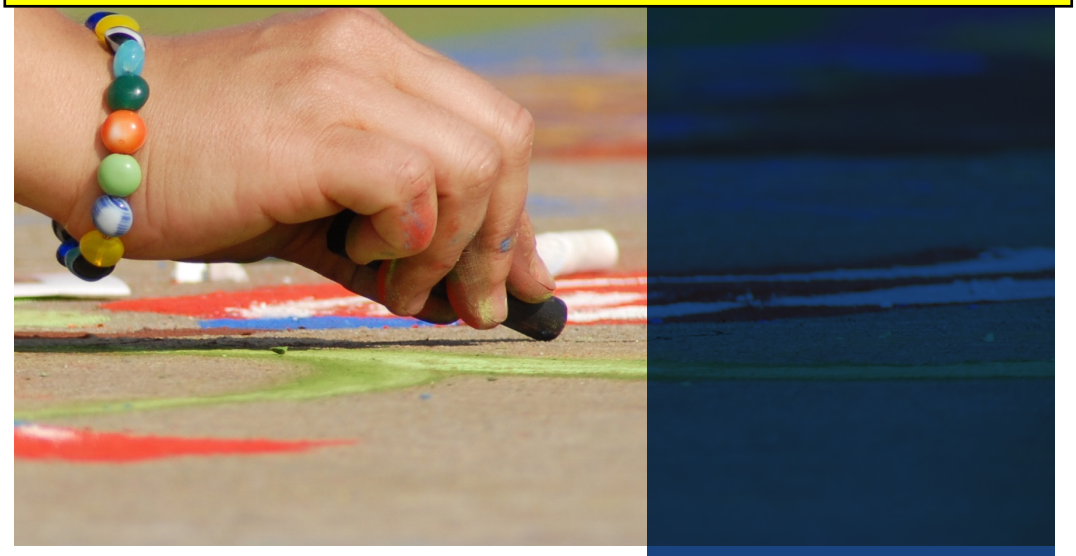
Treatment for Pediatric Gender Dysphoria

Appendix 4: Overview of Systematic Reviews

Methodology, Evidence Synthesis, Tables

Evidence tables on outcomes for PBs and CSHs in youth with gender dysphoria:
Tables 5.2 & 5.3
Tables 6.2 & 6.3



Department of Health and Human Services

November 19, 2025

5.12 Evidence tables on outcomes for PBs in youth with gender dysphoria

Table 5.2. Summary of evidence on gender dysphoria after PBs

Study ID	Included studies (study design)	Number of participants	Conclusion of the review	Certainty of evidence
Dopp 2024[151]	11 studies	-[152]	PBs lead to improved gender dysphoria. However, reductions in gender dysphoria were measured less consistently than pubertal changes, and almost never in the same study as pubertal changes	Very low certainty due to very serious risk of bias, very serious imprecision, and serious indirectness
Ludvigsson 2023[153]	2 prospective observational cohort studies	145 participants on hormones, 49 were evaluated	Two of these studies observed no statistically significant change in gender dysphoria	Very low certainty due to very serious risk of bias and very serious imprecision
Miroshnychenko 2025a[154]	2 before-after studies	59 participants	Standardised mean change: -0.1 (95% confidence interval: -0.4 to 0.19) (higher score indicating greater gender dysphoria)	Very low certainty due to serious inconsistency and very serious imprecision
Taylor 2024a[155]	2 before-after studies[156]	114 participants were treated	Two before-after studies measured gender dysphoria and reported no change before and after receiving treatment	No high-quality studies using an appropriate design were identified, and only four measured gender dysphoria as an outcome. No conclusions can be drawn
Evidence synthesis			No conclusion can be drawn on the impact of PBs on gender dysphoria.	The certainty of evidence is very low.

[151] Dopp et al. (2024).
[152] The number of participants is not calculated because this review included heterogeneous types of studies, including studies that were typically not included in a SR assessing treatment effect, for example, individual case studies (e.g., Pang et al. 2020), qualitative studies (e.g., Vrouenraets et al. 2022). One study included by this review, Grannis 2023, was about CSH not PBs.
[153] Ludvigsson et al. (2023).
[154] Miroshnychenko, Roldan, et al. (2025).
[155] Taylor, Mitchell, Hall, Heathcote, et al. (2024); This review only synthesized excluded low quality.
[156] Though four studies reported gender dysphoria, only two moderate quality before-after studies were included in evidence synthesis.

Table 5.3. Summary of evidence on mental health and well-being after PBs

Study ID	Included studies (study design)	Number of participants	Conclusion of the review	Certainty of evidence
Global function				
Dopp 2024[157]	21 studies	-[158]	Receipt of PBs was sometimes associated with improvements in mental health symptoms (depression, anxiety, quality of life, general functioning). In samples that started psychosocial supports prior to starting PBs, however, there were generally no changes in mental health outcomes after PBs began. Effects were also found less consistently for decreases in suicidality, potentially due to the outcome being relatively rare (which makes finding effects more difficult), especially in the younger age range seen in many of these studies.	Very low certainty due to very serious risk of bias, serious imprecision, inconsistency, and indirectness
Ludvigsson 2023[159]	Four observational cohort studies: one prospective and three retrospective	254 participants on hormones,[160] 113 were evaluated	Improved global function as assessed with the Children's Global Assessment Scale	Very low certainty[161] due to very serious risk of bias and very serious imprecision

[157] Dopp et al. (2024).
[158] The number of participants is not calculated because this review included heterogeneous types of studies, including studies that were typically not included in a SR assessing treatment effect, for example, individual case studies (e.g., Pang et al. 2020). One study included by this review, Grannis 2023, was about CSH not PBs.
[159] Ludvigsson et al. (2023).
[160] In two studies, the participants had also received CSH.
[161] Ludvigsson et al. concluded the review "cannot assess certainty of evidence," while according to the downgrading process in the review, the certainty of evidence was very low.

Study ID	Included studies (study design)	Number of participants	Conclusion of the review	Certainty of evidence
Miroshnychenko 2025a[162]	Two comparative observational studies, and two before-after studies	103 participants from two comparative observational studies, and 53 participants from two before-after studies	Difference in mean change from baseline 7.67 (95% confidence interval: -2 to 17.34) in PBs group compared with no PBs group from comparative observational studies; and Mean change 3.63 (95% confidence interval: 3.17 to 4.09) from before-after studies (higher score indicating greater global function)	Very low certainty due to serious risk of bias, inconsistency and imprecision
Taylor 2024a[163]	One cohort and two before-after studies	174 participants were treated	Three before-after studies reported no clinically significant change	No conclusions can be drawn.
Suicidality				
Ludvigsson 2023[164]	One prospective observational cohort study with mixed treatment (38 subjects with no pharmacological treatment)	42 participants on hormones, 28 were evaluated	No change in suicide ideation	Very low certainty[165] due to very serious risk of bias and very serious imprecision
Taylor 2024a[166]	One cross-sectional study and one before-after study	222 participants were treated	The cross-sectional study reported less self-harm/ suicidality in those treated, although similar to group with no gender dysphoria. The before-after study did not report change over time.	No conclusions can be drawn.

[162] Miroshnychenko, Roldan, et al. (2025).
[163] Taylor, Mitchell, Hall, Heathcote, et al. (2024); This review did not include low quality studies in the evidence synthesis.
[164] Ludvigsson et al. (2023).
[165] Ludvigsson et al concluded the review "cannot assess certainty of evidence," while according to the downgrading process in the review, the certainty of evidence was very low.
[166] Taylor, Mitchell, Hall, Heathcote, et al. (2024); This review did not include low quality studies in the evidence synthesis.

Study ID	Included studies (study design)	Number of participants	Conclusion of the review	Certainty of evidence
Depression				
Ludvigsson 2023[167]	Two prospective observational cohort studies of which one included mixed treatment	97 participants on hormones, 60 were evaluated	No change in depression	Very low certainty[168] due to very serious risk of bias and very serious imprecision
Miroshnychenko 2025a[169]	One comparative observational study, and one before-after study	26 participants from one comparative observational study, and 41 participants from two before-after study	The comparative observational study with 26 participants (88% using PBs) reported that PBs did not result in a statistically significant decrease in scores in female to male participants; but resulted in a statistically significant decrease in score in male to female participants; and Mean change -3.36 (95% confidence interval: -3.69 to -3.03) from one before-after study (higher score indicating greater depression)	Very low certainty due to very serious risk of bias, and serious imprecision
Taylor 2024a[170]	One before-after study	70 participants were treated	Reduction in depressive symptom	A single study, reporting small to moderate improvement. No conclusion can be drawn.
Anxiety				
Ludvigsson 2023[171]	Two prospective observational cohort studies	97 participants on hormones, 60 were evaluated	No change in anxiety	Very low certainty[172] due to very serious risk of bias and very serious imprecision

[167] Ludvigsson et al. (2023).
[168] Ludvigsson et al concluded the review "cannot assess certainty of evidence," while according to the downgrading process in the review, the certainty of evidence was very low.
[169] Miroshnychenko, Roldan, et al. (2025).
[170] Taylor, Mitchell, Hall, Heathcote, et al. (2024); This review did not include low quality studies in the evidence synthesis.
[171] Ludvigsson et al. (2023).
[172] Ludvigsson et al concluded the review "cannot assess certainty of evidence," while according to the downgrading process in the review, the certainty of evidence was very low.

Study ID	Included studies (study design)	Number of participants	Conclusion of the review	Certainty of evidence
Taylor 2024a[173]	One before-after study	70 participants were treated	No change in anxiety symptom	A single study, reporting small to moderate improvement. No conclusion can be drawn.
Quality of life				
Ludvigsson 2023[174]	Two prospective observational cohort studies, including one retrospective	98 participants on hormones, 46 were evaluated	One study reported improvement in quality of life most pronounced in subjects receiving puberty-blocking hormones, followed by CSH treatment. A second study reported some improvement in quality of life.	Very low certainty[175] due to very serious risk of bias and very serious imprecision.
Taylor 2024a[176]	One before-after study	44 participants were treated	No change in quality of life	A single study, reporting no improvement. No conclusion can be drawn.
Internalising problems				
Taylor 2024a[177]	One cross-sectional study, and two before-after studies	292 participants were treated	No change to small decrease in internalising problems	No high certainty evidence was identified. No conclusion can be drawn.
Externalising problems				
Taylor 2024a[178]	One cross-sectional study, and two before-after studies	292 participants were treated	No change to small decrease in externalising problems	No conclusion can be drawn.

[173] Taylor, Mitchell, Hall, Heathcote, et al. (2024); This review did not include low quality studies in the evidence synthesis.
[174] Ludvigsson et al. (2023).
[175] Ludvigsson et al concluded the review “cannot assess certainty of evidence,” while according to the downgrading process in the review, the certainty of evidence was very low.
[176] Taylor, Mitchell, Hall, Heathcote, et al. (2024); this review did not include low quality studies in the evidence synthesis.
[177] See above note,
[178] See above note.

Study ID	Included studies (study design)	Number of participants	Conclusion of the review	Certainty of evidence
Psychological functioning/psychopathology				
Taylor 2024a[179]	Two before-after studies	114 participants were treated	Two before-after studies reported no change to a small change	No conclusions can be drawn.
Evidence synthesis			No conclusion can be drawn on the impact of mental health outcomes by PBs.	The certainty of evidence is very low.

[179] See above note.

6.11 Evidence tables on outcomes for CSH in youth with gender dysphoria

Table 6.2. Summary of evidence on gender dysphoria after CSH

Study ID	Included studies (study design)	Number of participants	Conclusion of the review	Certainty of evidence
Dopp 2024[259]	Eight studies	-[260]	CSH was associated with reduced gender dysphoria and improved body satisfaction in adolescents and young adults with gender dysphoria. Effect sizes for gender dysphoria and body dissatisfaction were the largest observed for any intervention.	Very low certainty due to very serious risk of bias, very serious imprecision, and serious indirectness
Miroshnychenko 2025b[261]	One comparative observational study and one before-after study	146 participants from one comparative study and 36 from one before-after study	Mean difference: -0.4 (95% confidence interval: -0.24 to 0.16) from one comparative study and Standardised mean change 0.26 lower (1.64 lower to 1.13 higher) from one before-after study (higher score indicating higher levels of gender dysphoria)	Very low certainty due to very serious risk of bias (rating down by three levels) and serious imprecision
Taylor 2024b[262]	One before-after study	23 participants from the before-after study	The pre–post study reported a reduction in dysphoria with no participants in clinical range at follow-up.	There was insufficient evidence regarding changes to gender dysphoria
Evidence synthesis			No conclusion can be drawn on the impact of gender dysphoria by CSH.	The certainty of evidence is very low.

[259] Dopp et al. (2024).
[260] The number of participants is not calculated because this review included studies with heterogeneous study types, including studies that were typically not included in a SR assessing treatment effect, for example, individual case studies (e.g., Ristori et al, 2019).
[261] Miroshnychenko, Ibrahim, et al. (2025).
[262] Taylor, Mitchell, Hall, Langton, et al. (2024); This review did not include low quality studies in the evidence synthesis.

Table 6.3. Summary of evidence on mental health and well-being after CSH

Study ID	Included studies (study design)	Number of participants	Conclusion of the review	Certainty of evidence
Global function				
Dopp 2024[263]	21 studies	-[264]	CSH was associated with improved mental health outcomes in adolescents and young adults with gender dysphoria. The majority of reported outcomes showed statistical significance, and nearly all of those indicated clinical significance as well (measured by effect size or compared to norms). Benefits were most consistent for depression and general functioning or well-being, and sometimes also seen for reductions in outcomes like suicidality, anxiety, and substance use (although these less consistently showed statistically significant changes).	Very low certainty due to very serious risk of bias, serious imprecision, inconsistency, and indirectness
Miroshnychenko 2025b[265]	Two prospective observational cohort studies; two before-after studies	125 from two prospective observational cohort studies; 73 participants from two before-after studies	Standardised mean difference: 0.87 higher (95% confidence interval: 0.25 lower to 2 higher) from two prospective observational cohort studies; Standardised mean change 0.25 higher (0.09 higher to 0.4 higher) from two before-after studies (higher score indicating better global function)	Very low certainty due to very serious risk of bias, serious imprecision and indirectness

[263] Dopp et al. (2024)
[264] The number of participants is not calculated because this review included heterogeneous types of studies, including studies that were typically not included in a SR assessing treatment effect, for example, individual case studies. For this outcome of mental health, this systematic review included two reports on the same large cross-sectional study (2015 US Transgender Survey, Lee et al. (2024); Turban et al. (2022)). Cross-sectional studies were of limited value in assessing treatment effects, while the reviewers should not include two studies on the same survey.
[265] Miroshnychenko, Ibrahim, et al. (2025)

Study ID	Included studies (study design)	Number of participants	Conclusion of the review	Certainty of evidence
Death by suicide				
Miroshnychenko 2025b[266]	One case series study	315 participants	6 per 1000 (1 to 18)	Very low certainty due to very serious risk of bias (rating down by three levels)
Self-harm and suicide				
Taylor 2024b[267]	Two before-after, and two cross-sectional studies	99 participants from two before-after studies, and 293 participants from two cross-sectional studies	Two before-after studies found a reduction in treatment needs for (or lower levels of) suicidality/self-harm, while two cross-sectional studies found conflicting results: those receiving hormones were less likely to have seriously considered/attempted suicide compared with adolescents not receiving hormones, and in birth-registered females there was no difference between groups	Hormone treatment may in the short-term improve psychological health.[268]
Depression				
Miroshnychenko 2025b[269]	Two prospective observational cohort studies; two before-after studies	154 from two prospective observational cohort studies; 389 participants from two before-after studies	Standardised mean difference: 0.3 lower (95% confidence interval: 0.85 lower to 0.25 higher) from two prospective observational cohort studies; Standardised mean change 0.41 lower (0.65 lower to 0.17 lower) from two before-after studies (higher score indicating worse depression)	Very low certainty due to very serious risk of bias (rating down by three levels) and serious imprecision

[266] Miroshnychenko, Ibrahim, et al. (2025)
[267] Taylor, Mitchell, Hall, Langton, et al. (2024); This review did not include low quality studies in the evidence synthesis.
[268] This systematic review concluded that there is “moderate-quality evidence from mainly pre–post studies that hormone treatment may in the short-term improve psychological health”. This review had identified “moderate quality” observational studies. Including “moderate-quality observational studies” is not equivalent to “moderate quality (certainty) evidence.” According to the GRADE methodology, the starting point for evidence from observational studies (including high quality observational studies) is generally low quality. To have moderate certainty, the certainty of evidence needs to be upgraded. See Guyatt, Oxman, Sultan, et al. (2011).
[269] Miroshnychenko, Ibrahim, et al. (2025)