



GRADE and ROBINS – 1 Quality Appraisal Rating Tables for included studies

2024

Citation: Ministry of Health. 2024. *GRADE and ROBINS – 1 Quality Appraisal Rating Tables for included studies*. Wellington: Ministry of Health.

Published in November 2024 by the Ministry of Health
PO Box 5013, Wellington 6140, New Zealand

HP 9067



This document is available at health.govt.nz



This work is licensed under the Creative Commons Attribution 4.0 International licence. In essence, you are free to: share ie, copy and redistribute the material in any medium or format; adapt ie, remix, transform and build upon the material. You must give appropriate credit, provide a link to the licence and indicate if changes were made.

Contents

GRADE ratings of the certainty of the evidence (<i>GRADE Handbook</i> , 2013)	1
ROBINS – I Risk of Bias criteria (Sterne et al., 2016)	1
Costa, R., Dunsford, M., Skagerberg, E., Holt, V., Carmichael, P., & Colizzi, M. (2015). Psychological Support, Puberty Suppression, and Psychosocial Functioning in Adolescents with Gender Dysphoria. <i>The journal of sexual medicine</i> , 12(11), 2206-2214. doi: https://dx.doi.org/10.1111/jsm.13034	2
GRADE Evidence Profile	2
Risk of Bias ROBINS-I	3
De Vries, A. L., Steensma, T. D., Doreleijers, T. A., & Cohen-Kettenis, P. T. (2011). Puberty suppression in adolescents with gender identity disorder: A prospective follow-up study. <i>Journal of Sexual Medicine</i> , 8(8), 2276-2283. doi:10.1111/j.1743-6109.2010.01943.x	8
GRADE evidence profile	8
Risk of Bias ROBINS-1	11
de Vries, A. L. C., McGuire, J. K., Steensma, T. D., Wagenaar, E. C. F., Doreleijers, T. A. H., & Cohen-Kettenis, P. T. (2014). Young adult psychological outcome after puberty suppression and gender reassignment. <i>Pediatrics</i> , 134(4), 696-704. doi: https://dx.doi.org/10.1542/peds.2013-2958	15
GRADE evidence profile	15
Risk of Bias ROBINS – 1	18
Elkadi, J., Chudleigh, C., Maguire, A. M., Ambler, G. R., Scher, S., & Kozłowska, K. (2023). Developmental Pathway Choices of Young People Presenting to a Gender Service with Gender Distress: A Prospective Follow-Up Study. <i>Children</i> , 10(2), 314. Retrieved from https://www.mdpi.com/2227-9067/10/2/314	21
GRADE Evidence Profile	21
Risk of Bias Robins - 1	22
Lavender, R., Shaw, S., Maninger, J. K., Butler, G., Carruthers, P., Carmichael, P., & Masic, U. (2023). Impact of Hormone Treatment on Psychosocial Functioning in Gender-Diverse Young People. <i>LGBT health</i> . doi:10.1089/lgbt.2022.0201	26

GRADE Evidence Profile	26
Risk of Bias ROBINS – 1	28
López de Lara, D., Pérez Rodríguez, O., Cuellar Flores, I., Pedreira Masa, J. L., Campos-Muñoz, L., Cuesta Hernández, M., & Ramos Amador, J. T. (2020). Psychosocial assessment in transgender adolescents. <i>Anales de Pediatría</i>, 93(1), 41-48. doi:10.1016/j.anpedi.2020.01.019	34
GRADE Evidence Profile	34
Risk of Bias – ROBIN-1	36
Olsavsky, A. L., Grannis, C., Bricker, J., Chelvakumar, G., Indyk, J. A., Leibowitz, S. F., . . . Nahata, L. (2023). Associations Among Gender-Affirming Hormonal Interventions, Social Support, and Transgender Adolescents' Mental Health. <i>The Journal of adolescent health : official publication of the Society for Adolescent Medicine</i>, 72(6), 860-868. doi: https://dx.doi.org/10.1016/j.jadohealth.2023.01.031	42
GRADE Evidence profile	42
Risk of Bias ROBINS- 1	45
Tordoff, D. M., Wanta, J. W., Collin, A., Stepney, C., Inwards-Breland, D. J., & Ahrens, K. (2022). Mental Health Outcomes in Transgender and Nonbinary Youths Receiving Gender-Affirming Care. <i>JAMA network open</i>, 5(2), e220978. doi: https://dx.doi.org/10.1001/jamanetworkopen.2022.0978	49
GRADE Evidence Profile	49
Risk of Bias ROBINS-1	51
Turban, J. L., King, D., Carswell, J. M., & Keuroghlian, A. S. (2020). Pubertal Suppression for Transgender Youth and Risk of Suicidal Ideation. <i>Pediatrics</i>, 145(2). doi: https://dx.doi.org/10.1542/peds.2019-1725	57
GRADE Evidence Profile	57
Risk of Bias ROBINS- 1	58
Kuper, L. E., Stewart, S., Preston, S., Lau, M., & Lopez, X. (2020). Body Dissatisfaction and Mental Health Outcomes of Youth on Gender-Affirming Hormone Therapy. <i>Pediatrics</i>, 145(4). doi:10.1542/peds.2019-3006	61
GRADE Evidence Profile	61
Risk of Bias – ROBINS - 1	63
Achille, C., Taggart, T., Eaton, N. R., Osipoff, J., Tafuri, K., Lane, A., & Wilson, T. A. (2020). Longitudinal impact of gender-affirming	

endocrine intervention on the mental health and well-being of transgender youths: preliminary results. International journal of pediatric endocrinology, 2020, 8. doi: https://dx.doi.org/10.1186/s13633-020-00078-2	68
GRADE Evidence Profile	68
Risk of Bias ROBINS -1	69

GRADE ratings of the certainty of the evidence (*GRADE Handbook, 2013*)

Ratings	Definitions
High	This research provides a very good indication of the likely effect. The likelihood that the effect will be substantially different is low.
Moderate	This research provides a good indication of the likely effect. The likelihood that the effect will be substantially different is moderate.
Low	This research provides some indication of the likely effect. However, the likelihood that it will be substantially different (a large enough difference that it might have an effect on a decision) is high.
Very Low	This research does not provide a reliable indication of the likely effect. The likelihood that the effect will be substantially different (a large enough difference that it might have an effect on a decision) is very high.

ROBINS – I Risk of Bias criteria (Sterne et al., 2016)

Part of the ROBINS–I process is to identify key confounding factors to assess the studies against for bias. Confounding factors were chosen from common limitations identified in the literature that were likely to have an impact on outcomes associated with mental health and wellbeing of the participants. These were: socio-demographic situation, presence/absence of family support, recruitment of participants from specialised gender or endocrine clinics, lack of disclosure of public funding of the treatment provided, and puberty development (Tanner stage).

Response option	Criteria
Low risk of bias (the study is comparable to a well-performed randomized trial)	The study is judged to be at low risk of bias for all domains.
Moderate risk of bias (the study appears to provide sound evidence for a non-randomized study but cannot be considered comparable to a well-performed randomized trial)	The study is judged to be at low or moderate risk of bias for all domains.
Serious risk of bias (the study has some important problems)	The study is judged to be at serious risk of bias in at least one domain, but not at critical risk of bias in any domain.
Critical risk of bias (the study is too problematic to provide any useful evidence and should not be included in any synthesis)	The study is judged to be at critical risk of bias in at least one domain.
No information on which to base a judgement about risk of bias	There is no clear indication that the study is at serious or critical risk of bias and there is a lack of information in one or more key domains of bias (a judgement is required for this).

Costa, R., Dunsford, M., Skagerberg, E., Holt, V., Carmichael, P., & Colizzi, M. (2015). Psychological Support, Puberty Suppression, and Psychosocial Functioning in Adolescents with Gender Dysphoria. *The journal of sexual medicine*, 12(11), 2206-2214. doi: <https://dx.doi.org/10.1111/jsm.13034>

GRADE Evidence Profile

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication bias	+ve factors	Overall
Gender dysphoria	<p>↓ 1 level</p> <p>No control group</p> <p>Limited measurement of known confounding prognostic factors such as family and peer support</p>	<p>↓ 2 level</p> <p>Unknown heterogeneity due to UGDS not repeated in either cohort group</p>	<p>↓ 1 level</p> <p>No direct evidence presented about changes in gender dysphoria</p>	<p>↓ 1 levels</p> <p>few patients < 400 participants</p> <p>comparison group from different population (England vs Stockholm)</p>	<p>Not assessed</p> <p>Not enough evidence to justify a decrease in level</p>	<p>No change</p>	<p>Very low</p>
Suicidality	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Self-harm	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Anxiety	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Depression	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Life satisfaction/QoL	<p>↓ one level</p> <p>No control group</p> <p>Limited measurement of known confounding prognostic factors such as family and peer support</p>	<p>↓ 1 level</p> <p>Unknown heterogeneity due low follow up numbers at T3 for both cohort groups</p>	<p>↓ 1 level</p> <p>CGAS is an indirect measure of mental health and wellbeing and QoL/life satisfaction</p>	<p>↓ 1 levels</p> <p>few patients < 400 participants</p> <p>comparison group from different population (England vs Stockholm)</p>	<p>Not assessed</p> <p>Not enough evidence to justify a decrease in level</p>	<p>No</p>	<p>Very Low</p>

Risk of Bias ROBINS-I

Confounding domains: socio-demographic situation; family support; public funding available; enrolled in a specialised service; puberty development (Tanner stage)

Co-interventions likely to have impact: counselling, family therapy, school based support, peer support, community group engagement

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
<p>Bias due to confounding</p> <p>Critical risk of bias (the study is too problematic to provide any useful evidence on the effects of intervention)</p>	<p>Confounding from ethnicity, socioeconomic status, family support, enrolment in specialised service, Tanner stage inherently not controllable</p>	N/A	N/A	N/A	N/A	<p>Confounding from ethnicity, socioeconomic status, family support, enrolment in specialised service, Tanner stage inherently not controllable</p>	<p>Living with family, education, living in role and changed name were identified as potentially confounding characteristics but not controlled for. Other socio-demographic variables not reported for either cohort group</p>
<p>Bias in selection of participants into the study</p> <p>Serious risk of bias (the study has some important problems)</p>	<p>Selection into the study was related (but not very strongly) to intervention and outcome due to being enrolled in a specialised clinic and all participants were eligible and eventually received PBs</p> <p>and</p> <p>this could not be adjusted for in analyses</p>	N/A	N/A	N/A	N/A	<p>Selection into the study was related (but not very strongly) to intervention and outcome due to being enrolled in a specialised clinic and all participants were eligible and eventually received PBs</p> <p>and</p> <p>this could not be adjusted for in analyses</p>	<p>All participants and clinicians knew they could receive PBs and/or the likely outcome of this</p>

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
<p>Bias in classification of interventions</p> <p>Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial)</p>	<p>Intervention status is well defined between two cohorts but is unclear about process/intervention of introducing PBs to the non-receiving group who all went onto receive them and</p> <p>Some aspects of the assignments of intervention (psychological support) status were likely determined retrospectively.</p>	N/A	N/A	N/A	N/A	<p>Intervention status is well defined between two cohorts but is unclear about process/intervention of introducing PBs to the non-receiving group who all went onto receive them and</p> <p>Some aspects of the assignments of intervention (psychological support) status were likely determined retrospectively</p>	<p>Assignment into eligible cohort determined by WPATH Standards of Care</p> <p>Unclear what psychological support entailed and how this was assessed as received (retrospectively or case notes etc)</p>
<p>Bias due to deviations from intended Interventions</p> <p>Serious risk of bias (the study has some important problems)</p>	<p>Effect of starting and adhering to intervention:</p> <p>There was limited clarity about the important co-intervention of psychological support or if there were deviations from the intended interventions (in terms of implementation and/or adherence) that were likely to impact on the outcome;</p>	N/A	N/A	N/A	N/A	<p>Effect of starting and adhering to intervention:</p> <p>There was limited clarity about the important co-intervention of psychological support or if there were deviations from the intended interventions (in terms of implementation and/or adherence) that were</p>	<p>Unknown factors related to any deviation of usual practice and quality of psychosocial support received by the two cohorts.</p> <p>No information about adherence of either cohort group.</p> <p>Likely imbalance between co-interventions received</p>

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
	<p>and</p> <p>The analysis was not appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome.</p>					<p>likely to impact on the outcome;</p> <p>and</p> <p>The analysis was not appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome</p>	<p>Not re-assessing for gender dysphoria in follow up</p>
<p>Bias due to missing data</p> <p>Serious risk of bias (the study has some important problems)</p>	<p>Reasons for missingness are not well explained across cohorts and interventions; no explanation for why gender dysphoria was not re-assessed</p> <p>and</p> <p>The analysis is unlikely to have removed the risk of bias arising from the missing data;</p> <p>and</p> <p>The nature of the missing data means that the risk of bias cannot be removed</p>	N/A	N/A	N/A	N/A	<p>Reasons for missingness are not well explained across cohorts and interventions;</p> <p>and</p> <p>The analysis is unlikely to have removed the risk of bias arising from the missing data;</p> <p>and</p> <p>The nature of the missing data means that the risk of bias cannot be removed through appropriate analysis</p>	<p>No explanation or description of loss of participants over the three time series</p> <p>No follow up scores for gender dysphoria assessment</p>

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
	through appropriate analysis.						
Bias in measurement of outcomes Serious risk of bias (the study has some important problems)	<p>The outcome measure was subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants);</p> <p>and</p> <p>The outcome was assessed by assessors aware of the intervention received by study participants</p>	N/A	N/A	N/A	N/A	<p>The outcome measure was subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants);</p> <p>and</p> <p>The outcome was assessed by assessors aware of the intervention received by study participants</p>	<p>CGAS is a subjective assessment by clinicians.</p> <p>Unknown if same assessor completed the three assessments for each participants.</p> <p>Multiple assessors involved including Stockholm and England based practitioners</p> <p>Unknown if assessors knew about study and intervention, but probably likely</p> <p>UGDS is self-reported</p>
Bias in selection of the reported result Serious risk of bias (the study has some important problems)	There is a high risk of selective reporting from among multiple analyses	N/A	N/A	N/A	N/A	There is a high risk of selective reporting from among multiple analyses	<p>No reporting on outcomes for the loss of participants</p> <p>Delayed eligible participants all received puberty suppression treatment but psychological intervention they received not reported</p>

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
							No reporting of gender dysphoria only CGAF score
Critical risk of bias	Critical risk of bias	N/A	N/A	N/A	N/A	Critical risk of bias	

De Vries, A. L., Steensma, T. D., Doreleijers, T. A., & Cohen-Kettenis, P. T. (2011). Puberty suppression in adolescents with gender identity disorder: A prospective follow-up study. *Journal of Sexual Medicine*, 8(8), 2276-2283. doi:10.1111/j.1743-6109.2010.01943.x

GRADE evidence profile

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall
Gender dysphoria	<p>↓ 2 levels</p> <p>Failure to develop and apply appropriate eligibility criteria as no inclusion of control population</p> <p>Incomplete accounting of patients and outcome events</p> <p>No reporting on other psychological interventions</p>	<p>No change</p> <p>limited information about population and sub-groupings except AFAB/AMAB & all commenced GAHT</p>	<p>↓ 1 levels</p> <p>Difference of approx. one year between AFAB/AMAB ages for assessment and treatment with females presenting one year later and sex characteristics and menstruation already commenced</p> <p>Between-sex differences $P = < 0.001$</p>	<p>↓ 1 levels</p> <p>few patients < 400 participants</p> <p>No reporting of outcomes for participants not assessed pre-post treatment</p> <p>T0 n = 16</p> <p>T1 n=29</p>	<p>No change</p> <p>Systematic reviews performed early in the development of a body of research may be biased due to the tendency for positive results to be published sooner and for negative results to be published later or withheld.</p>	<p>↑ 1 level</p> <p>Consistent pre-post testing for 41/70 of cohort and has longitudinal value over several years</p>	low
Suicidality	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Self-harm	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Anxiety	<p>↓ 2 levels</p> <p>Use of unvalidated outcome measures (e.g. patient-reported outcomes)</p>	<p>↓ 1 levels</p> <p>Patients vary widely in their pre-intervention or baseline risk as high SD in baselines</p>	No Change	<p>↓ 1 level</p> <p>few patients < 400 participants</p> <p>no comparison group</p>	<p>No change</p> <p>Systematic reviews performed early in the development of a body of research may be biased due</p>	No change	Very low

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall
	Incomplete accounting of patients and outcome events No reporting on confounding interventions				to the tendency for positive results to be published sooner and for negative results to be published later or withheld.		
Depression	<p>↓ 2 levels</p> <p>Use of unvalidated outcome measures (e.g. patient-reported outcomes)</p> <p>Incomplete accounting of patients and outcome events</p> <p>No reporting on confounding interventions</p>	<p>No change</p> <p>Patients vary widely in their pre-intervention or baseline risk as high SD in baselines</p>	<p>↓ 2 levels</p> <p>Self-reporting mean scores lie within normal range at T0 & T1</p> <p>(1-10 These ups and downs are considered normal)</p> <p>However study implies change in depression</p> <p>“depressive symptom scores on the BDI-II significantly decreased”</p>	<p>↓ 1 level</p> <p>few patients < 400 participants</p> <p>no comparison group</p> <p>No reporting of outcomes for participants not assessed</p> <p>T0 n = 16</p> <p>T1 n=29</p>	<p>↓ 1 level</p> <p>Systematic reviews performed early in the development of a body of research may be biased due to the tendency for positive results to be published sooner and for negative results to be published later or withheld.</p>	No Change	Very low
Life satisfaction/QoL	<p>↓ 2 levels</p> <p>CBCL & YSR are unvalidated outcome measures (e.g. patient-reported outcomes)</p> <p>No description of CGAS administration</p>	<p>No change</p> <p>Patients vary widely in their pre-intervention or baseline risk as high SD baseline anxiety figure</p>	<p>↓ 2 levels</p> <p>CBCL and YSR clinical indication score is from >63 and it is unclear in results table how the conclusions have been made as mean</p>	<p>↓ 1 level</p> <p>few patients < 400 participants</p> <p>no comparison group</p>	<p>↓ 1 level</p> <p>Systematic reviews performed early in the development of a body of research may be biased due to the tendency for positive results to be</p>	No change	Very low

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall
	<p>at T0 & T1 – if same clinicians/process used</p> <p>Incomplete accounting of patients and outcome events</p> <p>No reporting on confounding interventions</p>		<p>participants scored in non-clinical range on CBCL</p> <p>T0 = 60.70</p> <p>TI = 54.46</p> <p>YSR</p> <p>T0 = 55.56</p> <p>T1= 50.00</p> <p>“Adolescents showed a significant decrease in behavioural and emotional problems over time on mean <i>T</i>-scores of the total problem scale, the internalizing and externalizing scale of both CBCL and YSR. In addition, the percentage of adolescents scoring in the clinical range significantly decreased between T0 and T1, on the CBCL total problem scale and the internalizing scale of the YSR.</p>		<p>published sooner and for negative results to be published later or withheld.</p> <p>No reporting of outcomes for participants not assessed pre-post treatment n = 16 and n=29</p>		

Risk of Bias ROBINS-1

Confounding domains: socio-demographic situation; family support; public funding available; enrolled in a specialised service; puberty development (Tanner stage)

Co-interventions likely to have impact: counselling, family therapy, school based support, peer support, community group engagement

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Bias due to confounding Critical risk of bias (the study is too problematic to provide any useful evidence on the effects of intervention)	Confounding inherently not controllable	N/A	N/A	Confounding inherently not controllable	Confounding inherently not controllable	Confounding inherently not controllable	All participants enrolled in specialised clinic with public funded healthcare No disaggregation of ethnicity
Bias in selection of participants into the study Critical risk of bias (the study is too problematic to provide any useful evidence on the effects of intervention)	Selection into the study was very strongly related to intervention and outcome; and This could not be adjusted for in analyses	N/A	N/A	Selection into the study was very strongly related to intervention and outcome; and This could not be adjusted for in analyses	Selection into the study was very strongly related to intervention and outcome; and This could not be adjusted for in analyses	Selection into the study was very strongly related to intervention and outcome; and This could not be adjusted for in analyses	Participants were selected from those who commenced GAHT after age 16years
Bias in classification of interventions	Intervention status is not well defined; and	N/A	N/A	Intervention status is not well defined;	Intervention status is not well defined;	Intervention status is not well defined; and	Intervention status is not well defined, it is unclear about what

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Serious risk of bias (the study has some important problems)	Major aspects of the assignments of intervention status were determined in a way that could have been affected by knowledge of the outcome.			and Major aspects of the assignments of intervention status were determined in a way that could have been affected by knowledge of the outcome.	and Major aspects of the assignments of intervention status were determined in a way that could have been affected by knowledge of the outcome.	Major aspects of the assignments of intervention status were determined in a way that could have been affected by knowledge of the outcome.	PB were used and how they were administered No reference to additional interventions alongside PB Participants were on a pre-defined treatment pathway
Bias due to deviations from intended Interventions Serious risk of bias (the study has some important problems)	Co-interventions were not balanced or mentioned in the intervention group, and there was no reporting on deviations from the intended interventions in terms of implementation and/or adherence to inform reader of the likelihood of these impacting outcomes; and The analysis was not appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were	N/A	N/A	Co-interventions were not balanced or mentioned in the intervention group, and there was no reporting on deviations from the intended interventions in terms of implementation and/or adherence to inform reader of the likelihood of these impacting outcomes; and	Co-interventions were not balanced or mentioned in the intervention group, and there was no reporting on deviations from the intended interventions in terms of implementation and/or adherence to inform reader of the likelihood of these impacting outcomes; and	Co-interventions were not balanced or mentioned in the intervention group, and there was no reporting on deviations from the intended interventions in terms of implementation and/or adherence to inform reader of the likelihood of these impacting outcomes; and The analysis was not appropriate to estimate the effect of starting and adhering	Little discussion on co-interventions provided at the clinic and what these involved e.g frequency, quality, what they were Analysis covers long period of time and does not account for deviations in treatments or how the large attrition of the sample might related to this

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
	likely to impact on the outcome.			The analysis was not appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome.	The analysis was not appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome.	to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome.	
Bias due to missing data Serious risk of bias (the study has some important problems)	The analysis is unlikely to have removed the risk of bias arising from the missing data; Missing data were addressed inappropriately [not addressed] in the analysis	N/A	N/A	The analysis is unlikely to have removed the risk of bias arising from the missing data; Missing data were addressed inappropriately [not addressed] in the analysis	The analysis is unlikely to have removed the risk of bias arising from the missing data; Missing data were addressed inappropriately [not addressed] in the analysis	The analysis is unlikely to have removed the risk of bias arising from the missing data; Missing data were addressed inappropriately [not addressed] in the analysis	Missing data is not addressed in either publication, large attrition by final follow up with no explanation of why or the outcomes for these participants
Bias in measurement of outcomes	The outcome measures were subjective (i.e. vulnerable to influence by knowledge of the	N/A	N/A	The outcome measures were subjective (i.e. vulnerable to	The outcome measures were subjective (i.e. vulnerable to	The outcome measures were subjective (i.e. vulnerable to	

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Serious risk of bias (the study has some important problems)	intervention received by study participants); and The outcome was assessed by assessors aware of the intervention received by study participants			influence by knowledge of the intervention received by study participants); and The outcome was assessed by assessors aware of the intervention received by study participants	influence by knowledge of the intervention received by study participants); and The outcome was assessed by assessors aware of the intervention received by study participants	influence by knowledge of the intervention received by study participants); and The outcome was assessed by assessors aware of the intervention received by study participants	
Bias in selection of the reported result Serious risk of bias (the study has some important problems)	There is a high risk of selective reporting from among multiple analyses; UGDS is not repeated and the CGAS is used instead as a proxy for gender dysphoria Removal of missing participants in analysis	N/A	N/A	There is a high risk of selective reporting from among multiple analyses; Selective reporting of the CBSL and YSR scale and Beck Depression Inventory – II Removal of missing participants in analysis	There is a high risk of selective reporting from among multiple analyses; Selective reporting of the CBSL and YSR scale and Beck Depression Inventory – II Removal of missing participants in analysis	There is a high risk of selective reporting from among multiple analyses; Selective reporting of the CBSL and YSR scale and Beck Depression Inventory – II Removal of missing participants in analysis	UGDS is not repeated and the CGAS is used instead as a proxy to measure gender dysphoria Selective reporting of the CBSL and YSR scale and Beck Depression Inventory – II results Missing data creates uncertainty of evidence
Critical risk of bias	Critical risk of bias	N/A	N/A	Critical risk of bias	Critical risk of bias	Critical risk of bias	Critical risk of bias

de Vries, A. L. C., McGuire, J. K., Steensma, T. D., Wagenaar, E. C. F., Doreleijers, T. A. H., & Cohen-Kettenis, P. T. (2014). Young adult psychological outcome after puberty suppression and gender reassignment. *Pediatrics*, 134(4), 696-704. doi: <https://dx.doi.org/10.1542/peds.2013-2958>

GRADE evidence profile

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall
Gender dysphoria	<p>↓ 2 levels</p> <p>Failure to develop and apply appropriate eligibility criteria as no inclusion of control population</p> <p>Incomplete accounting of patients and outcome events</p> <p>No reporting on other psychological interventions</p>	<p>No change</p> <p>limited information about population and sub-groupings except AFAB/AMAB & all commenced GAHT</p>	<p>↓ 2 levels</p> <p>Difference of approx. one year between AFAB/AMAB ages for assessment and treatment with females presenting one year later and sex characteristics and menstruation already commenced</p> <p>Between-sex differences $P = < 0.001$</p>	<p>↓ 1 levels</p> <p>few patients < 400 participants</p> <p>no comparison group</p> <p>No reporting of outcomes for 37 participants who also received PBs but did not progress to GRS</p>	<p>No change</p> <p>Systematic reviews performed early in the development of a body of research may be biased due to the tendency for positive results to be published sooner and for negative results to be published later or withheld.</p>	<p>No change</p> <p>Consistent pre-post testing for cohort and longitudinal value over several years</p>	Low
Suicidality	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Self-harm	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Anxiety	<p>↓ 2 levels</p> <p>Use of unvalidated outcome measures</p>	<p>↓ 2 levels</p> <p>Patients vary widely in their pre-intervention or baseline risk as high</p>	<p>No Change</p>	<p>↓ 1 level</p> <p>few patients < 400 participants</p>	<p>No change</p> <p>Systematic reviews performed early in the development of a</p>	<p>No change</p> <p>Consistent pre-post testing for cohort and</p>	Very low

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall
	(e.g. patient-reported outcomes) Incomplete accounting of patients and outcome events No reporting on confounding interventions	SD baseline anxiety figure Small sample size		no comparison group No reporting of outcomes for 38 participants who also received puberty blockers but did not progress to GRS	body of research may be biased due to the tendency for positive results to be published sooner and for negative results to be published later or withheld.	longitudinal value over several years	
Depression	↓ 2 levels Use of unvalidated outcome measures (e.g. patient-reported outcomes) Incomplete accounting of patients and outcome events No reporting on confounding interventions	Unable to assess due to minimal information about T0 & T1 administration and scoring	↓ 2 levels Self-reporting scores lie within normal range at T0 & T1 (1-10 These ups and downs are considered normal) However study states "depressive symptom scores on the BDI-II significantly decreased"	↓ 1 level few patients < 400 participants no comparison group No reporting of outcomes for 38 participants who also received puberty blockers but did not progress to GRS	No change Systematic reviews performed early in the development of a body of research may be biased due to the tendency for positive results to be published sooner and for negative results to be published later or withheld.	No Change Consistent pre-post testing for cohort and longitudinal value over several years	Very low
Life satisfaction/QoL	↓ 2 levels CBCL, YSR, WHOQOL-BREF, SWLS, SHS are unvalidated	No change Unable to assess due to minimal information about T0, T1, T2	↓ 2 levels CBCL and YSR clinical scores are from >63 and it is unclear in results table how the	↓ 1 level few patients < 400 participants no comparison group	No Change Systematic reviews performed early in the development of a body of research may be biased due to the	No change Consistent pre-post testing for cohort and longitudinal	Very low

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall
	<p>outcome measures (e.g. self-reported)</p> <p>No description of CGAS administration at T0 & T1 – if same clinicians/process used</p> <p>Incomplete accounting of patients and outcome events</p> <p>No reporting on confounding interventions</p> <p>WHOQOL-BREF, Satisfaction With Life Scale and Subjective Happiness Scale only administered in 2014, unable to be compared to pre PBs (T0)</p>	administration and scoring	<p>conclusions have been made as mean participants scored in non-clinical range on CBCL T0 = 60.70 T1 = 54.46</p> <p>YSR T0 = 55.56 T1= 50.00</p> <p>“Adolescents showed a significant decrease in behavioural and emotional problems over time on mean <i>T</i>-scores of the total problem scale, the internalizing and externalizing scale of both</p> <p>CBCL and YSR (see Table 2). In addition, the percentage of adolescents scoring in the clinical range significantly decreased between T0 and T1, on the CBCL total problem scale and the internalizing scale of the YSR.</p>	No reporting of outcomes for 30 participants who also received puberty blockers but did not progress to GRS	tendency for positive results to be published sooner and for negative results to be published later or withheld.	<p>value over several years</p> <p>WHOQOL-BREF is an extensively tested QoL instrument</p>	

Risk of Bias ROBINS – 1

[NOTE Risk of Bias is the same as the 2011 study as have used the same data. Additional data is assessed only in the table below]

Confounding domains: socio-demographic situation; family support; public funding available; enrolled in a specialised service; puberty development (Tanner stage)

Co-interventions likely to have impact: counselling, family therapy, school based support, peer support, community group engagement

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Bias due to confounding Critical risk of bias (the study is too problematic to provide any useful evidence on the effects of intervention)	Y	N/A	N/A	Y	Y	Y	All participants enrolled in specialised clinic with public funded healthcare No disaggregation for ethnicity
Bias in selection of participants into the study Critical risk of bias (the study is too problematic to provide any useful evidence on the effects of intervention)	Y	N/A	N/A	N	N	N	Participants were selected from those who progressed to gender reassignment surgery
Bias in classification of interventions Serious risk of bias (the study has some important problems)	Y	N/A	N/A	Y	Y	Y	

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Bias due to deviations from intended Interventions Serious risk of bias (the study has some important problems)	Y	N/A	N/A	Y	Y	Y	
Bias due to missing data Serious risk of bias (the study has some important problems)	Y	N/A	N/A	Y	Y	Y	
Bias in measurement of outcomes Serious risk of bias (the study has some important problems)	Y	N/A	N/A	Y	Y	Y	
Bias in selection of the reported result Serious risk of bias (the study has some important problems)	Y	N/A	N/A	There is a high risk of selective reporting from among multiple analyses; Selective reporting WHOQOL-BREF, Satisfaction With Life Scale and	There is a high risk of selective reporting from among multiple analyses; Selective reporting of the Selective reporting WHOQOL-BREF, Satisfaction With Life Scale	There is a high risk of selective reporting from among multiple analyses; Selective reporting WHOQOL-BREF, Satisfaction With Life Scale and Subjective Happiness Scale	WHOQOL-BREF, Satisfaction With Life Scale and Subjective Happiness Scale Were only administered in the 2014 study

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
				Subjective Happiness Scale Removal of missing participants in analysis	and Subjective Happiness Scale Removal of missing participants in analysis	Removal of missing participants in analysis	
Critical risk of bias	Critical risk of bias	N/A	N/A	Critical risk of bias	Critical risk of bias	Critical risk of bias	Critical risk of bias

Elkadi, J., Chudleigh, C., Maguire, A. M., Ambler, G. R., Scher, S., & Kozłowska, K. (2023). Developmental Pathway Choices of Young People Presenting to a Gender Service with Gender Distress: A Prospective Follow-Up Study. *Children, 10*(2), 314. Retrieved from <https://www.mdpi.com/2227-9067/10/2/314>

GRADE Evidence Profile

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall
Gender dysphoria	N/A	N/A	N/A	N/A	N/A	N/A	
Suicidality	N/A	N/A	N/A	N/A	N/A	N/A	
Self-harm	N/A	N/A	N/A	N/A	N/A	N/A	
Anxiety	<p>↓ 2 levels</p> <p>Use of unvalidated outcome measures (self-report via telephone call/medical note review using custom-made questionnaire)</p>	<p>↓ 2 levels</p> <p>No statistical analysis, % comparison only of pre-post treatment with no analysis or discussion of variables</p>	<p>↓ 2 levels</p> <p>Demographic and/or social differences within population cohort not discussed</p>	<p>↓ 1 levels</p> <p><400 participants</p>	<p>No change</p> <p>Unable to be determined</p>	<p>No change</p>	<p>Very low</p>
Depression	<p>↓ 2 levels</p> <p>Use of unvalidated outcome measures (self-report via telephone call/medical note review using custom-made questionnaire)</p>	<p>↓ 2 levels</p> <p>No statistical analysis, % comparison only of pre-post treatment with no analysis or discussion of variables</p>	<p>↓ 2 levels</p> <p>Demographic and/or social differences within population cohort not discussed</p>	<p>↓ 1 levels</p> <p><400 participants</p>	<p>No change</p> <p>Unable to be determined</p>	<p>No change</p>	<p>Very low</p>

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall
Life satisfaction/QoL	N/A	N/A	N/A	N/A	N/A	N/A	

Risk of Bias Robins - 1

Confounding domains: socio-demographic situation; family support; public funding available; enrolled in a specialised service; puberty development (Tanner stage)

Co-interventions likely to have impact: counselling, family therapy, school based support, peer support, community group engagement

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Bias due to confounding Critical risk of bias (the study is too problematic to provide any useful evidence on the effects of intervention)	N/A	N/A	N/A	Confounding inherently not controlled or discussed in analysis and discussion	Confounding inherently not controlled or discussed in analysis and discussion	N/A	No confounding factors mentioned or controlled for
Bias in selection of participants into the study Critical risk of bias (the study is too problematic to provide any useful evidence on the effects of intervention)	N/A	N/A	N/A	Selection into the study was very strongly related to intervention and outcome; and This could not be adjusted for in analyses; and	Selection into the study was very strongly related to intervention and outcome; and This could not be adjusted for in analyses; and	N/A	Inconsistent follow-up time period (4-9 years) Treatment pathway was pre-determined and known by participants

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
				A substantial amount of follow-up time is likely to be missing from analyses	A substantial amount of follow-up time is likely to be missing from analyses		
Bias in classification of interventions Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial)	N/A	N/A	N/A	Intervention status is well defined: long acting goserelin acetate (Zoladex) injections were given every 10 weeks. and (ii) Some aspects of the assignments of intervention status were determined retrospectively “the founding multidisciplinary team also became aware of the increase of presentations of what was termed late-onset, rapid-onset, or adolescent-onset GD.” (p.3)	Intervention status is well defined: long acting goserelin acetate (Zoladex) injections were given every 10 weeks. and (ii) Some aspects of the assignments of intervention status were determined retrospectively “the founding multidisciplinary team also became aware of the increase of presentations of what was termed late-onset, rapid-onset, or adolescent-onset GD.” (p.3)	N/A	
Bias due to deviations from intended interventions	N/A	N/A	N/A	Important co-interventions were not controlled for or discussed [pre-requisite for	Important co-interventions were not controlled for or discussed [pre-requisite for	N/A	

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Serious risk of bias (the study has some important problems)				psychologist/counselling involvement] across the intervention group. Any deviations from the intended interventions (in terms of implementation and/or adherence) were not discussed and the analysis was not appropriate to estimate the effect of starting and adhering to intervention on anxiety and depression self-reports	psychologist/counselling involvement] across the intervention group. Any deviations from the intended interventions (in terms of implementation and/or adherence) were not discussed and the analysis was not appropriate to estimate the effect of starting and adhering to intervention on anxiety and depression self-reports		
Bias due to missing data Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial)	N/A	N/A	N/A	Proportions of and reasons for missing participants declared but was a large number 29/70 and The analysis is unlikely to have removed the risk of bias arising from the missing data.	Proportions of and reasons for missing participants declared but was a large number 29/70 and The analysis is unlikely to have removed the risk of bias arising from the missing data.	N/A	No explanation or commentary on outcomes for missing participants

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Bias in measurement of outcomes Serious risk of bias (the study has some important problems)	N/A	N/A	N/A	The outcome measure was subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants); and The outcome was assessed by assessors aware of the intervention received by study participants	The outcome measure was subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants); and The outcome was assessed by assessors aware of the intervention received by study participants	N/A	
Bias in selection of the reported result No information on which to base a judgement about risk of bias for this domain.	N/A	N/A	N/A	–	–	N/A	No quotes or information about telephone interviews or medical note reviews to verify the anxiety and depression scores
Critical Risk of Bias	N/A	N/A	N/A	Critical Risk of Bias	Critical Risk of Bias	N/A	

Lavender, R., Shaw, S., Maninger, J. K., Butler, G., Carruthers, P., Carmichael, P., & Masic, U. (2023). Impact of Hormone Treatment on Psychosocial Functioning in Gender-Diverse Young People. *LGBT health*. doi:10.1089/lgbt.2022.0201

GRADE Evidence Profile

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall
Gender dysphoria	<p>↓1 level</p> <p>Failure to develop and apply appropriate eligibility criteria (inclusion of control population)</p>	<p>↓2 levels</p> <p>38/109 of those who received PBs were reported on with no exploration of the 71 not in the data</p> <p>Only 19 participants completed the gender dysphoria questionnaire in all 3 assessments</p>	<p>↓2 levels</p> <p>Differences between numbers of AFAB and AMAB in study populations were large (28 compared to 10).</p> <p>A year different in starting PBs between the sex. Tanner stage of puberty development not reported/aggregated for either sex</p>	<p>↓1 level</p> <p>Few participants <400 optimal information size</p>	Not enough information to assess	N/A	Very low
Suicidality	<p>↓1 level</p> <p>Failure to develop and apply appropriate eligibility criteria (inclusion of control population)</p>	<p>↓2 levels</p> <p>38/109 of those who received PBs were reported on with no exploration of the 71 not in the data</p> <p>Only 11 participants (young person and caregiver)</p>	<p>↓2 levels</p> <p>Differences between numbers of AFAB and AMAB in study populations were large (28 compared to 10).</p> <p>A year different in starting PBs between the sex. Tanner stage of puberty development not</p>	<p>↓1 level</p> <p>Few participants <400 optimal information size</p>	Not enough information to assess	N/A	Very low

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall
		completed the suicidality question questionnaire in all 3 assessments	reported/aggregated for either sex Not stated if outcome measured was of primary importance to participants				
Self-harm	<p>↓1 level</p> <p>Failure to develop and apply appropriate eligibility criteria (inclusion of control population)</p> <p>Baseline scores for young person and caregiver questionnaires were in normal range for self-report mental wellbeing and behaviour</p>	<p>↓2 levels</p> <p>38/109 of those who received PBs were reported on with no exploration of the 71 not in the data</p> <p>Only 11 participants (young person and caregiver) completed the self-harm question in the questionnaire in all 3 assessments</p>	<p>↓2 levels</p> <p>Differences between numbers of AFAB and AMAB in study populations were large (28 compared to 10).</p> <p>A year different in starting PBs between the sex. Tanner stage of puberty development not reported/aggregated for either sex</p> <p>Not stated if outcome measured was of primary importance to participants</p>	<p>↓1 level</p> <p>Few participants <400 optimal information size</p>	Not enough information to assess	N/A	Very low
Anxiety	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Depression	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Life satisfaction/QoL	<p>↓1 level</p> <p>Failure to develop and apply</p>	<p>↓2 levels</p> <p>38/109 of those who received PBs</p>	<p>↓2 levels</p> <p>Differences between numbers of AFAB and</p>	<p>↓1 level</p> <p>Few participants</p>	Not enough information to assess	N/A	Very low

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall
	appropriate eligibility criteria (inclusion of control population) Participants were all in 'normal' non clinical functioning range in all 3 assessments	were reported on with no exploration of the 71 not in the data Only 19 participants completed the social responsiveness scale questionnaire in all 3 assessments	AMAB in study populations were large (28 compared to 10). A year different in starting PBs between the sex. Tanner stage of puberty development not reported/aggregated for either sex Unclear if outcome measured was of primary importance to participants	<400 optimal information size			

Risk of Bias ROBINS – 1

Confounding domains: socio-demographic situation; family support; public funding available; enrolled in a specialised service; puberty development (Tanner stage)

Co-interventions likely to have impact: counselling, family therapy, school based support, peer support, community group engagement

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Bias due to confounding Critical risk of bias (the study is too problematic to provide any useful	Confounding inherently not controllable "therapeutic engagement with Gender Identity Development Service" occurred whilst treated.	Confounding inherently not controllable "therapeutic engagement with Gender Identity	Confounding inherently not controllable "therapeutic engagement with Gender Identity	N/A	N/A	Confounding inherently not controllable "therapeutic engagement with Gender Identity	29/38 White ethnicity (5 unknown) Family support not reported

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
evidence on the effects of intervention)	No explanation for what this entailed.	Development Service" occurred whilst treated. No explanation for what this entailed.	Development Service" occurred whilst treated. No explanation for what this entailed.			Development Service" occurred whilst treated. No explanation for what this entailed.	Enrolled at a specialised clinic No disaggregation by Tanner stage of puberty
Bias in selection of participants into the study Serious risk of bias (the study has some important problems)	Selection into the study was related (but not very strongly) to intervention and outcome; and This could not be adjusted for in analyses; Participation in study was only open to those at clinic and tied to them receiving PBs and GAH Intervention status is well defined, all completed comprehensive assessment, and received PBs and some aspects of the assignments of intervention status were determined retrospectively – baseline	Selection into the study was related (but not very strongly) to intervention and outcome; and This could not be adjusted for in analyses; Participation in study was only open to those at clinic and tied to them receiving PBs and GAH Intervention status is well defined, all completed comprehensive assessment, and received PBs	Selection into the study was related (but not very strongly) to intervention and outcome; and This could not be adjusted for in analyses; Participation in study was only open to those at clinic and tied to them receiving PBs and GAH Intervention status is well defined, all completed comprehensive assessment, and received PBs	N/A	N/A	Selection into the study was related (but not very strongly) to intervention and outcome; and This could not be adjusted for in analyses; Participation in study was only open to those at clinic and tied to them receiving PBs and GAH Intervention status is well defined, all completed comprehensive assessment, and received PBs and	

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
	psychological data was after comprehensive assessment	and some aspects of the assignments of intervention status were determined retrospectively – baseline psychological data was after comprehensive assessment	and some aspects of the assignments of intervention status were determined retrospectively – baseline psychological data was after comprehensive assessment			some aspects of the assignments of intervention status were determined retrospectively – baseline psychological data was after comprehensive assessment	
Bias in classification of interventions Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial)	Intervention status is well defined and Some aspects of the assignments of intervention status were determined retrospectively.	Intervention status is well defined and Some aspects of the assignments of intervention status were determined retrospectively.	Intervention status is well defined and Some aspects of the assignments of intervention status were determined retrospectively.	N/A	N/A	Intervention status is well defined and Some aspects of the assignments of intervention status were determined retrospectively.	Clear intervention groups and some explanation of co-interventions
Bias due to deviations from	important co-interventions (such as therapeutic engagement	important co-interventions (such as therapeutic	Y important co-interventions (such as	N/A	N/A	important co-interventions (such as therapeutic	

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
<p>intended Interventions</p> <p>Serious risk of bias (the study has some important problems)</p>	<p>from GIDS, family therapy) were not reported on across PBs time period</p> <p>meaning the analysis was not able to estimate the effect of starting and adhering to intervention, while allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome.</p>	<p>engagement from GIDS, family therapy) were not reported on across PBs time period</p> <p>meaning the analysis was not able to estimate the effect of starting and adhering to intervention, while allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome.</p>	<p>therapeutic engagement from GIDS, family therapy) were not reported on across PBs time period</p> <p>meaning the analysis was not able to estimate the effect of starting and adhering to intervention, while allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome.</p>			<p>engagement from GIDS, family therapy) were not reported on across PBs time period</p> <p>meaning the analysis was not able to estimate the effect of starting and adhering to intervention, while allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome.</p>	
<p>Bias due to missing data</p> <p>Serious risk of bias (the study has some important problems)</p>	<p>Reasons for missingness in questionnaires are not reported across PBs and GAH interventions and the analysis is unlikely to have removed the risk of bias arising from the missing data;</p>	<p>Reasons for missingness in questionnaires are not reported across PBs and GAH interventions and the analysis is unlikely to have</p>	<p>Reasons for missingness in questionnaires are not reported across PBs and GAH interventions and the analysis is</p>	N/A	N/A	<p>Reasons for missingness in questionnaires are not reported across PBs and GAH interventions and the analysis is unlikely to have removed the risk of</p>	

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
		removed the risk of bias arising from the missing data;	unlikely to have removed the risk of bias arising from the missing data;			bias arising from the missing data;	
Bias in measurement of outcomes Serious risk of bias (the study has some important problems)	<p>The outcome measures were subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants);</p> <p>and</p> <p>The outcomes were assessed by assessors aware of the intervention received by study participants</p>	<p>The outcome measures were subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants);</p> <p>and</p> <p>The outcomes were assessed by assessors aware of the intervention received by study participants</p>	<p>The outcome measures were subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants);</p> <p>and</p> <p>The outcomes were assessed by assessors aware of the intervention received by study participants</p>	N/A	N/A	<p>The outcome measures were subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants);</p> <p>and</p> <p>The outcomes were assessed by assessors aware of the intervention received by study participants</p>	
Bias in selection of the reported result Moderate risk of bias (the study is sound for a non-randomized study)	<p>The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent;</p>	<p>The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and</p>	<p>The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and</p>	N/A	N/A	<p>The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent;</p>	

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
with regard to this domain but cannot be considered comparable to a well-performed randomized trial)	and There is no indication of selection of the reported analysis from among multiple analyses; and There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results.	both internally and externally consistent; and There is no indication of selection of the reported analysis from among multiple analyses; and There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results.	externally consistent; and There is no indication of selection of the reported analysis from among multiple analyses; and There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results.			and There is no indication of selection of the reported analysis from among multiple analyses; and There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results.	
Critical risk of bias	Critical risk of bias	Critical risk of bias	Critical risk of bias	N/A	N/A	Critical risk of bias	

López de Lara, D., Pérez Rodríguez, O., Cuellar Flores, I., Pedreira Masa, J. L., Campos-Muñoz, L., Cuesta Hernández, M., & Ramos Amador, J. T. (2020). Psychosocial assessment in transgender adolescents. *Anales de Pediatría*, 93(1), 41-48. doi:10.1016/j.anpedi.2020.01.019

GRADE Evidence Profile

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall
Gender dysphoria	No change	No change Results consistently reported across participants and sub-groups	↓ 1 level Intervention is indirectly related to the study due to measuring impact of GAHT on gender dysphoria rather than PBs, however age range is the same as target population for the review & baseline T0 scores evidence gender dysphoria while receiving PBs	↓1 level Few participants <400 optimal information size	No change Not assessed	↑ 1 level Control group No missing participants Reported confounding factors of family support, socioeconomic and ethnicity Participants volunteered (high level of informed consent)	Moderate
Suicidality	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Self-harm	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Anxiety	No change	Results consistently reported across participants and sub-groups	↓ 1 level Intervention is indirectly related to the study due to measuring impact of	↓1 level Few participants <400 optimal information size	No change Not assessed	↑ 1 level Control group No missing participants	Moderate

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall
			GAHT on gender dysphoria rather than PBs, however age range is the same as target population for the review & baseline T0 scores evidence gender dysphoria while receiving PBs			Reported confounding factors of family support, socioeconomic and ethnicity Participants volunteered (high level of informed consent)	
Depression	No change	Results consistently reported across participants and sub-groups	↓ 1 level Intervention is indirectly related to the study due to measuring impact of GAHT on gender dysphoria rather than PBs, however age range is the same as target population for the review & baseline T0 scores evidence gender dysphoria while receiving PBs	↓ 1 level Few participants <400 optimal information size	No change Not assessed	↑ 1 level Control group No missing participants Reported confounding factors of family support, socioeconomic and ethnicity Participants volunteered (high level of informed consent)	Moderate
Life satisfaction/QoL	No change	Results consistently reported across	↓ 1 level Intervention is indirectly related to	↓ 1 level Few participants	No change Not assessed	↑ 1 level Control group	Moderate

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall
		participants and sub-groups	the study due to measuring impact of GAHT on gender dysphoria rather than PBs, however age range is the same as target population for the review & baseline T0 scores evidence gender dysphoria while receiving PBs	<400 optimal information size		No missing participants Reported confounding factors of family support, socioeconomic and ethnicity Participants volunteered (high level of informed consent)	

Risk of Bias – ROBIN-1

Confounding domains: socio-demographic situation; family support; public funding available; enrolled in a specialised service; puberty development (Tanner stage)

Co-interventions likely to have impact: counselling, family therapy, school based support, peer support, community group engagement

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Bias due to confounding Moderate risk of bias (the study is sound for a non-randomized study with regard to this)	Confounding expected, all known important confounding domains appropriately measured and controlled for (family support, socio-economic, ethnicity, enrolled in a clinic) with the exception	N/A	N/A	Confounding expected, all known important confounding domains appropriately measured and controlled for	Confounding expected, all known important confounding domains appropriately measured and controlled for	Confounding expected, all known important confounding domains appropriately measured and controlled for (family support, socio-	Extensive effort to control for multiple confounding variables

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
domain but cannot be considered comparable to a well-performed randomized trial)	of Tanner stage of participants when commenced PBs] and Reliability and validity of measurement of anxiety, gender dysphoria and depression were sufficient, such that we do not expect serious residual confounding.			(family support, socio-economic, ethnicity, enrolled in a clinic) with the exception of Tanner stage of participants when commenced PBs] and Reliability and validity of measurement of anxiety, gender dysphoria and depression were sufficient, such that we do not expect serious residual confounding.	(family support, socio-economic, ethnicity, enrolled in a clinic) with the exception of Tanner stage of participants when commenced PBs] and Reliability and validity of measurement of anxiety, gender dysphoria and depression were sufficient, such that we do not expect serious residual confounding.	economic, ethnicity, enrolled in a clinic) with the exception of Tanner stage of participants when commenced PBs] and Reliability and validity of measurement of anxiety, gender dysphoria and depression were sufficient, such that we do not expect serious residual confounding.	
Bias in selection of participants into the study Serious risk of bias (the study has some important problems)	Selection into the study was related (but not very strongly) to receiving GAHT and due to being volunteers likely to have impacted outcomes; and	N/A	N/A	Selection into the study was related (but not very strongly) to receiving GAHT and due to being volunteers likely	Selection into the study was related (but not very strongly) to receiving GAHT and due to being volunteers likely	Selection into the study was related (but not very strongly) to receiving GAHT and due to being volunteers likely to have impacted outcomes;	

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
	This could not be adjusted for in analyses			to have impacted outcomes; and This could not be adjusted for in analyses	to have impacted outcomes; and This could not be adjusted for in analyses	and This could not be adjusted for in analyses	
Bias in classification of interventions Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial)	Intervention status was well defined (all on PBs and then GAHT); and some aspects of the assignments of intervention status were determined	N/A	N/A	Intervention status was well defined (all on PBs and then GAHT); and some aspects of the assignments of intervention status were determined	Intervention status was well defined (all on PBs and then GAHT); and some aspects of the assignments of intervention status were determined	Intervention status was well defined (all on PBs and then GAHT); and some aspects of the assignments of intervention status were determined	retrospectively [unknown due to volunteers being participants and limited other information about selection process].
Bias due to deviations from intended Interventions Serious risk of bias (the study has some important problems)	Effect of assignment to intervention: There were deviations from usual practice, with GAHT provided after age 14 instead of after the usual 16 years and is likely to have affected outcomes	N/A	N/A	Effect of assignment to intervention: There were deviations from usual practice, with GAHT provided after age 14 instead of	Effect of assignment to intervention: There were deviations from usual practice, with GAHT provided after age 14 instead of	Effect of assignment to intervention: There were deviations from usual practice, with GAHT provided after age 14 instead of after the usual 16 years and is likely to	There were deviations from usual practice of administering GAHT, it was commenced by age 14 for all participants instead of after the usual 16 years

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
				after the usual 16 years and is likely to have affected outcomes	after the usual 16 years and is likely to have affected outcomes	have affected outcomes	
Bias due to missing data Low risk of bias (the study is comparable to a well-performed randomized trial with regard to this domain)	Data were reasonably complete, no missing data was reported in analysis	N/A	N/A	Data were reasonably complete, no missing data was reported in analysis	Data were reasonably complete, no missing data was reported in analysis	Data were reasonably complete, no missing data was reported in analysis	
Bias in measurement of outcomes Serious risk of bias (the study has some important problems)	The outcome measures were subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants); and The outcomes were assessed by assessors aware of the intervention received by study participants;	N/A	N/A	The outcome measures were subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants); and The outcomes were assessed by assessors aware of the intervention	The outcome measures were subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants); and The outcomes were assessed by assessors aware of the intervention	The outcome measures were subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants); and The outcomes were assessed by assessors aware of the intervention received by study participants;	

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
				received by study participants;	received by study participants;		
<p>Bias in selection of the reported result</p> <p>Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial)</p>	<p>The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent;</p> <p>and</p> <p>There is no indication of selection of the reported analysis from among multiple analyses;</p> <p>and</p> <p>There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results.</p>	N/A	N/A	<p>The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent;</p> <p>and</p> <p>There is no indication of selection of the reported analysis from among multiple analyses;</p> <p>and</p> <p>There is no indication of selection of the cohort or subgroups for analysis and reporting on the</p>	<p>The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent;</p> <p>and</p> <p>There is no indication of selection of the reported analysis from among multiple analyses;</p> <p>and</p> <p>There is no indication of selection of the cohort or subgroups for analysis and reporting on the</p>	<p>The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent;</p> <p>and</p> <p>There is no indication of selection of the reported analysis from among multiple analyses;</p> <p>and</p> <p>There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results.</p>	

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
				basis of the results.	basis of the results.		
Overall	Serious risk of bias	N/A	N/A	Serious risk of bias	Serious risk of bias	Serious risk of bias	

Olsavsky, A. L., Grannis, C., Bricker, J., Chelvakumar, G., Indyk, J. A., Leibowitz, S. F., . . . Nahata, L. (2023). Associations Among Gender-Affirming Hormonal Interventions, Social Support, and Transgender Adolescents' Mental Health. *The Journal of adolescent health : official publication of the Society for Adolescent Medicine*, 72(6), 860-868. doi: <https://dx.doi.org/10.1016/j.jadohealth.2023.01.031>

GRADE Evidence profile

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall quality
Gender dysphoria	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Suicidality	<p>↓2 level</p> <p>No follow up study</p> <p>36% more of the sample (from non-intervention group went onto receive hormone interventions so were exploring the idea at time of study)</p> <p>Sample size too small for treatment/no treatment case matching</p> <p>Suicidality high self-reporting rate but not using a validated assessment</p>	<p>↓1 level</p> <p>Heterogeneity from puberty blocker vs other hormone treatments; age group or by Tanner stage not explained</p>	<p>↓1 level</p> <p>Direct comparisons between treatment/non-treatment cohorts have occurred.</p> <p>Likely not applicable to community based populations, non-white and those not enrolled in a specialised clinic</p>	<p>↓1 level</p> <p>Few participants <400 optimal information size</p>	<p>No change</p> <p>Not assessed</p>	<p>No change</p>	<p>Very low</p>

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall quality
Self-harm	<p>↓2 level</p> <p>No follow up study</p> <p>36% more of the sample (from non-intervention group went onto receive hormone interventions so were exploring the idea at time of study)</p> <p>Sample size too small for treatment/no treatment case matching</p> <p>NSSI high self-reporting rate but not using a validated assessment</p>	<p>↓1 level</p> <p>heterogeneity from puberty blocker vs other hormone treatments; age group or by Tanner stage not explained</p>	<p>↓1 level</p> <p>Direct comparisons between treatment/non-treatment cohorts have occurred.</p> <p>Likely not applicable to community based populations, non-white and those not enrolled in a specialised clinic</p>	<p>↓1 level</p> <p>Few participants</p> <p><400 optimal information size</p>	<p>No change</p> <p>Not assessed</p>	<p>No change</p>	<p>Very low</p>
Anxiety	<p>↓2 level</p> <p>No follow up study</p> <p>36% more of the sample (from non-intervention group went onto receive hormone interventions so were exploring the idea at time of study)</p>	<p>↓1 level</p> <p>Heterogeneity from puberty blocker vs other hormone treatments; age group or by Tanner stage not explained</p>	<p>↓1 level</p> <p>Direct comparisons between treatment/non-treatment cohorts have occurred.</p> <p>Likely not applicable to community based populations, non-white and those not</p>	<p>↓1 level</p> <p>Few participants</p> <p><400 optimal information size</p>	<p>No change</p> <p>Not assessed</p>	<p>No change</p>	<p>No change</p> <p>Very low</p>

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall quality
	<p>Sample size too small for treatment/no treatment case matching</p> <p>Baseline anxiety, was in clinical indication range.</p>		enrolled in a specialised clinic				
Depression	<p>↓1 level</p> <p>No follow up study</p> <p>36% more of the sample (from non-intervention group went onto receive hormone interventions so were exploring the idea at time of study)</p> <p>Sample size too small for treatment/no treatment case matching</p>	<p>↓1 level</p> <p>Heterogeneity from puberty blocker vs other hormone treatments; age group or by Tanner stage not explained</p>	<p>↓1 level</p> <p>Direct comparisons between treatment/non-treatment cohorts have occurred.</p> <p>Likely not applicable to community based populations, non-white and those not enrolled in a specialised clinic</p>	<p>↓1 level</p> <p>Few participants <400 optimal information size</p>	<p>No change</p> <p>Not assessed</p>	<p>No change</p> <p>Direct evidence (P<0.05) of co-relation between depression symptoms improving with receiving/not receiving gender affirming hormone treatment</p>	Very low
Life satisfaction/QoL	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Risk of Bias ROBINS- 1

Confounding domains: socio-demographic situation; family support; public funding available; enrolled in a specialised service; puberty development (Tanner stage)

Co-interventions likely to have impact: counselling, family therapy, school based support, peer support, community group engagement

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Bias due to confounding Serious risk of bias (the study has some important problems)	N/A	Therapeutic intervention/counselling was not appropriately measured, or not controlled for;	Therapeutic intervention/counselling was not appropriately measured, or not controlled for;	Therapeutic intervention/counselling was not appropriately measured, or not controlled for;	Therapeutic intervention/counselling was not appropriately measured, or not controlled for;	N/A	Reliability or validity of suicidality and NSSI not validated assessments
Bias in selection of participants into the study Serious risk of bias (the study has some important problems)	N/A	Selection into the study was related to enrolment to multi-disciplinary clinic and likelihood of receiving PB treatment. This could not be adjusted for in analyses	Selection into the study was related to enrolment to multi-disciplinary clinic and likelihood of receiving PB treatment. This could not be adjusted for in analyses	Selection into the study was related to enrolment to multi-disciplinary clinic and likelihood of receiving PB treatment. This could not be adjusted for in analyses	Selection into the study was related to enrolment to multi-disciplinary clinic and likelihood of receiving PB treatment. This could not be adjusted for in analyses	N/A	
Bias in classification of interventions	N/A	Intervention status was not well defined between puberty	Intervention status was not well defined between puberty	Intervention status was not well defined between puberty	Intervention status was not well defined between puberty	N/A	

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Serious risk of bias (the study has some important problems)		blockers and other hormone treatments	blockers and other hormone treatments	blockers and other hormone treatments	blockers and other hormone treatments		
Bias due to deviations from intended interventions	N/A	N/A	N/A	N/A	N/A	N/A	Not a follow up study but was included due to limited literature available that assessed suicidality
Bias due to missing data Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial)	N/A	The analysis is unlikely to have removed the risk of bias arising from the missing data due to the age and developmental stage of mental health of those on puberty blockers compared to more advanced adolescents	The analysis is unlikely to have removed the risk of bias arising from the missing data due to the age and developmental stage of mental health of those on puberty blockers compared to more advanced adolescents	The analysis is unlikely to have removed the risk of bias arising from the missing data due to the age and developmental stage of mental health of those on puberty blockers compared to more advanced adolescents	The analysis is unlikely to have removed the risk of bias arising from the missing data due to the age and developmental stage of mental health of those on puberty blockers compared to more advanced adolescents	N/A	Proportions of puberty blockers compared to other hormone treatments and reasons for why 2 participants were on both were not reported across treatment and non-treatment groups
Bias in measurement of outcomes	N/A	The outcome measures were self-reports and subjective (i.e.	The outcome measures were self-reports and subjective (i.e.	The outcome measures were self-reports and subjective (i.e.	The outcome measures were self-reports and subjective (i.e.	N/A	

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Serious risk of bias (the study has some important problems)		vulnerable to influence by knowledge of the intervention received by study participants); and The assessments were scored by assessors aware of the intervention received by study participants	vulnerable to influence by knowledge of the intervention received by study participants); and The assessments were scored by assessors aware of the intervention received by study participants	vulnerable to influence by knowledge of the intervention received by study participants); and The assessments were scored by assessors aware of the intervention received by study participants	vulnerable to influence by knowledge of the intervention received by study participants); and The assessments were scored by assessors aware of the intervention received by study participants		
Bias in selection of the reported result Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed	N/A	The outcome measurements and analyses are consistent with an a priori plan and There was no indication of selection of the reported analysis from among multiple analyses; and There is no indication of	The outcome measurements and analyses are consistent with an a priori plan and There was no indication of selection of the reported analysis from among multiple analyses; and There is no indication of	The outcome measurements and analyses are consistent with an a priori plan and There was no indication of selection of the reported analysis from among multiple analyses; and There is no indication of	The outcome measurements and analyses are consistent with an a priori plan and There was no indication of selection of the reported analysis from among multiple analyses; and There is no indication of	N/A	

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
randomized trial)		selection of the cohort or subgroups for analysis and reporting on the basis of the results	selection of the cohort or subgroups for analysis and reporting on the basis of the results	selection of the cohort or subgroups for analysis and reporting on the basis of the results	selection of the cohort or subgroups for analysis and reporting on the basis of the results		
Serious risk of bias	N/A	Serious risk of bias	Serious risk of bias	Serious risk of bias	Serious risk of bias	N/A	

Tordoff, D. M., Wanta, J. W., Collin, A., Stepney, C., Inwards-Breland, D. J., & Ahrens, K. (2022). Mental Health Outcomes in Transgender and Nonbinary Youths Receiving Gender-Affirming Care. *JAMA network open*, 5(2), e220978. doi: <https://dx.doi.org/10.1001/jamanetworkopen.2022.0978>

GRADE Evidence Profile

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall quality
Gender dysphoria	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Suicidality	<p>↓2 levels</p> <p>Unclear surveillance for outcomes in exposed and unexposed in cohort studies due to incomplete follow-up and/or reporting on the proportion of intervention/non-intervention participants who dropped out of the study at 12 mths (6 mths n = 84, 12mths n = 65)</p>	<p>No change</p> <p>Pre-intervention baseline risk was reported</p>	<p>↓1 levels</p> <p>Most of the PB cohort (14/19) also reported receiving GAH as well</p>	<p>↓1 levels</p> <p>Few participants <400 optimal information size</p>	Not assessed	<p>↑ 1 level</p> <p>4 confounding variables were modelled and limitations included all confounding interventions</p>	low
Self-harm	<p>↓2 levels</p> <p>Unclear surveillance for outcomes in exposed and unexposed in cohort</p>	<p>No change</p> <p>Pre-intervention baseline risk was reported</p>	<p>↓1 levels</p> <p>Most of the PB cohort (14/19) also reported receiving GAH as well</p>	<p>↓1 levels</p> <p>Few participants <400 optimal information size</p>	Not assessed	<p>↑ 1 level</p> <p>4 confounding variables were modelled and limitations</p>	low

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall quality
	<p>studies due to incomplete follow-up and/or reporting on the proportion of intervention/non-intervention participants who dropped out of the study at 12 mths</p> <p>(6 mths n = 84, 12mths n = 65)</p>					included all confounding interventions	
Anxiety	<p>↓1 levels</p> <p>Unclear surveillance for outcomes in exposed and unexposed in cohort studies due to incomplete follow-up and/or reporting on the proportion of intervention/non-intervention participants who dropped out of the study at 12 mths</p> <p>Only one question used out of a standardised assessment to assess</p>	<p>No change</p> <p>Pre-intervention baseline risk was reported</p>	<p>↓1 levels</p> <p>Most of the PB cohort (14/19) also reported receiving GAH as well</p>	<p>↓1 levels</p> <p>Few participants <400 optimal information size</p>	Not assessed	<p>↑ 1 level</p> <p>4 confounding variables were modelled and limitations included all confounding interventions</p>	moderate

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall quality
Depression	<p>↓1 levels</p> <p>Unclear surveillance for outcomes in exposed and unexposed in cohort studies due to incomplete follow-up and/or reporting on the proportion of intervention/non-intervention participants who dropped out of the study at 12 mths</p>	<p>No change</p> <p>Pre-intervention baseline risk was reported</p>	<p>↓1 levels</p> <p>Most of the PB cohort (14/19) also reported receiving GAH as well</p>	<p>↓1 levels</p> <p>Few participants <400 optimal information size</p>	Not assessed	<p>↑ 1 level</p> <p>4 confounding variables were modelled and limitations included all confounding interventions</p>	moderate
Life satisfaction/QoL	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Risk of Bias ROBINS-1

Confounding domains: socio-demographic situation; family support; public funding available; enrolled in a specialised service; puberty development (Tanner stage)

Co-interventions likely to have impact: counselling, family therapy, school based support, peer support, community group engagement

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Bias due to confounding	N/A	Confounding expected, all known important	Confounding expected, all known important	Confounding expected, all known important	Confounding expected, all known important	N/A	Ethnicity, receiving mental health therapy, family support and

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial)		confounding domains appropriately measured and controlled for; and Reliability and validity of measurement of important domains were sufficient, such that we do not expect serious residual confounding.	confounding domains appropriately measured and controlled for; and Reliability and validity of measurement of important domains were sufficient, such that we do not expect serious residual confounding.	confounding domains appropriately measured and controlled for; and Reliability and validity of measurement of important domains were sufficient, such that we do not expect serious residual confounding.	confounding domains appropriately measured and controlled for; and Reliability and validity of measurement of important domains were sufficient, such that we do not expect serious residual confounding.		substance use all controlled for in analysis
Bias in selection of participants into the study Serious risk of bias (the study has some important problems)	N/A	Selection into the study was by enrolment in a clinic and likely related to intervention and outcome; and This could not be adjusted for in analyses;	Selection into the study was by enrolment in a clinic and likely related to intervention and outcome; and This could not be adjusted for in analyses;	Selection into the study was by enrolment in a clinic and likely related to intervention and outcome; and This could not be adjusted for in analyses;	Selection into the study was by enrolment in a clinic and likely related to intervention and outcome; and This could not be adjusted for in analyses;	N/A	Difficult to extract the PB from gender affirming hormone participants and clinic based sample

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
		<p>Start of follow up and start of intervention do not coincide – some participants started PB and/or GAH within the 12mth follow up period</p> <p>and</p> <p>a potentially important amount of follow-up time is missing from analyses as do not know the stop-start dates of PB interventions</p> <p>and</p> <p>the rate ratio is not constant over time as 12mth survey completion rate was different to 6 mths survey and age rate changed to 13-17 from 13-20 years.</p>	<p>Start of follow up and start of intervention do not coincide – some participants started PB and/or GAH within the 12mth follow up period</p> <p>and</p> <p>a potentially important amount of follow-up time is missing from analyses as do not know the stop-start dates of PB interventions</p> <p>and</p> <p>the rate ratio is not constant over time as 12mth survey completion rate was different to 6 mths survey and age rate changed to 13-17 from 13-20 years.</p>	<p>Start of follow up and start of intervention do not coincide – some participants started PB and/or GAH within the 12mth follow up period</p> <p>and</p> <p>a potentially important amount of follow-up time is missing from analyses as do not know the stop-start dates of PB interventions</p> <p>and</p> <p>the rate ratio is not constant over time as 12mth survey completion rate was different to 6 mths survey and age rate changed to 13-17 from 13-20 years.</p>	<p>Start of follow up and start of intervention do not coincide – some participants started PB and/or GAH within the 12mth follow up period</p> <p>and</p> <p>a potentially important amount of follow-up time is missing from analyses as do not know the stop-start dates of PB interventions</p> <p>and</p> <p>the rate ratio is not constant over time as 12mth survey completion rate was different to 6 mths survey and age rate changed to 13-17 from 13-20 years.</p>		

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Bias in classification of interventions Serious risk of bias (the study has some important problems)	N/A	Major aspects of the intervention was determined in a way that could have been affected by knowledge of the outcome.	Major aspects of the intervention was determined in a way that could have been affected by knowledge of the outcome	Major aspects of the intervention was determined in a way that could have been affected by knowledge of the outcome	Major aspects of the intervention was determined in a way that could have been affected by knowledge of the outcome	N/A	Longitudinal outcomes likely to be known from clinical experience/prior cases from clinic
Bias due to deviations from intended interventions Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial)	N/A	There were possibly deviations from intended intervention, but their impact on the outcome is expected to be slight.	There were possibly deviations from intended intervention, but their impact on the outcome is expected to be slight.	There were possibly deviations from intended intervention, but their impact on the outcome is expected to be slight.	There were possibly deviations from intended intervention, but their impact on the outcome is expected to be slight.	N/A	Adherence to PB regimes and starting GAH/stopping PB was not detailed for the 19 who received it
Bias due to missing data Serious risk of bias (the study has some important problems)	N/A	Proportions of missing participants is unknown across intervention cohorts;	Proportions of missing participants is unknown across intervention cohorts;	Proportions of missing participants is unknown across intervention cohorts;	Proportions of missing participants is unknown across intervention cohorts;	N/A	6 mth follow up n = 84 12 mth follow up n = 65 Unknown how many received PBs and how many were not

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
		and The nature of the missing data means that the risk of bias cannot be removed through appropriate analysis.	and The nature of the missing data means that the risk of bias cannot be removed through appropriate analysis.	and The nature of the missing data means that the risk of bias cannot be removed through appropriate analysis.	and The nature of the missing data means that the risk of bias cannot be removed through appropriate analysis.		
Bias in measurement of outcomes Serious risk of bias (the study has some important problems)	N/A	The outcome measures were subjective as all self-reported (i.e. vulnerable to influence by knowledge of the intervention received by study participants); and The outcome was assessed by assessors aware of the intervention received by study participants	The outcome measures were subjective as all self-reported (i.e. vulnerable to influence by knowledge of the intervention received by study participants); and The outcome was assessed by assessors aware of the intervention received by study participants	The outcome measures were subjective as all self-reported (i.e. vulnerable to influence by knowledge of the intervention received by study participants); and The outcome was assessed by assessors aware of the intervention received by study participants	The outcome measures were subjective as all self-reported (i.e. vulnerable to influence by knowledge of the intervention received by study participants); and The outcome was assessed by assessors aware of the intervention received by study participants	N/A	All measurement tools were self-reported surveys/questionnaires
Bias in selection of the reported result	N/A	The outcome measurements and	The outcome measurements	The outcome measurements	The outcome measurements	N/A	Cohort reporting clearly explained and limitations

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial)		analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent; and There is no indication of selection of the reported analysis from among multiple analyses; and There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results.	and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent; and There is no indication of selection of the reported analysis from among multiple analyses; and There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results.	and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent; and There is no indication of selection of the reported analysis from among multiple analyses; and There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results.	and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent; and There is no indication of selection of the reported analysis from among multiple analyses; and There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results.		and caveats provided e.g 13-17 years only for statistical analysis
Serious risk of bias	N/A	N/A	Serious risk of bias	Serious risk of bias	Serious risk of bias	N/A	

Turban, J. L., King, D., Carswell, J. M., & Keuroghlian, A. S. (2020). Pubertal Suppression for Transgender Youth and Risk of Suicidal Ideation. *Pediatrics*, 145(2). doi: <https://dx.doi.org/10.1542/peds.2019-1725>

GRADE Evidence Profile

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall quality
Gender dysphoria	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Suicidality	<p>↓1 level</p> <p>Not a follow up study but retrospective analysis which does provide longitudinal results</p> <p>No results for <18 years skewing results to ages a different social context</p>	<p>No change</p> <p>Investigators have explored explanations for heterogeneity, and offered several plausible explanations</p>	<p>↓1 levels</p> <p>Differences in interventions is not possible to verify as using self-reported uptake of PBs and in the past</p>	<p>↓2 levels</p> <p>Full context of findings is not adequate due to self-reporting of PB uptake and suicidality, no ability to analyse from social and geographic location and the v small intervention cohort group (89)</p>	<p>No change</p> <p>Not assessed due to limited information</p>	<p>No change</p> <p>>large sample size n = 3954</p>	Very low
Self-harm	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Anxiety	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Depression	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Life satisfaction/QoL	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Risk of Bias ROBINS- 1

Confounding domains: socio-demographic situation; family support; public funding available; enrolled in a specialised service; puberty development (Tanner stage)

Co-interventions likely to have impact: counselling, family therapy, school based support, peer support, community group engagement

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
<p>Bias due to confounding Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial)</p>	N/A	<p>Confounding expected, most known important confounding domains appropriately measured and controlled for;</p> <p>and</p> <p>Reliability and validity of measurement of important domains were sufficient, such that we do not expect serious residual confounding.</p>	N/A	N/A	N/A	N/A	family support, sexual orientation, education level, employment status, and total household income, age, gender identity, ethnicity and relationship status controlled for psychological support not controlled for
<p>Bias in selection of participants into the study Low risk of bias (the study is comparable to a well-performed randomized trial)</p>	N/A	All participants who were eligible for the inclusion in the survey responses were included in the study;	N/A	N/A	N/A	N/A	Statistical methods for inclusion ensured that all participants should have been captured from data set

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
with regard to this domain)							
Bias in classification of interventions No information on which to base a judgement about risk of bias for this domain	N/A	Not possible to assess	N/A	N/A	N/A	N/A	Unable to determine if it was PB or another GAH from self-reports, no information about prescription of intervention
Bias due to deviations from intended Interventions No information on which to base a judgement about risk of bias for this domain	N/A	Not possible to assess	N/A	N/A	N/A	N/A	Unable to determine adherence to PB regimes followed from information
Bias due to missing data No information on which to base a judgement about risk of bias for this domain	N/A	N/A	N/A	N/A	N/A	N/A	No information provided about missed Qs from the survey within the cohort groups

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Bias in measurement of outcomes Serious risk of bias (the study has some important problems)	N/A	The Survey used was subjective as self-reported	N/A	N/A	N/A	N/A	Survey used is not a validated assessment of suicidality; unclear how the K6+ mental health questionnaire is embedded in this study
Bias in selection of the reported result Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial)	N/A	The outcome measurements and analyses are consistent with an a priori plan; and There is no indication of selection of the reported analysis from among multiple analyses; and There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results.	N/A	N/A	N/A	N/A	Methods and analysis are consistent
Overall Serious risk of bias	N/A	Serious risk of bias	N/A	N/A	N/A	N/A	

Kuper, L. E., Stewart, S., Preston, S., Lau, M., & Lopez, X. (2020). Body Dissatisfaction and Mental Health Outcomes of Youth on Gender-Affirming Hormone Therapy. *Pediatrics*, 145(4). doi:10.1542/peds.2019-3006

GRADE Evidence Profile

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall quality
Gender dysphoria	<p>↓1 level</p> <p>No control/comparison</p>	<p>No change</p> <p>Limited explanation of the differences between PB and other cohort groups which may have had effects, some caveats provided in text and tables to account for this</p>	<p>↓1 level</p> <p>Body image scale has been used rather than a gender dysphoria assessment</p> <p>Limited explanation of the differences between PB and other cohort groups which may have had effects, some caveats provided in text and tables to account for this</p>	<p>↓1 level</p> <p>Few participants <400 optimal information size</p>	<p>No change</p> <p>Not assessed</p>	<p>No change</p>	<p>Very low</p>
Suicidality	<p>↓2 level</p> <p>No control/comparison</p> <p>No information about the PB sub-group in relation to suicidal ideation</p>	<p>↓1 level</p> <p>Heterogeneity not able to be evaluated due to no reporting of the differences between PB and other cohort groups</p>	<p>↓2 level</p> <p>No direct evidence for PB sub-group about suicidality</p>	<p>↓1 level</p> <p>Few participants <400 optimal information size</p>	<p>No change</p> <p>Not assessed</p>	<p>No change</p>	<p>Very low</p>

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall quality
Self-harm	<p>↓2 level</p> <p>No control/comparison</p> <p>No information about the PB sub-group in relation to self-harm</p>	<p>↓1 level</p> <p>Heterogeneity not able to be evaluated due to no reporting of the differences between PB and other cohort groups</p>	<p>↓2 level</p> <p>No direct evidence for PB sub-group about self-harm</p>	<p>↓1 level</p> <p>Few participants <400 optimal information size</p>	<p>No change</p> <p>Not assessed</p>	<p>No change</p>	<p>Very low</p>
Anxiety	<p>↓1 level</p> <p>No control</p>	<p>No change</p> <p>Limited explanation of the differences between PB and other cohort groups but heterogeneity evident in of results tables</p>	<p>No change</p> <p>Direct measurement of anxiety and reporting of subset anxiety results for all cohorts provided</p>	<p>↓1 level</p> <p>Few participants <400 optimal information size</p>	<p>No change</p> <p>Not assessed</p>	<p>No change</p>	<p>Moderate</p>
Depression	<p>↓1 level</p> <p>No control</p>	<p>No change</p> <p>Limited explanation of the differences between PB and other cohort groups but heterogeneity evident in of results tables</p>	<p>No change</p> <p>Direct measurement of depression including self-reported and clinically reported for all cohorts provided</p>	<p>↓1 level</p> <p>Few participants <400 optimal information size</p>	<p>No change</p> <p>Not assessed</p>	<p>No change</p>	<p>Moderate</p>
Life satisfaction/QoL	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Risk of Bias – ROBINS - 1

Confounding domains: socio-demographic situation; family support; public funding available; enrolled in a specialised service; puberty development (Tanner stage)

Co-interventions likely to have impact: counselling, family therapy, school based support, peer support, community group engagement

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Bias due to confounding Serious risk of bias (the study has some important problems)	At least one known important domain was not appropriately measured [family support] or not controlled for;	At least one known important domain was not appropriately measured [family support] or not controlled for;	At least one known important domain was not appropriately measured [family support] or not controlled for;	At least one known important domain was not appropriately measured [family support] or not controlled for;	At least one known important domain was not appropriately measured [family support] or not controlled for;	N/A	Many other confounding variables factored in study but not the key one of family support
Bias in selection of participants into the study Serious risk of bias (the study has some important problems)	Selection into the study was related to receiving PBs and outcome; and This could not be adjusted for in analyses;	Selection into the study was related to receiving PBs and outcome; and This could not be adjusted for in analyses;	Selection into the study was related to receiving PBs and outcome; and This could not be adjusted for in analyses;	Selection into the study was related to receiving PBs and outcome; and This could not be adjusted for in analyses;	Selection into the study was related to receiving PBs and outcome; and This could not be adjusted for in analyses;	N/A	PBs received prerequisite for study. Clinic based inclusion criteria only
Bias in classification of interventions Moderate risk of bias (the study is sound for a non-randomized study)	Intervention status is well defined with three cohort groups	Intervention status is well defined with three cohort groups	Intervention status is well defined with three cohort groups	Intervention status is well defined with three cohort groups	Intervention status is well defined with three cohort groups	N/A	Some aspects of the assignments of intervention status were unclear (i.e being on both PB and GAH and if this status changed

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
with regard to this domain but cannot be considered comparable to a well-performed randomized trial)							during follow up period)
<p>Bias due to deviations from intended Interventions</p> <p>Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial)</p>	<p>Effect of starting and adhering to intervention:</p> <p>Therapy support was not balanced across intervention groups, and</p> <p>The analysis was appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome.</p>	<p>Effect of starting and adhering to intervention:</p> <p>Therapy support was not balanced across intervention groups, and</p> <p>The analysis was appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely</p>	<p>Effect of starting and adhering to intervention:</p> <p>Therapy support was not balanced across intervention groups, and</p> <p>The analysis was appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to</p>	<p>Effect of starting and adhering to intervention:</p> <p>Therapy support was not balanced across intervention groups, and</p> <p>The analysis was appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely</p>	<p>Effect of starting and adhering to intervention:</p> <p>Therapy support was not balanced across intervention groups, and</p> <p>The analysis was appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely</p>	N/A	Frequency of therapy received was accounted for in analysis

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
		to impact on the outcome.	impact on the outcome.	to impact on the outcome.	to impact on the outcome.		
<p>Bias due to missing data</p> <p>Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial)</p>	<p>Proportions of and reasons for missing participants differ slightly across intervention groups; and</p> <p>The analysis is unlikely to have removed the risk of bias arising from the missing data.</p>	<p>Proportions of and reasons for missing participants differ slightly across intervention groups; and</p> <p>The analysis is unlikely to have removed the risk of bias arising from the missing data.</p>	<p>Proportions of and reasons for missing participants differ slightly across intervention groups; and</p> <p>The analysis is unlikely to have removed the risk of bias arising from the missing data.</p>	<p>Proportions of and reasons for missing participants differ slightly across intervention groups; and</p> <p>The analysis is unlikely to have removed the risk of bias arising from the missing data.</p>	<p>Proportions of and reasons for missing participants differ slightly across intervention groups; and</p> <p>The analysis is unlikely to have removed the risk of bias arising from the missing data.</p>	N/A	<p>Body dissatisfaction (10/25) and self-report depression (13/25) low responses compared to other measures</p>
<p>Bias in measurement of outcomes</p> <p>Serious risk of bias (the study has some important problems)</p>	<p>The outcome measure was subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants); and</p> <p>The outcome was assessed by assessors aware of the</p>	<p>The outcome measure was subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants); and</p> <p>The outcome was assessed by</p>	<p>The outcome measure was subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants); and</p> <p>The outcome was assessed by</p>	<p>The outcome measure was subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants); and</p> <p>The outcome was assessed by</p>	<p>The outcome measure was subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants); and</p> <p>The outcome was assessed by</p>	N/A	<p>All measures self-reported or clinician based with full knowledge of PB and therapy provided</p>

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
	intervention received by study participants;	assessors aware of the intervention received by study participants;	assessors aware of the intervention received by study participants;	assessors aware of the intervention received by study participants;	assessors aware of the intervention received by study participants;		
<p>Bias in selection of the reported result</p> <p>Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial)</p>	<p>The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent; and</p> <p>There is no indication of selection of the reported analysis from among multiple analyses;</p> <p>and</p> <p>There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results.</p>	<p>The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent; and</p> <p>(ii) There is no indication of selection of the reported analysis from among multiple analyses;</p> <p>and</p> <p>(iii) There is no indication of selection of the cohort or subgroups for analysis and</p>	<p>The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent; and</p> <p>(ii) There is no indication of selection of the reported analysis from among multiple analyses; and</p> <p>(iii) There is no indication of selection of the cohort or subgroups for analysis and reporting on the</p>	<p>The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent; and</p> <p>(ii) There is no indication of selection of the reported analysis from among multiple analyses; and</p> <p>(iii) There is no indication of selection of the cohort or subgroups for analysis and</p>	<p>The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent; and</p> <p>(ii) There is no indication of selection of the reported analysis from among multiple analyses; and</p> <p>(iii) There is no indication of selection of the cohort or subgroups for analysis and</p>	N/A	<p>Transparent reporting of cohort groups with response rates.</p> <p>Limited explanation for missing data but not why e.g. 2/25 PB cessation but very small number</p>

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
		reporting on the basis of the results.	basis of the results.	reporting on the basis of the results.	reporting on the basis of the results.		
Serious risk of bias	Serious Risk of bias	Serious Risk of bias	Serious Risk of bias	Serious Risk of bias	Serious Risk of bias	N/A	

Achille, C., Taggart, T., Eaton, N. R., Osipoff, J., Tafuri, K., Lane, A., & Wilson, T. A. (2020). Longitudinal impact of gender-affirming endocrine intervention on the mental health and well-being of transgender youths: preliminary results. *International journal of pediatric endocrinology*, 2020, 8. doi: <https://dx.doi.org/10.1186/s13633-020-00078-2>

GRADE Evidence Profile

Outcome	Limitations	inconsistency	indirectness	imprecision	Publication	+ve factors	Overall quality
Gender dysphoria	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Suicidality	Unable to assess due to no disaggregated data for PB cohort	Unable to assess due to no disaggregated data for PB cohort	Unable to assess due to no disaggregated data for PB cohort	Unable to assess due to no disaggregated data for PB cohort	Unable to assess due to no disaggregated data for PB cohort	Unable to assess due to no disaggregated data for PB cohort	Unable to assess due to no disaggregated data for PB cohort
Self-harm	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Anxiety	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Depression	Unable to assess due to no disaggregated data for PB cohort	Unable to assess due to no disaggregated data for PB cohort	Unable to assess due to no disaggregated data for PB cohort	Unable to assess due to no disaggregated data for PB cohort	Unable to assess due to no disaggregated data for PB cohort	Unable to assess due to no disaggregated data for PB cohort	Unable to assess due to no disaggregated data for PB cohort
Life satisfaction/QoL	Unable to assess due to no disaggregated data for PB cohort	Unable to assess due to no disaggregated data for PB cohort	Unable to assess due to no disaggregated data for PB cohort	Unable to assess due to no disaggregated data for PB cohort	Unable to assess due to no disaggregated data for PB cohort	Unable to assess due to no disaggregated data for PB cohort	Unable to assess due to no disaggregated data for PB cohort

Risk of Bias ROBINS -1

Confounding domains: socio-demographic situation; family support; public funding available; enrolled in a specialised service; puberty development (Tanner stage)

Co-interventions likely to have impact: counselling, family therapy, school based support, peer support, community group engagement

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Bias due to confounding Critical risk of bias (the study is too problematic to provide any useful evidence on the effects of intervention)	N/A	Confounding inherently not controllable	N/A	N/A	Confounding inherently not controllable	Confounding inherently not controllable	Medication and receiving counselling were controlled for but key factor of family support was referred to but not accounted for. No reporting of socio-demographic analysis
Bias in selection of participants into the study Critical risk of bias (the study is too problematic to provide any useful evidence on the effects of intervention)	N/A	Selection into the study was very strongly related to intervention and outcome; and This could not be adjusted for in analyses;	N/A	N/A	Selection into the study was very strongly related to intervention and outcome; and This could not be adjusted for in analyses;	Selection into the study was very strongly related to intervention and outcome; and This could not be adjusted for in analyses;	All participants attended clinic and were referred for endocrine assessment/treatment
Bias in classification of interventions	N/A	Major aspects of the assignments of intervention	N/A	N/A	Major aspects of the assignments of intervention	Major aspects of the assignments of intervention status	All PB participants started during the 12 mths of follow up

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Serious risk of bias (the study has some important problems)		status were determined in a way that could have been affected by knowledge of the outcome.			status were determined in a way that could have been affected by knowledge of the outcome	were determined in a way that could have been affected by knowledge of the outcome	
Bias due to deviations from intended Interventions Critical risk of bias (the study is too problematic to provide any useful evidence on the effects of intervention)	N/A	Effect of starting and adhering to intervention: There were substantial imbalances in important co-interventions across intervention groups, or there were substantial deviations from the intended interventions (in terms of implementation and/or adherence) that were likely to impact on the outcome; and	N/A	N/A	Effect of starting and adhering to intervention: (i) There were substantial imbalances in important co-interventions across intervention groups, or there were substantial deviations from the intended interventions (in terms of implementation and/or adherence) that were likely to impact on the outcome; and	Effect of starting and adhering to intervention: (i) There were substantial imbalances in important co-interventions across intervention groups, or there were substantial deviations from the intended interventions (in terms of implementation and/or adherence) that were likely to impact on the outcome; and (ii) The analysis was not appropriate to estimate the effect of starting and adhering to intervention,	Uncertain adherence with therapeutic support, not disaggregated by intervention groups “Most subjects were followed by mental health professionals. Those that were not were encouraged to see a mental health professional” While this was controlled for in the regression analysis there is no discussion of the impact on the outcomes of receive therapy

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
		(ii) The analysis was not appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome.			(ii) The analysis was not appropriate to estimate the effect of starting and adhering to intervention, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome.	allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome.	
Bias due to missing data Serious risk of bias (the study has some important problems)	N/A	Proportions of missing participants is unknown across the interventions and The nature of the missing data means that the risk of bias cannot be removed through appropriate analysis.	N/A	N/A	Proportions of missing participants is unknown across the interventions and The nature of the missing data means that the risk of bias cannot be removed through appropriate analysis.	Proportions of missing participants is unknown across the interventions and The nature of the missing data means that the risk of bias cannot be removed through appropriate analysis.	50/116 completed the follow up questionnaires, no analysis of missing 66 is provided

Domain	Outcome 1 Gender dysphoria	Outcome 2 Suicidality	Outcome 3 Self-harm	Outcome 4 Anxiety	Outcome 5 Depression	Outcome 6 Life satisfaction/QoL	Comments
Bias in measurement of outcomes Serious risk of bias (the study has some important problems)	N/A	The methods of outcome assessment were not comparable across intervention groups;	N/A	N/A	The methods of outcome assessment were not comparable across intervention groups;	The methods of outcome assessment were not comparable across intervention groups;	Incomplete measures for the different cohort groups i.e PB depression, suicidality and QoL scores, unable to compare
Bias in selection of the reported result Moderate risk of bias (the study is sound for a non-randomized study with regard to this domain but cannot be considered comparable to a well-performed randomized trial)	N/A	The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent;	N/A	N/A	The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent;	The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent;	Clear method and reporting alignment
Critical Risk of Bias		Critical Risk of Bias			Critical Risk of Bias	Critical Risk of Bias	

GRADE Handbook. (2013). In H. Schünemann, J. Brožek, G. Guyatt, & A. Oxman (Eds.). Retrieved from <https://gdt.gradeapro.org/app/handbook/handbook.html#h.svwngs6pm0f2>

Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., . . . Higgins, J. P. (2016). ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*, 355, i4919. doi:10.1136/bmj.i4919