6. Proportions of Youth Scoring in the Clinical Range for Depression and Anxiety at Each Timepoint

	Baseline	**6-month**	**12-month**	**18-month**	**24-month**
Beck Depression Inventory-II n (%)	*n*=307	*n*=281	*n*=248	*n*=210	*n*=219
Minimal Depression	149 (48.5)	152 (54.1)	143 (57.7)	125 (59.5)	126 (57.5)
Mild Depression	53 (17.3)	46 (16.4)	41 (16.5)	25 (11.9)	41 (18.7)
Moderate Depression	57 (18.6)	43 (15.3)	24 (9.7)	30 (14.3)	22 (10)
Severe Depression	48 (15.6)	40 (14.2)	40 (16.1)	30 (14.3)	30 (13.7)
Revised Children's Manifest Anxiety Scale 2	*n*=308	*n*=282	*n*=248	*n*=209	*n*=216
M (*SD*)	60.0 (11.5)	58.6 (11.6)	58.6 (11.3)	56.8 (11.4)	57.4 (12.1)
n (%) in Clinical range (*T*>60)	181 (58.8)	145 (51.4)	115 (46.4)	90 (43.1)	103 (47.7)

Note. *%* calculated as valid percent using the n for each timepoint as the denominator.

Table S7. Independent Samples *t*-tests Comparing Baseline Scores between Youth Initiating GAH in Early versus Late Puberty

	Total sample	Early gender-affirming care			
		Yes	No		
	N=315	n = 24	n = 291	*p*-value	effect size
Appearance congruence	2.36 (0.88)	3.08 (0.95)	2.31 (0.85)	<0.001	0.86
Depression	16.44 (12.11)	9.57 (8.26)	17.00 (12.21)	<0.001	0.71
Anxiety	60.03 (11.48)	51.54 (12.20)	60.75 (11.15)	<0.001	0.79
Positive affect	43.05 (10.78)	50.27 (12.08)	42.47 (10.49)	<0.001	0.69
Life satisfaction	39.76 (10.85)	44.90 (14.13)	39.35 (10.46)	0.08	0.45

Note. Variables are presented as mean (SD). Results are based on *t*-tests. Effect sizes are Cohen's *d* (ranges: 0.20, small; 0.50, medium; 0.80, large).

Table S8. Independent Samples *t*-tests Comparing Baseline Scores between Youth Initiating GAH in Early versus Late Puberty Among Youth Designated Male at Birth

	DMAB	Early gender-affirming care			
		Yes	No		
	n=111	n = 20	n = 91	*p*-value	Effect Size
Appearance congruence	2.27 (1.03)	3.09 (1.02)	2.10 (0.95)	<0.001	1.00
Depression	17.52 (13.35)	9.41 (8.70)	19.23 (13.56)	<0.001	0.86
Anxiety	59.12 (11.47)	52.30 (11.94)	60.67 (10.85)	0.008	0.73
Positive affect	42.06 (12.68)	51.24 (12.70)	40.14 (11.87)	0.002	0.90
Life satisfaction	38.82 (13.47)	45.71 (15.20)	37.38 (12.71)	0.04	0.59

Note. DMAB = designated male at birth. Variables are presented as mean (SD). Results are based on *t*-tests. Effect sizes are Cohen's *d* (ranges: 0.20, small; 0.50, medium; 0.80, large).

Table S9. Independent Samples *t*-tests Comparing Baseline Scores between Youth Initiating GAH in Early versus Late Puberty among Youth Designated Female at Birth

	DFAB	Early gender-affirming care			
		Yes	No		
	n=204	n = 4	n = 200	*p*-value	Effect Size
Appearance congruence	2.42 (0.78)	3.04 (0.56)	2.40 (0.77)	0.11	0.94
Depression	15.85 (11.36)	10.32 (6.69)	15.96 (11.42)	0.19	0.60
Anxiety	60.52 (11.48)	47.75 (14.66)	60.78 (11.30)	0.17	1.00
Positive affect	43.59 (9.59)	45.65 (8.19)	43.55 (9.62)	0.65	0.24
Life satisfaction	40.27 (9.10)	41.08 (7.43)	40.25 (9.14)	0.84	0.10

Note. DFAB = designated female at birth. Variables are presented as mean (SD). Results are based on *t*-tests. Effect sizes are Cohen's *d* (ranges: 0.20, small; 0.50, medium; 0.80, large).

Figure S1 Conceptual Model of Parallel Process Latent Growth Curve Models

Conceptual model of parallel process latent growth curve models. Rectangles indicate measured variables. Ovals represent model-based estimates of baseline scores (intercepts) and linear rates of change (slopes). Straight arrows indicate regression paths to model (1) moderating effects of baseline covariates on growth curve intercepts and slopes and (2) effects of intercepts on slopes. Curved arrows represent correlations between intercepts and slopes.

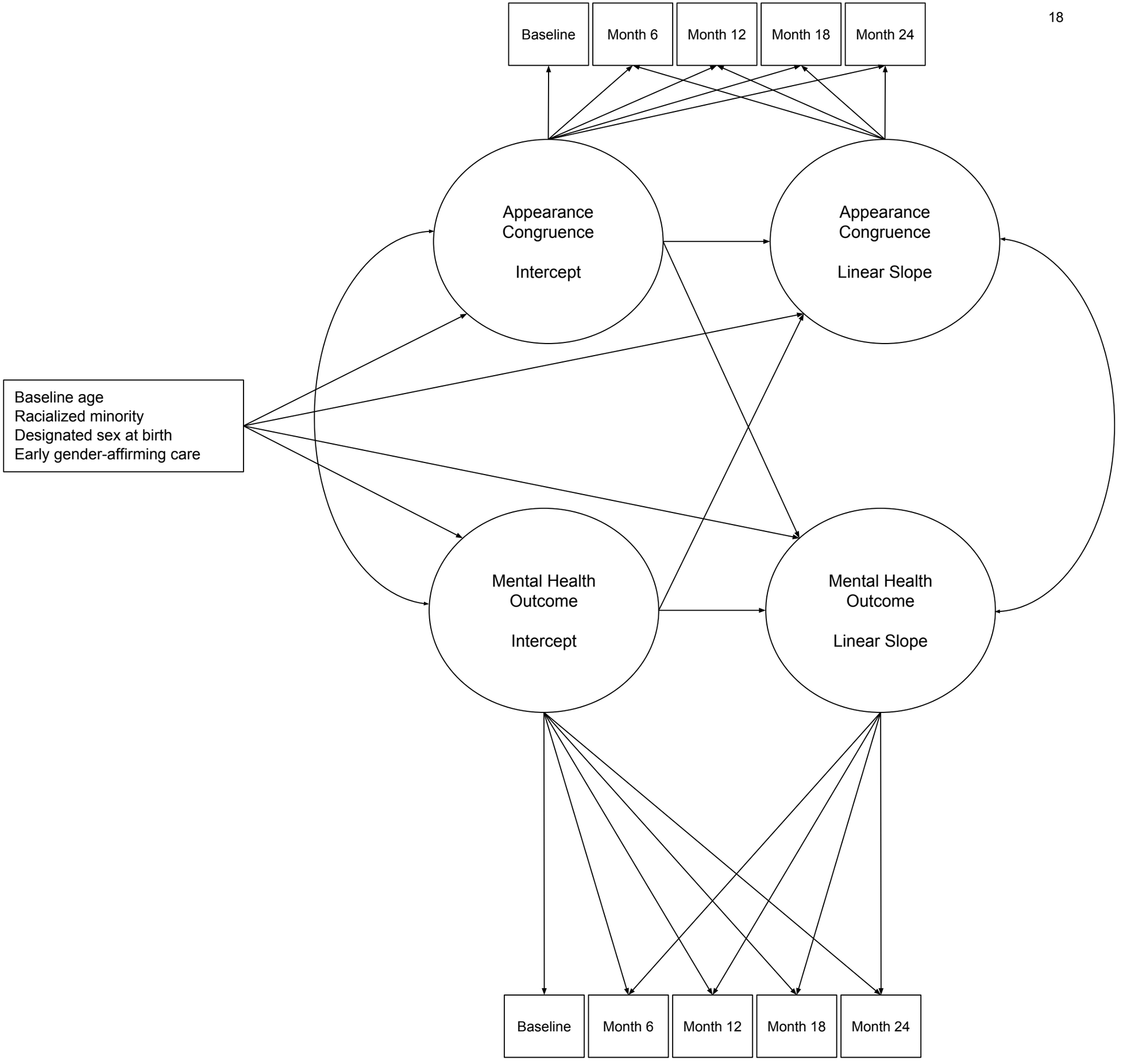

Baseline
Month 6
Month 12
Month 18
Month 24
Appearance Congruence Intercept
Appearance Congruence Linear Slope
Baseline age
Racialized minority
Designated sex at birth
Early gender-affirming care
Mental Health Outcome Intercept
Mental Health Outcome Linear Slope
Baseline
Month 6
Month 12
Month 18
Month 24

Figure S2 Consort Diagram

Flow diagram of the progress through the phases of a prospective, observational study, including enrollment, follow-up, and data analysis for latent growth curve models.

Figure S2. Consort diagram

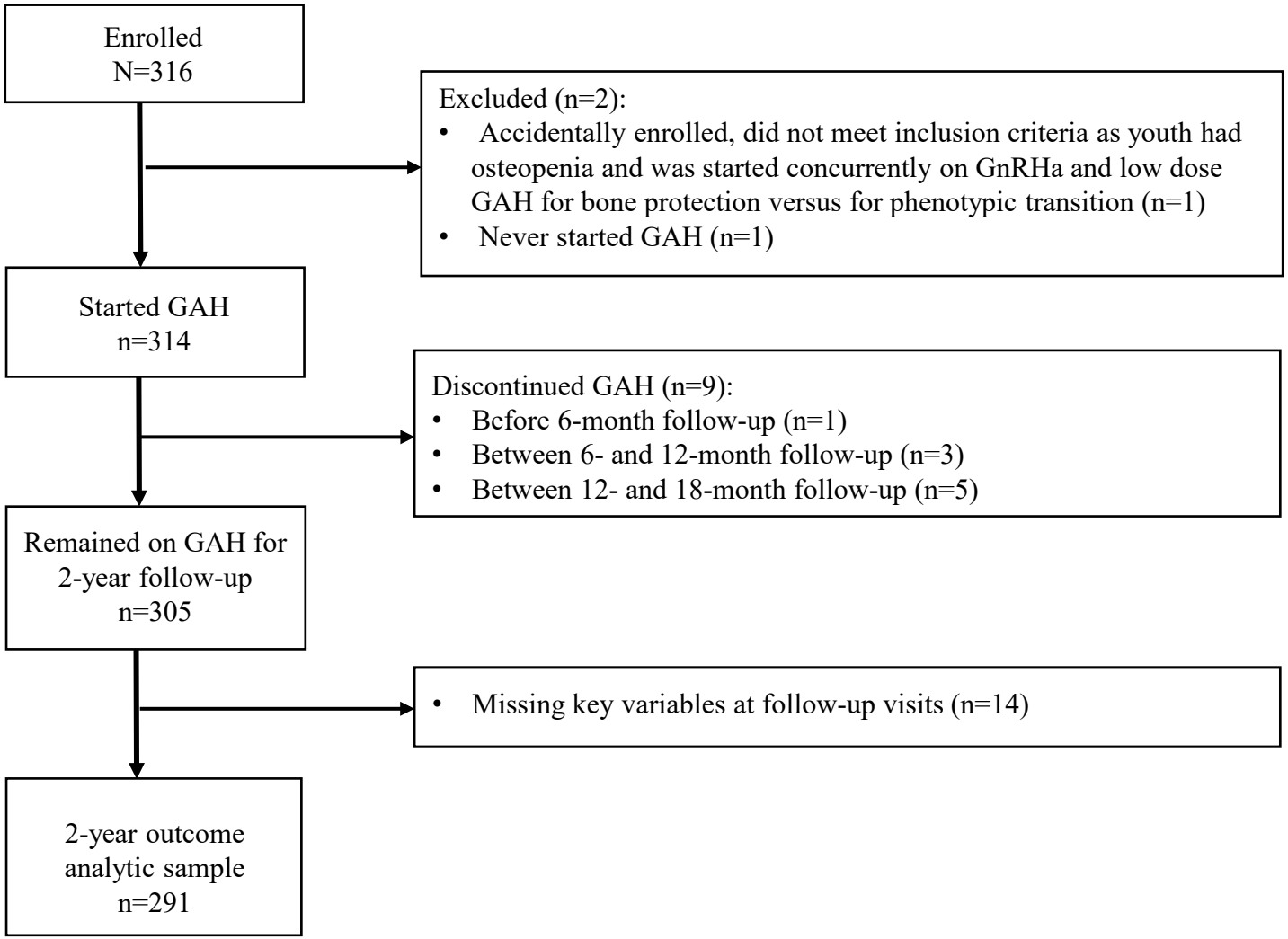
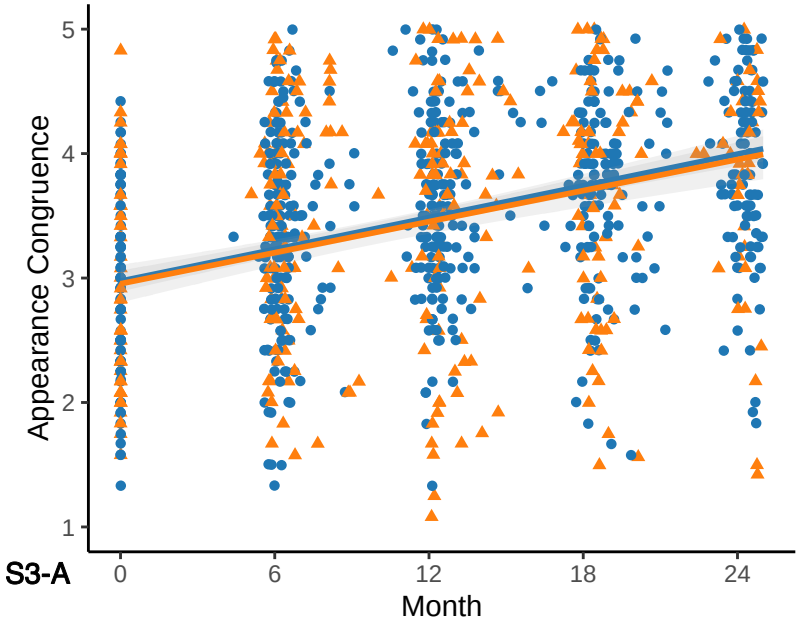
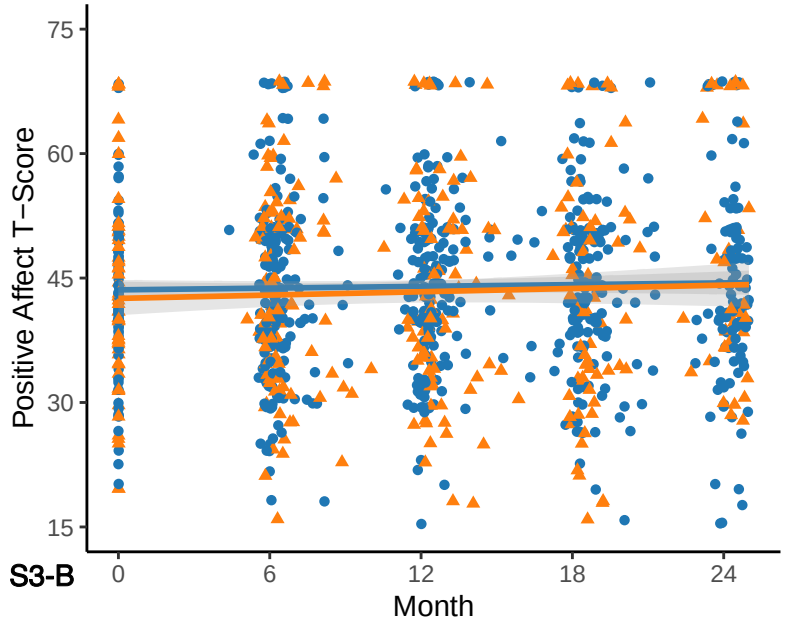
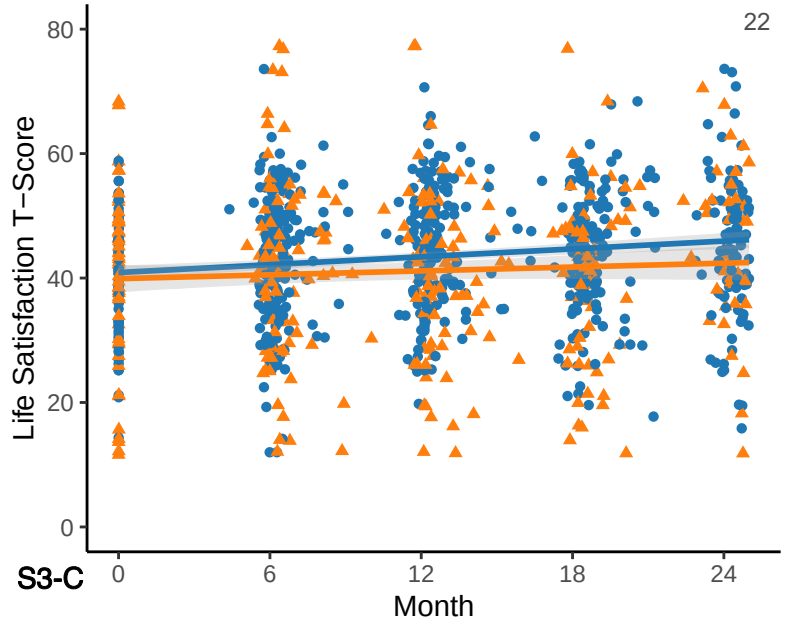


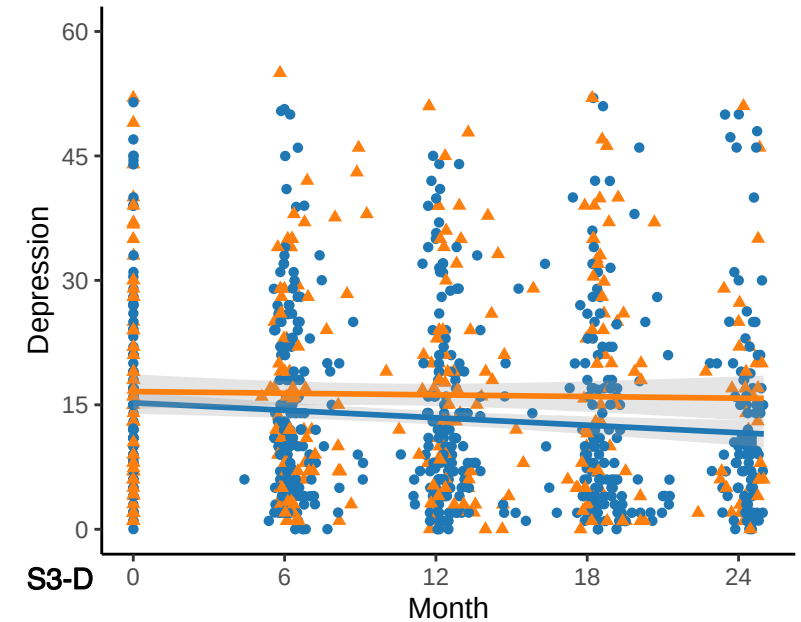
Figure S3 Change in Psychosocial Outcomes by Designated Sex at Birth

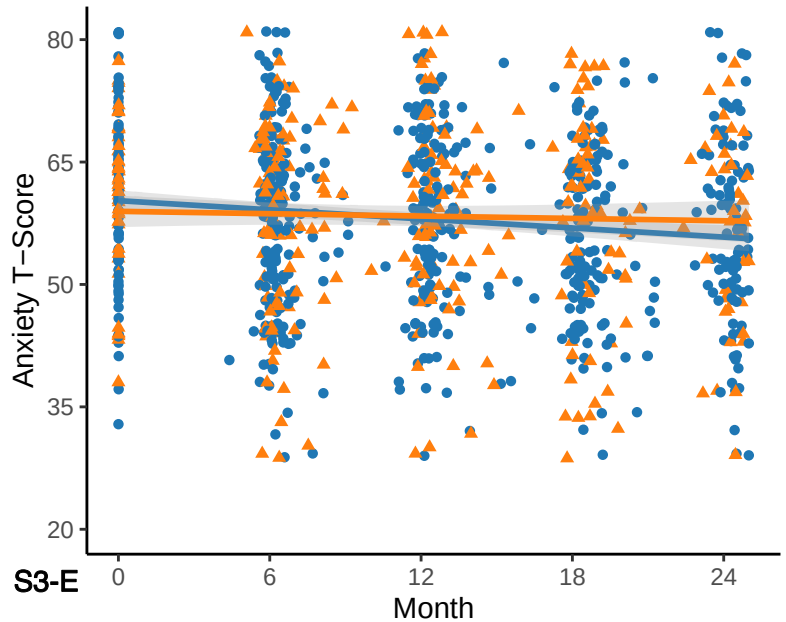
Figure panels display changes in psychosocial outcomes over two years of GAH by designated sex at birth (designated female at birth: blue circles; designated male at birth: orange triangles). Lines indicate mean scores for each group with gray shaded bands for 95% confidence intervals. Outcomes shown are as follows: (S3-A) Transgender Congruence Scale, range: 1-5; (S3-B) Positive Affect Scale T-Score (NIH Toolbox), range: 0-100; (S3-C) Life Satisfaction T-Score (NIH Toolbox), range 0-100); (S3-D) Beck Depression Inventory-II, range: 0-63; (S3-E) Revised Children's Manifest Anxiety Scale, Second Edition T-Score, range: 0-100.


Appearance Congruence
Month
S3-A


Positive Affect T−Score
Month
S3-B


Life Satisfaction T−Score
Month
S3-C


Depression
Month
S3-D


Anxiety T−Score
Month
S3-E

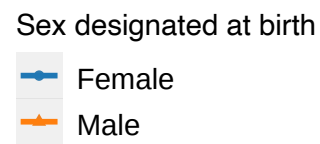
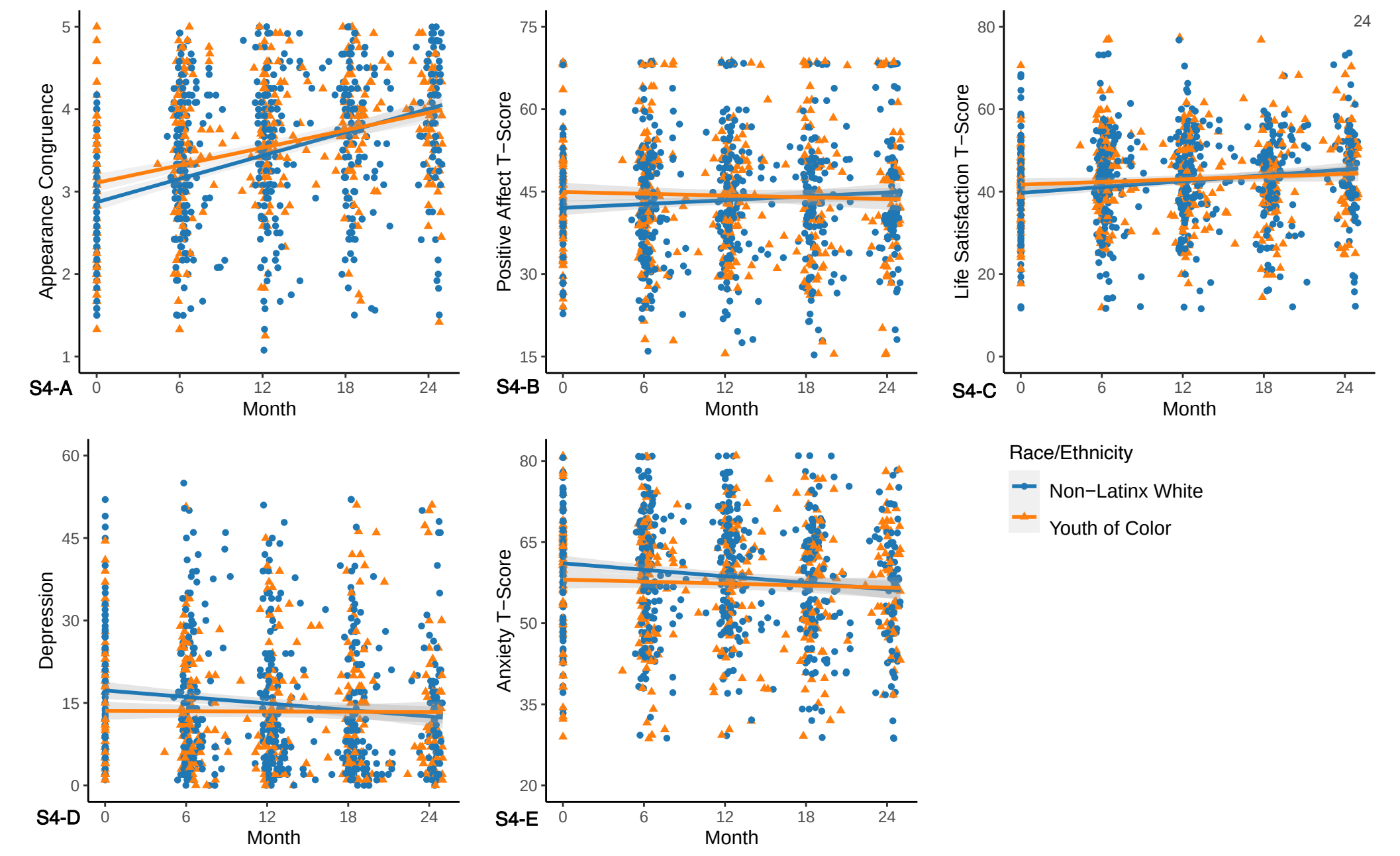

Sex designated at birth
Female
Male

Figure S4 Change in Psychosocial Outcomes by Racial/Ethnic Identity

Figure panels display changes in psychosocial outcomes over two years of GAH by racial/ethnic identity (Non-Latinx White: blue circles; youth of color: orange triangles). Lines indicate mean scores for each group with gray shaded bands for 95% confidence intervals. Outcomes shown are as follows: (S4-A) Transgender Congruence Scale, range: 1-5; (S4-B) Positive Affect Scale T-Score (NIH Toolbox), range: 0-100; (S4-C) Life Satisfaction T-Score (NIH Toolbox), range 0-100); (S4-D) Beck Depression Inventory-II, range: 0-63; (S4-E) Revised Children's Manifest Anxiety Scale, Second Edition T-Score, range: 0-100.


Appearance Congruence
Positive Affect T−Score
Life Satisfaction T−Score
Depression
Anxiety T−Score
Month
S4-A
S4-B
S4-C
S4-D
S4-E
24
Race/Ethnicity
Non−Latinx White
Youth of Color

Published Manuscripts Using TYC Data

1. Olson-Kennedy, J., Chan, Y.M., Garofalo, R., Spack, N., Chen, D., Clark, L., Ehrensaft, D., Hidalgo, M.A., Tishelman, A.C., & Rosenthal, S.M., (2019). Impact of early medical treatment for transgender youth: Protocol for the longitudinal, observational Trans Youth Care study. JMIR Research Protocols, 8(7): e14434.

2. Lee, J.Y., Finlayson, C., Olson-Kennedy, J., Garofalo, R., Chan, Y.M., Glidden, D.V., & Rosenthal, S.M. (2020). Low bone mineral density in early pubertal transgender/gender diverse youth: Findings from the Trans Youth Care Study. *Journal of the Endocrine Society, 4(9):* bvaa065. DOI: 10.1210/jendso/bvaa065.

3. Millington, K., Schulmeister, C., Finlayson, C., Grabert, R., Olson-Kennedy, J., Garofalo, R., Rosenthal, S.M., & Chan, Y.M. (2020). Physiological and metabolic characteristics of a cohort of transgender and gender-diverse youth in the United States. *Journal of Adolescent Health, 67(3):* 376-383. DOI: 10.1016/j.jadohealth.2020.03.028.

4. Chen, D., Abrams, M., Clark, L., Ehrensaft, D., Tishelman, A.C., Chan, Y.M., Garofalo, R., Olson-Kennedy, J., Rosenthal, S.M., & Hidalgo, M.A. (2021). Psychosocial characteristics of transgender youth seeking gender-affirming medical treatment: Baseline findings from the Trans Youth Care Study. *Journal of Adolescent Health, 68(6):* 1104-1111. DOI: 10.1016/j.jadohealth.2020.07.033.

5. Millington, K., Finlayson, C., Olson-Kennedy, J., Garofalo, R., Rosenthal, S.M., & Chan, Y.M. (2021). Association of high-density lipoprotein cholesterol with sex steroid treatment in transgender and gender-diverse youth. *JAMA Pediatrics, 175(5):* 520-521. DOI: 10.1001/jamapediatrics.2020.5620. DOI: 10.1089/trgh.2020.0055.

6. Olson-Kennedy, J., Streeter, L.H., Garofalo, R., Chan, Y.M., & Rosenthal, S.M. (2021). Histrelin implants for suppression of puberty in youth with gender dysphoria: A comparison of 50 mcg/day (Vantas) and 65 mcg/day (SupprelinLA). *Transgender Health, 6(1):* 36-42.

Supplemental Appendix References

1. Kozee HB, Tylka TL, Bauerband LA. Measuring transgender individuals' comfort with gender identity and appearance: Development and validation of the transgender congruence scale. *Psychol Women Q*. 2012;36(2):179-196.

2. Beck AT, Steer RA, Brown GK, others. Manual for the beck depression inventory-II. *San Antonio TX Psychol Corp*. 1996;1(82):10-1037.

3. Reynolds CR, Richmond BO. *Revised Children's Manifest Anxiety Scale*. Western Psychological Services Los Angeles; 1985.

4. Slotkin J, Nowinski C, Hays R, et al. NIH toolbox scoring and interpretation guide. *Wash DC Natl Inst Health*. Published online 2012:6-7.

5. Little RJ. A test of missing completely at random for multivariate data with missing values. *J Am Stat Assoc*. 1988;83(404):1198-1202.

6. Li C. Little's Test of Missing Completely at Random. *Stata J*. 13(4):795-809.

7. Muthén B, Muthén L. Mplus User's Guide. Eighth Edition. Published online 2017.

8. Ware JH, Harrington D, Hunter DJ, D'Agostino RB. Missing Data. *N Engl J Med*. 2012;367(14):1353-1354. doi:10.1056/NEJMsm1210043

9. McNeish D. On Using Bayesian Methods to Address Small Sample Problems. *Struct Equ Model Multidiscip J*. 2016;23(5):750-773. doi:10.1080/10705511.2016.1186549

10. Oravecz Z, Muth C. Fitting growth curve models in the Bayesian framework. *Psychon Bull Rev*. 2018;25(1):235-255. doi:10.3758/s13423-017-1281-0

11. Stefan AM, Oertzen T. Bayesian power equivalence in latent growth curve models. *Br J Math Stat Psychol*. 2020;73(S1):180-193. doi:10.1111/bmsp.12193

12. van de Schoot R, Depaoli S, King R, et al. Bayesian statistics and modelling. *Nat Rev Methods Primer*. 2021;1(1):1. doi:10.1038/s43586-020-00001-2

13. Zondervan-Zwijnenburg M, Depaoli S, Peeters M, van de Schoot R. Pushing the Limits: The Performance of Maximum Likelihood and Bayesian Estimation With Small and Unbalanced Samples in a Latent Growth Model. *Methodology*. 2019;15(1):31-43. doi:10.1027/1614-2241/a000162

14. Depaoli S, Clifton JP. A Bayesian Approach to Multilevel Structural Equation Modeling With Continuous and Dichotomous Outcomes. *Struct Equ Model Multidiscip J*. 2015;22(3):327-351. doi:10.1080/10705511.2014.937849

1. Olson-Kennedy J, Chan Y-M, Garofalo R, et al. Impact of Early Medical Treatment for Transgender Youth: Protocol for the Longitudinal, Observational Trans Youth Care Study. *JMIR Res Protoc.* 2019;8(7):e14434.
2. Achenbach TM. *Manual for the Youth Self-Report/4-18 and 1991 Profiles.* Burlington, VT: Department of Psychiatry: University of Vermont;1991.
3. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry.* 1998;59 Suppl 20:22-33;quiz 34-57.
4. Sheehan DV, Sheehan KH, Shytle RD, et al. Reliability and validity of the Mini International Neuropsychiatric Interview for Children and Adolescents (MINI-KID). *J Clin Psychiatry.* 2010;71(3):313-326.
5. de Vries AL, McGuire JK, Steensma TD, Wagenaar EC, Doreleijers TA, Cohen-Kettenis PT. Young adult psychological outcome after puberty suppression and gender reassignment. *Pediatrics.* 2014;134(4):696-704.
6. Berg D, Edwards-Leeper L, eds. *Child and Family Assessment.* Washington, DC: American Psychological Association; 2018. The Gender Affirmative Model: An Interdisciplinary Approach to Supporting Transgender and Gender Expansive Children.