

Supplementary Table S5 - Gender-related, body image, psychological, psychosocial and cognitive outcomes

Study ID	Country	Study design	Study quality	Treated sample	Average age CSH treatment started	Comparator	Intervention	Outcome (measure)	Follow-up specific to outcome	Study results
<b>Gender dysphoria</b>										
Lopez de Lara 2020	Spain	Cohort	Moderate	23 (16 brf, 7 brm)	Mean age 16 y (range 14-18 y)	30 (no GD)	Hormones (with GnRHa)	Gender dysphoria (UGDS)	12m	Decrease (improvement) in gender dysphoria score. Using clinical cut-off, no gender dysphoria at follow-up compared to all at baseline.
<b>Body image</b>										
Grannis 2021	US	Cross-sectional	Moderate	19 brf	Mean age at data collection 17 y (SD 1.2 y) Mean treatment duration 13.1 m (SD 10.3m)	23	Testosterone	Body image dissatisfaction (BIS)	N/A	Lower levels of dissatisfaction with body image in those treated compared to those not treated.
<b>Psychological / mental health</b>										
<b>Depression</b>										
Lopez de Lara 2020	Spain	Cohort	Moderate	23 (16 brf, 7 brm)	Mean age 16 y (range 14-18 y)	30 (no GD)	Hormones (with GnRHa)	Depressive symptoms (BDI II)	12m	Decrease in depression (~70% in normal range at follow-up) - symptoms higher compared to controls.
Kaltiala 2020	Finland	Pre-post	Moderate	52 (11 brm, 41 brf)	Mean age at diagnosis 18.1 y (SD 1.1y)	N/A	Hormones	Need for treatment due to depression	12m	Need for treatment decreased at follow-up.
Grannis 2021	US	Cross-sectional	Moderate	19 brf	Mean age at data collection 17 y (SD 1.2 y) Mean treatment duration 13.1 m (SD 10.3m)	23	Testosterone	Depressive symptoms (CDI)	N/A	Depressive symptoms were lower in the treated group compared to those not treated.
Green 2022	US	Cross-sectional	Moderate	274 (205 brf, 27 brm, 42 not reported)	Not reported	2961	Hormones	Frequency of depressed mood and anhedonia (PHQ-2)	N/A	Prevalence of recent depression lower in those treated (~61%) compared to those not treated (75%).
<b>Anxiety</b>										
Lopez de Lara 2020	Spain	Cohort	Moderate	23 (16 brf, 7 brm)	Mean age 16 y (range 14-18 y)	30 (no GD)	Hormones (with GnRHa)	Anxiety (STAI (state and trait subscales))	12m	Anxiety decreased over time for both subscales, but continued to be higher compared to controls at follow-up.
Kaltiala 2020	Finland	Pre-post	Moderate	52 (11 brm, 41 brf)	Mean age at diagnosis 18.1 y (SD 1.1y)	N/A	Hormones	Need for treatment due to anxiety	12m	Need for treatment decreased.
Grannis 2021	US	Cross-sectional	Moderate	19 brf	Mean age at data collection 17 y (SD 1.2 y) Mean treatment duration 13.1 m (SD 10.3m)	23	Testosterone	Anxiety disorders (CARED); social anxiety (LSAS)	N/A	Decreased levels of anxiety compared to those not treated.
<b>Self-harm and suicidality</b>										
Allen 2019	US	Pre-post	Moderate	47 (33 brf, 14 brm)	Mean age 16.5 y	N/A	Hormones (plus/minus GnRHa)	Suicidality (Ask Suicide-Screening Questions)	Single (mean 349d, range 113-1016d)	Suicidality scores decreased.
Kaltiala 2020	Finland	Pre-post	Moderate	52 (11 brm, 41 brf)	Mean age at diagnosis 18.1 y (SD 1.1y)	N/A	Hormones	Need for treatment due to suicidality/self-harm	12m	Need for treatment decreased.
Grannis 2021	US	Cross-sectional	Moderate	19 brf	Mean age at data collection 17 y (SD 1.2 y) Mean treatment duration 13.1 m (SD 10.3m)	23	Testosterone	Suicidality (SBQ-R), Lifetime suicidality)	N/A	No evidence for difference between groups in suicidality.
Green 2022	US	Cross-sectional	Moderate	274 (205 brf, 27 brm, 42 not reported)	Not reported	2961	Hormones	Suicidality/non-suicidal self-injury (SBQ-R)	N/A	Treated participants less likely to have seriously considered suicide or attempted suicide compared to untreated.
<b>Need for treatment</b>										

Kaltiala 2020	Finland	Pre-post	Moderate	52 (11 brm, 41 brf)	Mean age at diagnosis 18.1 y (SD 1.1y)	N/A	Hormones	Need for treatment (conduct problems, psychosis, substance abuse, autism, ADHD, eating disorders)	12m	No change in need for treatment for all outcomes, although baseline need for treatment was low.
<b>Internalising problems</b>										
Lopez de Lara 2020	Spain	Cohort	Moderate	23 (16 brf, 7 brm)	Mean age 16 y (range 14-18 y)	30 (no GD)	Hormones (with GnRH <sub>a</sub> )	Emotional symptoms (subscales of SDQ)	12m	Emotional symptoms decreased at follow-up and were similar at follow up to adolescents not experiencing GD.
<b>Externalising problems</b>										
Lopez de Lara 2020	Spain	Cohort	Moderate	23 (16 brf, 7 brm)	Mean age 16 y (range 14-18 y)	30 (no GD)	Hormones (with GnRH <sub>a</sub> )	Conduct problems and hyperactivity (SDQ subscales)	12m	Conduct problems and hyperactivity decreased and were lower at follow-up compared to adolescents not experiencing GD.
<b>Strengths and Difficulties</b>										
Lopez de Lara 2020	Spain	Cohort	Moderate	23 (16 brf, 7 brm)	Mean age 16 y (range 14-18 y)	30 (no GD)	Hormones (with GnRH <sub>a</sub> )	Mental health (SDQ)	12m	Difficulties decreased with fewer in borderline/abnormal range at follow-up, and were similar at follow up to adolescents not experiencing GD.
<b>Psychosocial</b>										
<b>Family functioning and peer relationships</b>										
Lopez de Lara 2020	Spain	Cohort	Moderate	23 (16 brf, 7 brm)	Mean age 16 y (range 14-18 y)	30 (no GD)	Hormones (with GnRH <sub>a</sub> )	Family functioning (FAT), peer problems (SDQ), prosocial (SDQ)	12m	No change in family functioning or peer problems but a small improvement in prosocial skills. At follow-up, family functioning similar to adolescents not experiencing GD, but fewer peer problems and more prosocial skills (small significant difference).
Kaltiala 2020	Finland	Pre-post	Moderate	52 (11 brm, 41brf)	Mean age at diagnosis 18.1 y (SD 1.1y)	N/A	Hormones	Peer relationships (normative vs restricted)	12m	Small decrease in proportion with normative peer relationships.
<b>Wellbeing</b>										
Allen 2019	US	Pre-post	Moderate	47 (33 brf, 14 brm)	Mean age 16.5 y	N/A	Hormones (plus/minus GnRH <sub>a</sub> )	Wellbeing (PQLI General Wellbeing Scale)	Single (mean 349d, range 113-1016d)	Increase in wellbeing.
<b>Other</b>										
Kaltiala 2020	Finland	Pre-post	Moderate	52 (11 brm, 41brf)	Mean age at diagnosis 18.1 y (SD 1.1y)	N/A	Hormones	Living arrangements, school / work participation, romantic involvement, competence in managing everyday matters	12m	Decrease in participants living with parents/guardians. No change in other outcomes.
<b>Cognitive</b>										
Beking 2020	Netherlands	Cohort	Moderate	21 brf	Mean age 16.1 y (SD 0.7)	41 (no GD)	Testosterone (with GnRH <sub>a</sub> )	Lateralisation index of the amygdala (during emotional face matching task with angry and fearful faces)	Single (mean 9.8m, range 5.6-14.8)	Mean lateralisation index decreased (more rightward), and at follow up was similar to controls.
Burke 2016	Netherlands	Cohort	Moderate	21 brf	Mean age at data collection 16.1 y (SD 0.8) Mean treatment duration 10 m (range 6-15)	41 (no GD)	Testosterone (with GnRH <sub>a</sub> )	Visuospatial working memory (brain activation pattern, performance on a mental rotation task)	Single (mean 10m, range 6-15)	No difference between groups in performance. The treated group and male controls showed stronger activations at follow-up.
Grannis 2021	US	Cross-sectional	Moderate	19 brf	Mean age at data collection 17 y (SD 1.2 y) Mean treatment duration 13.1 m (SD 10.3m)	23	Testosterone	Amygdala activation (face processing task)	N/A	Higher activation in left amygdala and cluster of stronger connectivity in right amygdala in treated compared to the controls.
Strang 2022	US	Cross-sectional	Moderate	52 (brs not reported)	Received GAH <1y: mean age 16.7 (SD 1.7) Received GAH >1y: mean age 16.1 (SD 1.4)	72	Hormones (plus/minus GnRH <sub>a</sub> )	Executive functioning (BRIEF Global Executive Composite)	N/A	Treated groups (<1y and >1y treated) had better executive functioning, cognitive flexibility and working memory compared to untreated.

---

Abbreviations: ADHD - Attention Deficit Hyperactivity Disorder; BDI - Beck's Depression Inventory; BIS - Body Image Scale; brf - birth-registered females; brm - birth-registered males; BRIEF - Behaviour Rating Inventory of Executive Function; CARED - Child Anxiety Related Emotional Disorders; CDI - Children's Depression Inventory; FAT - Family Apgar Test; GD - gender dysphoria; GnRH<sub>a</sub> - gonadotrophin releasing-hormone analogues; LSAS - Liebowitz Social Anxiety Scale; m = months; PHQ - Patient Health Questionnaire; PQLI - Paediatric Quality of Life Inventory; SBQ-R - Suicidal Behaviours Questionnaire - Revised; SD - standard deviation; SDQ - Strengths and Difficulties Questionnaire; STAI - State-Trait Anxiety Inventory; UGDS - Utrecht Gender Dysphoria Scale; y = years