



RESEARCH ARTICLE

Effectiveness of an Integrated Care Pathway for Adolescents with Depression: A Pilot Clinical Trial Protocol

Darren B. Courtney MD¹; Amy Cheung MD²; Joanna Henderson PhD³; Kathryn Bennett PhD⁴; Marco Battaglia MD⁵; John Strauss⁵; Rachel Mitchell⁶; Karen Wang⁶; Peter Szatmari⁷

Abstract

Objective: Our group aims to improve treatment response for adolescents with depression through the use of an Integrated Care Pathway (ICP) we developed using: (1) recommendations from a high quality Clinical Practice Guideline (CPG); and, (2) a measurement-based care framework. **Method:** Pre-specified criteria will identify eligible adolescents in two outpatient hospital study sites. Study group allocation, to the ICP versus treatment as usual (TAU), is based on site of presentation. The primary clinical outcome is reduction of depression symptoms, assessed using the Childhood Depression Rating Scale – Revised (CDRS-R). Measures will be taken at baseline and every four weeks until 20 weeks of treatment has been offered. **Results:** Our overall hypothesis is that the ICP will be associated with greater improvement in depressive symptoms compared to TAU. Feasibility targets for this pilot trial include the following: recruitment of 30 participants per site over a 21-month period, 95% baseline assessment completion rates, 90% clinician adherence to the ICP in the intervention arm and 80% completion of the scheduled CDRS-R measures over the 20-week interval. Focus-group feedback from youth and parents will also produce qualitative information. **Conclusions:** If feasibility targets are met, and preliminary results regarding clinical outcomes are promising, then a multi-center cluster RCT would be pursued.

Key Words: adolescent, depression, integrated care pathway, measurement-based care, pilot study

Résumé

Objectif: Notre groupe vise à améliorer la réponse au traitement pour les adolescents souffrant de dépression grâce à l'utilisation d'une trajectoire de soins intégrés (TSI) que nous avons élaborée à l'aide (1) des recommandations de Lignes directrices de pratique clinique (LDPC) de grande qualité et (2) d'un cadre de soins axés sur les mesures. **Méthode:** Des critères pré-spécifiés sélectionneront les adolescents admissibles à deux sites de l'étude en milieu hospitalier ambulatoire. La répartition du groupe de l'étude, vers la TSI par opposition au traitement habituel (TH), est basée sur le site de la présentation. Le principal résultat clinique est la réduction des symptômes dépressifs, évaluée à l'aide de l'échelle de dépression chez les enfants – révisée (CDRS-R). Les mesures seront prises au départ et à toutes les 4

¹Centre for Addiction and Mental Health (CAMH), Assistant Professor, University of Toronto, Toronto, Ontario

²Sunnybrook Health Sciences Centre, Associate Professor, University of Toronto, Toronto, Ontario

³Centre for Addiction and Mental Health (CAMH), Associate Professor, University of Toronto, Toronto Ontario

⁴Professor, Health Research Methods, Evidence and Impact (formerly Clinical Epidemiology and Biostatistics), McMaster University, Hamilton, Ontario

⁵Centre for Addiction and Mental Health (CAMH), Associate Professor, University of Toronto, Toronto, Ontario

⁶Sunnybrook Health Sciences Centre, Assistant Professor, University of Toronto, Toronto, Ontario

⁷Centre for Addiction and Mental Health (CAMH), Hospital for Sick Children, Professor, University of Toronto, Toronto, Ontario

Corresponding E-Mail: Darren.Courtney@camh.ca

Submitted: December 15, 2018; Accepted: September 1, 2019

semaine jusqu'à 20 semaines de traitement. **Résultats:** Notre hypothèse générale est que la TSI sera associée à une amélioration plus marquée des symptômes dépressifs, comparativement au TH. Les cibles de faisabilité pour cet essai pilote sont notamment le recrutement de 30 participants par site sur une période de 21 mois, des taux d'achèvement de 95 % de l'évaluation de départ, de 90 % d'adhésion du clinicien à la TSI durant le segment d'intervention, et de 80 % d'achèvement des mesures de la CDRS-R prévues dans l'intervalle de 20 semaines. [Commentaire supprimé sur le test d'efficacité préliminaire.]. Les commentaires des groupes de discussion des adolescents et des parents produiront aussi une information qualitative. **Conclusions:** Si les cibles de faisabilité sont atteintes, et que les résultats préliminaires à l'égard des résultats cliniques sont prometteurs, alors un groupe d'essais randomisés contrôlés (ERC) multicentrique serait réalisable.

Mots clés: adolescent, dépression, trajectoire de soins intégrés, soins basés sur les mesures, étude pilote

The lifetime prevalence of major depressive disorders in adolescents (MDD-A) in Canada is estimated at 7.6% (Cheung & Dewa, 2006). The World Health Organization found that, among youth aged 10-24 years, depressive disorders are the leading cause of global burden of disease as measured by Disability Adjusted Life Years (Gore et al., 2011). Depression in adolescents is also a leading risk factor for completed suicide (Renaud et al., 2008). Accordingly, optimizing treatment for MDD-A may be reasonably expected to decrease morbidity and mortality and improve function.

A gap between what is scientifically known to be effective in mental health care, and what is actually practiced, presents a missed opportunity to optimize treatment for MDD-A (Girlanda, Fiedler, Becker, Barbui, & Koesters, 2017; McLennan, Wathen, MacMillan & Lavis, 2006). Indeed, through an Ontario-wide survey of services, our group has found that there is a high degree of variability in treatment offered to young people with mental health needs (manuscript in process) – indicating a lack of adherence to evidence-based treatments for the disorder.

There are several steps involved in bridging this research-practice gap. The first step is to identify a broad consensus on evidence-based treatments for MDD-A. High quality clinical practice guidelines (CPGs) are designed to play such a role; they are defined as “statements that include recommendations intended to optimize patient care that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options” (Graham, Mancher, Wolman, & Greenfield, 2011).

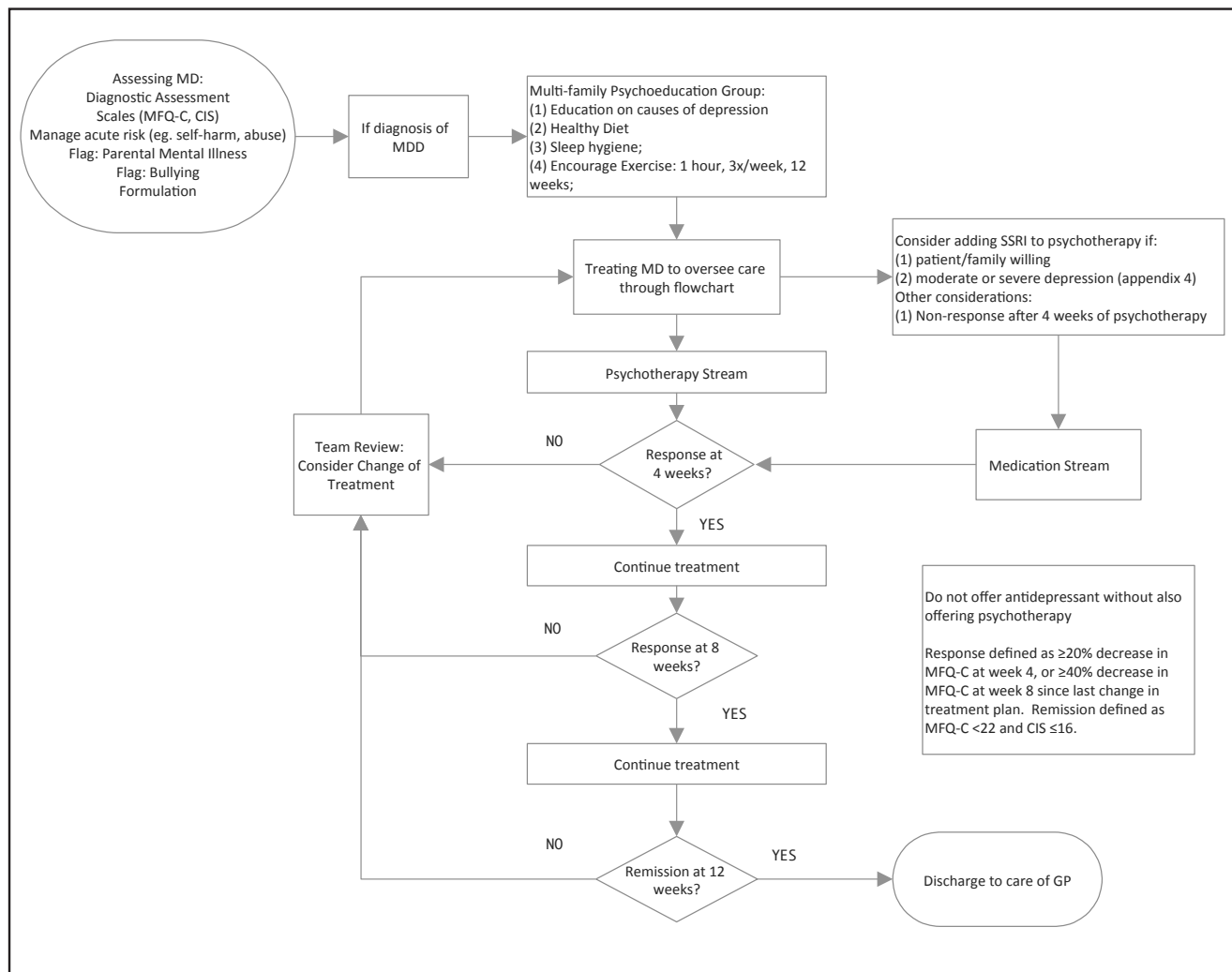
The second step in narrowing the research-practice gap is to implement the CPG recommendations in the clinical context. Integrated Care Pathways (ICPs) are one implementation tool: they are pre-set treatment decision aides based on CPG recommendations, designed to inform clinical choices throughout a patient's care (Campbell, Hotchkiss, Bradshaw, & Porteous, 1998).

The third step in closing the research-practice gap is to actively monitor the patients' progress as they carry out successive CPG recommendations, to verify that treatment is working towards improvement, or alerting clinicians that

treatment is not working and needs to be modified. Measurement-based care (MBC) offers such an approach; it “entails the systematic administration of symptom rating scales and uses the results to drive clinical decision-making at the level of the individual patient” (Fortney et al., 2017).

Our group has developed an ICP for the treatment of MDD-A that is based on the National Institute for Health and Care Excellence CPG for Depression in Children and Young People (hereafter, NICE-CPG) (NICE, 2015) and measurement-based care. For adolescents with depression, we anticipate that this ICP will lead to better symptom and functional outcomes relative to treatment as usual for this population.

Other research groups have studied, in part, the effectiveness of ICPs and MBC for the management of adolescent mental disorders, including MDD-A. Asarnow et al. (2006) conducted a RCT where a six-month quality improvement intervention aimed at improving *access* to evidence-based depression treatments for MDD-A resulted in significant improvement of depressive symptoms compared to usual care; however, this study did not involve reference to CPG recommendations nor did it involve MBC. Bickman, Kelley, Breda, de Andrade, and Riemer (2011) studied the effects of MBC in a home-based community intervention on multiple psychiatric symptoms and function in youth. They found that youth whose clinicians received feedback weekly on outcomes (high frequency MBC) had better outcomes than those who received feedback every 90 days (low frequency MBC). The effect size was small (Cohen's $d=0.18$). This study did not involve a treatment algorithm based on a CPG. In a pilot, Gunlicks-Stoessel et al. (2018) studied the efficacy of an “adaptive treatment strategy”, where measures are used at pre-determined time-points to make treatment decisions with regards to adolescent depression. They found that taking measurements at the four-week time point had greater effects than waiting until the eight-week time point to guide decisions making. The treatment options in the algorithm do not extend beyond interpersonal therapy for adolescents (IPT-A) in conjunction with fluoxetine treatment (Gunlicks-Stoessel et al., 2018). The Texas Children's Medication Algorithm Project (TMAP) (Emslie et al., 2004) used both an ICP-like treatment algorithm and

Figure 1. Integrated Care Pathway for Major Depressive Disorder in Adolescents

continued

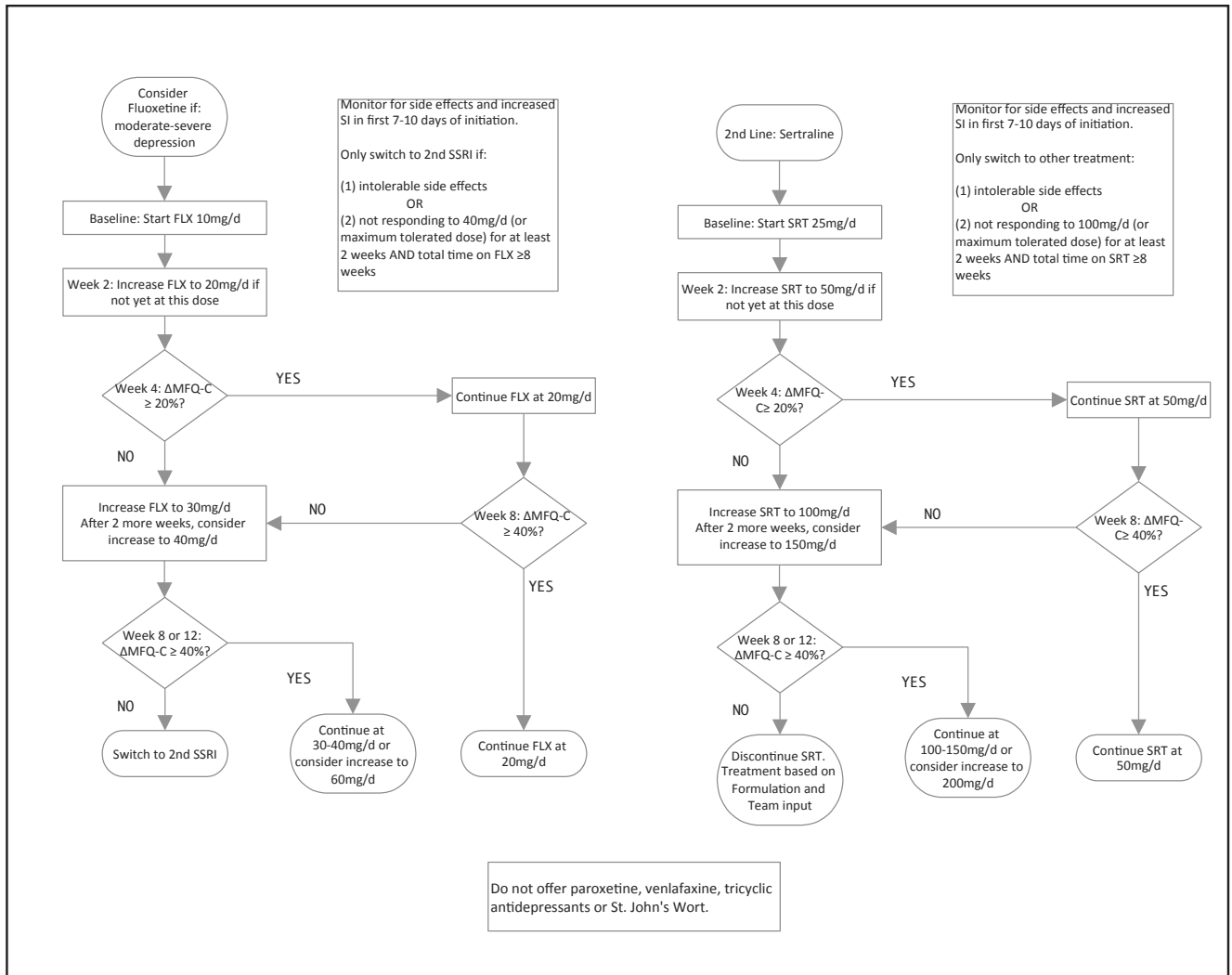
an MBC framework. They demonstrated superior response rates in depressive symptoms with the use of a medication treatment protocol (40%) compared to a historical cohort (23%), but this algorithm did not include psychotherapeutic interventions when psychotherapy is a first line treatment and key component in the NICE CPG (NICE, 2015). Moreover, the use of a historical control group limits the rigor of the TMAP study.

Most of what we know about CPG implementation, ICPs and MBC is extrapolated from the adult literature. In their meta-analysis, Girlanda et al. (2017) conclude that structured attempts to implement CPGs can lead to improved outcome for patients with mental health conditions. In a sample of adults with depression (n=1131), greater adherence to CPG recommendations was significantly associated with improved symptoms at 12 months, 18 months and 24 months (Hepner et al., 2007). In an RCT, Kane et al. (2016) have shown that implementation of an ICP for

young people with psychosis is effective in significantly improving psychopathology and function. In adults with mental health conditions, MBC has been shown to improve outcomes (Fortney et al., 2017). Shimokawa, Lambert, and Smart (2010) conducted a meta-analysis focusing on the use of MBC for psychotherapy interventions in adults and showed large effect sizes for those where clinicians were adherent to MBC (Hedge's $g=0.58$). Greater benefits are also noted for those who do not respond to initial treatment (Lambert et al., 2002).

High-quality CPGs for MDD-A exist but implementing these guidelines into clinical practice is difficult and complex. A key evidence gap is that we do not know what impact implementation will have on clinical outcomes. Our study aims to fill that evidence gap using an ICP and an MBC framework. The combined approach of implementing an ICP for MDD-A based on the broad range of recommendations from a high quality CPG and an MBC framework

Figure 1. continued: Medication Stream



has not yet been tested and would represent a major step forward in clinical care for such a common and debilitating disorder of adolescence.

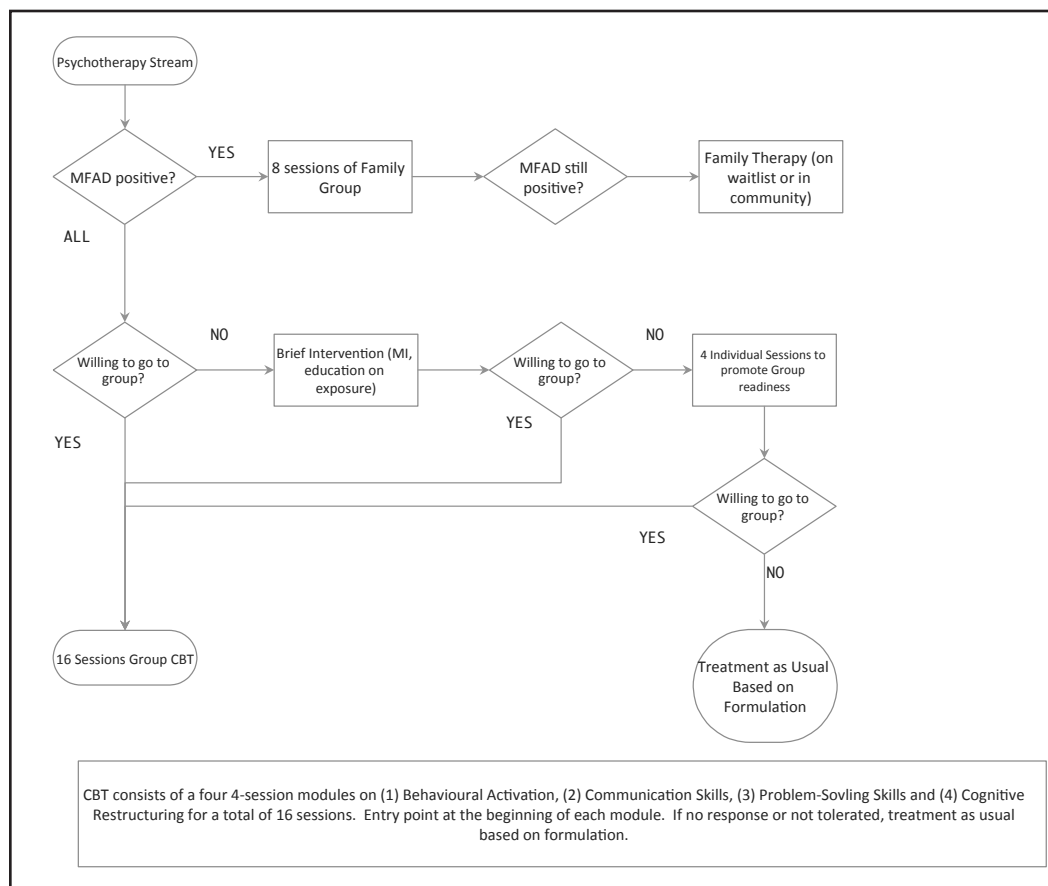
Preliminary Work

Our group has conducted a systematic review and quality appraisal of 21 CPGs relating to the management of MDD-A (Bennett et al, 2018). In this process we found that the NICE-CPG is the only up-to-date high quality CPG relating to the management of MDD-A. Subsequently, we translated the NICE-CPG recommendations into an ICP with collaboration from clinicians and youth with lived experience (manuscript on development methods in process). See Figure 1 for a visual representation of the resulting ICP.

Briefly, the ICP we developed offers all patients the following two treatment components which are all evidence-based and widely available: (1) A one-time multi-family “Mood

Foundations” group for patients and caregivers (ie. parents/guardians)– where we provide information on (a) the nature of depression, (b) sleep, (c) exercise and (d) diet; and, (2) a 16-session cognitive-behavioural group therapy treatment (CBT). The CBT group manual used is based off of the Adolescent Coping With Depression Course (Clarke et al, 2000), though, with permission, updated for current youth culture (available upon request). The following are conditional treatment components: (1) If the client is not willing to attend group CBT, four sessions of individual CBT is offered; (2) If the McMaster Family Assessment Device-general functioning subscale score (MFAD-GF; Epstein, Baldwin, & Bishop, 1983) is above an established cut-off at baseline (Akister & Stevenson-Hinde, 1991), caregivers are offered enrollment in an eight-session, “Caregivers Of Depressed Youth” group; (3) If measures indicate moderate-to-severe depression, medication options are offered; with an 8-12 week trial of fluoxetine as first line and an 8-12

Figure 1. continued: Psychotherapy Stream



week trial of sertraline as second line, consistent with the NICE-CPG recommendations.

To incorporate MBC, “team reviews” are held every four weeks over a span of 20 weeks. “Team review” members include: most-responsible physician, an allied health professional, the patient and, caregiver(s). All members of the “team review” discuss the change-scores in measures of depressive symptoms (Mood and Feeling Questionnaire Child Version– MFQ-C) (Angold, Costello, Messer, & Pickles, 1995), general functioning (the Columbia Impairment Scale – CIS) (Bird, Shaffer, Fisher, & Gould, 1993) and family relationships (MFAD) and decide to continue or change the current treatment plan at the indicated decision points. Staff are provided with a checklist of essential items to cover during the “team reviews”. In addition to discussion around change-scores, another essential item is to conduct a safety assessment. If the psychiatrist’s impression is that the youth is at high risk of self-harm or suicide, the psychiatrist has the option to see the youth more frequently between “team reviews”.

Study Objectives and Hypotheses

Overall study objectives

In our future definitive trial, we wish to answer the following questions:

1. In adolescents with depression presenting to an outpatient clinic, is the use of an ICP more effective in reducing symptoms of depression as measured by the Childhood Depression Rating Scale -Revised (CDRS-R; Poznanski & Mokros, 1996) compared to treatment as usual (TAU) over a 20-week interval?
2. Do adolescents with depression undergoing the ICP have greater improvements in functioning as measured by the WHO-Disability Assessment Schedule 2.0 for Children and Youth (WHODAS 2.0 CY; Scorza et al., 2013) compared to TAU over a 20-week interval?
3. Do caregivers of adolescents with depression report a greater change in internalizing symptoms on the Childhood Behaviour Checklist (CBCL; Achenbach, 1991) when their adolescent undergoes the ICP, relative to TAU over a 20-week interval?
4. This pilot study aims to conduct the feasibility assessments needed to inform the design of a pragmatic

controlled clinical trial (Roland & Torgerson, 1998) that would be sufficiently powered to provide a definitive answer.

Feasibility Objectives of this Pilot Study

In this pilot study, we wish to address the following questions:

1. Is our research team able to recruit a sufficient number of participants at each of the two sites over a 21-month span?
2. Do a sufficient number of youth complete our comprehensive baseline measurement battery within a reasonable time frame?
3. Do clinicians using the ICP and MBC framework adhere to the model?
4. Are a sufficient proportion of youth engaging in multiple longitudinal outcome assessments so as to provide enough data for a valid analysis?

Corresponding hypotheses are as follows

1. We will be able to recruit 30 participants at each of the two hospital study sites over a 21-month interval.
2. Ninety-five percent of youth will complete the baseline measurement battery within a span of 2 hours.
3. Within the ICP arm, the clinicians will have, on average, 90% adherence to the model as per a specified checklist.
4. Eighty percent of the scheduled data points have a corresponding completed CDRS-R at study end.

Methods and Analysis

We have used the CONSORT 2010 extension for pilot and feasibility studies (Elridge et al, 2016) and the SPRIRIT reporting guidelines for clinical trial protocols (Chan et al, 2013) to guide the reporting below (see Appendices A&B for corresponding checklists).

Study Design

This study is a pilot parallel non-randomized controlled clinical trial. Participants aged 14-18 in the ICP and TAU arms will be recruited from outpatient adolescent psychiatric clinics within two Toronto academic hospital sites in Canada: the Centre for Addiction and Mental Health (CAMH) and Sunnybrook Health Sciences Centre (SHSC), respectively. The site of clinical presentation determines allocation with a ratio of 1:1. Randomization at the individual level is not possible as the intervention is implemented at the clinic level; that is, if one patient receives the ICP at one clinic it would affect the care of other patients at the same clinic as provider behaviour has been cued to follow the treatment pathway. Blinding participants will not be possible in this pilot project as treatment assignment based

on site is too overt for blinding. Due to resource limitations, there is only one RA to co-ordinate recruitment as well as rate outcomes; as such, blinding the rater is also not possible. The trial is registered at ClinicalTrials.gov (NCT03428555). It was funded by the Cundill Centre for Child and Youth Depression. Funders are not involved in the design, analysis, interpretation or decision to publish of this study.

Participants

Participant recruitment. Youth are most often referred to the clinics by family doctors and/or pediatricians in the surrounding community. Recruitment will take place over a 21-month period. Up to 30 youth participants will be recruited for the ICP and TAU study groups at each location for a total of 60 participants; as this is a pilot study, we are not anticipating to be adequately powered for hypothesis-testing of the efficacy of the ICP. It is estimated that CAMH and SHSC see more than 120 and 200 new adolescent patients with MDD-A per year, respectively. Achieving our recruitment goal over a 21-month span is highly likely. For each youth, the participation of one caregiver (ie. parent or guardian) in the study process will be encouraged, but not required. Across both hospitals, the following procedures will be used as appropriate and feasible: psychiatrists at the respective outpatient clinics will assess patients referred to them as usual; as per this assessment, patients whose primary diagnosis is a depressive disorder will be asked if they are willing to be contacted by a research analyst (RA); those who agree will proceed to the consent process with the RA. Next, a visit will then be scheduled at the respective site with the youth and, if applicable, caregiver for screening assessments. The Mood and Feelings Questionnaire (MFQ-C) (Angold, Costello, Messer, & Pickles, 1995) will be used initially to screen candidates; those who score ≥ 22 , will proceed a comprehensive assessment using the Diagnostic Interview for Affective Symptoms for Children (DIAS-C) (Merikangas et al., 2014). The DIAS-C will be used to determine if candidates meet inclusion and exclusion criteria, listed below. Those who do not meet the criteria or who do not agree to participate in the study will proceed with care outside of the framework of the study at their respective sites.

Inclusion and exclusion criteria. Adolescents will be included at screening if: (1) the youth is aged 14 to 18 years, inclusive (i.e. up until their 19th birthday); (2) the youth is presenting to one of the participating hospital outpatient clinics; (3) the psychiatrist's impression is that depressive symptoms are a priority treatment target; and, (4) the MFQ-C score is ≥ 22 . Subsequently, youth are assessed with the DIAS-C for further inclusion and exclusion criteria. Those meeting DSM-5 (American Psychiatric Association, 2013) criteria for major depressive disorder or persistent depressive disorder will be included. Exclusion criteria are the following: active psychotic symptoms (delusions,

Table 1. Summary of assessments

	Screen	Baseline	4 weeks	8 weeks	12 weeks	16 weeks	20 weeks
Eligibility:							
MFQ-C	X						
DIAS-C	X						
Treatment Timeline							
ICP		■	-----	-----	-----	-----	■
Treatment as usual		■	-----	-----	-----	-----	■
Sample Description Measures							
Demographics Form		X					
Treatment History Form		X					
Feasibility Outcomes							
Time to complete baseline measures		X					
ICP Adherence measure		X	X	X	X	X	X
Overall CARIBOU Study Outcome Measures							
CDRS-R		X	X	X	X	X	X
WHODAS		X	X	X	X	X	X
CBCL		X	X	X	X	X	X
Measurement-based Care Measures							
MFQ-C		X	X	X	X	X	X
YCIS		X	X	X	X	X	X
PCIS		X	X	X	X	X	X
MFAD-GF		X	Xa	Xa	Xa	Xa	Xa

X = assessment to be completed; ■ = beginning or end of treatment; a = only if baseline MFAD above established cut-off.

hallucinations or disorganized speech that are persistent and affecting function), bipolar disorder (I or II), moderate to severe substance use disorder, autism spectrum disorder or intellectual disability, moderate to severe eating disorder, imminent risk of suicide requiring immediate intervention such as hospitalization; inability to read and write in English; and inability to provide informed consent to the study for any reason. Participants must not currently be receiving care at another academic hospital and must be referred to CAMH or SHSC. Community agency support is permitted in both arms to deliver a component of treatment (for example, individual psychotherapy). There are no restrictions on medications being prescribed at entry into the trial. Information on the extent of community agency support and medication use will be collected by the research assistant and be considered in the interpretation of the analysis.

Consent and withdrawal from the study and compensation. Informed, signed consent will be obtained from all study participants by the research assistant at the intake visit prior to study enrollment. The caregiver will consent for his/her own respective participation. Participation will be voluntary and participants will be free to withdraw from the study at any time without affecting their treatment. Through chart

review, the age, sex, gender, socio-economic status and depression severity will be compared between those consenting to the study and those refusing consent to examine for ascertainment bias. Should participants, family members and/or clinicians observe any clear harm brought about by the intervention, the participant will be withdrawn from the study. Participants will be compensated in the form of gift cards due to the substantial time required for research assessment (50\$CAD for each of the initial and final assessments, 25\$CAD for other assessments).

Study Interventions

Integrated care pathway. The ICP that was developed at CAMH was summarized in the introduction.

Treatment as usual. The TAU condition consists of the standard outpatient treatment provided at SHSC. There is no structured protocol and no MBC. TAU typically entails assessment and treatment planning by a psychiatrist, and may include medication management, various types of psychotherapy and/or internal and external referrals to treatment and other services, guided by local service standards. For comparison purposes, a research assistant (RA) will record the frequency and type of interventions received in both

groups for each participant using chart review and interview with the clinician. The RA will also document if clinicians for the TAU group are using measurement-based care. This comparator group was chosen as it represents the status quo and is most relevant to the question of whether or not ICPs should be developed and implemented on a larger scale.

Application of interventions. Both arms will have treatment delivered by an interdisciplinary team including licensed psychiatrists, nurses and social workers at each site. Clinicians running CBT groups at CAMH have had the following training: (1) the social worker and registered nurse each had a 20-hour course, “CBT fundamentals”, through the Ontario Institute for Studies in Education at the University of Toronto; and, (2) the two psychiatrists received training through their respective residency training programs.

Potential adverse events. It is possible that overreliance on automated and algorithmic processes may lead to ignoring important information for care; however, other data on MBC does not support this assumption (Fortney et al., 2017). We will be capturing patients’, families’ and clinicians’ impressions of the ICP after they move through the pathway through focus groups and at clinicians meetings. We will be similarly this collecting data for the TAU arm. Adverse events - including psychiatric hospitalizations, suicide attempts or deaths by any reason - will be clearly documented at CAMH in an Adverse Event Log immediately upon notification and duly reported to the REB.

Measures

The schedule of assessments is described in Table 1. To describe the basic characteristics of the sample recruited into the study, customized demographic information and treatment history forms will be administered to the youth and family member. The psychometric properties of measures for eligibility, research outcomes and MBC outcomes are described in Appendix C (see Supplement).

Feasibility outcome measures. Each feasibility hypothesis has a corresponding operationalized definition for the outcome. Hypothesis 1: The number of recruited participants is the number of candidates who complete baseline measures. The time interval will start upon the enrollment of the first participant and will end at 21 months from that time. Hypothesis 2: The baseline measurement battery will be timed by the research assistant conducting the battery from the time of sitting down in the assessment room to the time of completion of the final questionnaire. The proportion of participants completing baseline measures in less than two hours will be calculated. Hypothesis 3: The RA will complete a “clinician adherence form”, which provides a “percentage adherence score” for each participant that goes through the ICP (See Appendix D in the Supplement). Checklists completed by physicians and chart reviews by the research assistant will be conducted to see that the decision points listed are discussed using the MBC framework.

We will calculate the mean adherence score. Hypothesis 4: The proportion of completed scheduled time points for the primary outcome will be calculated as “number of CDRS-R ratings actually completed” ÷ “number of ideal ratings” (total number of participants x six time points each).

In order to get qualitative information on the youth and caregiver experience of the ICP and TAU, focus group responses will be collected and synthesized by psychology post-doctoral fellows. We will conduct a thematic content analysis using NVivo software (“NVivo 10 [software program]. Version 10. QSR International; 2012.” 2014) to synthesize the responses in the focus groups. The results will be used to either: (1) justify continuing its use in its current form; or, (2) making modifications that are more amenable to youth engagement.

Data collection procedures. In each arm, the follow-up visits coincide with clinic appointments. Data from assessments will be directly collected into the electronic data capture system REDCap with SSL encryption (Harris et al., 2009).

Statistical Plan

As described above, the feasibility outcomes involve counts and proportions. The sample size of 30 participants per group was derived from an estimate of previous research recruitment rates within our clinic and was thought to be optimal; enough to test our feasibility hypotheses while waiting to see the results of this small sample prior to investing more resources in a definitive trial.

We will use mixed effects modeling in STATA v.13.1 to describe the course of depressive symptoms (primary outcome and dependent variable) between treatment arms (independent variable) over time. The equation that will be used is:

$$Y_{ij} = \beta_0 + X'_{ij}\beta_1 + b_i + \epsilon_{ij}$$

where Y_{ij} is the severity of depression (as represented by the CDRS-R score) for a given participant (i) and time point (j) and participant, β_0 is the intercept term, β_1 is the fixed effect representing the treatment effect, b_i represents the random participant effect and ϵ_{ij} represents measurement error (Fitzmaurice, Laird & Ware, 2011). The analysis will use an intention-to-treat approach; data from all enrolled participants in the study will be analyzed. As the effect of interest is impact of the ICP as a decision aide, rather than the impact of specific components of pathway, variation in which elements of the ICP are actually used will not be the focus of the primary mixed effects model; however, will be described to assist in interpreting the outcomes. Multiple imputation will be used to address any missing data in the analysis of clinical outcomes within this pilot study. At the end of this pilot study, we will examine power for the definitive trial for a more precise estimate of the sample size needed in the larger definitive trial using Monte

Carlo simulation mixed effects modeling to account for site clustering.

Ethical Considerations

This study will be conducted according to the guidelines established by Good Clinical Practice (Food and Drug Administration, www.fda.gov/oc/gcp/regulations.html). The Research Ethics Board (REB) at CAMH has approved this study (reference number: 079_2017_01); and at SHSC (reference number 429_2017). Any protocol modifications will be communicated to all relevant parties (co-investigators, REBs, Clinicaltrials.gov, immediately as they are made).

Discussion

Should all four of our feasibility hypotheses be met, pursuing a future definitive cluster RCT is warranted. There is the risk that one or two of our feasibility targets will fall short of our hypothesized thresholds; this would not necessarily indicate the abandonment of the ICP research program, rather some minor adjustments may be required. If few or none of our feasibility hypotheses are met, our group will concede that a different type of design may be required to answer our overall research questions.

As this is a pilot trial, the limitations are important to acknowledge. Firstly, there is no randomization process. From this trial we will not be able to determine if randomization is tolerated. We had considered using statistical methods to adjust our mixed effects linear modeling to account for non-randomization (e.g. by using propensity score analysis); however, the sample size is too small to conduct any meaningful adjustment properly. Second, there is no blinding of the raters. We have attempted to decrease expectancy effects of participants by framing the ICP as “highly structured care” and the TAU as “flexible care” – so as not to frame one treatment as more desirable than the other; however, as with most psychosocial interventions, participants are aware of the treatment they are receiving. Our current funding scheme allows for only one research assistant who will be well aware of the treatment arm to which participants have been allocated. This blinding process in the future definitive trial (where a greater number of raters may be available) will have not been tested. Third, feasibility outcomes for participants at CAMH may be inflated due to this site being the “home base” for many members of the research team and the high engagement of clinical staff that will deliver the ICP. This finding may not generalize to other sites in a cluster RCT where there may be less investment in the program aims; whether other sites are community-based agencies or hospital settings. To address this issue, we anticipate that using the spirit of integrated knowledge translation and collaborative treatment development with future research sites will optimize mental investment in the overall project. Next, there were no adherence measures for the CBT group implementation; as this treatment is being

newly implemented at CAMH, our group wanted to allow for staff to develop experience with the treatment prior to testing adherence; we anticipate testing adherence rigorously in the definitive trial. Lastly, the window of observation in the research trial is only four weeks longer than the 16-weeks of CBT group, where the full effects of the group intervention may not be captured. The 20-week window was thought to be sufficient for our pilot feasibility outcomes. In the definitive trial, we intend to have a longer window of observation.

Should our overall ICP research program hypotheses be met, efforts to scale up the ICP development process to other treatment centers would be justified. Through this implementation strategy, there is great potential to unlock the knowledge acquired thus far on the optimal management of adolescent depression, and thus improving the current and future lives of many young people struggling with this debilitating disorder.

Acknowledgments / Conflicts of Interest:

We wish to thank Kirsten Neprily and Michelle Li (research analysts) for their work in helping to prepare this manuscript. Drs. Courtney, Cheung, and Strauss are paid by the Cundill Centre for their involvement in the study. The funding agency had no role in the study design; collection, management, analysis or interpretation of the data; writing of the report or decision to submit for publication. Peter Szatmari receives royalties from Guilford Press; research funding from CIHR; and support from the Patsy and Jamie Anderson Chair in Child and Youth Mental Health. Joanna Henderson receives research funding from CIHR; and support from the Margaret and Wallace McCain Family Foundation.

References

- Achenbach, T. M. (1991). Manual for the Child Behavior checklist/4-18 and 1991 Profile. Burlington, VT: University of Vermont, Department of Psychiatry.
- Akister, J., & Stevenson-Hinde, J. (1991). Identifying families at risk: Exploring the potential of the McMaster Family Assessment Device. *Journal of Family Therapy*, *13*(4), 411-421. doi: <http://doi.org/10.1046/j.1991.00437.x>
- Angold, A., Costello, E. J., Messer, S. C., & Pickles, A. (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research*, *5*(4), 237-249.
- Asamow, J. (2006). Effectiveness of a Quality Improvement Intervention for Adolescent Depression in Primary Care Clinics: A Randomized Controlled Trial. *JAMA*, *19*;293(3), 311-319. doi: [10.1001/jama.293.3.311](http://doi.org/10.1001/jama.293.3.311)
- American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorder (5th ed.). Arlington, VA: American Psychiatric Publishing.
- Bickman, L., Kelley, S. D., Breda, C., Regina de Andrade, A., & Riemer, M. (2011). Effects of Routine Feedback to Clinicians on Mental Health Outcomes of Youths: Results of a Randomized Trial. *Psychiatric Services*, *62*(12), 1423-1429. doi: <http://doi.org/10.1176/appi.ps.002052011>

- Bird, H. R., Andrews, H., Schwab-Stone, M., Goodman, S., Dulcan, M., Richters, J.,...Hoven, C. (1996). Global measures of impairment for epidemiologic and clinical use with children and adolescents. *International Journal of Methods in Psychiatric Research*, 6(3), 295-307.
- Bird, H. R., Shaffer, D., Fisher, P., & Gould, M. S. (1993). The Columbia Impairment Scale (CIS): Pilot findings on a measure of global impairment for children and adolescents. *International Journal of Methods in Psychiatric Research*, 3(3), 167-176.
- Brent, D., Emslie, G., Clarke, G., Wagner, K. D., Asarnow, J. R., Keller, M.,...Zelazny, J. (2008). Switching to another SSRI or to venlafaxine with or without cognitive behavioral therapy for adolescents with SSRI-resistant depression: The TORDIA randomized controlled trial. *JAMA*, 299(8), 901-913. doi: 10.1001/jama.299.8.901.
- Campbell, H., Hotchkiss, R., Bradshaw, N., & Porteous, M. (1998). Integrated care pathways. *BMJ (Clinical Research Ed.)*, 316(7125), 133-137. doi: http://doi.org/10.1136/bmj.316.7125.133
- Chan, A. W., Tetzlaff, J. M., Altman, D. G., Laupacis, A., Gotzsche, P. C., Krleza-Jerić, K.,...Doré, C. J. (2013). SPIRIT 2013 statement: Defining standard protocol items for clinical trials. *Annals of Internal Medicine*, 158(3), 200-207.
- Cheung, A. H., & Dewa, C. S. (2006). Canadian community health survey: Major depressive disorder and suicidality in adolescents. *Health Policy*, 2(2), 76-89.
- Clarke, G. N., Lewinsohn, P. M., Hops, H., & Grossen, B. (1990). Adolescent coping with depression course: Leader's manual for adolescent groups. Kaiser Permanente. Accessed on April 14th, 2019 from: https://research.kpchr.org/Research/Research-Areas/Mental-Health/Youth-Depression-Programs#Downloads.
- Daviss, W. B., Birmaher, B., Melhem, N. A., Axelson, D. A., Michaels, S. M., & Brent, D. A. (2006). Criterion validity of the Mood and Feelings Questionnaire for depressive episodes in clinic and non-clinic subjects. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 47(9), 927-934. doi: http://doi.org/10.1111/j.1469-7610.2006.01646.x
- Eldridge, S. M., Chan, C. L., Campbell, M. J., Bond, C. M., Hopewell, S., Thabane, L., & Lancaster, G. A. (2016). CONSORT 2010 statement: Extension to randomised pilot and feasibility trials. *Pilot and feasibility studies*, 2(1), 64.
- Emslie, G. J., Hughes, C. W., Crismon, M. L., Lopez, M., Pliszka, S., Toprac, M. G., & Boemer, C. (2004). A feasibility study of the childhood depression medication algorithm: The Texas Children's Medication Algorithm Project (CMAP). *Journal of the American Academy of Child & Adolescent Psychiatry*, 43(5), 519-527. doi: 10.1097/00004583-200405000-00005
- Epstein, N. B., Baldwin, L. M., & Bishop, D. S. (1983). The McMaster family assessment device. *Journal of Marital and Family Therapy*, 9(2), 171-180. doi: https://doi.org/10.1111/j.1752-0606.1983.tb01497.x
- Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2011). Applied longitudinal analysis. John Wiley & Sons; Hoboken, New Jersey. p190.
- Fortney, J. C., Unützer, J., Wrenn, G., Pyne, J. M., Smith, G. R., Schoenbaum, M., & Harbin, H. T. (2017). A Tipping Point for Measurement-Based Care. *Psychiatric Services*, 68(2), 179-188. doi: 10.1176/appi.ps.201500439
- Girlanda, F., Fiedler, I., Becker, T., Barbui, C., & Koesters, M. (2017). The evidence-practice gap in specialist mental healthcare: Systematic review and meta-analysis of guideline implementation studies. *British Journal of Psychiatry*, 210(1), 24-30. doi: 10.1192/bjp.bp.115.179093
- Goodyer, I. M., Herbert, J., Secher, S. M., & Pearson, J. (1997). Short-term outcome of major depression: I. Comorbidity and severity at presentation as predictors of persistent disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36(2), 179-187. doi: http://doi.org/10.1097/00004583-199702000-00008
- Gore, F. M., Bloem, P. J. N., Patton, G. C., Ferguson, J., Joseph, V., Coffey, C.,...Mathers, C. D. (2011). Global burden of disease in young people aged 10-24 years: A systematic analysis. *The Lancet*, 377(9783), 2093-2102. doi: 10.1016/S0140-6736(11)60512-6.
- Graham R., Mancher, M., Wolman, D. M., Greenfield, S., & Steinberg, E. (2011). Clinical Practice Guidelines We Can Trust. Washington, DC: National Academies Press.
- Gunlicks-Stoessel, M., Mufson, L., Bernstein, G., Westervelt, A., Reigstad, K., Klimes-Dougan, B.,...Vock, D. (2018). Critical Decision Points for Augmenting Interpersonal Psychotherapy for Depressed Adolescents: A Pilot SMART. *Journal of the American Academy of Child & Adolescent Psychiatry*. doi: http://doi.org/10.1016/j.jaac.2018.06.032
- Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, 42(2), 377-381. doi: 10.1016/j.jbi.2008.08.010.
- Hepner, K. A., Rowe, M., Rost, K., Hickey, S. C., Sherbourne, C. D., Ford, D. E.,...Rubenstein, L. V. (2007). The effect of adherence to practice guidelines on depression outcomes. *Annals of Internal Medicine*, 147(5), 320-329. doi: http://doi.org/10.7326/0003-4819-147-5-200709040-00007
- Kabacoff, R. I., Miller, I. W., Bishop, D. S., Epstein, N. B., & Keitner, G. I. (1990). A psychometric study of the McMaster Family Assessment Device in psychiatric, medical, and nonclinical samples. *Journal of Family Psychology*, 3(4), 431-439. doi: http://dx.doi.org/10.1037/h0080547
- Kane, J. M., Robinson, D. G., Schooler, N. R., Mueser, K. T., Penn, D. L., Rosenheck, R. A.,...Marcy, P. (2015). Comprehensive versus usual community care for first-episode psychosis: 2-year outcomes from the NIMH RAISE early treatment program. *American Journal of Psychiatry*, 173(4), 362-372.
- Kent, L., Vostanis, P., & Feehan, C. (1997). Detection of Major and Minor Depression in Children and Adolescents: Evaluation of the Mood and Feelings Questionnaire. *Journal of Child Psychology and Psychiatry*, 38(5), 565-573. doi: http://dx.doi.org/10.1111/j.1469-7610.1997.tb01543.x
- Lambert, M. J., Whipple, J. L., Vermeersch, D. A., Smart, D. W., Hawkins, E. J., Nielsan, S. L., & Goates, M. (2002). Enhancing psychotherapy outcomes via providing feedback on client progress: A replication. *Clinical Psychology and Psychotherapy*, 9(2), 91-103. doi: http://doi.org/10.1002/cpp.324
- March, J., Silva, S., Petrycki, S., Curry, J., Wells, K., Fairbank, J.,... Treatment for Adolescents With Depression Study, T. (2004). Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression: Treatment for Adolescents With Depression Study (TADS) randomized controlled trial. *JAMA*, 292(7), 807-820. doi:10.1001/jama.292.7.807
- Mayes, T. L., Bernstein, I. H., Haley, C. L., Kennard, B. D., & Emslie, G. J. (2010). Psychometric properties of the Children's Depression Rating Scale-Revised in adolescents. *Journal of Child and Adolescent Psychopharmacology*, 20(6), 513-516. doi: 10.1089/cap.2010.0063.
- McLennan, J. D., Wathen, C. N., MacMillan, H. L., & Lavis, J. N. (2006). Research-Practice Gaps in Child Mental Health. *Journal of the American Academy of Child & Adolescent Psychiatry*, 45(6), 658-665. doi: http://doi.org/10.1097/01.CHI.0000215153.99517.80
- Merikangas, K. R., Cui, L., Heaton, L., Nakamura, E., Roca, C., Ding, J.,...Angst, J. (2014). Independence of familial transmission of mania and depression: Results of the NIMH family study of affective spectrum disorders. *Molecular Psychiatry*, 19(2), 214-219. doi: http://doi.org/10.1038/mp.2013.116
- Neufeld, S. A. S., Dunn, V. J., Jones, P. B., Croudace, T. J., & Goodyer, I. M. (2017). Reduction in adolescent depression after contact with mental health services: A longitudinal cohort study in the UK.

- The Lancet Psychiatry*, 4(2), 120-127. doi: [http://doi.org/10.1016/S2215-0366\(17\)30002-0](http://doi.org/10.1016/S2215-0366(17)30002-0)
- NICE. (2015). Depression in children and young people: Clinical Guideline 28, (May), 1–271. Retrieved November 20, 2018, from <https://www.nice.org.uk/guidance/cg28>
- NVivo 10 [software program]. Version 10. QSR International; 2012. (2014). Ontario, Canada. Retrieved from <https://www.qsrinternational.com/nvivo/enabling-research/thematic-analysis>
- Poznanski, E. O., & Mokros, H. B. (1996). Children's Depression Rating Scale, revised (CDRS-R). Los Angeles, CA: Western Psychological Services.
- Renaud, J., Berlim, M. T., McGirr, A., Tousignant, M., & Turecki, G. (2008). Current psychiatric morbidity, aggression/impulsivity, and personality dimensions in child and adolescent suicide: A case-control study. *Journal of Affective Disorders*, 105(1-3), 221-228.
- Roland, M., & Torgerson, D. J. (1998). Understanding controlled trials: What are pragmatic trials? *BMJ: British Medical Journal*, 316(7127), 285. doi: <https://doi.org/10.1136/bmj.316.7127.285>
- Scorza, P., Stevenson, A., Canino, G., Mushashi, C., Kanyanganzi, F., Munyanah, M., & Betancourt, T. (2013). Validation of the "World Health Organization disability assessment schedule for children, WHODAS-child" in Rwanda. *PLOS One*, 8(3), e57725. doi: <https://doi.org/10.1371/journal.pone.0057725>
- Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology*, 78(3), 298-311. doi: <http://doi.org/10.1037/a0019247>
- Wilkinson, P., Kelvin, R., Roberts, C., Dubicka, B., & Goodyer, I. (2011). Clinical and psychosocial predictors of suicide attempts and nonsuicidal self-injury in the Adolescent Depression Antidepressants and Psychotherapy Trial (ADAPT). *American Journal of Psychiatry*, 168(5), 495-501. doi: <http://doi.org/10.1176/appi.ajp.2010.10050718>
- Wood, A., Kroll, L., Moore, A., & Harrington, R. (1995). Properties of the mood and feelings questionnaire in adolescent psychiatric outpatients: A research note. *Journal of Child*

Appendix A: CONSORT 2010 checklist of information to include when reporting a pilot or feasibility randomized trial in a journal or conference abstract

Item	Description	Reported on line number
Title	Identification of study as randomised pilot or feasibility trial	Page 115
Authors *	Contact details for the corresponding author	Title page
Trial design	Description of pilot trial design (eg, parallel, cluster)	Page 120
Methods		
Participants	Eligibility criteria for participants and the settings where the pilot trial was conducted	Pages 120-121
Interventions	Interventions intended for each group	Pages 118-122
Objective	Specific objectives of the pilot trial	Pages 119-120
Outcome	Prespecified assessment or measurement to address the pilot trial objectives**	Pages 122
Randomization	How participants were allocated to interventions	Not randomized; allocation described on page 120 .
Blinding (masking)	Whether or not participants, care givers, and those assessing the outcomes were blinded to group assignment	Lack of blinding described on page 10.
Results		
Numbers randomized	Number of participants screened and randomised to each group for the pilot trial objectives**	Planned number page 120.
Recruitment	Trial status†	
Numbers analysed	Number of participants analysed in each group for the pilot objectives**	«Intention to treat» - pg 122.
Outcome	Results for the pilot objectives, including any expressions of uncertainty**	To be determined
Harms	Important adverse events or side effects	Page 122
Conclusions	General interpretation of the results of pilot trial and their implications for the future definitive trial	To be determined
Trial registration	Registration number for pilot trial and name of trial register	Page 120
Funding	Source of funding for pilot trial	Page 120
*this item is specific to conference abstracts		
**Space permitting, list all pilot trial objectives and give the results for each. Otherwise, report those that are a priori agreed as the most important to the decision to proceed with the future definitive RCT.		
†For conference abstracts.		

Appendix B:			
SPIRIT 2013 Checklist: Recommended items to address in a clinical trial protocol and related documents*			
Section/item	Item No	Description	Addressed on page number
Administrative information			
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	Page 115
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry	Page 120
	2b	All items from the World Health Organization Trial Registration Data Set	Separate Checklist
Protocol version	3	Date and version identifier	N/A
Funding	4	Sources and types of financial, material, and other support	Page 120
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors	Page 115
	5b	Name and contact information for the trial sponsor	N/A
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	Page 120
	5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	N/A
Introduction			
Background and rationale	6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	Page 116-118
	6b	Explanation for choice of comparators	Page 121-122
Objectives	7	Specific objectives or hypotheses	Page 119-120
Trial design	8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, noninferiority, exploratory)	Page 120
Methods: Participants, interventions, and outcomes			
Study setting	9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained	Page 120
Eligibility criteria	10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	Page 120-121
Interventions	11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	Page 118-122
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving/worsening disease) ¹¹	Page 121
	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return, laboratory tests)	Page 122 and Appendix D
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	Page 121

continued on page 128

Appendix B: continued			
Section/item	Item No	Description	Addressed on page number
Outcomes	12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	Page 122, Table 1 and appendix C
Participant timeline	13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	Table 1
Sample size	14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	Page 122
Recruitment	15	Strategies for achieving adequate participant enrolment to reach target sample size	Page 122
Methods: Assignment of interventions (for controlled trials)			
Allocation:			
Sequence generation	16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	N/A
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	N/A
Implementation	16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	N/A
Blinding (masking)	17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	N/A
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	N/A
Methods: Data collection, management, and analysis			
Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol	Page 122-Appendix C
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	Page 121
Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	Page 122
Statistical methods	20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	Page 122
	20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	N/A
	20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	Page 122

continued on page 129

Appendix B: continued			
Section/item	Item No	Description	Addressed on page number
Methods: Monitoring			
Data monitoring	21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	N/A – not randomized or blinded
	21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	N/A – not randomized or blinded
Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	Page 122
Auditing	23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	N/A – brief feasibility trial
Ethics and dissemination			
Research ethics approval	24	Plans for seeking research ethics committee/institutional review board (REC/IRB) approval	Page 123
Protocol amendments	25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC/IRBs, trial participants, trial registries, journals, regulators)	Page 123
Consent or assent	26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	Page 120
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	N/A
Confidentiality	27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	REB protocol only
Declaration of interests	28	Financial and other competing interests for principal investigators for the overall trial and each study site	Page 120
Access to data	29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	In REB protocol only
Ancillary and post-trial care	30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	Included in REB protocol only
Dissemination policy	31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	In REB protocol only
	31b	Authorship eligibility guidelines and any intended use of professional writers	N/A
	31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	N/A
Appendices			
Informed consent materials	32	Model consent form and other related documentation given to participants and authorised surrogates	In REB protocol only.
Biological specimens	33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable	N/A

Appendix C: Psychometric Properties of Measures

Eligibility assessments. The Mood and Feelings Questionnaire – Child version (MFQ-C) is a 33 item self-report measure completed by the adolescent regarding depressive symptoms experienced over the prior 2 weeks (Angold, Costello, Messer, & Pickles, 1995). It is scored on a 3 point Likert scale (0-2) providing a range of scores from 0 to 66; with higher scores representing more severe depression. It has been found to have a high discriminatory ability in late adolescents as well (Daviss et al., 2006; Kent, Vostanis, & Feehan, 1997). Internal consistency has been found to be good ($\alpha=0.92-0.94$) (Neufeld, Dunn, Jones, Croudace, & Goodyer, 2017) and good test-retest reliability (Pearson's $r=0.78$) (Wood, Kroll, Moore, & Harrington, 1995). A cut-off of a score of 22 (Goodyer, Herbert, Secher, & Pearson, 1997) was found to be optimal for distinguishing clinical from non-clinical populations. The MFQ-C has been found to be sensitive to change (Goodyer et al., 1997). The MFQ-C will be used as both a screening measure for inclusion as well as a MBC measure within the ICP. There is no consensus on a cut-off to distinguish mild from moderate depression (which is required for a decision point on the ICP). We have mapped items onto the DSM-5 using the strategy listed in Appendix E (see below) in order to account for this.

The Diagnostic Interview for Affective Symptoms for Children (DIAS-C) is a comprehensive systematic diagnostic tool. It ascertains diagnostic criteria for current and lifetime DSM-IV-TR disorders. It has been used in large NIMH studies (Merikangas et al., 2014) and is used across other studies within CAMH Child, Youth and Family Services. This will be used to establish inclusion and exclusion criteria as well as describe the sample.

Overall research program outcome measures. These outcomes will not be disclosed to the clinicians and the families to remain independent of the MBC feedback.

The overall research program objective is to see whether the ICP is more effective in reducing symptoms of depression over the 20 weeks. To test this hypothesis, the RA will administer the Childhood Depression Rating Scale-revised score (CDRS-R) (Poznanski & Mokros, 1996) with youth participants every 4 weeks until the 20-week interval is complete. The CDRS-R is a 17-item measure rated by an evaluator after a semi-structured interview with the adolescent relating to symptoms of depression over the past 2 weeks. Mayes, Bernstein, Haley, Kennard, and Emslie (2010) evaluated its psychometric properties in a sample of adolescents receiving fluoxetine over time. They found good internal consistency ($\alpha=0.74-0.92$) and correlated significantly with measures of global severity of illness ($r=0.80-0.93$, $p<0.01$), functioning ($r=0.52-0.77$, $p<0.01$) and a diagnosis of major depressive disorder on the K-SADS-PL ($r=0.64$, $p<0.01$). Change in the score was also correlated with measures of global improvement with treatment ($r=0.83$, $p<0.01$). The CDRS-R has been used as the primary outcome in major RCTs for treatment of MDD-A (Brent et al., 2008; March et al., 2004).

Our secondary research program objectives are to assess clinical improvement in functioning and overall psychopathology in the ICP relative to TAU. This will be measured using the WHO Disability Assessment Schedule 2.0 for Children and Youth scores (WHODAS-2.0-CY) (Scorza et al., 2013) administered by the RA. The WHODAS-2.0-CY is a 36-item measure administered by an evaluator to assess function including domains regarding: understanding and communication, getting around, self-care, getting along with people, life activities, going to school and participation in society. Items are scored on a 5 point Likert scale (0-4) providing a range of overall scores from 0 to 100%; with higher scores representing impairment in functioning. It has been studied in adolescent populations (Scorza et al., 2013).

Childhood Behaviour Checklist (CBCL) is to be completed by a caregiver (Achenbach, 1991). The CBCL provides information from the caregiver. The Child Behaviour Checklist – Parent Report form (CBCL) (Achenbach, 1991) is a 118 item measure that is rated by caregivers assessing the child's behaviour. Each item is scored on a 3 point Likert scale (0-2). We are interesting in the total subscale scores representing internalizing problems, which combines anxious/depressed, withdrawn-depressed, and somatic complaints scores. The CBCL is a widely used measure with known population norms. One-week test-retest reliability was found to be 0.80-0.94 (Achenbach, 1991). Internal consistency is reported to be high; inter-rater reliability

(eg. between two caregivers) was found to be moderate to high (Achenbach, 1991).

The RA will be trained in all the assessment scales by the PI and Master's level staff in the research team at CAMH. Inter-rater agreement between the PI and RA will be described.

Measurement-based care measures. The results of these measures are discussed with patient, caregiver and clinician as part of the intervention: The MFQ-C is already described in the section on eligibility measures. This measure was chosen for the MBC as it is available at no cost, takes minimal time to complete and has established psychometric properties, making our protocol usable by other clinics should it be found to be effective.

The Columbia Impairment Scale (Bird et al., 1993) is a 13-item questionnaire, with likert scale of 0-4 for each item (range 0-52). There is a parent-rated version (PCIS) and a youth-rated version (YCIS). It asks questions regarding relationships to family members and peers, assesses engagement in hobbies and engagement in school "within the past week or two". It has been shown to have high internal consistency, excellent test-retest reliability and good correlation to other established measures of functioning (Bird et al., 1996). A cut-off score of ≥ 16 has been found to be associated with clinically significant impairment (Bird et al., 1996). This measure was chosen as it is available at no cost, takes minimal time to complete and has established psychometric properties.

The McMaster Family Assessment Device (Epstein et al., 1983) –General Functioning Subscale (MFAD-GF) is a 12 item questionnaire that is completed by all available family members above the age of 12. It takes less than 5 minutes to complete. Scores are averaged across family members. There is an established procedure to use the score to determine which family requires intervention (Akister & Stevenson-Hinde, 1991). The internal consistency has been reported to be good for the specific domain subscales ($\alpha=0.83-0.86$) in a clinical sample (Kabacoff, Miller, Bishop, Epstein, & Keitner, 1990). This measure was chosen for its known psychometric properties and availability for use free of charge. A score of 2.0 or more represents that clinical intervention is warranted (Akister & Stevenson-Hinde, 1991). Baseline MFAD-GF scores were significantly associated with suicide attempts in an RCT for treatment of adolescent depression (Wilkinson, Kelvin, Roberts, Dubicka, & Goodyer, 2011).

Appendix D: Integrated Care Pathway Adherence Measure			
Component	Criteria	Applicable?	Done?
Psychoeducation	<input type="checkbox"/> Offered "Mood Foundations" group	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No
Psychotherapy Group	<input type="checkbox"/> Offered "Group CBT" or "4-session individual CBT"	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No
Caregiver Group	<input type="checkbox"/> If MFAD positive, offered caregiver group.	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No
Medication	<input type="checkbox"/> If no previous medication trial and moderate-severe depression, fluoxetine offered as first-line.	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No
Medication	<input type="checkbox"/> If failed fluoxetine, sertraline offered as second-line.	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No
Medication	<input type="checkbox"/> If tolerated, medication allowed to continue until "team review corresponding to 8 weeks since medication initiation" even if no response.	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No
Medication	<input type="checkbox"/> If no response at "team review corresponding to 12 weeks since medication initiation", discussion around switching medication.	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No
Medication	<input type="checkbox"/> Not offered St. John's Wort, Venlafaxine or Tricyclic Antidepressant	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No
Team Review	<input type="checkbox"/> Offered every 4 weeks	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No
Team Review	Physician and client discussed MBC measure scores at:		
	<input type="checkbox"/> 4-week team review	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No
	<input type="checkbox"/> 8-week team review	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No
	<input type="checkbox"/> 12-week team review	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No
	<input type="checkbox"/> 16-week team review	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No
	<input type="checkbox"/> 20-week team review	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No
Totals		Total # applicable:	Total # completed:
Percentage adherence = # completed/ #applicable: _____			

Appendix E: Derivation of “Moderate to Severe Depression” Classification

There are no established cut-offs on the MFQ-C to determine “moderate to severe depression”. As such we are aiming follow the DSM-5 definition of moderate severity:

“Severity is based on the number of criterion symptoms, the severity of those symptoms, and the degree of functional disability.

Mild: Few, if any, symptoms in excess of those required to make the diagnosis are present, the intensity of the symptoms is distressing but manageable, and the symptoms result in minor impairment in social or occupational functioning.

Moderate: The number of symptoms, intensity of symptoms, and/or functional impairment are between those specified for “mild” and “severe.”

Severe: The number of symptoms is substantially in excess of that required to make the diagnosis, the intensity of the symptoms is seriously distressing and unmanageable, and the symptoms markedly interfere with social and occupational functioning.”

With this interpretation, there need to be at least 6 symptoms present and significant functional impairment to be considered moderate. The following mapping system

In order to approximate functional impairment, we will use the cut-off of 16 on the Youth Columbia Impairment Scale (either parent or child version).

If ≥ 6 symptoms are present and the CIS ≥ 16 , we this would represent “Moderate to Severe Depression” for the purposes of the pathway.

DSM-5 “A” criteria for Major Depressive Disorder	MFQ items that would count toward the DSM-5 symptoms (if any of the corresponding items scored at “2”)
1. Depressed mood or irritable: Depressed mood most of the day, nearly every day, as indicated by either subjective report (e.g., feels sad, empty, hopeless) or observation made by others (e.g., appears tearful). (Note: In children and adolescents, can be irritable mood.)	1. I felt miserable or unhappy 11. I felt grumpy and cross with my parents 14. I cried a lot 15. I thought there was nothing good for me in the future
2. Decreased interest or pleasure: Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day (as indicated by either subjective account or observation).	2. I didn't enjoy anything at all 20. I didn't want to see my friends 29. I didn't have any fun in school.
3. Significant weight change (5%) or change in appetite: Significant weight loss when not dieting or weight gain (e.g., a change of more than 5% of body weight in a month), or decrease or increase in appetite nearly every day. (Note: In children, consider failure to make expected weight gain.)	3. I was less hungry than usual 4. I ate more than usual.
4. Change in sleep: Insomnia or hypersomnia nearly every day.	32. I didn't sleep as well as I usually sleep 33. I slept more than usual.
5. Change in Activity: Psychomotor agitation or retardation nearly every day (observable by others, not merely subjective feelings of restlessness or being slowed down).	6. I was moving and walking more slowly than usual 7. I was very restless 13. I was talking slower than usual
6. Fatigue: Fatigue or loss of energy nearly every day.	5. I felt so tired I just sat around and did nothing.
7. Guilt/ Worthlessness: Feelings of worthlessness or excessive or inappropriate guilt (which may be delusional) nearly every day (not merely self-reproach or guilt about being sick).	8. I felt I was no good anymore 9. I blamed myself for things that weren't my fault 23 I hated myself 24. I felt I was a bad person 28. I thought no one really loved me 30. I thought I could never be as good as other kids 31. I did everything wrong.
8. Concentration: Diminished ability to think or concentrate, or indecisiveness, nearly every day (either by subjective account or as observed by others).	10. It was hard for me to make up my mind 21. I found it hard to think properly or concentrate.
9. Suicidality: Recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide.	16. I thought that life wasn't worth living 17. I thought about death or dying 18. I thought my family would be better off without me 19. I thought about killing myself