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

Ministry of Health logo

This document is available at health.govt.nz

|  |  |
| --- | --- |
| **CCBY** | 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 (*GRADE Handbook*, 2013) 1](#_Toc170398049)

[ROBINS – I Risk of Bias criteria (Sterne et al., 2016) 1](#_Toc170398050)

[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 2](#_Toc170398051)

[GRADE Evidence Profile 2](#_Toc170398052)

[Risk of Bias ROBINS-I 3](#_Toc170398053)

[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 8](#_Toc170398054)

[GRADE evidence profile 8](#_Toc170398055)

[Risk of Bias ROBINS-1 11](#_Toc170398056)

[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 15](#_Toc170398057)

[GRADE evidence profile 15](#_Toc170398058)

[Risk of Bias ROBINS – 1 18](#_Toc170398059)

[Elkadi, J., Chudleigh, C., Maguire, A. M., Ambler, G. R., Scher, S., & Kozlowska, 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 21](#_Toc170398060)

[GRADE Evidence Profile 21](#_Toc170398061)

[Risk of Bias Robins - 1 22](#_Toc170398062)

[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 26](#_Toc170398063)

[GRADE Evidence Profile 26](#_Toc170398064)

[Risk of Bias ROBINS – 1 28](#_Toc170398065)

[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 Pediatria, 93*(1), 41-48. doi:10.1016/j.anpedi.2020.01.019 34](#_Toc170398066)

[GRADE Evidence Profile 34](#_Toc170398067)

[Risk of Bias – ROBIN-1 36](#_Toc170398068)

[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 42](#_Toc170398069)

[GRADE Evidence profile 42](#_Toc170398070)

[Risk of Bias ROBINS- 1 45](#_Toc170398071)

[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 49](#_Toc170398072)

[GRADE Evidence Profile 49](#_Toc170398073)

[Risk of Bias ROBINS-1 51](#_Toc170398074)

[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 57](#_Toc170398075)

[GRADE Evidence Profile 57](#_Toc170398076)

[Risk of Bias ROBINS- 1 58](#_Toc170398077)

[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 61](#_Toc170398078)

[GRADE Evidence Profile 61](#_Toc170398079)

[Risk of Bias – ROBINS - 1 63](#_Toc170398080)

[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](#_Toc170398081)

[GRADE Evidence Profile 68](#_Toc170398082)

[Risk of Bias ROBINS -1 69](#_Toc170398083)

# 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**](https://dx.doi.org/10.1111/jsm.13034)

## GRADE Evidence Profile

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

## 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** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Bias due to confounding**  **Critical risk of bias** (the study is too problematic to provide any useful evidence on the effects of intervention) | Confounding from ethnicity, socioeconomic status, family support, enrolment in specialised service, Tanner stage inherently not controllable | **N/A** | **N/A** | **N/A** | **N/A** | Confounding from ethnicity, socioeconomic status, family support, enrolment in specialised service, Tanner stage inherently not controllable | 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 |
| **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 due to being enrolled in a specialised clinic and all participants were eligible and eventually received PBs  and  this could not be adjusted for in analyses | **N/A** | **N/A** | **N/A** | **N/A** | 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  and  this could not be adjusted for in analyses | All participants and clinicians knew they could receive PBs and/or the likely outcome of this |
| **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 between two cohorts but is unclear about process/intervention of introducing PBs to the non-receiving group who all went onto receive them  and  Some aspects of the assignments of intervention (psychological support) status were likely determined retrospectively. | **N/A** | **N/A** | **N/A** | **N/A** | 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  Some aspects of the assignments of intervention (psychological support) status were likely determined retrospectively | Assignment into eligible cohort determined by WPATH Standards of Care  Unclear what psychological support entailed and how this was assessed as received (retrospectively or case notes etc) |
| **Bias due to deviations from intended Interventions**  **Serious risk of bias** (the study has some important problems) | Effect of starting and adhering to intervention:  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;  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 likely to impact on the outcome. | **N/A** | **N/A** | **N/A** | **N/A** | Effect of starting and adhering to intervention:  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;  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 likely to impact on the outcome | Unknown factors related to any deviation of usual practice and quality of psychosocial support received by the two cohorts.  No information about adherence of either cohort group.  Likely imbalance between co-interventions received  Not re-assessing for gender dysphoria in follow up |
| **Bias due to missing data**  **Serious risk of bias** (the study has some important problems) | Reasons for missingness are not well explained across cohorts and interventions; no explanation for why gender dysphoria was not re-assessed  and  The analysis is unlikely to have removed the risk of bias arising from the missing data;  and  The nature of the missing data means that the risk of bias cannot be removed through appropriate analysis. | **N/A** | **N/A** | **N/A** | **N/A** | Reasons for missingness are not well explained across cohorts and interventions;  and  The analysis is unlikely to have removed the risk of bias arising from the missing data;  and  The nature of the missing data means that the risk of bias cannot be removed through appropriate analysis | No explanation or description of loss of participants over the three time series  No follow up scores for gender dysphoria assessment |
| **Bias in measurement of outcomes**  **Serious risk of bias** (the study has some important problems) | 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** | **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 | CGAS is a subjective assessment by clinicians.  Unknown if same assessor completed the three assessments for each participants.  Multiple assessors involved including Stockholm and England based practitioners  Unknown if assessors knew about study and intervention, but probably likely  UGDS is self-reported |
| **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 | No reporting on outcomes for the loss of participants  Delayed eligible participants all received puberty suppression treatment but psychological intervention they received not reported  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 | **↓ 2 levels**  Failure to develop and apply appropriate eligibility criteria as no inclusion of control population  Incomplete accounting of patients and outcome events  No reporting on other psychological interventions | **No change**  limited information about population and sub-groupings except AFAB/AMAB & all commenced GAHT | **↓ 1 levels**  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  Between-sex differences P= < 0.001 | **↓ 1 levels**  few patients < 400 participants  No reporting of outcomes for participants not assessed pre-post treatment  T0 n = 16  T1 n=29 | **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. | **↑ 1 level**  Consistent pre-post testing for 41/70 of cohort and has longitudinal value over several years | **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 | **↓ 2 levels**  Use of unvalidated outcome measures (e.g. patient-reported outcomes)  Incomplete accounting of patients and outcome events  No reporting on confounding interventions | **↓ 1 levels**  Patients vary widely in their pre-intervention or baseline risk as high SD in baselines | **No Change** | **↓ 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 tendency for positive results to be published sooner and for negative results to be published later or withheld. | **No change** | **Very low** |
| 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 | **No change**  Patients vary widely in their pre-intervention or baseline risk as high SD in baselines | **↓ 2 levels**  Self-reporting mean scores lie within normal range at T0 & T1  (1-10 These ups and downs are considered normal)  However study implies change in depression  “depressive symptom scores on the BDI-II significantly decreased” | **↓ 1 level**  few patients < 400 participants  no comparison group  No reporting of outcomes for participants not assessed  T0 n = 16  T1 n=29 | **↓ 1 level**  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** | **Very low** |
| **Life satisfaction/QoL** | **↓ 2 levels**  CBCL & YSR are unvalidated outcome measures (e.g. patient-reported outcomes)  No description of CGAS administration at T0 & T1 – if same clinicians/process used  Incomplete accounting of patients and outcome events  No reporting on confoundinginterventions | **No change**  Patients vary widely in their pre-intervention or baseline risk as high SD baseline anxiety figure | **↓ 2 levels**  CBCL and YSR clinical indication score is from >63 and it is unclear in results table how the conclusions have been made as mean participants scored in non-clinical range on CBCL  T0 = 60.70  TI = 54.46  YSR  T0 = 55.56 T1= 50.00  “Adolescents showed a significant decrease in behavioural and emotional problems over time on mean *T*-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. | **↓ 1 level**  few patients < 400 participants  no comparison group | **↓ 1 level**  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 reporting of outcomes for participants not assessed pre-post treatment n = 16 and n=29 | **No change** | **Very low** |

## 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**  **Serious risk of bias** (the study has some important problems) | Intervention status is not well defined;  and  Major aspects of the assignments of intervention status were determined in a way that could have been affected by knowledge of the outcome. | **N/A** | **N/A** | Intervention status is not well defined;  and  Major aspects of the assignments of intervention status were determined in a way that could have been affected by knowledge of the outcome. | Intervention status is not well defined;  and  Major aspects of the assignments of intervention status were determined in a way that could have been affected by knowledge of the outcome. | Intervention status is not well defined;  and  Major aspects of the assignments of intervention status were determined in a way that could have been affected by knowledge of the outcome. | Intervention status is not well defined, it is unclear about what 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 likely to impact on the outcome. | **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  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. | 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 likely to impact on the outcome. | 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 likely to impact on the outcome. | 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 |
| **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**  **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 outcome was 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 outcome was assessed by assessors aware of the intervention received by study participants | The outcome measures were 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 measures were 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 |  |
| **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 dysphoira  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 | **↓ 2 levels**  Failure to develop and apply appropriate eligibility criteria as no inclusion of control population  Incomplete accounting of patients and outcome events  No reporting on other psychological interventions | **No change**  limited information about population and sub-groupings except AFAB/AMAB & all commenced GAHT | **↓ 2 levels**  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  Between-sex differences P= < 0.001 | ↓ **1 levels**  few patients < 400 participants  no comparison group  No reporting of outcomes for 37 participants who also received PBs 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 | **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 | **↓ 2 levels**  Use of unvalidated outcome measures (e.g. patient-reported outcomes)  Incomplete accounting of patients and outcome events  No reporting on confounding interventions | **↓ 2 levels**  Patients vary widely in their pre-intervention or baseline risk as high SD baseline anxiety figure  Small sample size | **No Change** | **↓ 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** |
| 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 outcome measures (e.g. self-reported)  No description of CGAS administration at T0 & T1 – if same clinicians/process used  Incomplete accounting of patients and outcome events  No reporting on confoundinginterventions  WHOQOL-BREF,Satisfaction With Life Scale and Subjective Happiness Scale only administered in 2014, unable to be compared to pre PBs (T0) | **No change**  Unable to assess due to minimal information about T0, T1, T2 administration and scoring | **↓ 2 levels**  CBCL and YSR clinical scores are from >63 and it is unclear in results table how the conclusions have been made as mean participants scored in non-clinical range on CBCL T0 = 60.70  TI = 54.46  YSR T0 = 55.56 T1= 50.00  “Adolescents showed a significant decrease in behavioural and emotional problems over time on mean *T*-scores of the total problem scale, the internalizing and externalizing scale of both  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. | **↓ 1 level**  few patients < 400 participants  no comparison group  No reporting of outcomes for 30 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  WHOQOL-BREF is an extensively tested QoL instrument | **Very low** |

## 

## 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** |  |
| **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 Subjective Happiness Scale  Removal of missing participants in analysis | There is a high risk of selective reporting from among multiple analyses;  Selective reporting of the Selective reporting WHOQOL-BREF,Satisfaction With Life Scale and Subjective Happiness Scale  Removal of missing participants in analysis | There is a high risk of selective reporting from among multiple analyses;  Selective reporting WHOQOL-BREF,Satisfaction With Life Scale and Subjective Happiness Scale  Removal of missing participants in analysis | WHOQOL-BREF,Satisfaction With Life Scale and Subjective Happiness Scale  Were only administered in the 2014 study |
| **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., & Kozlowska, 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 | **↓ 2 levels**  Use of unvalidated outcome measures (self-report via telephone call/medical note review using custom-made questionnaire) | **↓ 2 levels**  No statistical analysis, % comparison only of pre-post treatment with no analysis or discussion of variables | **↓ 2 levels**  Demographic and/or social differences within population cohort not discussed | **↓ 1 levels**  <400 participants | **No change**  Unable to be determined | **No change** | **Very low** |
| Depression | **↓ 2 levels**  Use of unvalidated outcome measures (self-report via telephone call/medical note review using custom-made questionnaire) | **↓ 2 levels**  No statistical analysis, % comparison only of pre-post treatment with no analysis or discussion of variables | **↓ 2 levels**  Demographic and/or social differences within population cohort not discussed | **↓ 1 levels**  <400 participants | **No change**  Unable to be determined | **No change** | **Very low** |
| 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  A substantial amount of follow-up time is likely to be missing from analyses | Selection into the study was very strongly related to intervention and outcome;  and  This could not be adjusted for in analyses;  and  A substantial amount of follow-up time is likely to be missing from analyses | N/A | Inconsistent follow-up time period (4-9 years)  Treatment pathway was pre-determined and known by participants |
| 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  **Serious risk of bias** (the study has some important problems) | N/A | N/A | N/A | Important co-interventions were not controlled for or discussed [pre-requisite for 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 | Important co-interventions were not controlled for or discussed [pre-requisite for 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 | N/A |  |
| 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 |
| 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 | **↓1 level**  Failure to develop and apply appropriate eligibility criteria (inclusion of control population) | **↓2 levels**  38/109 of those who received PBs were reported on with no exploration of the 71 not in the data  Only 19 participants completed the gender dysphoria questionnaire in all 3 assessments | **↓2 levels**  Differences between numbers of AFAB and 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 | **↓1 level**  Few participants  <400 optimal information size | Not enough information to assess | **N/A** | **Very low** |
| Suicidality | **↓1 level**  Failure to develop and apply appropriate eligibility criteria (inclusion of control population) | **↓2 levels**  38/109 of those who received PBs were reported on with no exploration of the 71 not in the data  Only 11 participants (young person and caregiver) completed the suicidality question questionnaire in all 3 assessments | **↓2 levels**  Differences between numbers of AFAB and 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  Not stated if outcome measured was of primary importance to participants | **↓1 level**  Few participants  <400 optimal information size | Not enough information to assess | **N/A** | **Very low** |
| Self-harm | **↓1 level**  Failure to develop and apply appropriate eligibility criteria (inclusion of control population)  Baseline scores for young person and caregiver questionnaires were in normal range for self-report mental wellbeing and behaviour | **↓2 levels**  38/109 of those who received PBs were reported on with no exploration of the 71 not in the data  Only 11 participants (young person and caregiver) completed the self-harm question in the questionnaire in all 3 assessments | **↓2 levels**  Differences between numbers of AFAB and 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  Not stated if outcome measured was of primary importance to participants | **↓1 level**  Few participants  <400 optimal information size | 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 | **↓1 level**  Failure to develop and apply appropriate eligibility criteria (inclusion of control population)  Participants were all in ‘normal’ non clinical functioning range in all 3 assessments | **↓2 levels**  38/109 of those who received PBs 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 | **↓2 levels**  Differences between numbers of AFAB and 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 | **↓1 level**  Few participants  <400 optimal information size | Not enough information to assess | **N/A** | **Very low** |

## 

## 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 “therapeutic engagement with Gender Identity Development Service” occurred whilst treated. No explanation for what this entailed. | Confounding inherently not controllable “therapeutic engagement with Gender Identity Development Service” occurred whilst treated. No explanation for what this entailed. | Confounding inherently not controllable “therapeutic engagement with Gender Identity Development Service” occurred whilst treated. No explanation for what this entailed. | N/A | N/A | Confounding inherently not controllable “therapeutic engagement with Gender Identity Development Service” occurred whilst treated. No explanation for what this entailed. | 29/38 White ethnicity (5 unknown)  Family support not reported  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 psychological data was after comprehensive assessment | 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 psychological data was after comprehensive assessment | 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 psychological data was after comprehensive assessment | 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  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 intended Interventions**  **Serious risk of bias** (the study has some important problems) | important co-interventions (such as therapeutic engagement from GIDS, family therapy) were not reported on across PBs time period  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. | important co-interventions (such as therapeutic engagement from GIDS, family therapy) were not reported on across PBs time period  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. | Y important co-interventions (such as therapeutic engagement from GIDS, family therapy) were not reported on across PBs time period  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. | N/A | N/A | important co-interventions (such as therapeutic engagement from GIDS, family therapy) were not reported on across PBs time period  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. |  |
| **Bias due to missing data**  **Serious risk of bias** (the study has some important problems) | 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; | 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; | 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; | N/A | N/A | 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; |  |
| **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 | 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 | 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 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 randomized trial) | The outcome measurements 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. | The outcome measurements 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. | The outcome measurements 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. | 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;  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 Pediatria, 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 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** |
| 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 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** |

## 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 domain but cannot be considered comparable to a well-performed randomized trial) | Confounding expected, all known important confounding domains appropriately measured and controlled for (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. | N/A | | N/A | Confounding expected, all known important confounding domains appropriately measured and controlled for (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. | Confounding expected, all known important confounding domains appropriately measured and controlled for (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. | Confounding expected, all known important confounding domains appropriately measured and controlled for (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. | Extensive effort to control for multiple confounding variables |
| **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  This could not be adjusted for in analyses | N/A | | N/A | Selection into the study was related (but not very strongly) to receiving GAHT and due to being volunteers likely to have impacted outcomes;  and  This could not be adjusted for in analyses | Selection into the study was related (but not very strongly) to receiving GAHT and due to being volunteers likely to have impacted outcomes;  and  This could not be adjusted for in analyses | Selection into the study was related (but not very strongly) to receiving GAHT and due to being volunteers likely to have impacted outcomes;  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 after the usual 16 years and is likely to have affected outcomes | 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 | 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 | 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 |
| **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 received by study participants; | 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; | 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; |  |
| **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) | The outcome measurements 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. | 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;  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. | The outcome measurements 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. | The outcome measurements 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. |  |
| 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 | **↓2 level**  No follow up study  36% more of the sample (from non-intervention group went onto receive hormone interventions so were exploring the idea at time of study)  Sample size too small for treatment/no treatment case matching  Suicidality high self-reporting rate but not using a validated assessment | **↓1 level**  Heterogeneity from puberty blocker vs other hormone treatments; age group or by Tanner stage not explained | **↓1 level**  Direct comparisons between treatment/non-treatment cohorts have occurred.  Likely not applicable to community based populations, non-white and those not enrolled in a specialised clinic | **↓1 level**  Few participants  <400 optimal information size | **No change**  Not assessed | **No change** | **Very low** |
| Self-harm | **↓2 level**  No follow up study  36% more of the sample (from non-intervention group went onto receive hormone interventions so were exploring the idea at time of study)  Sample size too small for treatment/no treatment case matching  NSSI high self-reporting rate but not using a validated assessment | **↓1 level**  heterogeneity from puberty blocker vs other hormone treatments; age group or by Tanner stage not explained | **↓1 level**  Direct comparisons between treatment/non-treatment cohorts have occurred.  Likely not applicable to community based populations, non-white and those not enrolled in a specialised clinic | **↓1 level**  Few participants  <400 optimal information size | **No change**  Not assessed | **No change** | **Very low** |
| Anxiety | **↓2 level**  No follow up study  36% more of the sample (from non-intervention group went onto receive hormone interventions so were exploring the idea at time of study)  Sample size too small for treatment/no treatment case matching  Baseline anxiety, was in clinical indication range. | **↓1 level**  Heterogeneity from puberty blocker vs other hormone treatments; age group or by Tanner stage not explained | **↓1 level**  Direct comparisons between treatment/non-treatment cohorts have occurred.  Likely not applicable to community based populations, non-white and those not enrolled in a specialised clinic | **↓1 level**  Few participants  <400 optimal information size | **No change**  Not assessed | **No change** | **No change**  **Very low** |
| Depression | **↓1 level**  No follow up study  36% more of the sample (from non-intervention group went onto receive hormone interventions so were exploring the idea at time of study)  Sample size too small for treatment/no treatment case matching | **↓1 level**  Heterogeneity from puberty blocker vs other hormone treatments; age group or by Tanner stage not explained | **↓1 level**  Direct comparisons between treatment/non-treatment cohorts have occurred.  Likely not applicable to community based populations, non-white and those not enrolled in a specialised clinic | **↓1 level**  Few participants  <400 optimal information size | **No change**  Not assessed | **No change**  Direct evidence (P<0.05) of co-relation between depression symptoms improving with receiving/not receiving gender affirming hormone treatment | **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**  **Serious risk of bias** (the study has some important problems) | N/A | Intervention status was not well defined between puberty blockers and other hormone treatments | Intervention status was not well defined between puberty blockers and other hormone treatments | Intervention status was not well defined between puberty blockers and other hormone treatments | Intervention status was not well defined between puberty blockers and other hormone treatments | N/A |  |
| **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**  **Serious risk of bias** (the study has some important problems) | N/A | The outcome measures were self-reports and subjective (i.e. 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 | The outcome measures were self-reports and subjective (i.e. 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 | The outcome measures were self-reports and subjective (i.e. 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 | The outcome measures were self-reports and subjective (i.e. 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 | N/A |  |
| **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 was 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 | 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 selection of the cohort or subgroups for analysis and reporting on the basis of the results | 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 selection of the cohort or subgroups for analysis and reporting on the basis of the results | 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 selection of the cohort or subgroups for analysis and reporting on the basis of the results | N/A |  |
| **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 | **↓2 levels**  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) | **No change**  Pre-intervention baseline risk was reported | **↓1 levels**  Most of the PB cohort (14/19) also reported receiving GAH as well | **↓1 levels**  Few participants  <400 optimal information size | Not assessed | **↑ 1 level**  4 confounding variables were modelled and limitations included all confounding interventions | **low** |
| Self-harm | **↓2 levels**  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) | **No change**  Pre-intervention baseline risk was reported | **↓1 levels**  Most of the PB cohort (14/19) also reported receiving GAH as well | **↓1 levels**  Few participants  <400 optimal information size | Not assessed | **↑ 1 level**  4 confounding variables were modelled and limitations included all confounding interventions | **low** |
| Anxiety | **↓1 levels**  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  Only one question used out of a standardised assessment to assess | **No change**  Pre-intervention baseline risk was reported | **↓1 levels**  Most of the PB cohort (14/19) also reported receiving GAH as well | **↓1 levels**  Few participants  <400 optimal information size | Not assessed | **↑ 1 level**  4 confounding variables were modelled and limitations included all confounding interventions | **moderate** |
| Depression | **↓1 levels**  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 | **No change**  Pre-intervention baseline risk was reported | **↓1 levels**  Most of the PB cohort (14/19) also reported receiving GAH as well | **↓1 levels**  Few participants  <400 optimal information size | Not assessed | **↑ 1 level**  4 confounding variables were modelled and limitations included all confounding interventions | **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**  **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 | Confounding expected, all known important 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 expected, all known important 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 expected, all known important 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 expected, all known important 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. | N/A | Ethnicity, receiving mental health therapy, family support and 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;  Start of follow up and start of intervention do not coincide – some participants started PB and/or GAH within the 12mth follow up period  and  a potentially important amount of follow-up time is missing from analyses as do not know the stop-start dates of PB interventions  and  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. | 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;  Start of follow up and start of intervention do not coincide – some participants started PB and/or GAH within the 12mth follow up period  and  a potentially important amount of follow-up time is missing from analyses as do not know the stop-start dates of PB interventions  and  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. | 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;  Start of follow up and start of intervention do not coincide – some participants started PB and/or GAH within the 12mth follow up period  and  a potentially important amount of follow-up time is missing from analyses as do not know the stop-start dates of PB interventions  and  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. | 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;  Start of follow up and start of intervention do not coincide – some participants started PB and/or GAH within the 12mth follow up period  and  a potentially important amount of follow-up time is missing from analyses as do not know the stop-start dates of PB interventions  and  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. | N/A | Difficult to extract the PB from gender affirming hormone participants and clinic based sample |
| **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;  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 intervention cohorts;  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 intervention cohorts;  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 intervention cohorts;  and  The nature of the missing data means that the risk of bias cannot be removed through appropriate analysis. | N/A | 6 mth follow up n = 84  12 mth follow up n = 65  Unknown how many received PBs and how many were not |
| **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**  **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;  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. | The outcome measurements 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. | The outcome measurements 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. | The outcome measurements 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. | N/A | Cohort reporting clearly explained and limitations 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 | **↓1 level**  Not a follow up study but retrospective analysis which does provide longitudinal results  No results for <18 years skewing results to ages a different social context | **No change**  Investigators have explored explanations for heterogeneity, and offered several plausible explanations | **↓1 levels**  Differences in interventions is not possible to verify as using self-reported uptake of PBs and in the past | **↓2 levels**  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) | **No change**  Not assessed due to limited information | **No change**  >large sample size n = 3954 | **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** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **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) | N/A | Confounding expected, most known important 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. | 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 |
| **Bias in selection of participants into the study**  **Low risk of bias** (the study is comparable to a well-performed randomized trial with regard to this domain) | 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 |
| **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 |
| **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 | **↓1 level**  No control/comparison | **No change**  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 | **↓1 level**  Body image scale has been used rather than a gender dysphoria assessment  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 | **↓1 level**  Few participants  <400 optimal information size | **No change**  Not assessed | **No change** | **Very low** |
| Suicidality | **↓2 level**  No control/comparison  No information about the PB sub-group in relation to suicidal ideation | **↓1 level**  Heterogeneity not able to be evaluated due to no reporting of the differences between PB and other cohort groups | **↓2 level**  No direct evidence for PB sub-group about suicidality | **↓1 level**  Few participants  <400 optimal information size | **No change**  Not assessed | **No change** | **Very low** |
| Self-harm | **↓2 level**  No control/comparison  No information about the PB sub-group in relation to self-harm | **↓1 level**  Heterogeneity not able to be evaluated due to no reporting of the differences between PB and other cohort groups | **↓2 level**  No direct evidence for PB sub-group about self-harm | **↓1 level**  Few participants  <400 optimal information size | **No change**  Not assessed | **No change** | **Very low** |
| Anxiety | **↓1 level**  No control | **No change**  Limited explanation of the differences between PB and other cohort groups but heterogeneity evident in of results tables | **No change**  Direct measurement of anxiety and reporting of subset anxiety results for all cohorts provided | **↓1 level**  Few participants  <400 optimal information size | **No change**  Not assessed | **No change** | **Moderate** |
| Depression | **↓1 level**  **No control** | **No change**  Limited explanation of the differences between PB and other cohort groups but heterogeneity evident in of results tables | **No change**  Direct measurement of depression including self-reported and clinically reported for all cohorts provided | **↓1 level**  Few participants  <400 optimal information size | **No change**  Not assessed | **No change** | **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**  **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 with regard to this domain but cannot be considered comparable to a well-performed randomized trial) | 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 during follow up period) |
| **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) | Effect of starting and adhering to intervention:  Therapy support was not balanced across intervention groups,  and  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. | Effect of starting and adhering to intervention:  Therapy support was not balanced across intervention groups,  and  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. | Effect of starting and adhering to intervention:  Therapy support was not balanced across intervention groups,  and  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. | Effect of starting and adhering to intervention:  Therapy support was not balanced across intervention groups,  and  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. | Effect of starting and adhering to intervention:  Therapy support was not balanced across intervention groups,  and  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. | N/A | Frequency of therapy received was accounted for in analysis |
| **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) | Proportions of and reasons for missing participants differ slightly across intervention groups;  and  The analysis is unlikely to have removed the risk of bias arising from the missing data. | Proportions of and reasons for missing participants differ slightly across intervention groups;  and  The analysis is unlikely to have removed the risk of bias arising from the missing data. | Proportions of and reasons for missing participants differ slightly across intervention groups;  and  The analysis is unlikely to have removed the risk of bias arising from the missing data. | Proportions of and reasons for missing participants differ slightly across intervention groups;  and  The analysis is unlikely to have removed the risk of bias arising from the missing data. | Proportions of and reasons for missing participants differ slightly across intervention groups;  and  The analysis is unlikely to have removed the risk of bias arising from the missing data. | N/A | Body dissatisfaction (10/25) and self-report depression (13/25) low responses compared to other measures |
| **Bias in measurement of outcomes**  **Serious risk of bias** (the study has some important problems) | 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; | 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; | 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 | All measures self-reported or clinician based with full knowledge of PB and therapy provided |
| **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) | The outcome measurements 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. | The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent;  and  (ii) There is no indication of selection of the reported analysis from among multiple analyses;  and  (iii) There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results. | The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent;  and  (ii) There is no indication of selection of the reported analysis from among multiple analyses;  and  (iii) There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results. | The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent;  and  (ii) There is no indication of selection of the reported analysis from among multiple analyses;  and  (iii) There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results. | The outcome measurements and analyses are consistent with an a priori plan; or are clearly defined and both internally and externally consistent;  and  (ii) There is no indication of selection of the reported analysis from among multiple analyses;  and  (iii) There is no indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results. | N/A | Transparent reporting of cohort groups with response rates.  Limited explanation for missing data but not why e.g. 2/25 PB cessation but very small number |
| **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-demo-graphic 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**  **Serious risk of bias** (the study has some important problems) | N/A | Major aspects of the assignments of intervention status were determined in a way that could have been affected by knowledge of the outcome. | N/A | N/A | 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 | All PB participants started during the 12 mths of follow up |
| **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  (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. | 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  (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. | 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, allowing for deviations (in terms of implementation, adherence and co-intervention) that were likely to impact on the outcome. | Uncertain adherance 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 receign therapy |
| **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 |
| **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. Brożek, G. Guyatt, & A. Oxman (Eds.). Retrieved from <https://gdt.gradepro.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