Depression symptom clusters in adolescents: a latent class analysis in a clinical sample

Article (Accepted Version)

Depression symptom clusters in adolescents: A latent class analysis in a clinical sample

Dr Maria E Loades[1], Dr Michelle St Clair[1], Dr Faith Orchard[2] , Prof Ian Goodyer[3], Prof Shirley Reynolds[4] and the IMPACT consortium†

†Members of the IMPACT Consortium: Ian M Goodyer, Shirley Reynolds, Barbara Barrett, Sarah Byford, Bernadka Dubicka, Jonathan Hill, Fiona Holland, Raphael Kelvin, Nick Midgley, Chris Roberts, Rob Senior, Mary Target, Barry Widmer, Paul Wilkinson, and Peter Fonagy.

Affiliations

[1] Department of Psychology, University of Bath, United Kingdom

[2] Department of Psychology, University of Sussex, United Kingdom

[3] Department of Psychiatry, Douglas House, 18b Trumpington Rd, Cambridge

[4] School of Psychology and Clinical Language Sciences, University of Reading, United Kingdom

Corresponding author contact details

Maria Loades, Department of Psychology, University of Bath, Bath, BA2 7AY, England.

Email m.e.loades@bath.ac.uk (+44) 01225 385249; BA(Cantab), DClinPsy

Co-author contact details

Michelle St Clair m.c.st.clair@bath.ac.uk

Faith Orchard f.orchard@sussex.ac.uk

Ian Goodyer ig04@cam.ac.uk

Shirley Reynolds s.a.reynolds@reading.ac.uk

Conflicts of Interest

Nothing to declare

Acknowledgements/Funding

The IMPACT study was funded by the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) programme (project number 06/05/01). Dr Maria Loades is funded by the National Institute for Health Research (Doctoral Research Fellowship, DRF-2016-09-021). This report is independent research. The views expressed in this publication are those of the authors(s) and not necessarily those of the HTA programme, NHS, The National Institute for Health Research or the Department of Health.

Contributors

IMG, SR and the IMPACT consortium designed and conducted the IMPACT study. ML and FO, with input from SR and MSC were responsible for designing this secondary data analysis. MSC undertook the statistical analyses. ML and MSC wrote the first draft of the manuscript. All authors contributed to and approved the final manuscript.

Clinical and methodological relevance

Depression is heterogeneous. This paper addresses questions about different symptom presentations and whether these are associated with outcomes in psychological therapy, including at up to 86 week follow up using data from the largest adolescent depression trial to date. Three classes of participants are identified using a data-driven approach, which differed on symptom severity, functional impairment, and risk. This classification could inform decisions about allocation to mental health services and help with managing resources.

Depression symptom clusters in adolescents: A latent class analysis in a clinical sample

Abstract

Background: Major depression is clinically heterogeneous. We aimed to identify classes of depressed adolescents with different symptom presentations and examine if these were differentially associated with illness severity, functioning, engagement with treatment and clinical outcomes.

Method: Baseline depression symptoms of 454 depressed adolescents (age 11-17) from the IMPACT trial were subjected to latent class analysis. We compared classes on self-reported symptoms and social impairment at baseline and follow-up, and their engagement in treatment.

Results: We identified three classes of participants which differed in the number and pattern of depression symptoms; Class 1 – Severe- (37.2%) - endorsed almost all symptoms and were most functionally impaired; Class 2 – Moderate- (41.9%) - endorsed fewer symptoms with high suicidal ideation, self-harm and worthlessness; Class 3 – Somatic (20.9%) - endorsed fewest symptoms, with high somatic symptoms. Groups did not differ on engagement, therapeutic alliance, or post-treatment symptom reduction. Adolescents in the severe and moderate subgroups reported symptom reductions after treatment ended, whilst those in the somatic subgroup did not.

Conclusions: At presentation, high somatic features in depressed adolescents, rather than severity or impairment levels, may indicate lower liability for responding to psychological treatment.

Keywords: depression, adolescents, subtyping, symptoms, latent class analysis

Depressive disorders are relatively common amongst young people – worldwide around 2.6% of young people meet the diagnostic criteria for a depressive disorder at any one time (Polanczyk et al., 2015), and approximately 15% of young people have an episode of depression by young adulthood (Kessler et al., 2003). Depressive disorders are associated with long term negative consequences for health, wealth, relationships, and psychological well-being. For example, recent systematic reviews have highlighted many adverse outcomes including dropping out of education, being unemployed, early pregnancy and childbirth, drug and alcohol use, and lifetime physical and mental health problems (Clayborne et al., 2019; Johnson et al., 2018).

Depression can be hard to recognise because it is characterised by a diverse range of symptoms. For example, in adolescents, Major Depressive Disorder (MDD) is often characterised by core symptoms of low mood and/or irritability and/or lack of enjoyment and interest in activities. Other symptoms include somatic problems of fatigue, insomnia/hypersomnia, changes in appetite, psychomotor agitation/retardation, and problems of guilt or worthlessness, suicidal ideation, and difficulties with attention, concentration, or memory. To meet diagnostic criteria for MDD an individual needs to experience at least one core mood symptom and a total of at least 5 symptoms, which impact significantly on functioning for at least the last weeks (A.P.A., 2013). From the 9 core symptom domains, more than 1000 symptom combinations can be computed that meet the diagnostic criteria (Fried, 2015). Amongst depressed adolescents, sleep problems (92%) can be more commonly reported than low mood (84%) per se but whether different combinations of symptoms at presentation are sufficient to form distinct subgroups of patients is not known (Goodyer et al., 2017).

Classifications within diagnostic systems like the DSM (A.P.A., 2013) have resulted from 'top-down' approaches where experts have decided how to group symptoms together, rather than bottom-up exploration, driven by symptom clusters (Krueger & Bezdjian, 2009). Arguably, the best approach would be to use both sources of information to improve the validity of the diagnostic system. A promising approach using data to identify subtypes of depression which has been used in

several studies in adult samples is latent class analysis (LCA) (Ulbricht et al., 2018). Most classes identified were distinguished by depression severity, although there have been variable numbers of classes identified. These classes (i.e. data-driven subtypes) have been found to be predictive of response to treatment, for example in CBT (Catarino et al., 2020; Simmonds-Buckley et al., 2021), which makes them potentially valuable in knowing what works for whom and in personalising treatment. However, few studies have attempted to subtype symptoms of adolescent depression using a data-driven approach. A 2 generation family study which used data from diagnostic interviews found that vegetative symptoms (appetite and weight change, loss of energy and insomnia) were more common in adolescent depression than adult depression, and that anhedonia and concentration problems were less common (Rice et al., 2019). However, inclusion was based on parental mental health status, which may have led to bias. Other efforts to subtype symptoms in adolescents have focused on general distress (Herman et al., 2007), or have investigated distinctions between diagnostic classifications like anxiety or depression (Ferdinand et al., 2005; van Lang et al., 2006). Using a data-driven approach to classify symptoms of major depressive disorder in adolescents could aid recognition and may be useful in predicting what treatment approaches work for whom, if found to meaningfully relate to outcome.

Prompt recognition and referral to evidence-based treatments may help improve the short and long term outcomes for young people who develop depression, both by helping reduce distressing symptoms, and the length of time that symptoms interfere with education, development and functioning. There is compelling evidence that prompt treatment for young people with psychosis improves their long-term well-being (Malla & McGorry, 2019). Similarly, treatment for depression in early adolescence improves future functioning and mental health (Catania et al., 2011). Thus it is reasonable to infer that an accelerated route into assessment and treatment for young people with depression would reduce their current disability and markedly improve their future prospects (Thapar et al., 2012).

In the UK, specialist treatment for adolescent depression typically takes place in a multidisciplinary Child and Adolescent Mental Health service (CAMHs) with access to psychological therapists as well as child and adolescent psychiatrists. However, access to CAMHs has been problematic and resources have historically failed to meet demand which has resulted in access being restricted, for example, to only young people who are 'at risk'. This means that children and young people who do not have suicidal thoughts, or plans, wait longer or are not accepted for treatment (Murphy, 2016). To address this problem of limited mental health resources for children and young people, recent health policy in England has increased funding and introduced more diverse mental health services within a stepped care model (NICE, 2019). This includes the expansion of community services, development of school based services, delivery of low intensity treatments by non-specialists, and the recruitment and development of a new workforce of low intensity clinicians (Care, 2015). However, this new configuration of services can only be effective and efficient if young people with mental health problems are identified accurately and then referred to the service that can best meet their needs. Given the diverse ways in which adolescent depression can present, it seems likely that some depressed young people will benefit from relatively brief, evidence-based psychological therapy and require no further specialist support, whereas others will require the full support of a multi-disciplinary team. Currently there is no good evidence to guide this clinical decision-making and therefore our aim is to address three research questions:

1. Are there different symptom presentations in young people with depression?
2. Are these presentations associated with young peoples' symptom severity and functioning?
3. Are these presentations associated with young peoples' engagement in treatment (alliance, drop out) and their outcome from therapy (reduction in symptoms and improvement in functioning)?

Method

Setting

This paper reports on the exploratory post-hoc analyses of data from a randomised controlled trial for adolescent depression. Young people in this study took part in a large, multi-centre, randomised controlled trial in the UK. The HTA-funded IMPACT (Improving Mood through Psychoanalytic and Cognitive-Behavioural Therapy) Study (Goodyer et al., 2017; Goodyer et al., 2011) was a pragmatic superiority randomised controlled trial comparing the clinical and cost effectiveness of three psychological treatments for adolescent MDD. Participants were recruited and treated at 15 National Health Service child and adolescent mental health service (CAMHs) clinics in three regions of England (North London, East Anglia and North-West England). Thus, participants were recruited from routine publicly funded mental health services.

Participants

Young people (N = 465) were randomised to one of three treatment approaches; Short-Term Psychoanalytic Psychotherapy (STPP); Cognitive Behavioural Therapy (CBT); and a manualised form of routine specialist clinical care termed Brief Psychosocial Intervention (BPI), which was chosen as the active control treatment. The three treatments were delivered in routine services by specialist clinicians according to pre-planned treatment manuals. Planned treatment duration varied for the three approaches: up to 28 sessions over 30 weeks for STPP; up to 20 sessions over 30 weeks for CBT; and up to 12 sessions over 20 weeks for BPI.

The final sample for this analysis was those 454 adolescents (341, 75% female) who met clinical criteria for Major Depressive Disorder, (Goodyer et al., 2017; Goodyer et al., 2011) 178 from the East Anglia region, 123 from North London, and 153 from the Northwest of England. Their average age was 15.63 years (SD = 1.42 years; range 11.30-17.99 years). Twelve cases were not included because at presentation they reported only 4 symptoms of depression but due to high self-reported depression sum scores they were included in the trial.

Measures

Diagnostic Interview

Depression diagnoses were assessed using the Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS; Kaufman et al, 1997), a semi-structured interview for diagnosing psychiatric disorders in children and adolescents (Simmons et al., 2015). All MDD symptoms were assessed: depressed mood, irritable mood, anhedonia, appetite/weight change, insomnia/hypersomnia, psychomotor agitation/retardation, fatigue or loss of energy, feelings of worthlessness or excessive/inappropriate guilt, decreased concentration or slow thinking or indecisiveness, recurrent thoughts of death or suicidal ideation or suicide attempts. As is conventional, the interview was conducted with adolescents and caregivers separately, and symptoms and diagnoses were based on the information obtained from both interviews. Consensus meetings were held throughout the study with clinical principal investigators to resolve uncertainties on scoring and assignment of clinical status of any item.

Trained research assistants, who were psychology postgraduates, conducted the interviews. They were closely supervised, including using audio recorded interviews. Reliability checks were conducted on a sample of 30 randomly selected cases and found to be good (100% for depression diagnosis and 95% for individual symptoms between pairs of research assistants), with high inter-item agreement (Goodyer et al., 2017).

Self-Report Measures

Multiple measures were completed at multiple time points in the IMPACT study (Goodyer et al., 2017; Goodyer et al., 2011), although we opted to draw on certain measures at certain time points in this paper to address our research questions. To be consistent with measures across the IMPACT study, the majority of measures reported here, unless otherwise indicated, were measured on a four-point scale from 'Never' to 'Always' (0 -3). We combined the "mostly" and "always" categories to be consistent with other population level studies (St Clair et al., 2017a) and other papers investigating this dataset (Davies et al., 2019).

We included the following measures from baseline assessment and again at the nominal timepoints of , 6, 12 weeks and 36 weeks when >95% of treatment episodes were complete, and 52 and 86 week follow-up post treatment.

The Mood and Feelings Questionnaire (MFQ) (Costello & Angold, 1988) is a 33-item self-report measure examining depression symptoms over the past two weeks. There was a total possible score of 66 with higher scores indicating more depressive symptoms. The MFQ has good test–retest reliability (Pearson's r=0.78), an α coefficient of 0.82 and discriminant validity for detecting an episode of Major Depressive Disorder in clinical adolescent samples. The reliability of this measure within this sample was excellent (α = .96). The self-reported depression symptoms sum score had been the primary end point in the IMPACT trial: by the end of study 84% of the participants reported >50% reduction in their MFQ from baseline.

The Revised Children's Manifest Anxiety Scale (RCMAS) (Reynolds & Richmond, 1978) is a 28 item self-report questionnaire of anxiety symptoms. There was a possible score of 56 with higher scores indicating more anxiety symptoms. The reliability of this measure within this sample was excellent (α = .91).

We also included the following measures from the baseline assessment only:

The short Leyton Obsessional Inventory (LOI) for adolescents (Bamber et al., 2002) is an 11 item measure of obsessionality. There was a possible total score of 22 with higher scores indicating more obsessionality symptoms. The reliability of the LOI in this sample was good α = .89.

The Anti-social Behaviours Questionnaire (ABQ) (St Clair et al., 2017b) an 11-item questionnaire measuring antisocial behaviours. The reliability of this measure in this sample was acceptable (α = .76). There was a possible total score of 22 with higher scores indicating more antisocial behaviour.

The Rosenberg Self Esteem Scale (RSES) is a 10-item (Rosenberg, 1965). Five items were negatively worded and were reversed for this analysis. There was a possible score of 20 with higher

scores indicating better self-esteem. The reliability of this measure in this sample is excellent (α = .89).

The Risk-Taking and Self-Harm Inventory for Adolescents (RTSHIA) (Vrouva et al., 2010). has two subscales with 7 items measuring risk-taking behaviours and 18 items measuring self-harm behaviours. All items were measured on a four-point scale (never, once, more than once, many times) and measured lifetime history. The reliability of the Risk-Taking subscale was acceptable (α = .77) and the reliability of the Self-Harm subscale was excellent (α = .91). There was a total score of 21 for the risk-taking subscale and 54 for the self-harm subscale with higher scores indicating more risk taking and self-harm behaviours.

The Rumination Response Scale (RRS) (Treynor et al., 2003) is a 22 item scale measuring ruminative thoughts related to depression. This scale was measured on a four-point scale (Almost Never (1), Sometimes, Often, Almost Always (4)). The reliability of the RRS was excellent in this sample (α = .93). The possible range of responses varied from 22 to 88, with higher scores indicating more ruminative thoughts.

The Health of the Nation Outcomes Scales for Children and Adolescence (HONOSCA) was rated by trained research assistants at all assessments (Gowers et al., 2002). This 13 item questionnaire measures current general health and social functioning of children and adolescents and is sensitive to change over relatively brief periods. All items were measured on a five-point scale (no problem, minor problem, mild problem, moderate problem, severe problem). The endorsements of severe problems were rarely endorsed for many items, so were combined with the adjacent moderate problems groups. The reliability of this measure was acceptable in this sample (α = .78). There was a total score of 39 for this questionnaire, with higher scores indicating more difficulties.

The Working Alliance Inventory Short form (WAI-S) was completed by young people at the 36 week follow-up (Hatcher & Gillaspy, 2006). The 12 items were measured on a seven-point scale (Never (1), Rarely, occasionally, sometimes, often, very often, always (7)). The reliability of this

measure was excellent (α = .95). The possible range of responses varied from 12 to 84, with higher scores indicating better working alliance between the patient and the therapist.

Procedure

The procedure for the IMPACT study has been detailed elsewhere (Goodyer et al., 2017; Goodyer et al., 2011). Ethical approval for the IMPACT study was granted by the Cambridgeshire 2 Research Ethics Committee (reference 09/H0308/137) and local NHS provider trusts.

If participants appeared to meet the eligibility criteria at their routine NHS assessment (i.e. age 11-17, with depression, not currently pregnant), the clinician informed them about the study and offered them the opportunity to be contacted by the research team. Those who opted to proceed gave written consent/assent with parental consent for those under 16 years, and then were interviewed using the KSADS to ensure that they met criteria for a MDD. All interviews were conducted face-to-face. Subsequently, participants were randomised to one of three psychological treatments, Brief Psychosocial Intervention, Cognitive Behavioural Therapy or Short-term Psychoanalytical Psychotherapy.

Data Analysis

Symptom generation. Symptoms were combined following the scoring procedure in the K-SADS. For diagnostic criteria, which could be met by several questions (appetite/weight change, insomnia/hypersomnia, psychomotor agitation/retardation, feelings of worthlessness or excessive/inappropriate guilt), any threshold level symptom on any one of the items was sufficient for an overall threshold level symptom. We examined data from all questions encompassing these composite symptoms before concluding that the symptoms was/was not present. If there was missing data on one question, the overall symptom was not calculated.

Latent Class Analysis. Mplus (version 8) was used to evaluate differing latent class solutions. The 10 symptoms were set as categorical (symptom present/absent). The estimator was maximum likelihood robust. There were 100 random starts with 25 final stage optimisations. A series of different class solutions were evaluated, from a single class to four classes. Each solution was

evaluated with a series of metrics. Firstly, the Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC) and Sample Size Adjusted BIC were evaluated, with lower values indicating better model fit to the observed data. Secondly, the Vuong-Lo-Mendell-Rubin Likelihood Ratio Test (LRT) and the Lo-Mendell-Rubin adjusted LRT were evaluated, with significant findings indicating the higher-class solution was a significantly better fit to the observed results. Finally, for specific comparisons we also used the parametric bootstrapped LRT for the final comparison choice. This was only used for the final model selection due to the higher computational time necessary for this LRT. Finally, entropy was consulted with a higher entropy being preferred, but this was not one of the model selection criteria. All model selection criteria were considered in parallel when considering how many class solutions to evaluate. However, there was an increased emphasis on the parametric bootstrapping approach for the final decision, as this has been shown to be a more reliable indicator of the best class in latent class solutions (Nylund et al., 2007).

Statistical analysis. The best class solution and class probabilities from each class were imported in Stata 15 (StataCorp, 2017). Descriptive statistics and further analysis were conducted. We compared the prevalence of each symptom, gender and total number of symptoms within each class to the overall rates within the full sample. This was done with chi square tests for the categorical variables and a t-test for the total symptom count.

When looking at the self-reported symptoms, we used regression techniques to examine differences in items between the classes. This involved using the class variable as a categorical variable and comparing item response level between the different classes.

For the longitudinal regression we implemented mixed effects models (xtmixed) that took into account the non-independence of the datapoints across time. Participant ID was the random effect and time (linear and quadratic) were the fixed effects, alongside class membership for some models. Figure 1 was created from fitted values derived from a mixed effects model with linear and quadratic time as well as an interactive fixed effect with class membership. Class specific

longitudinal analyses were conducted using the time variable as a categorical variable, comparing each specific time point for each class.

Results

Research Question 1: Can we identify different patterning of depression symptom presentations in young people with MDD?

Latent Class Solution

We evaluated between one and four classes. The final solution selected was the three-class solution. As can be seen in Table 1, there was only a marginal decrease in the AIC and BICSSA between the three and four class solution, while there was an increase in the BIC. Similarly, there was a nonsignificant LRT across all the statistical comparisons between the three and four class solution, indicating that the fourth class did not create a more parsimonious solution. The three-class solution was an improvement on the two-class solution across all likelihood ratio indices. The entropy, however, was lower than recommended levels of .80 (Ramaswamy et al., 1993). This indicates that the three-class solution had lower separability between the distinct classes than recommended and this should be kept in mind while interpreting the results. Each adolescent in the sample was classified by how they best fit into each of the three classes, with all three probabilities adding to 1.0. Individuals were assigned the class with the highest probability.

[TABLE 1 HERE]

Three Class Solution*.* See Table 2 for the symptom profile and demographic information for each class. Approximately 37.2% of the sample were best categorised by Class 1, with 41.9% fitting best into Class 2 and the remaining 20.9% within Class 3. To evaluate the low entropy levels, we also included the average probability of being included in each of the class for all individuals best classified into Class 1, 2 and 3, respectively (see online supplementary materials). We evaluated whether the three treatment options (CBT, BPI, STPP) different by subgroup and found no differences in any of the subgroups (*p*s > .81).

[TABLE 2 HERE]

Class 1. is characterised by a higher rate of endorsed symptoms than for the entire sample. All of the adolescents in this class reported threshold level depressed mood symptoms and problems with sleep, as well as decreased concentration, slow thinking and indecisiveness; these rates were significantly higher than the overall rates for the entire sample (depressed mood: χ = 29.73, p<.001, V = .22; problems with sleep: χ = 12.60, p<.001, V = .14; decreased concentration, slow thinking and indecisiveness: χ = 33.77, p<.001, V = .23). Over 90% of participants in this group also reported anhedonia, appetite or weight change, fatigue and worthlessness/excessive or inappropriate guilt, which was higher than full sample (anhedonia: χ = 36.02, p<.001, V = .24; appetite/weight change: χ = 26.67, p<.001, V = .21; fatigue: χ = 25.25, p<.001, V = .20; worthlessness/excessive/inappropriate guilt: χ = 19.33, p<.001, V = .18). Compared to the full sample there were also higher rates of irritable mood, psychomotor agitation or retardation, and suicidal thoughts or ideation (irritable mood: χ = 7.44, p<.01, V = .11; psychomotor agitation: χ = 12.60, p<.001, V = .14; suicidal thoughts/ideation: χ = 32.99, p<.001, V = .23). There was no difference in the gender distribution when compared to the overall sample (p = .35). This subgroup was labelled as “Severe ”.

Class 2. Participants in this group reported fewer symptoms than the average of the whole group. No symptom was endorsed by every member of the class. Over 90% reported depressed mood, and over 80% reported sleep difficulties and worthlessness and excessive/inappropriate guilt; depressed mood and guilt were significantly more often reported by participants in this class (depressed mood: χ = 6.02, p<.05, V = .10; worthlessness/excessive/inappropriate guilt: χ = 6.56, p<.05, V = .10). More than 60% of participants in this class reported suicidal ideation, similar to the sample average (p = .42). Most symptoms were less commonly reported than in the overall sample, i.e. irritable mood: χ = 10.78, p<.01, V = -.13; anhedonia: χ = 26.68, p<.001, V = -.20; appetite/weight change: χ = 26.80, p<.001, V = -.20; hypersomnia/insomnia: χ = 14.24, p<.001, V = -.15; fatigue: χ =

7.78, p<.01, V = -.11; decreased concentration/slow thinking/indecisiveness: χ = 31.81, p<.001, V = -.22. There was no difference in the other symptoms or the gender distribution when compared to the overall sample (ps > .13). This subgroup was labelled as “Moderate ”.

Class 3. One hundred percent of this class experienced insomnia or hypersomnia symptoms and over 90% reported problems with concentration, slow thinking or indecisiveness, this was significantly higher than in the overall sample (insomnia/hypersomnia: χ = 7.03, p<.01, V = .11; decreased concentration/slow thinking/indecisiveness: χ = 5.72, p<.05, V = .10). However, this group reported very low rates of suicidal thoughts or ideas (fewer than 10%) and significantly lower rates of suicidal thoughts/ideation (χ = 80.35, p<.001, V = -.38), depressed mood (χ = 79.07, p<.001, V = -.38), fatigue (χ = 4.26, p<.05, V = -.09) and worthlessness/excessive/inappropriate guilt (χ = 74.38, p<.001, V = -.37). There was a lower total number of symptoms within this class and no difference in the gender distribution (p = .18). This subgroup was labelled as “Somatic”.

Research Question 2: Are these presentations associated with young peoples’ symptom severity and functioning?

Baseline and follow-up data can be found in table 3.

Depression Symptoms

The Severe subgroup had higher levels of baseline depressive symptoms than the Moderate subgroup (β = -5.95, 95%CI(-7.90, -4.00), p < .001) and the Somatic subgroup (β = -13.83, 95%CI(-16.33, -11.34), p < .001). The Moderate subgroup had higher levels of baseline depression symptoms than the Somatic subgroup (β = -7.88, 95%CI(-10.56, -5.21), p < .001). There were no differences in the percentage of young people in the severe subgroup and the moderate subgroup who scored above the clinical threshold (>= 27) on the MFQ (p = .06). More severe (Fisher’s exact p <.001, V = -.33) and moderate (OR = 6.14, 95%CI(2.11,17.89), p < .005) participants scored above the clinical threshold than the somatic sub-group.

Anxiety symptoms

The severe subgroup reported more anxiety symptoms than the moderate subgroup (β = -2.97, 95%CI(-4.41, -1.52), $p < .001$) and the somatic subgroup (β = -6.70, 95%CI(-8.43, -4.94), $p < .001$). The moderate subgroup had higher levels of anxiety symptoms than the somatic subgroup (β = -3.73, 95%CI(-5.47, -2.00), $p < .001$).

Obsessionality Symptoms

The severe subgroup reported more obsessional symptoms than the moderate subgroup (β = -1.63, 95%CI(-2.75, -0.51), $p < .01$) and the somatic subgroup (β = -3.46, 95%CI(-4.70, -2.21), $p < .001$). The moderate subgroup had higher levels of obsessional symptoms than the somatic subgroup (β = -1.83, 95%CI(-3.10, -0.56), $p < .01$).

Antisocial Behaviour

The severe subgroup had higher levels of antisocial behaviour symptoms than the moderate subgroup (β = -0.27, 95%CI(-0.48, -0.07), $p < .05$) but did not differ from the somatic subgroup ($p = .09$). There were no differences in antisocial behaviour between the moderate and somatic subgroups ($p = .70$).

Self Esteem

The severe subgroup had lower self-esteem than the moderate subgroup (β = 1.64, 95%CI(0.82, 2.45), $p < .001$) and the somatic subgroup (β = 5.44, 95%CI(4.43, 6.44), $p < .001$). The moderate subgroup had lower self-esteem than the somatic subgroup (β = 3.80, 95%CI(2.82, 4.78), $p < .001$).

Risk Taking

There were no differences between the subgroups in risk taking behaviour ($ps > .07$).

Self-Harm

There was no difference in rates of self-harm between the severe and moderate subgroups ($p = .08$). Both groups had more self-harm symptoms than the somatic subgroup (severe, β = -0.84, 95%CI(-1.07, -0.61), $p < .05$; moderate, β = -0.67, 95%CI(-0.90, -0.45), $p < .001$).

HONOSCA

The severe subgroup had higher levels of functional impairment than the moderate subgroup (β = -2.81, 95%CI(-3.97, -1.65), $p < .001$) and the somatic subgroup (β = -2.91, 95%CI(-4.28, -1.54), $p < .001$). There were no differences between the moderate and somatic subgroups ($p = .88$).

Rumination Symptoms

The severe and moderate subgroups had similar levels of rumination ($p = .26$) and both were higher than the somatic subgroup (Severe vs. somatic: β = -8.51, 95%CI(-11.67, -5.36), $p < .001$; moderate vs somatic: β = -7.01, 95%CI(-10.09, -3.94), $p < .001$).

Antidepressant (SSRI) use

There was no differences between the three subgroups in SSRI use.

[TABLE 3 HERE]

Research Question 3: Are these presentations associated with young peoples' engagement in treatment (alliance, drop out) and their outcome from therapy (i.e. reduction in symptoms and improvement in functioning)?

Longitudinal analysis

No subgroup had a different rate of attrition when compared to the overall rate of attrition at 36, 52 or the 86-week follow-up. There was no difference in any baseline self-report measure between those who did and did not complete the follow-up self-report questionnaires at 36, 52 or 86 months. We also investigated whether there were any attrition differences within each subgroup on baseline characteristics. Overall, no differences found at the 86-week follow-up and a few differences were found at the 36- and 52-week follow-up in Class 1 and 3. Please see the supplementary materials for details. We also evaluated whether the type of treatment differentially impacted the main outcomes, depressive symptoms, within each subgroup. There was no interactive effect of the three subgroups and the treatment type on depressive outcomes at 36, 52 and 86 months.

Depression Symptoms.

There were no differences between the subgroups in the absolute change in depression symptoms from baseline to 36 weeks (*p*s > .47). At 36 weeks, the severe subgroup reported significantly more symptoms of depression than the moderate subgroup (β = -4.74, 95%CI (-9.78,-0.70), $p < .05$) and the somatic subgroup (β = -9.43, 95%CI(-14.53,-4.32), $p < .001$). There was no difference in severity of depression between the moderate and somatic subgroups ($p = .06$). There was no difference between the severe and moderate subgroups in the proportion of participants over the clinical cut-off on the MFQ ($p = .83$). However, both the severe and moderate subgroups were more likely to score above the clinical cut-off on the MFQ at the nominal 36 week follow-up than the somatic subgroup (severe subgroup: OR = 3.27, 95%CI (1.59,6.73), $p < .005$; moderate subgroup: OR = 3.10, 95%CI(1.53,6.25), $p < .005$).

There were no between-group differences in the change in depressive symptoms from baseline to 52 weeks (*p*s > .30). At 52 weeks, the severe subgroup did not differ in depressive syptoms when compared to the moderate group (*p* = .07) but reported more depressive symptoms than the somatic group (β =-5.87, 95%CI(-11.29,-.45), *p* < .05). There was no difference between the moderate and somatic groups (*p* = .45). There was no difference in the percentage of participants in the severe and moderate subgroups who scored above the clinical cut-off on the MFQ (*p* = .21) or between the moderate and somatic subgroups (= .12) scoring above MFQ cut-off, but a higher proportion of severe participants than somatic participants scored above the MFQ cut-off (OR = 2.42, 95%CI(1.20,4.88), *p* < .05).

There were no between-group differences in the change in depressive symptoms from baseline to 86 weeks (*p*s > .51). At 86 weeks, the severe subgroup did not differ in depression symptoms compared to the moderate subgroup (*p* < .18), but the severe subgroup changed more than the somatic (β =-6.18, 95%CI(-10.85,-1.50), *p* < .05) group. The moderate and somatic subgroups did not differ (*p* = .13). There were no differences in the percentage of participants scoring above the clinical cut-off on the MFQ between the severe and moderate subgroups (*p* = .07), but there were such differences between the severe group compared to the somatic group (OR = 3.53, 95%CI(1.73,7.20), *p* < .005) as well as increased rates in moderate group scored compared to the somatic group (OR = 2.20, 95%CI(1.09,4.45), *p* < .05).

Anxiety Symptoms.

At 36 weeks the severe subgroup reported more anxiety symptoms than the moderate subgroup (β = -4.65, 95%CI(-8.03,-1.27), *p* < .01) and the somatic subgroup (β = -6.65, 95%CI(-11.05,-2.25), *p* < .01). There was no difference between the moderate and somatic subgroups (*p* = .36).

HONOSCA.

At the 36 weeks there were no differences in functional impairment between the subgroups.

Therapeutic Working Alliance.

At 36 weeks there were no differences in participant reported working alliance (*p*s > .30).

Trajectory Analysis

We next looked at how the main outcome, self-reported depression symptoms, varied between the subgroups from baseline to the three selected follow-up time points. There was a rapid drop in depression severity in all subgroups from baseline to 36 weeks, with a smaller change to the 86 week follow-up (see figure 1). We evaluated the overall rate of change in the entire sample and found that it was best explained within a linear and quadratic time trend (Linear: β = -0.72, 95%CI(-0.79, -0.65), $p < .001$; Quadratic: β = 0.005, 95%CI(0.004,0.006), $p < .001$). There was also an overall significant interaction with subgroup membership and the quadratic time trend, $\chi(2) = 9.27$, $p < .01$. Upon further examining this interaction, it was found that the quadratic time trend did not differ between the severe and moderate subgroups ($p = .07$). However, there was a significant difference between the quadratic trends of the severe and somatic groups (β = 0.001, 95%CI(0.0003,0.002), $p < .005$) and between the moderate and somatic subgroups (β = 0.001, 95%CI(0.0001,0.001), $p < .05$). This was because depression severity significantly reduced between 36 and 52 weeks in both the severe and moderate subgroups (Severe: β = -3.53, 95%CI(-6.27, -0.81), $p < .05$; Moderate: β = -3.26, 95%CI(-5.23, -1.30), $p < .005$) whereas in the somatic subgroup there was no difference in depression severity between the 36, 52 and 86 week follow up (ps > .19). Similarly, in the severe and moderate subgroups there was a reduction in depression severity from 36 to 86 weeks (Severe: β = -4.91, 95%CI(-7.84, -1.98), $p < .005$; Moderate: β = -3.86, 95%CI(-6.68, -1.04), $p < .01$). For all groups there was a highly significant drop in depressive symptoms from baseline to all follow-up timepoints (ps < .001).

[FIGURE 1 HERE]

Discussion

Our data-driven approach identified three subgroups of depressed adolescents, based on symptom patterns, which we called 'severe', 'moderate' and 'somatic'. We note that the analysis indicates low separability between these groups.

Although these subgroups reported different symptom severity and functioning at baseline, they were not differentially related to engagement in therapy or the young people's experience of the therapeutic alliance. All three subgroups reported similar reductions in depression symptom severity over the course of treatment, but beyond treatment, the severe and moderate subgroups continued to show depression symptom reduction, whilst the somatic subgroup did not. Despite this, at the final 86 week follow-up assessment young people in the severe and moderate subgroups were more likely to report depressive symptoms that remained above the clinical cut-off. This may be influenced by the higher levels of these symptoms at baseline.

This above finding should be interpreted with caution given the low separability of the groups and in combination with the GMM trajectory analysis by Davies et al. (2020) who found two trajectory groups based on self-reported depression sum scores in the same sample over the same timeframe – one (84%) with continued improvement (e.g., decline in depressive symptoms) and one (16%) "halted improvers" whose improvement halted and then slightly raised. The differing results is likely reflective of the different subgroup analysis conducted. Here we conducted a latent class analysis of depression diagnostic symptoms at baseline, whereas the Davies et al. (2020) paper investigated distinct groups with differences in the trajectory of self-reported depressive symptoms across time. Therefore, the differing findings in relation to the longitudinal self-reported depressive symptoms is likely to reflect the different analytic approaches employed.

Consistent with other work (Fried & Nesse, 2014), our findings suggests that 'homogenising' heterogeneous conditions like depression may obscure meaningful differences in how young people experience depression and then progress in treatment. Our findings are consistent with those of data-driven latent class approaches to subtyping symptoms in adults with depression, which have generally found that classes are distinguished by depression severity (Ulbricht et al., 2018). Here

however, the low separability suggest that separability on depression items alone is lower than might have been expected from the adult literature. Whether this is a function of age, recurrent liability being higher in adult cases or other factors we cannot conclude form these results. Previous efforts to group symptoms in adolescents had focused on general distress (St Clair et al., 2017a) or on co-morbid symptoms (Ferdinand et al., 2005; Herman et al., 2007). Network analysis has recently been applied to explore symptoms amongst adolescents recruited from community settings (McElroy et al., 2019; Mullarkey et al., 2019) and a preliminary network analysis was conducted on the IMPACT study dataset using self-reported depressive symptoms (Schweren et al., 2018). However, despite calls to do so (Yu et al., 2010), ours is amongst the first study to specifically evaluate subgroups of adolescents with formally diagnosed DSM major depression, based on their depressive symptomatology obtained from face to face gold standard diagnostic interviews. The face validity of these subgroups was confirmed by demonstrating that the individuals in each of the three groups also differed somewhat in the extent to which their symptoms interfered with their functioning on the HONOSCA. Further studies could explore this using other functioning measures, as the HONOSCA includes symptoms as well as functioning.

Data on depression symptoms rated by clinicians in relation to participants in the Treatment for Adolescents with Depression Study (TADS) RCT identified two symptom clusters. One of these showed differential improvements in symptoms in the first 12 weeks of the trial depending on the treatment allocation (Bondar et al., 2020). In the current study, there was no interaction between the three sub-groups and treatment allocation on participants' response to treatment. There was also no association between the subgroups and their engagement in treatment assessed by self-reported therapeutic alliance.

Strengths and Limitations

Depression symptoms and diagnoses were identified by a gold standard diagnostic interview, and participants were recruited through 15 clinics within the UK National Health Service and followed up over a considerable length of time. However, those with more severe depression,

and/or those for whom their depression is interfering significantly with their functioning, are most likely to receive referrals and access into specialist Child and Adolescent Mental Health Services (CAMHS). As only adolescents accessing services were included, other types of depression presentations may be characteristic of young people who do not present to services at all, or who are less severely depressed and are treated outside NHS funded mental health services e.g. school based counselling.

Although a strength of latent class analysis is the data-driven approach, previous findings using LCA in adults with depression have been inconsistent (Ulbricht et al., 2018) and may just capture differences in severity and are limited to static classes, rather than looking at change in subtypes over the course of treatment. Novel innovations such as Markov modelling may enable future studies do refine our understanding of depressive symptoms and how they change over the course of treatment (Catarino et al., 2020).

Clinical and Research Implications

Our results suggest that there may be a distinct group of young people with depression who present with moderately severe symptoms (somatic) but not high levels of risk. This somatic group may not be recognised as being depressed with the clinical presentation being dominated by sleep dysregulation and concerns about weight and appetite and with less mood and cognitive impairments than usually expected. This group may also be less likely to get access to specialist child and adolescent mental health services, as currently, referral criteria tend to stipulate risk to self. We suggest referrers such as primary care, schools and other services should be made aware that depression may present in different ways and a more somatic form can occur in about 1 in 5 cases being considered for referral to specialist CAMHS.

We also identified a group of young people whose clinical presentation was depression symptoms of moderate severity (like the somatic group) and also elevated suicidality and self-harm. At the end of psychological therapy, most had improved but still reported symptoms that remained above the clinical cut off. These residual symptoms carry an elevated risk for relapse which suggests

that a post treatment monitoring/surveillance program might be of value to enhance early detection of recurrence and activate further treatment.

A third group of young people reported high severity of depression, high suicidality, and self-harm, and were most likely to report residual symptoms above the clinical cut off at the end of psychological treatment. Our data suggest that for this group the mean number of symptoms continued to reduce during follow up which implies that prognosis may be communicated as likely to be good. This is emphasised as a follow up of all participants using growth mixture modelling across all self-report data showed that their outcomes continued to improve without further input throughout the follow up period (Aitken et al., 2020). Whether good prognosis is actually more likely in the longer term for the severe group based on clinical interview data is not known.

The current findings are secondary analyses and as such provide hypothesis forming, rather than confirming, results. Future research should seek to replicate these findings in other samples of young people, including those referred to a wider range of services, including primary care, and community and school counselling to include mild cases of depression not likely to be referred to specialist CAMHS. The results of this study suggest a hypothesis that the clinical interview method may differentiate subgroups based on depressive clinical characteristics at presentation. This may form the clinical basis for therapeutic decision making. From the service research perspective further research on depressive subtypes can include which evidence-based psychological therapies may be safely and effectively provided. We speculate that including non-depressive features may be of value, including comorbid disorders. For example, depression comorbid with higher antisocial behaviour difficulties may respond somewhat better to a brief psychosocial. Intervention involving psychoeducation and social. therapy than to CBT or short-term psychoanalytic therapy (Aitken et al., 2020). Using data driven subtypes may determine that that some young people referred to multi-disciplinary mental health services could be ‘stepped down’ to community psychological therapies services. Equally it may be that some young people with severe symptoms and elevated risk who are referred to community psychological therapy services might benefit from being ‘stepped up’ to

multi-disciplinary services. From these preliminary results we conjecture that severe clinical cases could be immediately stepped up specialist CAMHS whereas moderate and somatic cases might be stepped down to community teams.

Conclusion

We used a data driven method to identified three potentially distinct clinical presentations amongst depressed adolescents treated in specialist CAMHS. This classification if replicated, could be investigated as providing a contributory factor in therapeutic decision making. At the population and service level, delineating clinical depression subgroups may assist health services allocation. We note that the current findings are from secondary analyses and further work and replication in other samples is needed.

Figure 1. Change in self-reported depression symptoms (MFQ) in Classes 1, 2 and 3

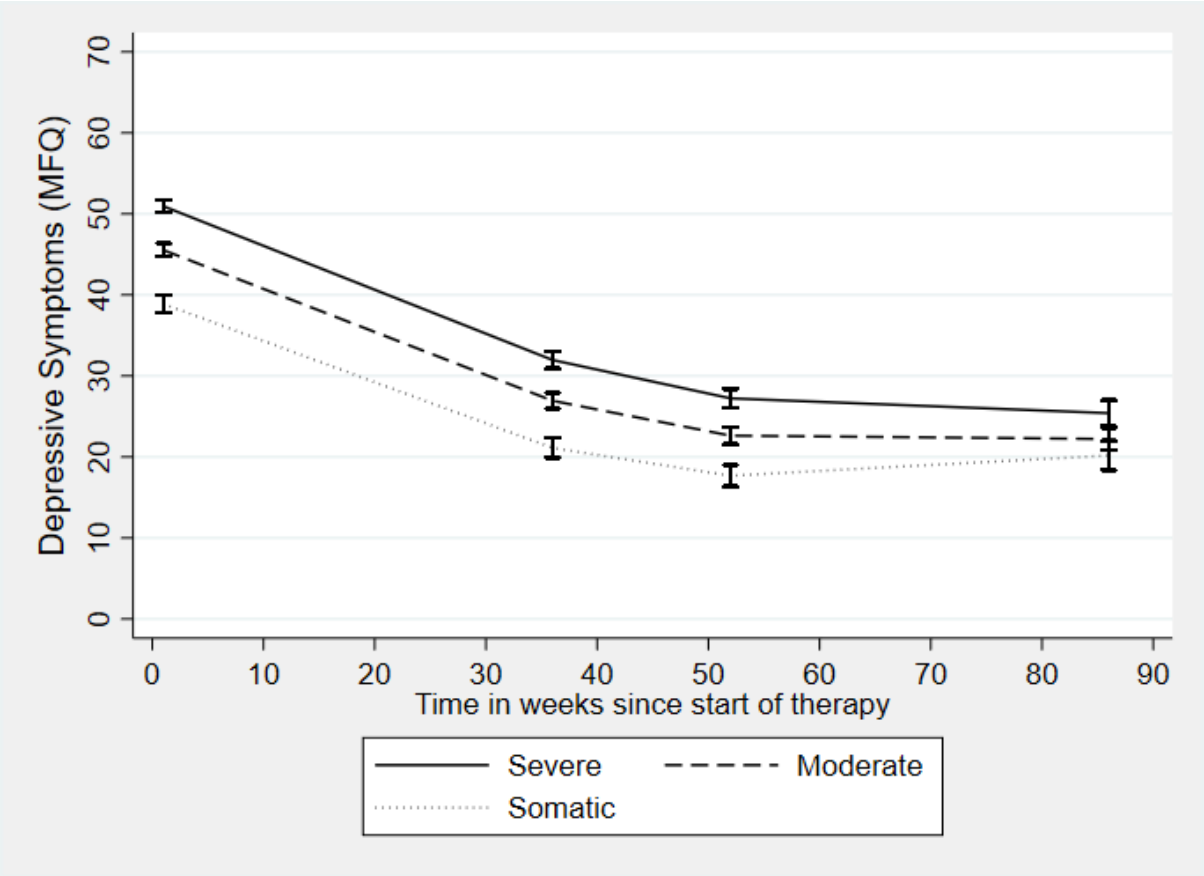

Table 1. Fit indices for the four class solutions evaluated for the IMPACT only analysis.

	AIC	BIC	BICSSA	Entropy	VLMH LRT	LMH ajd LRT	PBLRT	Class size
1 class	4965.75	5006.93	4975.20	--	--	--	--	454
2 class	4911.23	4997.71	4931.07	0.487	0.007	0.008		280/174
3 class	**4885.73**	**5017.51**	**4915.95**	**0.521**	**0.008**	**0.008**	**<.001**	**169/190 / 95**
4 class	4886.77	5063.85	4927.38	0.57	0.713	0.719	0.6667	202/49/ 111/92

*AIC = Akaike Information Criteria; BIC = Bayesian Information Criteria; BICSSA = Bayesian Information Criteria Sample Size Adjusted; VLMH = Vuong-Lo-Mendell-Rubin Likelihood Ratio Test; LMH ajd LRT = Lo-Mendell-Rubin adjusted Likelihood Ratio Test; PBLRT = parametric bootstrapped Likelihood Ratio Test.

Table 2. The three Class solution for the IMPACT sample with symptom prevalence, class probabilities, percentage female and average number of symptoms within each class. Each class compared to the overall rate of the full sample.

Symptom	**Class 1 – 'Severe-' (N = 169; 37.2%)**	**Class 2 – 'Moderate' (N = 190; 41.9%)**	**Class 3 – 'Somatic' (N = 95; 20.9%)**	**Full Sample (N = 454)**
Depressed Mood	100%***	91.6%*	42.1%***	84.3%
Irritable Mood	74.6%**	49.0%**	70.2%	62.9%
Anhedonia	90.5%***	44.4%***	67.4%	66.4%
Appetite Weight Change	90.5%***	48.3%***	77.0%	70.4%
Insomnia/Hypersomnia	100%***	83.1%***	100%**	92.9%
Psychomotor Agitation/Retardation	59.4%*	43.9%	47.9%	50.5%
Fatigue	92.3%***	63.3%**	63.8%*	74.3%
Worthlessness/excessive or inappropriate guilt	93.4%***	87.0%*	33.3%***	78.2%
Decreased concentration, slow thinking, indecisiveness	100%***	61.6%***	92.5%*	82.6%
Suicidal thoughts/ideation	83.9%***	62.8%	8.5%***	59.3%
Probability in Class 1	**.74**	.09	.04	--
Probability in Class 2	.20	**.81**	.18	--
Probability in Class 3	.06	.09	**.78**	--
% female	78.7%	75.3%	68.4%	75.1%
Average number of symptoms	8.8 (.77)***	6.3 (1.1)***	6.0 (1.3)***	7.25 (1.6)

*** $p < .001$; ** $p < .01$; * $p < .05$; note discrepancies reflect differences in power

Table 3. Baseline, 36 and 86 week follow-up self report symptoms and therapeutic alliance within each class.

Symptom	Severe subgroup (N = 169; 37.2%)	Moderate subgroup (N = 190; 41.9%)	Somatic subgroup (N = 95; 20.9%)	Full Sample (N = 454)
Baseline				
Depression Symptoms (MFQ)	51.44 (7.52)	45.49 (10.12)	46.15 (10.50)	46.15 (10.50)
% above clinical cutoff (MFQ)	100%	97.02%	84.2%	95.58%
Anxiety Symptoms (RCMAS)	43.64 (5.76)	40.67 (6.30)	36.93 (9.00)	41.02 (7.18)
Leyton Obsessional Inventory (LOI)	11.49 (4.98)	9.86 (5.59)	8.03 (4.66)	10.09 (5.33)
Antisocial Behaviour	3.86 (3.19)	2.94 (3.16)	3.09 (3.05)	3.31 (3.17)
Self Esteem (Rosenberg)	5.87 (3.63)	7.51 (4.05)	11.31 (3.37)	7.67 (4.32)
Risk Taking (RTSHIA*)*	6.70 (5.17)	5.63 (5.05)	5.28 (4.19)	5.97 (4.96)
Self Harm (RTSHIA*)*	17.18 (11.24)	14.54 (10.70)	7.43 (7.31)	14.03 (10.90)
HONOSCA	19.67 (4.72)	16.86 (5.12)	16.76 (4.50)	17.83 (5.02)
Rumination (RRS)	56.90 (10.38)	55.40 (11.64)	48.39 (11.51)	54.45 (11.60)
SSRI	21.1%	20.5%	14.9%	19.6%
36 Week Follow-Up	N = 105	N = 127	N = 55	N = 287
Still in sample	62.1%	66.8%	57.9%	63.2%
Working Alliance (WAIS)	53.66 (14.74)	54.66 (14.26)	54.37 (15.53)	54.23 (14.61)
Depression Symptoms (MFQ)	30.18 (16.11)	25.44 (15.43)	20.76 (13.99)	26.28 (15.75)
Reduction in Depression Symptoms	21.98 (15.65)	20.58 (15.39)	14.67 (14.15)	19.98 (15.43)
% above clinical cutoff (MFQ)	53.4%	52.0%	25.9%	47.5%
Anxiety Symptoms (RCMAS)	32.61 (12.95)	27.96 (13.08)	25.96 (13.72)	29.28 (13.38)
HONOSCA	10.40 (7.00)	9.24 (7.13)	9.6 (6.11)	9.76 (6.88)
52 Week Follow-Up	N = 104	N = 124	N = 59	N = 287
Still in sample	61.5%	65.3%	62.1%	63.2%
Depression Symptoms (MFQ)	26.41 (17.54)	22.48 (15.10)	20.54 (16.56)	23.51 (16.42)
Reduction in Depression Symptoms	26.52 (16.88)	23.83 (16.06)	16.29 (17.65)	23.32 (17.02)
% above clinical cutoff (MFQ)	45.2%	37.1%	25.4%	37.6%
86 Week Follow-Up	N = 112	N = 134	N = 66	N =312
Still in sample	66.3%	70.5%	69.5%	68.7%
Depression Symptoms (MFQ)	24.81 (17.06)	21.99 (15.64)	18.64 (14.22)	22.29 (16.00)

Reduction in Depression Symptoms	26.77 (17.07)	24.41 (16.28)	17.55 (15.83)	23.85 (16.76)
% above clinical cutoff (MFQ)	46.4%	35.1%	19.7%	35.9%

References

A.P.A. (2013). *Diagnostic and statistical manual of mental disorders : DSM-5*. American Psychiatric Association.

Aitken, M., Haltigan, J. D., Szatmari, P., Dubicka, B., Fonagy, P., Kelvin, R., Midgley, N., Reynolds, S., Wilkinson, P. O., & Goodyer, I. M. (2020). Toward precision therapeutics: general and specific factors differentiate symptom change in depressed adolescents. *J Child Psychol Psychiatry*, *61*(9), 998-1008. https://doi.org/10.1111/jcpp.13194

Bamber, D., Tamplin, A., Park, R. J., Kyte, Z. A., & Goodyer, I. M. (2002). Development of a short leyton obsessional inventory for children and adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*(10), 1246-1252.

Bondar, J., Caye, A., Chekroud, A. M., & Kieling, C. (2020). Symptom clusters in adolescent depression and differential response to treatment: a secondary analysis of the Treatment for Adolescents with Depression Study randomised trial. *The Lancet Psychiatry*, *7*(4), 337-343. https://doi.org/10.1016/s2215-0366(20)30060-2

[Record #3443 is using a reference type undefined in this output style.]

Catania, L. S., Hetrick, S. E., Newman, L. K., & Purcell, R. (2011). Prevention and early intervention for mental health problems in 0–25 year olds: Are there evidence-based models of care? *Advances in Mental Health*, *10*(1), 6-19. https://doi.org/10.5172/jamh.2011.10.1.6

Catarino, A., Fawcett, J. M., Ewbank, M. P., Bateup, S., Cummins, R., Tablan, V., & Blackwell, A. D. (2020). Refining our understanding of depressive states and state transitions in response to cognitive behavioural therapy using latent Markov modelling. *Psychol Med*, 1-10. https://doi.org/10.1017/S0033291720002032

Clayborne, Z. M., Varin, M., & Colman, I. (2019). Systematic Review and Meta-Analysis: Adolescent Depression and Long-Term Psychosocial Outcomes. *J Am Acad Child Adolesc Psychiatry*, *58*(1), 72-79. https://doi.org/10.1016/j.jaac.2018.07.896

Costello, E. J., & Angold, A. (1988). Scales to assess child and adolescent depression: checklists, screens, and nets. *J Am Acad Child Adolesc Psychiatry*, *27*, 726-737. https://doi.org/10.1097/00004583-198811000-00011

Davies, S. E., Neufeld, S. A. S., van Sprang, E., Schweren, L., Keivit, R., Fonagy, P., Dubicka, B., Kelvin, R., Midgley, N., Reynolds, S., Target, M., Wilkinson, P., van Harmelen, A. L., & Goodyer, I. M. (2019). Trajectories of depression symptom change during and following treatment in adolescents with unipolar major depression. *J Child Psychol Psychiatry*. https://doi.org/10.1111/jcpp.13145

Ferdinand, R. F., de Nijs, P. F., van Lier, P., & Verhulst, F. C. (2005). Latent class analysis of anxiety and depressive symptoms in referred adolescents. *J Affect Disord*, *88*(3), 299-306. https://doi.org/10.1016/j.jad.2005.08.004

Fried, E. I. (2015). Problematic assumptions have slowed down depression research: why symptoms, not syndromes are the way forward. *Front Psychol*, *6*, 309. https://doi.org/10.3389/fpsyg.2015.00309

Fried, E. I., & Nesse, R. M. (2014). The impact of individual depressive symptoms on impairment of psychosocial functioning. *PLoS One*, *9*(2), e90311. https://doi.org/10.1371/journal.pone.0090311

Goodyer, I. M., Reynolds, S., Barrett, B., Byford, S., Dubicka, B., Hill, J., Holland, F., Kelvin, R., Midgley, N., & Roberts, C. (2017). Cognitive-behavioural therapy and short-term psychoanalytic psychotherapy versus brief psychosocial intervention in adolescents with unipolar major depression (IMPACT): a multicentre, pragmatic, observer-blind, randomised controlled trial. *Health Technol Assess*, *21*(12), 1.

Goodyer, I. M., Tsancheva, S., Byford, S., Dubicka, B., Hill, J., Kelvin, R., Reynolds, S., Roberts, C., Senior, R., Suckling, J., Wilkinson, P., Target, M., & Fonagy, P. (2011). Improving mood with psychoanalytic and cognitive therapies (IMPACT): a pragmatic effectiveness superiority trial to investigate whether specialised psychological treatment reduces the risk for relapse in

adolescents with moderate to severe unipolar depression: study protocol for a randomised controlled trial. *Trials*, *12*, 175. https://doi.org/10.1186/1745-6215-12-175

Gowers, S., Levine, W., Bailey-Rogers, S., Shore, A., & Burhouse, E. (2002). Use of a routine, self-report outcome measure (HoNOSCA-SR) in two adolescent mental health services. *British Journal of Psychiatry*, *180*, 266-269. <Go to ISI>://WOS:000174355900015

Hatcher, R. L., & Gillaspy, J. A. (2006). Development and validation of a revised short version of the working alliance inventory. *Psychotherapy Research*, *16*(1), 12-25. https://doi.org/10.1080/10503300500352500

Herman, K. C., Ostrander, R., Walkup, J. T., Silva, S. G., & March, J. S. (2007). Empirically derived subtypes of adolescent depression: latent profile analysis of co-occurring symptoms in the Treatment for Adolescents with Depression Study (TADS). *J Consult Clin Psychol*, *75*(5), 716-728. https://doi.org/10.1037/0022-006X.75.5.716

Johnson, D., Dupuis, G., Piche, J., Clayborne, Z., & Colman, I. (2018). Adult mental health outcomes of adolescent depression: A systematic review. *Depress Anxiety*, *35*(8), 700-716. https://doi.org/10.1002/da.22777

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., Rush, A. J., Walters, E. E., & Wang, P. S. (2003). The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA*, *289*(23), 3095-3105.

Krueger, R. F., & Bezdjian, S. (2009). Enhancing research and treatment of mental disorders with dimensional concepts: toward DSM-V and ICD-11. *World Psychiatry*, *8*(1), 3.

Malla, A., & McGorry, P. (2019). Early Intervention in Psychosis in Young People: A Population and Public Health Perspective. *American Journal of Public Health*, *109*(S3), S181-S184. https://doi.org/10.2105/ajph.2019.305018

McElroy, E., Napoleone, E., Wolpert, M., & Patalay, P. (2019). Structure and Connectivity of Depressive Symptom Networks Corresponding to Early Treatment Response. *EClinicalMedicine*, *8*, 29-36. https://doi.org/10.1016/j.eclinm.2019.02.009

Mullarkey, M. C., Marchetti, I., & Beevers, C. G. (2019). Using network analysis to identify central symptoms of adolescent depression. *Journal of Clinical Child & Adolescent Psychology*, *48*(4), 656-668.

Murphy, R. (2016). Child and adolescent mental health: trends and key issues.

[Record #3442 is using a reference type undefined in this output style.]

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(4), 535-569. https://doi.org/10.1080/10705510701575396

Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., & Rohde, L. A. (2015). Annual Research Review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *Journal of Child Psychology and Psychiatry*, *56*(3), 345-365.

Ramaswamy, V., DeSarbo, W. S., Reibstein, D. J., & Robinson, W. T. (1993). An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Marketing Science*, *12*(1), 103-124.

Reynolds, C. R., & Richmond, B. O. (1978). What I Think and Feel - Revised Measure of Childrens Manifest Anxiety. *Journal of Abnormal Child Psychology*, *6*, 271-280. https://doi.org/Doi 10.1007/Bf00919131

Rice, F., Riglin, L., Lomax, T., Souter, E., Potter, R., Smith, D. J., Thapar, A. K., & Thapar, A. (2019). Adolescent and adult differences in major depression symptom profiles. *J Affect Disord*, *243*, 175-181. https://doi.org/10.1016/j.jad.2018.09.015

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton University Press.

Schweren, L., van Borkulo, C. D., Fried, E., & Goodyer, I. M. (2018). Assessment of symptom network density as a prognostic marker of treatment response in adolescent depression. *JAMA Psychiatry*, *75*(1), 98-100.

Simmonds-Buckley, M., Catarino, A., & Delgadillo, J. (2021). Depression subtypes and their response to cognitive behavioral therapy: A latent transition analysis. *Depress Anxiety*, *38*(9), 907-916. https://doi.org/10.1002/da.23161

Simmons, M., Wilkinson, P., & Dubicka, B. (2015). Measurement Issues: Depression measures in children and adolescents. *Child Adolesc Ment Health*, *20*(4), 230-241. https://doi.org/10.1111/camh.12106

St Clair, M. C., Neufeld, S., Jones, P. B., Fonagy, P., Bullmore, E. T., Dolan, R. J., Moutoussis, M., Toseeb, U., & Goodyer, I. M. (2017a). Characterising the latent structure and organisation of self-reported thoughts, feelings and behaviours in adolescents and young adults. *PLoS One*, *12*(4), e0175381. https://doi.org/10.1371/journal.pone.0175381

St Clair, M. C., Neufeld, S., Jones, P. B., Fonagy, P., Bullmore, E. T., Dolan, R. J., Moutoussis, M., Toseeb, U., & Goodyer, I. M. (2017b). Characterising the latent structure and organisation of self-reported thoughts, feelings and behaviours in adolescents and young adults. *Plos One*, *12*(4). https://doi.org/ARTN e0175381

10.1371/journal.pone.0175381

StataCorp. (2017). Stata 15. *Stata Cooperation, College Station, TX*.

Thapar, A., Collishaw, S., Pine, D. S., & Thapar, A. K. (2012). Depression in adolescence. *The Lancet*, *379*(9820), 1056-1067. https://doi.org/10.1016/s0140-6736(11)60871-4

Treynor, W., Gonzalez, R., & Nolen-Hoeksema, S. (2003). Rumination reconsidered: A psychometric analysis. *Cognitive Therapy and Research*, *27*(3), 247-259. <Go to ISI>://WOS:000183205900002

Ulbricht, C. M., Chrysanthopoulou, S. A., Levin, L., & Lapane, K. L. (2018). The use of latent class analysis for identifying subtypes of depression: A systematic review. *Psychiatry Res*, *266*, 228-246. https://doi.org/10.1016/j.psychres.2018.03.003

van Lang, N. D., Ferdinand, R. F., Ormel, J., & Verhulst, F. C. (2006). Latent class analysis of anxiety and depressive symptoms of the Youth Self-Report in a general population sample of young adolescents. *Behav Res Ther*, *44*(6), 849-860. https://doi.org/10.1016/j.brat.2005.06.004

Vrouva, I., Fonagy, P., Fearon, P. R. M., & Roussow, T. (2010). The Risk-Taking and Self-Harm Inventory for Adolescents: Development and Psychometric Evaluation. *Psychological Assessment*, *22*(4), 852-865. <Go to ISI>://WOS:000285231300014

Yu, Z. J., Kratochvil, C. J., Weller, R. A., Mooreville, M., & Weller, E. B. (2010). From TADS and SOFTADS to TORDIA and Beyond: What's New in the Treatment of Adolescent Depression? *Current Psychiatry Reports*, *12*(2), 88-95. https://doi.org/10.1007/s11920-010-0094-x

Supplementary Materials

We investigated differences in baseline characteristics for each subgroup based on whether they remained in the sample or not at 36, 52 and 86 weeks. In total, there were four significant differences at the p <.05 level from over 90 comparisons. There were no consistent patterns and if multiple comparison adjustments were to be made, none of the significant findings would remain.

36 week follow-up

Within the moderate subgroup, we found a significant increase in obsessionality in those who dropped out of the study at 36 months compared to those who remained in the sample (B = 1.77; 96%CI(.05;3.48;), p < .05; β = .15). There were no other differences in any subgroup across any of the baseline characteristics. Please see Table S1-S3 for means and significance levels for each subgroup respectively.

Table S1. Rates of baseline measures in the severe subgroup dependent on whether the participants remained in the sample at 36 months or dropped out

Symptom	**Severe subgroup – in sample at 36 months (N = 105)**	**Severe subgroup – not in sample at 36 months (N = 64)**	**P value**
Baseline			
Depression Symptoms (MFQ)	51.82 (6.98)	50.78 (8.43)	.44
Anxiety Symptoms (RCMAS)	43.99 (6.09)	43.03 (5.16)	.31
Leyton Obsessional Inventory (LOI)	11.25 (4.80)	11.90 (5.29)	.44
Antisocial Behaviour	3.47 (3.12)	4.48 (3.22)	.07
Self Esteem (Rosenberg)	5.87 (3.78)	5.87 (4.03)	.99
Risk Taking (RTSHIA*)*	6.05 (4.84)	7.84 (5.55)	.10
Self Harm (RTSHIA*)*	17.02 (10.43)	17.44 (12.55)	.82
HONOSCA	19.10 (4.64)	20.63 (4.74)	.08
Rumination (RRS)	57.22 (10.37)	56.38 (10.49)	.65
SSRI	22.1%	19.4%	.67

Table S2. Rates of baseline measures in the Moderate subgroup dependent on whether the participants remained in the sample at 36 months or dropped out

Symptom	**Moderate subgroup – in**	**Moderate subgroup – not in**	**P value**

	sample at 36 months (N = 127)	sample at 36 months (N =63)	
Baseline			
Depression Symptoms (MFQ)	45.52 (9.78)	45.44 (10.94)	.97
Anxiety Symptoms (RCMAS)	40.49 (6.48)	41.00 (5.97)	.61
Leyton Obsessional Inventory (LOI)	*9.28 (5.52)*	*11.05 (5.58)*	*.04*
Antisocial Behaviour	2.89 (3.02)	3.05 (3.46)	.74
Self Esteem (Rosenberg)	7.41 (4.19)	7.71 (3.78)	.62
Risk Taking (RTSHIA*)*	5.42 (4.69)	6.09 (5.79)	.50
Self Harm (RTSHIA*)*	14.24 (10.03)	15.13 (12.01)	.61
HONOSCA	17.41 (5.23)	15.68 (4.72)	.06
Rumination (RRS)	54.74 (11.91)	56.80 (11.06)	.31
SSRI	18.7%	24.2%	.38

Table S3. Rates of baseline measures in the Somatic subgroup dependent on whether the participants remained in the sample at 36 months or dropped out

Symptom	Somatic subgroup – in sample at 36 months (N = 55)	Somatic subgroup – not in sample at 36 months (N = 40)	P value
Baseline			
Depression Symptoms (MFQ)	36.17 (8.73)	39.65 (11.66)	.14
Anxiety Symptoms (RCMAS)	36.84 (8.75)	37.05 (9.43)	.93
Leyton Obsessional Inventory (LOI)	7.82 (4.50)	8.32 (4.93)	.63
Antisocial Behaviour	3.08 (3.45)	3.11 (2.41)	.96
Self Esteem (Rosenberg)	11.5 (2.96)	11.00 (3.98)	.53
Risk Taking (RTSHIA*)*	5.04 (4.38)	5.63 (3.94)	.57
Self Harm (RTSHIA*)*	7.55 (7.54)	7.28 (7.08)	.86
HONOSCA	16.35 (4.23)	17.39 (4.90)	.32
Rumination (RRS)	49.86 (11.26)	45.82 (11.70)	.14
SSRI	16.4%	12.8%	.64

52 week follow-up

Within the severe subgroup, we found higher baseline risk taking in with those who dropped out of the study at 52 months than those who remained in the sample (B = .29;

96%CI(.05;.54), $p < .05$). Those in the severe subgroup who dropped out also had higher baseline therapeutic alliance than those who remained in the sample (B = 1.95; 96%CI(.23;3.66), $p < .05$; $\beta = .04$). Within the somatic subgroups, those who remained in the sample had higher baseline rumination than those who dropped out by 52 months (B = 5.33; 96%CI(.04;10.61), $p < .05$; $\beta = .23$). There were no other differences in any subgroup across any of the baseline characteristics. Please see Table S4-S6 for means and significance levels for each subgroup respectively.

Table S4. Rates of baseline measures in the severe subgroup dependent on whether the participants remained in the sample at 52 months or dropped out

Symptom	**Severe subgroup – in sample at 52 months (N = 104)**	**Severe subgroup – not in sample at 52 months (N = 65)**	**P value**
Baseline			
Depression Symptoms (MFQ)	51.67 (7.53)	51.08 (7.56)	.64
Anxiety Symptoms (RCMAS)	43.91 (5.77)	43.17 (5.77)	.43
Leyton Obsessional Inventory (LOI)	11.21 (4.76)	11.95 (5.33)	.37
Antisocial Behaviour	3.72 (3.28)	4.06 (3.05)	.54
Self Esteem (Rosenberg)	6.05 (3.80)	5.58 (3.98)	.46
Risk Taking (RTSHIA)	*5.90 (4.73)*	*7.92 (5.59)*	*.02*
Self Harm (RTSHIA)	17.17 (10.65)	17.18 (12.23)	.99
HONOSCA	*18.96 (4.70)*	*20.91 (4.53)*	*.03*
Rumination (RRS)	56.68 (10.12)	57.27 (10.91)	.75
SSRI	17.5%	27.0%	.15

Table S5. Rates of baseline measures in the Moderate subgroup dependent on whether the participants remained in the sample at 52 months or dropped out

Symptom	**Moderate subgroup – in sample at 52 months (N = 124)**	**Moderate subgroup – not in sample at 52 months (N =66)**	**P value**
Baseline			
Depression Symptoms (MFQ)	45.71 (9.56)	45.04 (11.28)	.71
Anxiety Symptoms (RCMAS)	40.92 (6.23)	40.18 (6.45)	.46

Leyton Obsessional Inventory (LOI)	9.68 (5.54)	10.22 (5.70)	.53
Antisocial Behaviour	2.69 (2.68)	3.4 (3.88)	.15
Self Esteem (Rosenberg)	7.20 (3.88)	8.08 (4.32)	.18
Risk Taking (RTSHIA*)*	5.78 (4.98)	5.33 (5.23)	.60
Self Harm (RTSHIA*)*	14.48 (10.45)	14.65 (11.24)	.93
HONOSCA	17.19 (5.23)	16.18 (4.87)	.26
Rumination (RRS)	55.85 (11.45)	54.50 (12.08)	.50
SSRI	16.7%	27.7%	.08

Table S6. Rates of baseline measures in the Somatic subgroup dependent on whether the participants remained in the sample at 52 months or dropped out

Symptom	**Somatic subgroup – in sample at 52 months (N = 59)**	**Somatic subgroup – not in sample at 52 months (N = 36)**	**P value**
Baseline			
Depression Symptoms (MFQ)	38.10 (9.75)	36.81 (10.83)	.59
Anxiety Symptoms (RCMAS)	36.95 (9.16)	36.91 (8.86)	.99
Leyton Obsessional Inventory (LOI)	7.65 (4.71)	8.58 (4.60)	.36
Antisocial Behaviour	2.64 (2.41)	3.85 (3.83)	.08
Self Esteem (Rosenberg)	11.44 (3.24)	11.09 (3.63)	.65
Risk Taking (RTSHIA*)*	5.02 (3.73)	5.68 (4.87)	.51
Self Harm (RTSHIA*)*	6.73 (6.71)	8.58 (8.17)	.31
HONOSCA	16.89 (4.29)	16.48 (5.01)	.71
Rumination (RRS)	*50.40 (10.85)*	*45.07 (11.99)*	*<.05*
SSRI	16.95%	11.43%	.47

86 week follow-up

There were no differences in baseline characteristics in those who dropped out compared to those who remained in the sample at 86 months in any subgroup.

Table S7. Rates of baseline measures in the severe subgroup dependent on whether the participants remained in the sample at 86 months or dropped out

Symptom	**Severe subgroup – in sample at 86 months (N = 112)**	**Severe subgroup – not in sample at 86**	**P value**

		months (N = 57)	
Baseline			
Depression Symptoms (MFQ)	51.37 (7.68)	51.58 (7.27)	.87
Anxiety Symptoms (RCMAS)	43.47 (5.68)	43.98 (5.99)	.60
Leyton Obsessional Inventory (LOI)	11.26 (4.94)	11.94 (5.08)	.42
Antisocial Behaviour	3.94 (3.17)	3.68 (3.26)	.64
Self Esteem (Rosenberg)	5.96 (3.76)	5.67 (4.10)	.67
Risk Taking (RTSHIA*)*	6.72 (5.01)	6.67 (5.51)	.96
Self Harm (RTSHIA*)*	17.05 (10.74)	17.42 (12.28)	.85
HONOSCA	19.31 (4.99)	20.32 (4.15)	.26
Rumination (RRS)	57.08 (10.15)	56.55 (10.93)	.78
SSRI	18.9%	25.5%	.33

Table S8. Rates of baseline measures in the Moderate subgroup dependent on whether the participants remained in the sample at 86 months or dropped out

Symptom	**Moderate subgroup – in sample at 86 months (N = 134)**	**Moderate subgroup – not in sample at 86 months (N =56)**	**P value**
Baseline			
Depression Symptoms (MFQ)	45.60 (10.16)	45.2 (10.12)	.82
Anxiety Symptoms (RCMAS)	41.21 (6.36)	39.42 (6.02)	.08
Leyton Obsessional Inventory (LOI)	9.73 (5.75)	10.18 (5.20)	.62
Antisocial Behaviour	3.04 (3.28)	2.70 (2.87)	.50
Self Esteem (Rosenberg)	7.54 (3.91)	7.43 (4.40)	.87
Risk Taking (RTSHIA*)*	5.54 (4.83)	5.87 (5.64)	.75
Self Harm (RTSHIA*)*	13.95 (10.64)	15.95 (10.80)	.23
HONOSCA	6.68 (5.10)	17.32 (5.21)	.50
Rumination (RRS)	55.55 (11.86)	55.02 (11.18)	.80
SSRI	17.4%	28.3%	.10

Table S9. Rates of baseline measures in the Somatic subgroup dependent on whether the participants remained in the sample at 86 months or dropped out

Symptom	**Somatic subgroup – in sample at 86 months (N =66)**	**Somatic subgroup – not in**	**P value**

		sample at 86 months (N = 29)	
Baseline			
Depression Symptoms (MFQ)	36.81 (10.15)	39.54 (10.01)	.26
Anxiety Symptoms (RCMAS)	35.84 (9.58)	39.41 (7.08)	.09
Leyton Obsessional Inventory (LOI)	7.40 (4.69)	9.39 (4.39)	.06
Antisocial Behaviour	2.66 (2.64)	4.07 (3.70)	.06
Self Esteem (Rosenberg)	11.30 (3.40)	11.32 (3.38)	.98
Risk Taking (RTSHIA*)*	4.86 (3.90)	6.30 (4.77)	.21
Self Harm (RTSHIA*)*	6.71 (7.31)	9.07 (7.17)	.13
HONOSCA	16.47 (4.59)	17.42 (4.31)	.40
Rumination (RRS)	48.66 (10.96)	47.67 (13.72)	.74
SSRI	19.7%	3.6%	.08