Effectiveness of an Integrated Care Pathway for Adolescent Depression: A Quasi -experimental Multi -site, Cluster Controlled Trial

NCT Number: NCT05142683

Protocol Identifying Number: 019/2021

Principal Investigator: Dr. Darren Courtney

Version Number: 5.2

Version Date: 03-Jul-2026

Sponsor:	Centre for Addiction and Mental Health Cundill Centre for Child & Youth Depression 80 Workman Way
Co-Principal Investigator s:	Darren Courtney (Principal Investigator) Centre for Addiction and Mental Health (Toronto, ON) Darren.Courtney@camh.ca Melanie Barwick (Co-Principal Investigator) The Hospital for Sick Children (Toronto, ON) melanie.barwick@sickkids.ca
Co-Investigators	Bahar Amani (Co -Investigator) Centre for Addiction and Mental Health (Toronto, ON) Andrea Greenblatt (Co -Investigator) The Hospital for Sick Children (Toronto, ON) Madison Aitken (Co -Investigator) Centre for Addiction and Mental Health (Toronto, ON) Karolin Krause (Co -Investigator) Centre for Addiction and Mental Health (Toronto, ON) Brendan Andrade (Co -Investigator) Centre for Addiction and Mental Health (Toronto, ON) Kathryn Bennett (Co -Investigator) McMaster University (Hamilton, ON) Kristin Cleverley (Co -Investigator) University of Toronto (Toronto, ON) Claire de Oliveira (Co -Investigator) Centre for Addictions and Mental Health (Toronto, ON) Lisa Hawke (Co -Investigator) Centre for Addictions and Mental Health (Toronto, ON) Jo Henderson (Co -Investigator) Centre for Addictions and Mental Health (Toronto, ON) Wei Wang (Co -Investigator) Centre for Addictions and Mental Health (Toronto, ON)

Priya Watson (Co -Investigator)
Centre for Addictions and Mental Health (Toronto, ON)

Amy Gajaria (Co -Investigator)
Centre for Addictions and Mental Health (Toronto, ON)

Mandi Newton (Co -Investigator)
University of Alberta (Calgary, AB)

Vasudha Gidugu (Co -Investigator)
The Hospital for Sick Children (Toronto, ON)

Carly Albaum (Co -Investigator)
Centre for Addictions and Mental Health (Toronto, ON)

Kiefer Cowie (Co -Investigator)
Centre for Addictions and Mental Health (Toronto, ON)

Victoria Iris Rae Anderson (Co -Investigator)
Centre for Addictions and Mental Health (Toronto, ON)

Jia Hong Dai (Co -Investigator)
Centre for Addictions and Mental Health (Toronto, ON)

Stephanie Ameis (Co -Investigator)
Centre for Addictions and Mental Health (Toronto, ON)

Peter Szatmari (Co -Investigator)
Centre for Addictions and Mental Health (Toronto, ON)

Funders:

Canadian Institutes of Health Research (CIHR)
160 Elgin Stret, 9th Floor
Address Locator 4809A
Ottawa, Ontario, K1A 0W9

Cundill Centre for Child & Youth Mental Health, CAMH
80 Workman Way, 6th Floor
Toronto, ON M6J 1H4

Revision History:

Version #	Version Date	Brief Summary of Changes
2.7	03-Jun-2021	• New
2.9	01-Nov-2021	• Addition of Appendices K and J. • Revision of Figure 1.
3.0	01-Jan-2022	• Revised section 4.3 and 12.3, and Fig. 2. • Updated co-investigators list. • Revised Appendix F. and added 8 new measures.
3.1	02-Apr-2022	• Revised Appendix B. Stepped Wedge Design Chart. • Revised Appendix F. Revised schedule of measures. • Removed/added measure. • Updated co-investigators list. • Removed Autism Spectrum Disorder as exclusion criterion.
3.3	15-Jun-2023	• The CARIBOU-2 Study title, formally known as *"Effectiveness of an Integrated Care Pathway for* Adolescent Depression: A Multi-site Stepped-Wedge, Cluster-*Randomized Controlled Trial" has changed to "Effectiveness of an Integrated Care* Pathway for Adolescent Depression: A Quasi-experimental, Multi-site, Cluster-*Controlled Trial".* • Made important changes to study design from a stepped-wedge design to a quasi-experimental, multi-site cluster controlled clinical trial design. • Sample size revised from 648 youth to 300 youth. • End of study date revised to February 2029. • Updated co-investigators list. • Revised measures.
3.5	13-Mar-2026	• *Changed site number from 6 to a range of "6-8."* • Revision of measures. • Use of EverythingCard for compensation. • Addition of swag for participants. • Revision to youth qualitative interview methods.
3.7	30-Jul-2025	• New clinician demographic form. • Revision to caregiver qualitative interview method. • Updated co-investigators list.
3.9	31-Dec-2025	• Change in collaborator contact information for the CAMH Youth Engagement Initiative. • Revised section 8.2.3.

4.0	07-Jan-2026	• Revised section 8.2.3.2 • Updated co-investigator list.
5.0	29-Apr-2026	• Included distinction between clinical trial sites and implementation-only sites. • Updated co-investigator list.
5.2	03-Jul -2026	• Update protocol template according to new GCP R3 guidelines. • Included E-Consent Framework for site staff ICF for implementation-only sites

Table of Contents

STATEMENT OF COMPLIANCE 9

CLINICAL TRIAL SUMMARY 10

LIST OF ABBREVIATIONS 11

1 – ABSTRACT 12

2 - INTRODUCTION 13
- 2.1 Background
- 2.2 CARIBOU-2 Intervention, in brief
- 2.3 Literature Review

3 - STUDY OBJECTIVES, RESEARCH QUESTIONS AND HYPOTHESES 15
- 3.1 Study Objectives:
- 3.2 Research Questions and Hypotheses

4 - METHODS: DESIGN, IMPLEMENTATION PROCESS AND RECRUITMENT 17
- 4.1 Study Design:
- 4.2 Implementation Process (Quality Implementation Framework)
- 4.3 Primary Participant Recruitment (Youth)
- 4.4 Secondary Participant Recruitment (Caregivers)
- 4.5 Tertiary Participant Recruitment
- 4.6 Anticipated sample size
- 4.7 Participant Retention and Withdrawal

5 - METHODS: INTERVENTIONS 22
- 5.1 Treatment as Usual
- 5.2 CARIBOU-2 Intervention, In detail
  - 5.2.1 CARIBOU-2 Intervention Core Components
  - 5.2.2 Acceptable adaptations
  - 5.2.3 Implementation Strategies used for CARIBOU-2 (ERIC taxonomy)
  - 5.2.4 CARIBOU-2 Intervention Measurement-Based Care measures

6 - METHODS: YOUTH PARTICIPANT RESEARCH MEASURES 27
- 6.1 Baseline measures
- 6.2 Primary outcome measure
- 6.3 Secondary clinical outcome measures
- 6.4 Qualitative information
- 6.5 Youth informant measures collected as part of sub study examining psychometric properties of MFQ, CADLIS-Y
- 6.6 Timing of Youth Participant Measures

7 - METHODS: CAREGIVER PARTICIPANT RESEARCH MEASURES 32
- 7.1 Baseline measures
- 7.2 Secondary clinical outcome measures
- 7.3 Qualitative information

- 7.4 Caregiver informant measures collected as part of sub study examining psychometric properties of MFQ, CADLIS
- 7.5 Timing of Youth Participant Measures

8 - METHODS: IMPLEMENTATION RESEARCH MEASURES 35
- 8.1 Feasibility
- 8.2 Fidelity
  - 8.2.1 Fidelity to the Implementation Process:
  - 8.2.2 Fidelity to the Overall CARIBOU-2 intervention:
  - 8.2.3 Fidelity to Components of the ICP:
  - 8.2.3.1 Fidelity to the Mood Foundations Session, Medication, and Caregiver group
  - 8.2.3.2 Fidelity to CBT
  - 8.2.3.3 Fidelity to BPI
- 8.3 Economic Evaluation:
- 8.4 Acceptability, barriers, and facilitators of implementation:
- 8.5 Timing of Implementation Outcome Capture:

9 - METHODS: STATISTICAL PLAN 37
- 9.1 Clinical Effectiveness analysis
- 9.2 Power Calculation
- 9.3 Implementation outcomes analysis

10 - PROJECT TIMELINE 39

11 - RESEARCH TEAM 39
- 11.1 Steering Committee
- 11.2 Others Co-investigators and Collaborators
- 11.3 Knowledge users
- 11.4 International expert guidance

12 - ETHICAL, CONFIDENTIALITY AND SAFETY CONSIDERATIONS 41
- 12.1 Potential adverse events
- 12.2 Ethics, REB Approvals and Legal Contracts
- 12.3 Data Management and Sharing
- 12.4 Data Safety Monitoring Board

13 - KNOWLEDGE TRANSLATION PLAN 42
- 13.1 Integrated Knowledge Translation
- 13.2 Post-study Knowledge Translation

14 - RELEVANCE AND IMPACT 43
- 14.1 Service Development
- 14.2 Informing Future trials

15 – STUDY MONITORING, AUDITING, AND INSPECTING 43
- 15.1 Critical to Quality (CtQ) Factors
- 15.2 Study Monitoring Plan

16 – CLINICAL TRIAL FINANCES 44

17 – REFERENCES 45

Appendices

- Appendix A. Grant Approval
- Appendix B. Plans, Tables, Charts, Diagrams
- Appendix C. Site Selection
- Appendix D. Consent Forms
- Appendix E. SWPs
- Appendix F. Measures
- Appendix G. Scripts
- Appendix H. DSMB
- Appendix I. Engagement and Fidelity Checklists
- Appendix J. Protocols for Psychometric Sub Studies
- Appendix K. Recruitment Materials
- Appendix L. Fully Executed Contracts

STATEMENT OF COMPLIANC E

This clinical trial will be carried out in accordance with the following:

- International Council for Harmonisation Good Clinical Practice (ICH GCP)
- Tri-Council Policy Statement 2022 (TCPS 2)
- Personal Health Information Protection Act (PHIPA), 2004; Chapter 3 Schedule A (PHIPA) and applicable regulations
- Institutional and REB policies and procedures



_____	July 3, 2026
Signature of PI	_____ Date

CLINICAL TRIAL SUMMARY

Title	Effectiveness of an Integrated Care Pathway for Adolescent Depression: A Quasi-experimental Multi-site, Cluster Controlled Trial
Short Title	CARIBOU-2 Study
Methodology	Quasi-experimental Multi-site, Cluster Controlled Trial
Clinical trial Duration	~ 84 months

Participating site(s)	Total Sites = 6 -8
	• Clinical Trials Sites (Sites 1-3)
	  ○ Maltby Centre
	  ○ Reach Out Centre for Kids (ROCK)
	  ○ SickKids Community Centre for Mental Health (CCMH)
	• Implementation-only Sites
	  ○ Ontario Shores Centre for Mental Health Sciences

Objectives	1- Measure the clinical effectiveness of CARIBOU -2 compared to TAU in improving depressive symptoms for adolescents with depression. 2- Measure the implementation feasibility, fidelity and cost of delivering the CARIBOU-2 intervention.
Number of Participants	• Youth = Up to 150 • Caregivers = ~ 50 • Clinicians = ~ 50
Study Intervention and Comparator	• Control: Treatment-as-usual for youth depression in community mental health agencies and hospitals. • Intervention : CARIBOU-2, a treatment algorithm for youth depression that includes: (1) structured assessment, including safety assessment; (2) education on depression, sleep, exercise and diet; (3) psychotherapy (with 1st line Cognitive Behavioural Therapy, 2nd line “Brief Psychosocial Intervention”); (4) a caregiver structured support group; (5) medication options (1st line fluoxetine, 2nd line sertraline); (6) “team reviews” every four weeks, (meeting with the youth and involved clinicians to review measures and discuss treatment changes); and, (7) discharge and follow-up planning
Duration of Intervention	Up to 52-weeks
Statistical Methodology	Linear mixed-effects model, qualitative analysis, descriptive

LIST OF ABBREVIATIONS

AE	Adverse Event
BPI	Brief Psychosocial Intervention
CARIBOU-2	Care for Adolescents who Receive Information ‘Bout Outcomes 2nd Iteration
CBT	Cognitive Behavioural Therapy
CO-I	Co-investigator
EDI	Equity, Diversity, and Inclusion
GBA+	Gender-Based Analysis Plus
ICF	Informed consent form
ICP	Integrated Care Pathway
IYS	Integrated Youth Service(s)
KT	Knowledge Translation
MBC	Measurement-Based Care
MHA	Mental Health Agency(s)
NICE	National Institute for Health and Care Excellence
NIRN	National Implementation Research Network
PHI	Personal Health Information
PI	Principal Investigator
REB	Research Ethics Board
RCT	Randomized Controlled Trial
SAE	Severe Adverse Event
TAU	Treatment As Usual
YEI	Youth Engagement Initiative

1 - ABSTRACT

Background: Depression is the leading cause of disability in adolescents and a potent risk factor for suicide. Evidence-based treatments are available; however, many clinics do not provide guidelines-based treatments. Integrated Care Pathways (ICPs) are treatment algorithms based on the highest quality practice guidelines intended to facilitate the delivery of evidence-based treatment at the clinic level. Our group has already tested the feasibility of ICP for adolescent depression at an academic setting. There is still uncertainty regarding whether ICPs lead to improved outcomes in depression in adolescents in community settings relative to typical care.

Objective: The current study aim is to test the effectiveness of an ICP for depression in adolescence, called the CARIBOU-2 intervention, versus treatment-as-usual (TAU) in community and hospital settings. This study will also examine important implementation outcomes.

Method: The primary participants are adolescents (Up to N=150), between the ages of 13 to 18 with depressive symptoms, presenting to one of the participating sites. Through a quasi-experimental, multi-site cluster controlled clinical trial design, sites began in the TAU condition and transitioned to the ICP condition once local enrollment to TAU has reached up to 25 participants. The primary clinical outcome of interest is the difference between treatment groups in the rate of change of depressive symptoms from baseline to 24-week endpoint as measured by the Mood and Feelings Questionnaire. Secondary outcomes include rate of change of functional improvement, as measured by the Childhood Anxiety and Depression Life Interference Scale, and caregiver-rated internalizing symptoms as rated by the Childhood Behaviour Checklist. This study will also be examining the following implementation outcomes: feasibility, fidelity cost and acceptability.

Relevance: Should our results be consistent with our hypotheses, systematic implementation of the CARIBOU-2 intervention to other community mental health agencies and hospitals would be indicated.

Key Words: Adolescent, Depression, Integrated Care Pathway, Measurement-Based Care, Implementation

2 - INTRODUCTION

2.1 Background

Adolescent depression is prevalent, debilitating and a potent risk factor for suicide. Our group conducted a province-wide survey of services for depressed children and youth showing marked heterogeneity in treatments offered and that evidence-based treatments are not consistently implemented in the community (Watson et al, 2019). A gap between what is scientifically known to be effective in mental health care, and what is actually practiced in the real world, presents a missed opportunity to optimize treatment for youth with depression.

As a first step to bridging the research-practice gap, our group has already identified current consensus on evidence-based treatments for youth with depression. A high quality and up-to-date clinical practice guideline (hereafter referred to as, "guideline") for the management of adolescent depression sets the current standard for what is effective, evidence-based care (Courtney et al., 2019). We conducted a systematic review and quality appraisal using the AGREE critical appraisal framework and found that the UK's National Institute for Health and Care Excellence Clinical Practice Guideline for Depression in Children and Young People (hereafter referred to as the "NICE guideline") is the only up-to-date, high-quality guideline available for the management of child and adolescent depression (Bennett et al., 2018).

As a second step in narrowing the research-practice gap, we have developed and implemented an Integrated Care Pathway (ICP) in the Canadian clinical context, based on the NICE guideline recommendations. An ICP is a service delivery strategy whereby guideline recommendations are used to design a pre-set treatment algorithm intended to inform clinical decision-making throughout a patient's care. ICPs are intended to standardize coordination of multidisciplinary care teams and timing of evidence-based treatment components, matched to the clinical profile of the patient. Our group has translated the NICE guideline recommendations into a preliminary flow diagram. We shared this flow diagram with clinical staff and youth collaborators at the Centre for Addictions and Mental Health (CAMH) academic outpatient psychiatry clinic. Together, we modified the flow diagram and contextualized the recommendations (Courtney et al., 2019; https://onlinelibrary.wiley.com/doi/epdf/10.1111/eip.12918) to form an ICP called the CARIBOU-1 intervention ("Care for Adolescents who Receive Information 'Bout OUtcomes). Youth collaborators provided extensive feedback and revisions to various materials, including patient handouts and psychotherapy manual materials (https://t.co/7ImxwTPLot?amp=1). We have piloted the CARIBOU-1 intervention in a non-randomized controlled clinical trial in an academic setting. Thirty-five youth participants with depression at CAMH received the CARIBOU-1 intervention; and 31 youth participants at Sunnybrook Health Sciences Centre (SHSC) received treatment as usual. Our pilot study objectives were met, including recruitment targets (>60 participants across 21 months), data collection efficiency (>80% of intended data collected on the primary outcome) and retention (over 80% of participants who started the trial, completed the protocol). The protocol has been published (Courtney et al., 2019; https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6863574/pdf/ccap28_p0115.pdf); the manuscript with results for the CARIBOU-1 study is currently being prepared.

As a third step in closing the research-practice gap, it is important to demonstrate that treatment improves health outcomes and is acceptable for individual patients. Our group is combining principles of Measurement-Based care (MBC) and Shared Decision-Making (SDM) to enhance personalization of the delivery of evidence-based care. Measurement-Based Care refers to "the systematic administration of symptom rating scales that uses outcomes to drive clinical decision-making at the level of the individual patient" (Fortney et al., 2016). It is thought to optimize outcomes by rendering clinicians aware of non-response early in the course of treatment, so that treatment can be changed accordingly. The MBC framework fits well with the ICP's algorithm structure as sequential guideline recommendations are often based on treatment response and tolerability. MBC also promotes individually tailored care, which can get lost in the context of standardized treatment approaches. MBC was incorporated into the CARIBOU-1 intervention that was piloted. Shared Decision-Making refers to

a process whereby a clinician and patient share information with each other, and work to build consensus on the preferred treatment course. The premise has recently been emphasized in the updated NICE guidelines for depression in adults. The SDM approach has been proposed to improve outcomes by enhancing patient satisfaction, treatment adherence, quality of life and well-being (Charles et al., 1997 & Joosten et al., 2008).

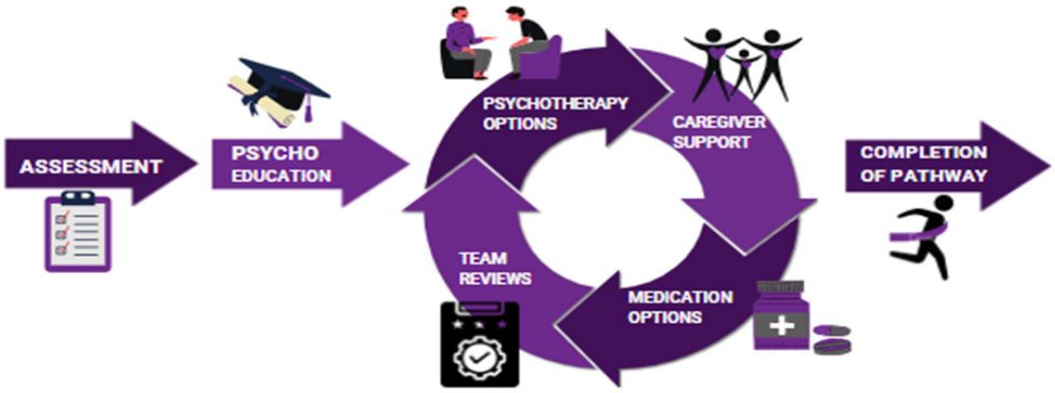
As a fourth step we aim to ensure that the ICP can be effectively implemented externally—in academic hospital, and community mental health agency settings. This will pave the way for future scale up within the system of care, toward improving health outcomes for youth with depression. In the CARIBOU-1 study, clinicians at an academic setting (CAMH) showed good fidelity to the ICP; that is, >80% fidelity on a locally-developed checklist with respect to the overall ICP. The extent to which an ICP for adolescent depression can be implemented externally—in community mental health agencies or hospitals has yet to be determined.

An ICP using the principles above is intended to find the intersection between research evidence, clinician judgment and patient/caregiver values, all within the context of finite resources; and thus deliver a feasible model of high quality and evidence-based care. In the pursuit of this point of intersection, it is hoped that the lives of adolescents with depression will improve (Sackett et al., 1996).

2.2 CARIBOU-2 Intervention, in brief.

After the CARIBOU-1 study, we have since revised the ICP to render it more applicable to other organizations as well as expand the psychosocial component for youth who do not engage with, or respond to, cognitive-behavioural therapy. The revised version is called the CARIBOU-2 intervention. Details on the CARIBOU-2 intervention are included below in section 5.2 of this protocol. The current iteration of the pathway involves a series of steps: (1) structured assessment, including safety assessment; (2) education on depression, sleep, exercise and diet; (3) psychotherapy (with 1st line Cognitive Behavioural Therapy, 2nd line "Brief Psychosocial Intervention"); (4) a caregiver structured support group; (5) medication options (1st line fluoxetine, 2nd line sertraline); (6) "team reviews" every four weeks, (meeting with the youth and involved clinicians to review measures and discuss treatment changes); and, (7) discharge and follow-up planning. Figure 1 outlines a schematic of these steps.

Figure 1: Schematic of CARIBOU -2 intervention:



2.3 Literature Review

Our group has conducted a related, but separate, scoping review of Randomized Controlled Trials (RCTs) (ref, Courtney et al 2022, JAMA Open)Through this review we have identified 6 research programs relating to the effectiveness of ICPs and MBC for the management of adolescent depression. None of these studies involved efforts to specifically implement an ICP based on the highest quality evidence on treatment of adolescent depression (derived from guideline

recommendations), nor did they extensively involve collaborative efforts with clinicians and youth collaborators in the development and implementation of the treatment algorithms.

The research programs identified are as follows: (1) Richardson and colleagues (2014) have conducted an RCT of a collaborative care model (that offers education, medication options and CBT in a structured format with MBC) for depression in adolescents presenting to primary care. Youth who received collaborative care had a significantly greater decrease in depression severity scores at both 6- and 12-month time points, with a response rate odds ratio of >3, relative to treatment as usual (TAU). This study did not achieve the intended sample size and within-site contamination effects may have affected results as participants were randomized at the individual level; (2) Asarnow and colleagues (2005) conducted an RCT where a quality improvement intervention aimed at improving access to evidence-based depression treatments for adolescents. The intervention included: education regarding treatment of depression, use of expert leader teams, care management support of primary care clinicians, and training of care managers in CBT. The intervention resulted in significant improvement of depressive symptoms compared to TAU at 6-months. This study did not involve MBC and randomized participants at the individual level rendering results vulnerable to within-site contamination effects. (3) Gunlicks-Stoessel and colleagues (2019) demonstrated the feasibility of an “adaptive treatment strategy” for adolescent depression with a focus on MBC. Options in the protocol did not extend beyond interpersonal therapy and fluoxetine treatment. (4) The Texas Children’s Medication Algorithm Project (Hughes et al, 2007) used both a medication algorithm and an MBC framework. They demonstrated superior response rates in depressive symptoms with the use of the algorithm compared to a historical cohort. This was not an RCT and the algorithm did not include psychotherapeutic interventions. (5) Weisz and colleagues (2012) have studied a modular psychotherapy treatment (MATCH) for children that relied on MBC, finding significant improvement in overall psychopathology for MATCH relative to TAU. Chorpita and colleagues (2017) found similar effects in a trial of children receiving MATCH. The trials testing MATCH involved younger children with a broad range of psychopathology. (6) Tam and Ronan (2017) conducted a meta-analysis regarding MBC in youth. They identified 3 RCTs of MBC in youth targeting broad psychopathology and found a small effect size overall; how MBC can affect depressive outcomes in adolescence specifically is unknown.

The adult literature on guideline implementation, ICPs and MBC is more developed. For example, in their meta-analysis of 19 studies, Girlanda and colleagues (2017) conclude that strategies to implement guideline recommendations can lead to improved outcomes for patients with mental health conditions. Shimokawa and colleagues (2010) used meta-analytic strategies on 6 studies 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). In an RCT by Kane and colleagues (2016), implementation of a community-based ICP for young people with psychosis was effective in significantly improving psychopathology and function. Despite the wealth of evidence in adult populations, testing this approach in adolescents is vital to knowing if the results are generalizable and feasible. A collaboratively-developed ICP using MBC that is based on high quality guidelines for adolescent depression has not yet been tested in an RCT for clinical or implementation effectiveness.

3 – STUDY OBJECTIVES, RESEARCH QUESTIONS AND HYPOTHESES

3.1 Study Objectives

Based on our work-to-date, and the existing literature, we believe that the CARIBOU-2 intervention will help close the research-practice gap in community mental health agencies and hospitals, and optimize outcomes for adolescents with depression. The primary objective of this study is to establish the clinical effectiveness of the CARIBOU-2 intervention delivered to depressed adolescents in the community and hospital setting, with respect to reducing evaluator-rated depressive symptoms relative to Treatment as Usual (TAU). Our secondary objectives are to explore changes in clinician-rated function and caregiver-rated symptoms for youth receiving the CARIBOU-2 intervention relative to

TAU. The third objective is to explore and compare implementation effectiveness of the CARIBOU-2 intervention in community and hospital settings, namely investigating: feasibility, fidelity, cost, and acceptability.

3.2 Research Questions and Hypotheses

A. Clinical Effectiveness Outcomes

1. Is the use of the CARIBOU-2 intervention more effective than TAU with regards to rate of improvement of depressive symptoms, for adolescents with depression?

   Hypothesis A: Assuming the CARIBOU-2 intervention is implemented and delivered with fidelity (see hypotheses G/H/I), there will be a greater rate of reduction in depressive symptoms in the group receiving the CARIBOU-2 intervention relative to TAU over 24 weeks as per the scores on the Mood and Feelings Questionnaire.

2. Do caregivers report greater rates of reduction in youth internalizing psychopathology when youth receive the CARIBOU-2 intervention, relative to TAU?

   Hypothesis C: There will be a greater rate of reduction in caregiver-reported youth internalizing symptoms psychopathology over a 24-week period as per the Child Behaviour Checklist Internalizing Broadband subscale in the group receiving the CARIBOU-2 intervention, relative to TAU.

Longitudinal changes in suicidal ideation and self-harm behaviour outcomes will be examined in an exploratory manner.

B. Implementation Outcomes

1. Feasibility: Do youth participants end up receiving and engaging with indicated components of the CARIBOU-2 intervention?

   Hypotheses D: We anticipate that youth will receive and engage in ≥75% of indicated treatment components while they are in the pathway as per a locally-developed checklist.

2. Fidelity: Did implementation teams and clinicians deliver and demonstrate fidelity to the CARIBOU-2 intervention model?

   Hypothesis E: The implementation process will be followed with ≥75% fidelity for each of the 3 sites as per a locally-developed checklist.

   Hypothesis F: The CARIBOU-2 intervention will be delivered with >75% fidelity for the overall ICP (i.e. how and when components are offered) as per a locally-developed checklist.

   Hypothesis G: The CARIBOU-2 intervention will be delivered with >75% fidelity for each separate component of the ICP (e.g., fidelity to Cognitive Behavioural Therapy) using established fidelity checklists where available, and locally-developed checklists where no established checklist is available.

3. Cost: What is the cost-effectiveness of the CARIBOU-2 intervention?
   Hypothesis H: The CARIBOU-2 intervention will be cost-effective compared to TAU.

Barriers and facilitators to implementation, as well as clinician acceptability of implementing the CARIBOU-2 intervention, will be explored through qualitative methods. There are no associated hypotheses with regards to these outcomes.

We will also explore the mediating role of adolescent-reported shared decision-making and youth-reported coping skill development as a mediator of the primary outcome (i.e. Hypothesis A).

Furthermore, as a sub study, we will be examining certain psychometric properties of the Mood and Feelings Questionnaire, the Child Anxiety and Depression Life Interference Scale-Youth Version, Child Anxiety and Depression Life Interference Scale-Caregiver Version, and Childhood Depression Rating Scale-Revised. See Appendix J for the detailed protocols.

4 – METHODS: DESIGN, IMPLEMENTATION PROCESS & RECRUITMENT

4.1 Study Design

The study uses a Type 1 Hybrid Implementation Effectiveness design (Curran et al., 2012) that focuses on the effectiveness of the clinical intervention (ICP) while exploring the implementation effectiveness of the intervention. A hybrid type 1 design is indicated because the clinical effectiveness evidence for ICP is yet to be determined and the study conditions offer an ideal opportunity to explore implementation issues and plan implementation strategies.

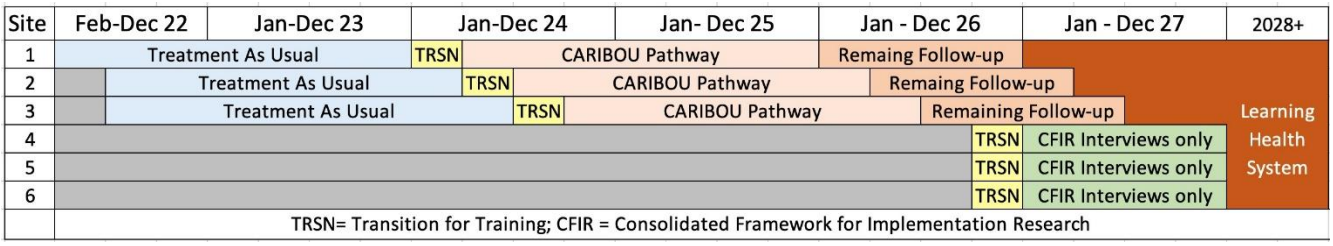
A cluster-controlled trial is indicated as the intervention is applied at the clinic-level. Within site contamination of treatment precludes the use of a trial where the individual is assigned to intervention or control.

A stepped wedge design was initially chosen over a parallel design; however, this design was not feasible for the following reasons: (1) enrollment rates at the 6 initially chosen sites were highly variable (for example, in the first year of enrollment, one site had enrolled 25 participants and another only one, (2) Two sites have withdrawn from the study). It is taking time to recruit new sites and we cannot expect existing sites to wait for the onboarding of other sites, to be able to implement the intervention. In an effort to both optimize feasibility and minimize threats to validity, we have opted to change the design to a quasi-experimental, multi-site cluster controlled clinical trial design. This design is similar to a stepped wedge trial; however, the sequence of transitioning from TAU to the ICP is not randomized.

This study is recruiting six community mental health agencies and/or hospitals (herein referred to as sites). Three sites will be used to assess clinical effectiveness of the pathway relative to treatment as usual, as well as implementation outcomes. Additionally, another three sites will be trained in the pathway without a comparator group, wherein implementation outcomes only will be explored.

In our design, for the three sites assessing clinical effectiveness of the pathway (i.e., sites 1-3), each site is the “cluster” unit. For the first phase of data collection, each of the three sites will remain in the Treatment as Usual (TAU) condition. Once a site has enrolled up to 25 youth participants into TAU, we will begin the transition from the TAU condition to the ICP condition. In order to stagger implementation, there will be a minimum of 3 months between one site beginning to transition relative to the next one. Please refer to Figure 2 for the site sequencing.

Figure 2: CARIBOU-2 Study Timeline



4.2 Implementation Process (Quality Implementation Framework)
The Quality Implementation Framework identifies the critical phases in the implementation process along with specific actions related to these steps that can be utilized to achieve quality implementation (Meyers et al., 2012). The four phases are:

Phase 1: Initial considerations regarding hosting setting
Phase 2: Creating a structure for implementation
Phase 3: Ongoing support once implementation begins
Phase 4: Improving future applications

Phase 1: This phase involves setting up governance, site invitation and site selection.

(a) Setting up governance: A leadership implementation team (Courtney, Barwick, Henderson, and others) and operational implementation teams at the sites (clinical manager, clinicians, and local implementation specialists) will guide the implementation process.

(b) Site invitation: Community mental health agencies and hospitals across Ontario and other provinces (e.g., Alberta) will be invited to a Webinar wherein a detailed description of the study will be discussed. Interested site staff (i.e., managers and directors) will be asked to complete an application form to assess further for eligibility, including: i) high volume of adolescents with depression presenting to the agency, ii) number of clinicians (social workers, psychologists, occupational therapists, registered nurses or registered therapists with >6 months experience working with adolescents with depression) who can provide the psychosocial interventions, iii) existing models of care, and iv) anticipated implementation of new models of care. Subsequently, members of our research team will hold unique meetings with staff from each interested site (e.g. managers) to further clarify details of the study.

(c) Site selection: Agencies that have continued to express interest in the CARIBOU-2 study will be assessed for implementation readiness, using the Checklist to Assess Organizational Readiness (CARI) and the NIRN Hexagon Tool (Appendix C). Six sites will be selected for study participation based on their level of readiness, anticipated volume of appropriate youth participants, current resources and potential for confounding from other interventions being implemented at the site. Equity of access to resources will also be considered.

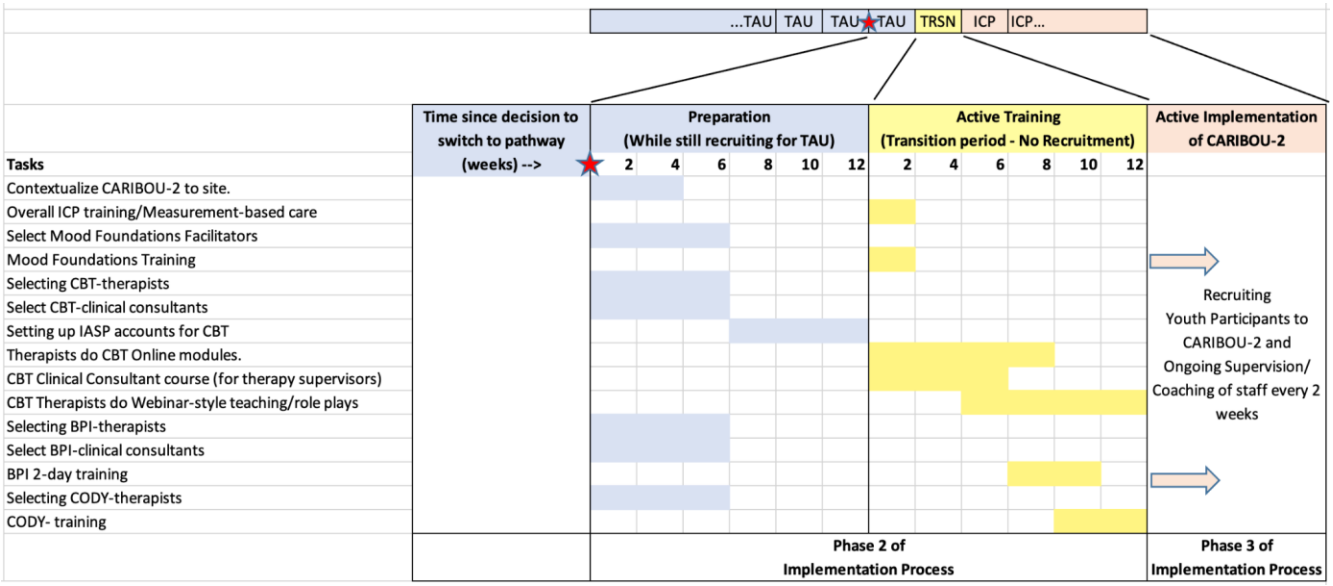
Phase 2: The sites' operational implementation teams will lead the establishment of organizational and personal competencies to ensure the successful implementation of the intervention. Selection of site staff who will be trained to deliver the CARIBOU-2 intervention will take place here. Discussion will touch on the need for adaptations to how core components of the CARIBOU-2 intervention are delivered, engaging and maintaining explicit buy-in from organizational stakeholders. During

"transition" periods, where sites are switching over to the CARIBOU-2 intervention, training of therapies will occur. See Figure 3 below.

Phase 3: This phase involves the initiation of the CARIBOU-2 intervention and its availability to youth participants. The key focus is on continuous improvement of the newly learned skills (i.e., CBT, BPI, and ICP). Ongoing support will be provided through technical supervision of the specific therapies, continuous process evaluation and supportive feedback. Implementation teams place a heavy emphasis on the systematic review of fidelity data during this stage to ensure that any changes to the model or approach are purposeful and planned, rather than reactionary or opportunistic. Meetings between implementation teams and participant-facing clinicians will be held every 2 weeks for these tasks to take place.

Phase 4: This phase will involve synthesizing the results from the clinical and implementation outcomes and examining potential modifications to the pathway and implementation process for scaling up.

Figure 3: Gantt chart outlining steps to take place in Phases 2 and 3 in Implementation Process



4.3 Primary Participant Recruitment (Youth)

Youth participant recruitment will only occur in the three sites with a TAU control condition (i.e., sites 1-3) over 4.5 years, from February 2022 to December 2026. All participants will be recruited after their intake at a participating site. The following recruitment procedures will be used: (1) clinicians will conduct a standard mental health assessment of the youth, including use of the Mood and Feelings Questionnaire (Angold & Costello, 1987), (2) adolescent patients meeting inclusion/exclusion criteria will be asked if they are willing to meet with a Research Analyst (RA) to further discuss the study. If the youth is agreeable to meet with the RA, the clinician will collect the youth's contact information and provide it to the CAMH research team, so that the team can reach out to the youth to set up a meeting to further discuss the study.

Additionally, at sites 1-3 will be given the option to have a CAMH RA join the first or last 5 minutes of a virtual session with a youth in their caseload, so that the RA can briefly review the study to the youth. This option will be provided at the discretion of the sites, and participating clinicians. All sites will be given this option, but they will decide whether it fits with their current practices. Further, individual clinicians—from sites that are agreeable to this option—will also decide if they would like to implement this recruitment strategy. In working on this study, we have learned that practices are vast across organizations. Ensuring that there's various options is essential. For example, a clinician that specializes in working with high-risk youth may feel it would be inappropriate to ask youth in their caseload if a RA can briefly join to explain the study. They may feel more comfortable using the previous option—broadly asking youth if they want to learn more about a study, whether or not they would like to meet with the RA, and if they are agreeable to providing contact information to the CAMH team.

Once the youth has met with the RA, the informed consent process will ensue (refer to Appendix D and G for relevant consent forms and scripts). Those who do not meet criteria, or who do not agree to participate in the study, will proceed with the standard care, provided by the site, without further research visits pertaining to this study.

As depression has preponderance in girls, we will be monitoring the girl-to-boy ratio (as per current gender identity) over the course of the trial to ensure that approximately one third of participants identify as a boy. This will allow for enhanced generalizability of results. We will also be monitoring diversity of gender identification of participants enrolled in the study and look for opportunities to optimize involvement of youth who identify as gender diverse (e.g. transgender or nonbinary) to ensure that their experiences are also represented.

Inclusion and exclusion criteria

Inclusion Criteria

- Youth is aged 13 to 18 years, inclusive.
- Youth and/or their caregiver is expressing that 'depression" (or some synonym) is a concern
- Clinician agrees that depressive symptoms are a treatment target.
- MFQ score is ≥22 at two sequential visits (screening and baseline assessment)
- Youth must be new to the site (in past 3 months) or have a period of no treatment for 3 months

Exclusion Criteria

- Known or highly suspected presentations of psychotic symptoms that are persistent, affect functioning, and have observable effects on behaviour.
- Severe substance use disorder, bipolar disorder, intellectual disability, severe eating disorder, imminent risk of suicide requiring hospitalization as per judgment of the assessing clinician. Inability to provide informed consent to the study for any reason. Youth in, or applying to, a Day Treatment Program

The requirement of a formal diagnostic assessment is intentionally omitted from the inclusion/exclusion criteria as participating sites often do not have access to this process; and we would like the results to be generalizable to typical clinical processes. The RA will conduct a diagnostic assessment at baseline to describe the sample, but this information will not be fed back to clinicians unless an acute safety issue arises.

If there is ambiguity about whether or not the client meets any of the above exclusion criteria, the decision to include/exclude can be postponed until the youth is assessed. There are no restrictions on medications being prescribed at entry into the trial.

Our age limits for adolescents are age 13-18, inclusive. The CARIBOU-2 intervention was designed for young people aged 13-18 in mind; with evidence pertaining to this age group as a base for the decision aide.

Ability to speak and read English for adolescents, caregivers, and clinicians is also required. The CBT and psychoeducation materials are all provided in English. Moreover, the clinicians involved in this study only conduct practice in English. Care delivered in another language would be inappropriate with the current resources available.

4.4 Secondary Participant Recruitment (Caregivers): For sites 1-3, youth will be able to decide the level of involvement their caregiver will have in their treatment process and study participation. Caregivers that are fully involved in the treatment process will be asked to participate in the study, with youth's permission.

4.5 Tertiary Participant Recruitment (Local implementation team members, Supervisors and Clinicians): For all six sites, supervisors and clinicians interested in being a research participant will be recruited after legal contracts have been fully executed. Other than fluency in English, and capacity to make decisions with regards to consenting to research, there are no inclusion or exclusion criteria for caregiver or supervisor/clinician participation.

Site staff research participants at implementation-only sites (i.e., sites 4-6) will provide informed consent using the REDCap e-Consent Framework. Site staff participants will be provided with a read-only copy of the ICF via REDCap prior to conducting the consent discussion. Following the consent discussion, the prospective site staff participant will be sent a link to the e-consent via email or chat feature in WebEx. The participant will complete the e-consent and be provided with the option to download and/or email themselves the signed ICF. If email is chosen, the email will only be used for this purpose (it is not retained by REDCap). Following the participant signature, the person conducting the consent discussion will complete the Person Conducting Consent Discussion Attestation Page. PDF copies of the signed ICFs and Attestation pages will be retained in the REDCap File Repository. The research team will provide the site staff research participant with a copy of the fully signed ICF via email.

4.6 Anticipated Sample Size
Youth: 150 (from sites 1-3)
Caregivers: ~50 (from sites 1-3; based on preliminary recruitment results where one-third of youth agreed to have caregiver involvement and caregiver consents)
Clinicians: ~50 (from all six sites)

Give our recruitment rates early in the study, we expect to recruit another 17 participants at site 3 to reach our current sample size target by December 2026.

4.7 Participant retention and withdrawal:
Several strategies will be used to maintain engagement throughout the 12 months; mainly informed by the retention strategies used by the ABCD study (Ewing et al., 2018) and suggestions from the youth in our Youth Engagement Initiative. Strategies include (a) rapport building between the RAs and the youth, (b) $50 Everything Card compensation for the first and last research visit, and $25 Everything Card compensation for interim visits, (c) the option of using their time in research as part of volunteer hours for high school credit, (d) obtaining multiple points of contact (e.g., cell phone numbers, home numbers, caregiver numbers, and email addresses), (e) using multiple modalities to follow-up on participants (i.e., phone calls, emails, and texts), (f) sending thank you messages after each time point of participation,, (g) reminder messages prior to their upcoming assessment, (h) birthday cards, and (i)

anticipate supports that participants might need (i.e., letters for school absences for study visits). Youth collaborators will assist in providing continual feedback as to how best to retain participants.

Caregiver participants will receive a $50 Everything Card compensation at their first and last research visit, and $25 Everything Card compensation for all other participation time points.

Site staff, including program leads, supervisors and clinicians will not receive study compensation. Their participation is dependent on their interest in implementing the CARIBOU-2 intervention, which includes various professional development benefits (i.e., free treatment training, access to new resources, and coaching and consultation).

Participation is voluntary, and participants are free to withdraw from the study at any time. Withdrawing from the study will not affect youth's treatment. Should participants, family members and/or clinicians observe any clear harm brought about by the intervention, the participant will be withdrawn from the study.

In our pilot of the original CARIBOU pathway, 66 participants completed the study protocol, and 11 participants dropped out prior to study completion. With increased supports, we anticipate that attrition from the research protocol will be lower in the present study.

5 – METHODS: INTERVENTIONS

5.1 Treatment as Usual:

Study involvement for sites 1-3 will begin in the TAU condition, which is the typical treatment at the participating sites. A clinician on staff at the implementing site (usually a social worker) will assess and plan treatment with clients. A range of treatments observed in our previous survey of sites will be on offer in these agencies, depending on the preferences and context of the local agency. There is typically no structured treatment algorithm to guide treatment and no pre-arranged multidisciplinary "team reviews" every four weeks. Clinicians, youth, and caregivers negotiate the clinical appointment frequency. In typical TAU in community mental health agencies, depressive symptoms and function are not systematically monitored via standardized rating scales. TAU may or may not include referral to psychotherapy and/or parent support. Psychiatric care and the use of psychotropic medication is allowed. There are no prompts to prescribe specific medications, and/or internal and external referrals to treatment and other services, guided by local service standards. An RA will record the treatment received in each site for each participant using chart review and interviews with the participant and their primary clinician to clarify the difference in service provision between the CARIBOU-2 intervention and TAU.

This type of comparator group was selected with consideration of the NIH expert panel's recommendations for selecting comparator groups in behavioural interventions, particularly as it relates to the overall objective of our study (Freeland et al, 2019). TAU represents the status quo and is most relevant to the question of whether or not CARIBOU-2 is more effective than TAU and should be developed and implemented on a larger scale.

5.2 CARIBOU-2 Intervention, in detail

Sites 1-3 will transition to the CARIBOU-2 intervention according to enrollment numbers, as described above. After presentation to a site, consenting participants will be offered the CARIBOU-2 intervention. Sites 4-6 will also be trained in the fall of 2026. The Core Components of the CARIBOU-2 intervention are shared in the flow diagram in Appendix B and described briefly here. "Clinicians" refer to social workers, psychologists, occupational therapists, registered nurses or registered therapists with >6 months experience working with adolescents with depression.

5.2.1 CARIBOU-2 Intervention Core Components.
Core Component 1. Participant and caregiver assessment : Patients will undergo an assessment by a clinician at the mental health agency that includes the completion of various measures related to depression (Mood and Feelings Questionnaire - MFQ), anxiety, (Revised Children's Anxiety and Depression Scale –15 item anxiety subscale), function (Child Anxiety and Depression Life Interference Scale- CADLIS), caregiver conflict (Conflict Behaviour Questionnaire, Adolescent version) and Goals Based Outcome (GBO); see below for details on these measures. Youth indicating a risk of suicide (answering at least a "sometimes" on items 16-19 that relate to suicidal ideation on the MFQ), will undergo a safety assessment. This safety assessment includes use of the Columbia Suicide Severity Rating Scale (C-SSRS). If the C-SSRS indicates low-moderate suicidal ideation (e.g., score of 1-3 on the C-SSRS suicidal ideation subscale), participants will be offered a telepsychiatry appointment with a study psychiatrist to assess risk management and/or medication options with an 8-12-week trial of fluoxetine as first line and an 8-12-week trial of sertraline as second line. If the assessment and C-SSRS indicates more active suicidal ideation (e.g., score of 4-5) or recent suicide attempt, imminent review with senior staff at the site will take place to decide on next steps; which may or may not include hospitalization.

Core Component 2. Education. All participants and caregivers will be offered a one-time multi-family psychoeducation session, called Mood Foundations, where a clinician will provide information on the nature of depression, improving sleep quality, increasing exercise and healthy diet.

Core Component 3a. Psychotherapy options: Cognitive Behaviour Therapy . Clinicians will offer participants 8-sessions of CBT with two 4- session modules: activation and cognitive restructuring. In order to further personalize care, participants whose depression has not remitted will be offered another two 4-session modules of CBT, relating to anxiety, insomnia, social skills and problem-solving. Module choice will depend on the youth's treatment goals, symptom profile and clinician impression. The CARIBOU CBT manual is an updated version of Lewinsohn and Clarke's Coping with Depression for Adolescent course (Lewinsohn & Clarke, 1984). The original manual has been shown to be effective in multiple randomized controlled trials (Weersing et al., 2016). Through input from the Youth Engagement Team at CAMH, it was adapted with examples and language more in keeping with current youth culture. It was also separated into 4 modules for more flexible use (i.e. Behaviour Activation, Cognitive Restructuring, Problem-solving and Communication/Relationships). The adaptation was done with permission from the authors for use and dissemination. It has been piloted in our CARIBOU-1 study. The manual is currently available at https://edc.camhx.ca/redcap/surveys/?s=DK8YEYWMEWN7ED3J.

Core Component 3b. Psychotherapy options: Brief Psychosocial Intervention . Youth who do not respond to or do not engage with CBT will be offered up to 12 sessions of "Brief Psychosocial Intervention" (BPI). BPI is a manualized and evidence-based treatment strategy for the treatment of depression in adolescents recommended as a viable treatment option in the NICE guidelines. It involves supportive approaches as well as directly addressing factors thought to be contributing to the youth's depression as per the formulation. The manual is available at: http://dev.psychiatry.cam.ac.uk/files/2016/06/BPI-Manual-1.pdf

Core Component 4. Caregiver support. For youth who identify conflict with their caregiver(s) (CBQ-A score of ≥9) (Rengasamy et al., 2013) clinicians will also offer the youth's caregivers an 8-session group of CBT-based strategies for interacting with young people with depression. This intervention was also based on the caregiver component of the Coping with Depression course and adapted for current cultural context. Dr. Madison Aitken (co-Investigator) is actively conducting research on this component at CAMH. The original manual is available here: https://research.kpchr.org/Portals/0/Docs/project%20websites/ACWD/CWDA_parent_manual.pdf?ver=2016-04-07-083508-963.

Core Component 5. Medication options. For youth who initially present with severe depression (by mapping DSM-5 criteria mapped from MFQ items as per Appendix F; by clinical impression, or presence of self-harm or suicidal ideation), psychiatry appointments will be offered. Study psychiatrists may include psychiatrists already affiliated with selected sites or affiliated with CAMH. Psychiatry appointments will also be offered to youth who are not responding to 8 weeks of psychotherapy. At these appointments, psychiatrists will assess if medication options are warranted and discuss the option of medication management for whom it is indicated. Should it be decided to start medication for depression, the psychiatrists will follow the recommendations of medication stream flow diagram based on the NICE guideline (see Appendix B). Diagnostic clarification and recommendations for safety management may also take place at these appointments.

Core Component 6. Measurement Based Care/Team Review. Measurement-based care (MBC) “team reviews” consisting of the most-responsible clinician, other involved health professionals (e.g., psychiatrist, nurse, or social worker), the patient +/- caregiver(s) will be held every 4 weeks throughout the intervention. All members of the “team review” will discuss the change-scores in measures of the MBC package (see section 5.2.4 and Appendix F) and decide to continue or change the current treatment plan at the indicated decision points. Extra psychiatry follow-up between team reviews will be permitted when indicated (for example, when first starting a medication). Clinicians will be reminded that the MBC guides decisions, but do not rigidly prescribe treatment.

Core Component 7. Relapse prevention and pathway graduation planning. This will be an ongoing item of discussion "team reviews"; with more emphasis once remission is achieved (MFQ<22, and clinician, youth +/- caregiver impression). Communication with local resources (both formal and informal –e.g., family doctors, family members) to look for indicators of relapse will take place. The following are 5 situations that may lead to discharge:

(a) If the youth is in remission and a full course of therapy is not yet complete, discharge would occur after therapy is done - as it may help prevent relapse.

(b) If the youth is in remission , but prefers not to continue with therapy, discharge could take place immediately with proper supports in place

(c) If the youth is remains depressed , but has had an improvement in symptoms and function that is satisfactory for them, a collaborative plan to discharge could also take place.

(d) If a youth remains depressed and all treatment components offered in the pathway have been completed/refused or 52 weeks have passed since initiation of the pathway and the youth remains depressed, referral to more intensive/specialist care could be initiated; in which case discharge from the pathway would be equivalent to transfer of care.

(e) If youth are lost to follow -up (e.g., missed 3 sessions in a row, or not responding to calls for appointments), discharge could also take place.

Wait time will be defined as the number of days between "Date of referral to depression treatment at agency" and "Date of first treatment session". Length of stay will be defined as number of days between "Date of first treatment session" and "Date of last treatment session" within the 52-week interval of the study. As mentioned, participants may receive the CARIBOU-2 intervention for up to 52 weeks to allow for a course of up to 16 weeks in CBT and 12 weeks in BPI, if needed, as well as a window of time to allow for arrangement of various treatment components. As per the Intent-to-Treat principle, participants may leave the intervention earlier and still have the scheduled follow-up research visits.

5.2.2 Acceptable Adaptations:
The following table outlines adaptations that sites may decide to make to contextualize the pathway to their resources, needs and culture. Table 1. Acceptable Adaptations

Component	Individual or Group?	In-person or Online?
1. Assessment	Individual (+/- Caregiver)	Either
2. Education	Either	Either
3. Team Reviews	Individual (+/- Caregiver)	Either
4a. CBT	Either	Either
4b. BPI	Individual	Either
4c. Caregiver group	Group	Either
5. Medication	Individual (+/- Caregiver)	Online
6. Measurement Based Care	Individual (+/- Caregiver)	Either
7. Relapse prevention/ Pathway graduation planning	Individual (+/- Caregiver)	Either

5.2.3 Implementation Strategies:
Throughout the implementation of the CARIBOU-2 intervention, multiple implementation strategies will be used. See Appendix B for the ERIC taxonomy of various strategies anticipated.

5.2.4 CARIBOU-2 Intervention Measurement -Based Care Measures:
These measures are fed back to clinicians and youth participants in the CARIBOU-2 intervention arm just prior to team reviews every 4-weeks as an integral part of the CARIBOU-2 intervention. The CARIBOU-2 flow diagram suggests treatment options (refer to Appendix B, for the flow diagram) and shared decision-making around treatment changes occur based on the measures, youth participant values as well as clinician expertise.

The Mood and Feelings Questionnaire child, long version (MFQ; Angold & Costello, 1987): The MFQ is a 33-item self-report measure for youth regarding depressive symptoms experienced over the prior 2 weeks. It was specifically recommended in the NICE guideline, given its strong psychometric properties; namely, it has high discriminatory ability in adolescents, good internal consistency (α=0.92-0.94) and good test-retest reliability (Pearson's r=0.78). It has also been used in a large trial of psychotherapy with adolescents where it was sensitive to change (Goodyer et al, 2017).

Revised'Children's Anxiety and Depression Scale-25**–** Anxiety Subscale (15 items) (RCADS -anx-15; Chorpita et al, 2000; Ebesutani et al. 2012): In our CARIBOU-1 pilot study, 69% of youth entering the intervention also met criteria for an anxiety disorder. Treatment of anxiety overlaps substantially with treatment of depression (including the use of SSRIs and CBT), so targeting these symptoms, along with depression, is logical to pursue. We will use 15-item subscale of the RCADS-25 (Chorpita et al, 2000; Ebesutani et al. 2012) pertaining to anxiety. After a thorough review, the

RCADs were recommended by the International Consortium for Health Outcomes Measurement (ICHOM) for clinical use due to its feasibility and good psychometric properties (Krause et al., 2021).

The Child Anxiety and Depression Life Interference Scale -Youth Version (CADLIS -Y; Lyneham et al., 2013): This 9-item self-reported impairment measure evaluates the extent to which symptoms of anxiety and/or depression interfere with activities of daily living (e.g. "being with friends outside of school"). All items, are rated on a five-point Likert scale (0 = not at all, 4 = a great deal), with higher scores indicating higher impairment from anxiety and/or depression. This measure was chosen by our youth partners due to its face validity and acceptability. In collaboration with the instrument's developers, we will make minor changes to the question wording to ensure its suitability in the Canadian context, as the tool was originally developed in Australia. Within this study, we will examine the psychometric properties of the CADLIS-Y in Canadian youth, and we will statistically explore meaningful indicators of change for the interpretation of CADLIS-Y change scores (see Appendix J).

Clinical Global Impression Scale – both the Improvement Subscale (CGI -I) and Severity Subscale (CGI -S) (Busner & Targum, 2007) : The CGI-I is a one-item assessment of the clinician's sense of change in a patient's symptom burden and overall functioning. Seven response options are available with scores ranging from 1 through 7: "very much improved", "much improved", "a little improved", "no change", "a little worse", "much worse", and "very much worse". In a previous study, the CGI-I has demonstrated good inter-rater reliability at identifying responders – defined as cases rated to be "very much improved" or "much improved" (Asarnow, Emslie, Clarke et al. 2009). This measure is used broadly across RCTs for adolescent depression (Courtney et al, 2021) and will be used to select the clinical significance of our finding and render findings comparable to other studies.

The Columbia Suicide Severity Rating Scale (C -SSRS; Public Health Agency of Canada, 2006): This measure contains items related self-injurious thoughts and behaviours. Interviewers rate 3 subscales: Suicidal Ideation Severity, Suicidal Behaviour (which includes suicide attempts and an item for non-suicidal self-injury that is recorded as distinct from suicidal behaviour), and Lethality. Favourable psychometric properties have been described in both adolescent and young adult samples (Posner et al., 2011). It is a commonly used measure to quantify these thoughts and behaviours and has been recommended in treatment trials by the FDA. When used as a part of measurement-based care, the frame of reference is "since last assessment". Active planning of suicide (score of 4 or 5 on the suicidal ideation subscale) or recent suicide attempt would prompt further acute safety management which may involve consultation with senior staff at the agency as well as engagement of caregivers in safety planning.

Conflict Behaviour Questionnaire -Child About Caregiver (CBQ-A; Robin & Foster, 1989): A 20-item true/false scale that assesses general conflict between parents and their children. The CBQ-A has been used extensively as a potential predictor of outcome in multiple RCTs of adolescent depression (Brent et al., 1997; Brent et al., 2008; Mufson et al., 2004; March et al., 2004) and has adequate internal consistency (Robin & Foster, 1989). The cut-off reference is a total score of ≥9 (Rengasamy et al., 2013). This measure is only applied at baseline to see if the caregiver support group is indicated.

Goals Based Outcomes (GBO; Law & Jacob, 2015): A measure used to evaluate the progress a client has made towards a goal they have set for themselves at the beginning of an intervention. The client is asked to identify 3 goals within the first 3 sessions and is asked to rate their progress towards the goal (0 meaning no the goal is not met in anyway, 5 meaning halfway towards the goal, 10 is reaching the goal). The participant tracks their progress towards the goal before and after the intervention.

Table 2 outlines details regarding the timing of the measurement-based care measures.

Table 2. Assessment Schedule for Measurement -Based Care Measures

Measures	Time (mins)	Week													
		0	4	8	12	16	20	24	28	32	36	40	44	48	52
MFQ*	5	X	X	X	X	X	X	X	X	X	X	X	X	X	X
RCADS-anx-15	3	X	X	X	X	X	X	X	X	X	X	X	X	X	X
CADLIS-Y	3	X	X	X	X	X	X	X	X	X	X	X	X	X	X
PGI-I	1		X	X	X	X	X	X	X	X	X	X	X	X	X
C-SSRS	5	X[a]	X[a]	X[a]	X[a]	X[a]	X[a]	X[a]	X[a]	X[a]	X[a]	X[a]	X[a]	X[a]	X[a]
CBQ-A	3	X													
GBO	10	X	X	X	X	X	X	X	X	X	X	X	X	X	X

*includes 4 items related to suicidal ideation
[a]Only completed if MFQ indicates some suicidal ideation.
[b]Two weeks after starting therapy (either CBT or BPI)

NB: If there is 8 weeks of remission (MFQ<22), participant may discontinue MBC measures while continuing research measures and still be considered adherent to the research protocol.

At the beginning of the assessment package, youth will be made aware that the results of these assessments are not monitored in real-time and that if they consider themselves to be a risk to their own safety or the safety of someone else, to consider going to the emergency room for assessment.

6 – METHODS: YOUTH INFORMANT RESEARCH MEASURES

For sites 1-3 only, these measures will be either completed by youth as self-report or measures completed by an evaluator (e.g., research analyst or clinician) after an interview where the youth is the main informant in both arms of the study (i.e., CARIBOU-2 intervention and TAU). They will be used to test our hypotheses and exploratory aspects with respect to clinical effectiveness of the CARIBOU-2 intervention, assuming it is delivered with good fidelity. Blinding youth, families, or clinicians to the type of treatment provided is not feasible.

6.1 Baseline Measures:

The Mood and Feelings Questionnaire child, long version (MFQ; Angold & Costello, 1987): As previously described. In addition to an MBC measure, it will also be used a screening measure to determine eligibility for the study.

Demographic Form: A standard demographic form will be used to capture multiple aspects of 'he youth's profile at baseline. These include questions around sex assignment at birth, current gender identity, sexual orientation, caregiving arrangements (i.e., parents/guardians), socio-economic status, and ethnicity.

Kiddie Schedule for Affective Disorders and Schizophrenia - Lifetime version (K -SADS-PL DSM-5 November 2016; Kaufman et al., 2016): The KSADS is a semi-structured interview that combines dimensional and categorical assessment approaches to diagnose current and past episodes of psychopathology in children and adolescents (ages 6-18 years old) according to DSM-5 criteria. This measure will be used to describe the participants in the study.

The Childhood Interview for Borderline Personality Disorder (CI -BPD; Zanarini, 2003): This is a semi-structured RA-rated interview which assesses for the 9 DSM-IV-TR criteria for Borderline Personality Disorder (identical to DSM-5 criteria) requiring that symptoms be "definitely present" for 2 or more years to be considered to meet the threshold. It has been validated in adolescent samples (Sharp et al, 2012). In a longitudinal study, it was found that symptoms of BPD in adolescents were associated with poor outcome (Winograd, Cohen, & Chen, 2008). It will be collected to describe our baseline sample and potentially explored of a predictor or moderator of treatment outcome.

Beck Hopelessness Scale -Revised (BHS -**R; N'ufeld, O'Rourke, & Donnelly, 2010):** A 20-item self-report revision of the original Beck Hopelessness Scale (BHS; Beck & Steer, 1998), that measures negative attitudes about one's future and perceived inability to avert negative life occurrences. The three aspects of hopelessness (negative feelings about the future; loss of motivation; and pessimistic expectations) are measured on a four-point Likert-type scale. It will be collected to describe our baseline sample and potentially explored as a predictor or moderator of treatment outcome.

Revised Children's Anxiety and Depression Scale -25– Anxiety Subscale (15 items) (RCADS -anx-15; Chorpita et al, 2000; Ebesutani et al. 2012): As previously described. In addition to being an MBC measure in the ICP arm, the RCADS-anx-15 will be administered at baseline in both treatment arms as a potential predictor of treatment response.

Conflict Behaviour Questionnaire -Child About Caregiver (CBQ-A; Robin & Foster, 1989): As previously described. In addition to being an MBC measure in the ICP arm, we will administer the CPQ-A at baseline in both treatment arms to examine child-caregiver conflict as a potential predictor of treatment response.

6.2 Primary outcome measure

The Mood and Feelings Questionnaire child, long version (MFQ; Angold & Costello, 1987): As previously described. In addition to an MBC measure, it will also be used a screening measure to determine eligibility for the study. It is also the primary outcome measure at 24-weeks.

6.3 Secondary clinical outcome measures

The Columbia Suicide Severity Rating Scale (C -SSRS; Public Health Agency of Canada, 2006): As previously described. In addition to being an MBC measure, it is also an important exploratory outcome with respect to our research. As a research measure, it will be captured at baseline with the highest lifetime rating on each subscale will be obtained as outcomes to explore potential as a risk factor for various outcomes. At baseline and 6-month follow-up, highest rating in the past 6 months on each subscale will also be captured in order to assess longitudinal patterns.

Depression Rating Scale (DRS; Kaufman et al., 2016): An RA-rated 13-item subscale found in the K-SADS-PL DSM-5 November 2016 used to assess for DSM-5 criteria of Major Depressive Disorder. The RA will be blind to treatment arm when conducting this assessment. This will be used to compare with the MFQ over time to examine how much the scale scores overlap with DSM-5 criteria.

The Youth Quality of Life Scale Research Version (Edwards et al., 2002; Patrick et al., 2002): The YQOL-R is a 41 item self-report scale measuring the broad array of constructs including sense of self-worth, quality of relationships, sense of agency and life satisfaction. The YQOL-R showed reliability, content validity, construct validity and criterion validity (Edwards et al, 2002; Patrick et al, 2002). This measure will be used as part of the economic evaluation (see below). There is no time reference window in the stem in the original version. We have added a 12-week time reference window for the purposes of the study.

Clinical Global Impression Scale – both the Improvement Subscale (CGI -I) and Severity Subscale (CGI -S) (Busner & Targum, 2007) : The CGI-I is a one-item assessment of the clinician's sense of change in a patient's symptom burden and overall functioning. Seven response options are

available with scores ranging from 1 through 7: “very much improved”, “much improved”, “a little improved”, “no change”, “a little worse”, “much worse”, and “very much worse”. In a previous study, the CGI-I has demonstrated good inter-rater reliability at identifying responders – defined as cases rated to be “very much improved” or “much improved” (Asarnow, Emslie, Clarke et al. 2009). This measure is used broadly across RCTs for adolescent depression (Courtney et al, 2021) and will be used to select the clinical significance of our finding and render findings comparable to other studies.

Health and Social Service Utilization (HSSU) (Browne et al., 1990; 1994): A 10-item, self-report questionnaire that collects data regarding services and medications used by the patient.

Ontario Perception of Care Tool for Mental Health and Addictions (CAMH, 2019): This evidence-based tool standardizes how substance use, mental health, and concurrent disorder services obtain client perception of care feedback, which can be used to make valuable service improvements. The tool was developed by CAMH and there are 2 versions of the OPOC-MHA (one for registered patients and one for family members of clients) and typically consists of 38 items on a Likert scale.

Adolescent Alcohol and Drug Involvement Scale Grid (AADIS) (Moberg, 2000): Scale used for adolescent substance abuse frequency. It includes both alcohol and other drug abuse. This will be used to describe the sample.

Cognitive Behavioural Therapy Skills Questionnaire (CBTSQ) (Jacob et al., 2011): A 16-item measure utilizing a Likert scale from 1 (I don’t do this) to 5 (I always do this) that measures an individual’s cognitive behavioral skills. The measure has demonstrated good internal consistency (Jacob et al., 2011) and is used to assess an individual’s cognitive restructuring and behavioral activation skills (Webb et al., 2016). This measure was added as youth have identified that coping strategies are an important outcome to them. It will also be explored as a mediator of the primary outcome.

CollaboRATE (Elwyn, 2013) A 3 item measure on a 5-point Likert Scale that measures the level of shared decision making in the clinical encounter from the patient's perspective, as part of assessing health care quality and provider performance. It will be explored as a mediator of the primary outcome. It was designed to be a brief patient-reported measure based on the two core elements of shared-decision making: i) provision of information or explanation to the patient about the relevant health issues or possible treatment options and ii) elicitation of the patient’s preferences related to the health issues or treatment options. Key terms such as ‘decisions’ and ‘preferences’ were deliberately avoided as these terms are viewed as barriers, potentially being unfamiliar to youth and also indirectly assuming the patients take an active role in the decision-making process

- Psychometric properties
    - ColllaboRATE demonstrates discriminative validity, intra-rater reliability, and sensitivity to change (Barr et al., 2015)
- Relation to existing literature:
    - Its three items correspond to match the three items in SDM-Q-9 to best predict patient satisfaction and match the existing conceptualization of SDM
- Practice implications:
    - The ability for completion in less than 30 seconds, and across a range of routine settings
    - Will serve as a performance metric or feedback tool at the provider level by aggregating a large number of responses

NSSI: The NSSI comes from the Self-Injurious Thoughts and Behaviors Interview (SITBI). The SITBI is a structured interview that assesses the presence, frequency, and characteristics of a wide range of self-injurious thoughts and behaviors, including suicidal ideation, suicide plans, suicide gestures, suicide attempts, and nonsuicidal self-injury (NSSI). The SITBI has high validity compared to other

measures of NSSI (average kappa = .87). The NSSI section of the SITBI has been adapted for the purposes of the CARIBOU-2 study.

6.4 Qualitative information

Youth Participant Interviews : For sites 1-3, individual interviews will be conducted with youth participants enrolled in the intervention arm of the study and who completed at least one psychotherapy session, and one measurement-based care team review, as part of the pathway. Using a purposive sampling matrix, interview participants will be selected and invited to participate in a 60-minute interview, conducted by a trained Research Analyst (RA), using a semi-structured interview guide. Sample size will be ultimately determined by the purposive sampling matrix and by thematic saturation; estimated sample size to achieve these goals is expected to be 15-20 participants (Guest, Bunce and Johnson, 2006; Hennik, Kaiser and Marconi, 2017). Interviews will be audio/video recorded and transcribed through WebEx.

6.5 Youth Informant Measures collected as part of an embedded sub study examining psychometric properties of MFQ, CADLIS -Y (see Appendix J for details on how this information will be used).

The Mood and Feelings Questionnaire child, long version (MFQ; Angold & Costello, 1987): As previously described. In addition to being an MBC measure, it will also be collected in the TAU group, without feedback to the clinicians, in order to examine its psychometric properties (see Appendix J).

Revised'Children's Anxiety and Depression Scale-25**–** Anxiety Subscale (15 items) (RCADS -anx-15; Chorpita et al, 2000; Ebesutani et al. 2012): As previously described. In addition to being an MBC measure, it will also be collected in the TAU group, without feedback to the clinicians, in order to examine the psychometric properties of the CADLIS measures (see Appendix J).

The Child Anxiety and Depression Life Interference Scale -Youth Version (CADLIS -Y; Lyneham et al., 2013): As previously described. In addition to being an MBC measure, it will also be collected in the TAU group, without feedback to the clinicians, in order to examine its psychometric properties (see Appendix J).

Adapted versions of the PGI -I and CGI-I improvement items: Within the TAU group, we will administer an adapted version of the PGI-I Improvement to ask separately about change in symptoms and functioning. These versions will be administered right after the MFQ, and after the CADLIS, respectively, and the question wording will be adjusted to refer to the construct captured in each of these measures (i.e., symptomatology and functional impairment). After completing the MFQ and CADLIS, respectively, Youth will be asked to indicate: "Check the number that best describes how you feel you have changed in the areas described on this page, comparing how things are now with how they were when you started treatment". Similarly, we will use an adapted version of the matching CGI, where we will ask clinicians their perspective on improvement as it pertains specifically to depressive symptoms, and separately, with respect to functioning.

Adapted versions of the PGI Severity and CGI Severity items: The CGI Severity (CGI-S) is a one-item assessment that asks the clinicians "Considering your total clinical experience with this particular population, how mentally ill is the patient at this time?". Seven response options range from "not at all" to "among the most extremely ill patients". Clinicians are instructed to base their rating upon observed and reported symptoms, behavior, and function in the past seven days. We will adapt the CGI-S to ask about severity of symptomatology and severity of functional impairment separately. Within the TAU group, we will administer the CGI-S as an anchor to calibrate severity cut-offs on the MFQ and CADLIS to facilitate the interpretation of symptom and impairment scores. We will review and adapt the wording of the CGI-S jointly with clinicians as needed, to adjust wording that may be perceived as stigmatizing (e.g., "mentally ill"). Similarly, the PGI Severity (PGI-S) is a one-item assessment asking the patient to rate the severity of their symptoms and impairment. We will review and adapt the

wording of the PGI-S jointly with youth advisors as needed and will administer the PGI-S in the TAU group separately with reference to depressive symptoms and functional impairment.

6.6 Timing of Youth Participant Informant Research Measures
Refer to Table 3 for details on the administration of the effectiveness research measures. Customized demographic information and treatment history forms will be administered to youth participants to describe the basic characteristics of the recruited sample. Refer to Table 4 for details on the administration of additional measures for the psychometric sub studies.

Table 3. Assessment Schedule for Summative Youth -Informant Research Measures for Sites 1 - 3 (RA, Youth, or Clinician -Rated)

Measures	Rater	Time (min)	Weeks					
			0	4	12	24	36	52
Baseline only								
MFQ	RA	5	ALL					
K-SADS-PL DSM-5	RA	60	ALL					
CI-BPD	RA	15	ALL					
Lifetime C-SSRS	RA	10	ALL					
Demographic Form	Youth	5	ALL					
BHS	Youth	10	ALL					
CBQ-A	Youth	10	ALL					
AADIS grid	Youth	5	ALL					
NSSI (Lifetime)	RA	5	ALL					
Longitudinal								
CGI-I overall	Clinician	1	ALL	ALL	ALL	ALL	ALL	ALL
6-month C-SSRS	RA	10	ALL			ALL	ALL	ALL
DRS	RA	5	ALL (in KSADS)		ALL	ALL	ALL	ALL
HSSU	RA	5	ALL		ALL	ALL	ALL	ALL
Youth Interview	RA/RC	90						~20 ICP youth
Youth Quality of Life	Youth	15	ALL		ALL	ALL	ALL	ALL
Collaborate	Youth	2		ALL	ALL	ALL	ALL	ALL
OPOC	Youth	5			ALL	ALL	ALL	ALL
CBTSQ	Youth	5	ALL	ALL	ALL	ALL	ALL	ALL
COVID-19	Youth	5	ALL	ALL	ALL	ALL	ALL	ALL
NSSI (6mnth & 1 year)	RA	5				ALL		ALL

The following are justifications for our time points:

- 4-week time point: Early response predicts outcome (Gunlicks Stoessel, 2019).
- 12-week time point: Major studies like TADS (March et al., 2004), TORDIA (Brent et al., 2008), IMPACT (Goodyer, 2017) have this time point to facilitate data comparison.
- 24 weeks is the primary end time point: To allow enough time for youth to proceed through much of the pathway, without concern about contamination effects.
- 36 weeks: To allow further exploration of effects of the CARIBOU-2 intervention over time, though contamination effects may be problematic.
- 52-week time point: To allow further time to explore effects of the CARIBOU-2 intervention, though contamination effects may be problematic.

Table 4. Assessment Schedule for Additional Youth Informant Measures Informing the Psychometric Sub Studies for Sites 1 -3 (Appendix J)

Measures	Rater	Time (min)	Eligibility Assessment		Weeks					
				0	2	4	12	24	36	52
	Longitudinal									
MFQ	Youth	5	X	TAU[a]		TAU[a]	TAU[a]	TAU[a]	TAU[a]	TAU[a]
RCADS-15-anx	Youth	5		TAU[a]		TAU[a]	TAU[a]	TAU[a]	TAU[a]	TAU[a]
CADLIS-Y	Youth	5		TAU[a]	N=100 TAU[b]	TAU[a]	TAU[a]	TAU[a]	TAU[a]	TAU[a]
PGI-S depression	Youth	1		TAU[a]		TAU[a]	TAU[a]	TAU[a]	TAU[a]	TAU[a]
PGI-S function	Youth	1		TAU[a]		TAU[a]	TAU[a]	TAU[a]	TAU[a]	TAU[a]
PGI-I depression	Youth	1				TAU[a]	TAU[a]	TAU[a]	TAU[a]	TAU[a]
PGI-I function	Youth	1			N=100 TAU[b]	TAU[a]	TAU[a]	TAU[a]	TAU[a]	TAU[a]
CGI-S depression	Clinician	1		TAU[a]		TAU[a]	TAU[a]	TAU[a]	TAU[a]	TAU[a]
CGI-S function	Clinician	1		TAU[a]		TAU[a]	TAU[a]	TAU[a]	TAU[a]	TAU[a]
CGI-I depression	Clinician	1				TAU[a]	TAU[a]	TAU[a]	TAU[a]	TAU[a]
CGI-I function	Clinician	1				TAU[a]	TAU[a]	TAU[a]	TAU[a]	TAU[a]

[a] Recruitment into the psychometric sub study for analysis of longitudinal measurement properties will be from TAU only, to manage respondent burden in ICP.

[b] The CADLIS-Y will be re-administered to 100 youth in the TAU group at a 2-week time point along with the PGI-I-function to assess its test-retest reliability as part of a psychometric sub study (see Appendix J).

7 – METHODS: CAREGIVER INFORMANT MEASURES

For sites 1-3, these measures are administered to participating caregivers (e.g., Parents/Guardians) in both study arms (i.e., TAU and ICP). The data will be used to further test our hypotheses and the

exploratory aspects of the clinical effectiveness of the CARIBOU-2 intervention, assuming it is delivered with good fidelity.

7.1 Baseline measures

Demographic Form: A standard demographic form will be used to capture multiple aspects of the caregiver's profile at baseline. These include questions around sex assignment at birth, current gender identity, sexual orientation, socio-economic status, and ethnicity.

Kiddie Schedule for Affective Disorders and Schizophrenia - Lifetime version (K -SADS-PL DSM-5 November 2016; Kaufman et al., 2016): The KSADS is a semi-structured interview that combines dimensional and categorical assessment approaches to diagnose current and past episodes of psychopathology in children and adolescents (ages 6-18 years old) according to DSM-5 criteria. This measure will be used to describe the participants in the study.

Health and Social Service Utilization (HSSU; Browne et al., 1990 & Browne et al., 1994): A 10-item, self-report questionnaire that collects data regarding services and medications used by the youth.

Conflict Behaviour Questionnaire -20-Parent (CBQ -20-P; Robin & Foster, 1989): A 20-item measure of perceived parent/caregiver-child communication conflict behavior at home. Caregivers retrospectively rate (i.e., “True” or “False”) their interactions with their adolescent children over the 2 or 3 weeks preceding the assessment session. The cut-off reference is a total score of 11 or greater (Rengasamy et al., 2013). This measure will be compared with the CBQ-A at baseline to see if there are discrepancies.

Collaborate -Parent Version (Elwyn et al., 2013): A 3 item measure on a 5-point Likert Scale that measures the level of shared decision making in the clinical encounter from the patients’ caregiver perspective, as part of assessing health care quality and provider performance. This will be explored as a mediator of the primary outcome.

7.2 Secondary clinical outcome measure

The Child Behaviour Checklist – Parent Report Form (CBCL; Achenbach, 2001): A 118-item caregiver-rated measure assessing the youth’s behaviour and general psychopathology. It is a widely used measure with known population norms. One-week test-retest reliability was found to be 0.80-0.94. Internal consistency is reported to be high; inter-rater reliability (e.g., between two parents) was found to be moderate to high (reference needed). All subscales will be used at baseline to describe general psychopathology. The internalizing broadband scale of the CBCL (i.e. anxious-depressed, depressed-withdrawn, and somatic subscales combined) will be measured longitudinally to get an impression of how the caregiver is observing any changes in mood or anxiety with treatment.

Depression Rating Scale (DRS; Kaufman et al., 2016): An RA-rated 13-item subscale found in the K-SADS-PL DSM-5 November 2016 used to assess for DSM-5 criteria of Major Depressive Disorder. The RA will be blind to treatment arm when conducting this assessment. This will be used to compare with the MFQ over time to examine how much the scale scores overlap with DSM-5 criteria

7.3 Qualitative Information

Caregiver Individual Interview : For sites 1-3, caregiver individual interviews will be conducted to further explore the experience of caregivers during CARIBOU-2 intervention. Caregivers who consented to participate in the CARIBOU-2 Study will be eligible for a 60-minute, semi-structured individual interviews once the following conditions are met:

1. The youth in their care has completed at least one psychotherapy session.
2. The youth has participated in at least one team review.
3. The youth has exited the CARIBOU-2 pathway.

From the pool of eligible caregivers, participants will be selected using a purposive sampling matrix. This approach—monitored and implemented on a rolling basis by the Research Coordinator—ensures inclusion of individuals with relevant lived and living experiences (Campbell et al., 2020; Palinkas et al., 2015). Participants will be stratified across the following categories, with the final sample aiming for one to two individuals per category (with expected overlap):

a. Caregiver gender identity (Cis-Male, Cis-Female, Gender Expansive)
b. Caregiver race and ethnicity (White, Non-White)
c. Participating agency (Site)
d. Caregiver participation in Mood Foundation (Component 2: Psychoeducation) (Yes or No)
e. Caregiver participation in CARE Groups (Component 4: Caregiver Support) (Yes or No)
f. Youth's prior service access (Yes or No)
g. Youth's case severity (Moderate or High)
h. Psychotherapeutic modality received by youth (CBT Group, CBT Individual, BPI)
i. Youth's pathway completion (Graduated, Exited Before Graduation)

Caregiver interviews will be conducted by a trained Research Analyst (RA), using a semi-structured interview guide. Sample size will be ultimately determined by the purposive sampling matrix and by thematic saturation; estimated sample size to achieve these goals is expected to be 15-20 participants (Guest, Bunce and Johnson, 2006; Hennik, Kaiser and Marconi, 2017). Interviews will be audio/video recorded and transcribed through WebEx.This information will be used in consideration of future iterations of the ICP. Guides for these interviews are included in Appendix F.

7.4 Caregiver Information Measure collected as part of an embedded sub study examining psychometric properties of MFQ, and CADLIS (see Appendix J for details on how this information will be used).

The Child Anxiety and Depression Life Interference Scale -Parent/Caregiver Version (CADLIS -P; Lyneham et al., 2013): This measure is a 9-item complement to the youth version, looking at similar constructs from the caregiver (e.g., parent) perspective. Since the CADLIS has not been validated with Canadian caregivers to date, we will undertake a comprehensive examination of its psychometric properties in the study sample in order to appraise its validity and reliability as an outcome measure (see Appendix J). It will be collected longitudinally in the TAU group and in the ICP at baseline only, without feedback to the clinicians, for this purpose

Adapted version of the PGI -I (PGI-I-function -cg): Within the TAU group, we will administer an adapted version of the PGI-I Improvement to ask the caregiver participant specifically about change youth functioning. This version will be administered right after the CADLIS-P and the question wording will be adjusted to reflect that the caregiver is the informant, not the youth. After completing the CADLIS-P, the caregiver will be asked to indicate: "Check the number that best describes how you feel you have changed in the areas described on this page, comparing how things are now with how they were when you started treatment". Response options correspond to the original PGI-I.

7.5 Timing of Caregiver Participant Measures Refer to Table 5 for details on the administration of these caregiver participant research measures.

Table 5. Assessment Schedule for Caregiver Informant Research Measures for Sites 1 -3

Measures	Rater	Time (Mins)	Weeks						
			0	2	4	12	24	36	52
K-SADS-PL DSM-5	RA	60	ALL						
HSSU	RA	10	ALL			ALL	ALL	ALL	ALL
DRS	RA	10				ALL	ALL		
Individual Interview	RC/RA	60							15-20 Caregivers
Demographic Form	Caregiver	5	ALL						
CBCL	Caregiver	10	ALL				ALL		ALL
CADLIS-P	Caregiver	5	ALL	N=100 (TAU)[a]	TAU	TAU	TAU	TAU	TAU
CBQ-20	Caregiver	5	ALL						
PGI-I-Function	Caregiver			N=100 (TAU)[a]	TAU	TAU	TAU	TAU	TAU
Collaborate-P	Caregiver	2			ALL	ALL	ALL	ALL	ALL

[a] The CADLIS-P will be re-administered to 100 caregivers in the TAU group at a 2-week time point to assess its test-retest reliability as part of a psychometric sub study (see Appendix J).

8 – METHODS: IMPLEMENTATION RESEARCH MEASURES

To assess the implementation outcomes, assessments at the level of the clinic will be held continually throughout the study period. The implementation outcomes are based on the Implementation Outcome Taxonomy outlined by Proctor and colleagues (2011). Implementation outcomes that will be examined in the current study are feasibility, fidelity, cost-effectiveness, and acceptability. Implementation outcomes that are outside the scope of evaluation for this particular study include: adoption, penetration and sustainability, though some may be explored in future studies.

8.1 Feasibility:

For sites 1-3, feasibility is “the extent to which a new treatment, or an innovation, can be successfully used or carried out within a given agency or setting” (Procter et al, 2011). If youth are engaging with the various components of the pathway, assuming fidelity to the model, it can clearly be said that the intervention is feasible. Using the youth participant engagement checklist, prospective chart review, and RA communication with clinicians, the proportion of indicated components the youth participant attends will be tracked over time. We anticipate that youth will engage in ≥75% of indicated treatment components while they are in the pathway.

8.2 Fidelity

Applicable to sites 1-3. Appendix I includes each of the fidelity measures listed below.

8.2.1 Fidelity to the Implementation Process:
In order for our results to be replicable, the implementation process set out in section 4.2 will need to be followed with fidelity. We have created a checklist for this process to ensure replicability. This fidelity measure is with reference to tasks completed by the implementation teams.

8.2.2 Fidelity to the Overall CARIBOU -2 intervention:
Fidelity to the overall CARIBOU-2 intervention will be assessed by a checklist for each youth participant as they proceed through the pathway. This fidelity measure is with reference to tasks to be completed by direct clinicians Items include whether or not youth were offered the treatment components of the ICP as indicated, whether or not measurement-based care was offered and whether or not team reviews were offered every 4 weeks until exit from the pathway. Items are similar to the Youth Participant Engagement checklist; however, these items are monitoring factors under the clinicians' control. We anticipate that clinicians will have an average of ≥75% fidelity to the CARIBOU-2 intervention.

8.2.3 Fidelity to Components of the ICP:
For measuring fidelity to CBT, all sessions will be audio recorded. For Mood Foundations, at least 5 sessions will be chosen randomly to be assessed for fidelity. For other psychotherapy sessions, 20% of all sessions will be selected randomly to assess for fidelity. RAs will score the fidelity based on the recordings. ≥75% fidelity on average is targeted for each component.

8.2.3.1 Fidelity to the Mood Foundations Session, Medication algorithm, and Caregiver support group will be measured by locally-created checklists, developed a priori.

8.2.3.2 Fidelity to CBT will be assessed using a locally developed scale to assess the delivery of content that corresponds to the CARIBOU CBT manual, based on the Cognitive Therapy Rating Scale (Young et al., 1981) for process elements of CBT.

8.2.3.3 Fidelity to BPI will be assessed using the Brief Psychosocial Intervention Adherence Scale. This is an 8-item scale created by the BPI developers. It was also used in the IMPACT trial (Goodyer et al, 2017) to assess fidelity to the therapy.

8.3 Economic Evaluation :
One of the objectives of the trial is to determine the cost-effectiveness of the CARIBOU model at sites 1-3 compared to TAU. To that end, we will perform two analyses, a cost-effective analysis (CEA) and a cost-utility analysis (CUA), to determine the incremental costs of the CARIBOU-2 intervention relative to TAU in improving health outcomes as measured by the MFQ (for the CEA) and quality-adjusted life years (QALYs) measured with a validated utility instrument, the YQOL-R (for the CUA). The HSSU tool and custom data collection tools have been developed to measure direct costs to the patient, as well as indirect costs (e.g., lost income due to appointments). These data will be collected at all assessment times. Additional data regarding direct costs to the health care system will be obtained through ICES, an independent non-profit research institute which holds administrative health care and cost data for health services, such as physician billings and emergency departments visits, like for the province of Ontario. The associated analyses are outlined in section 9.3.

8.4 Acceptability, barriers, and facilitators of implementation:
In order to qualitatively assess acceptability of the CARIBOU-2 intervention at all six sites, as well as facilitators and barriers of implementation, we will assess the 39 implementation constructs from the Consolidated Framework for Implementation Research (CFIR) outlined by Damschroder et al 2009. One-to-one interviews or focus groups between research staff and clinical staff using the CFIR interview (see Appendix F) will be conducted 6 months to one year after initial implementation to assess these constructs. Three staff at each site, each with high involvement in the delivery of the CARIBOU-2 intervention, will undergo this interview process.

The CFIR is composed of five major domains, each of which may affect an intervention's implementation:

a) Intervention characteristics: Features of an intervention that might influence implementation. It consists of eight constructs (e.g., stakeholders' perceptions about the relative advantage of implementing the intervention, complexity).

b) Inner setting: Features of the implementing organization that might influence implementation. It consists of twelve constructs (e.g., implementation climate, leadership engagement).

c) Outer setting: Features of the external context or environment that might influence implementation. It consists of four constructs (e.g., external policy and incentives).

d) Characteristics of individuals: The characteristics of the individuals involved in implementation that might influence implementation. It consists of five constructs (e.g., knowledge and beliefs about the intervention).

e) Implementation process: Strategies or tactics that might influence implementation. It consists of eight constructs (e.g., engaging appropriate individuals in the implementation and use of the intervention, reflecting, and evaluating).

Clinician participants at all six sites will be invited to complete a short demographics questionnaire administered through REDCap. The survey is self-administered and is designed to be completed at the participant's convenience. Completion time is approximately 5 minutes.

8.5 Timing of Implementation Outcome Capture:
Implementation outcomes will follow the timeline outlined in table 6 below.

Table 6. Assessment Schedule for Implementation Outcomes Sites 1 -3

	Month Since Initiation					
	-6	-3	0	12	24	52
TAU						
Economic Evaluation			x	------	------	x
ICP						
Feasibility: Youth participant engagement checklist.				x	------	□
Phase I Fidelity	x	X				
Phase II Fidelity		X	x			
Phase III Fidelity				x	------	□
ICP Fidelity				x	------	□
Mood Foundations Fidelity				x	------	□
CBT Fidelity				x	------	□
BPI Fidelity				x	------	□

Caregiver Group Fidelity				x	------	□
Medication Algorithm Fidelity				x	------	□
Cost	x	------	------	------	------	□
CFIR constructs: Qualitative Interview*						X

□ = until end of study period.
* =will be completed at all six sites

9 - METHODS: STATISTICAL PLAN

9.1 Clinical Effectiveness analysis:

For sites 1-3, descriptive data analysis will be first conducted to examine distribution of collected measures and evaluate whether there are significant differences across arms of assigned sites. Missing data patterns and outliers will also be carefully examined to provide insight for the next steps of analyses. Generalized linear mixed-effects model will be the primary analytic tool for evaluating whether the CARIBOU-2 intervention is more effective than TAU for adolescents with depression presenting to care with regards to improvement of depressive symptoms (Hypothesis A), self-reported functioning (Hypothesis B), caregiver-reported internalizing psychopathology (Hypothesis C), and suicidal ideation and behaviours (exploratory). Generalized linear mixed-effects model is a proper choice to control for covariates (e.g. demographics and baseline clinical measures), to accommodate multiple forms of the outcome (e.g. continuous, categorical and count type), and to account for clustering at individual level (for repeated measures) and at site level. Time, treatment assignment and their interactions will serve as the primary predictors for the analyses. As an example, if we let Y_{ijt} to denote a continuous outcome of the j-th participant of the i-th site measured at time t, a linear model for Y_{ijt} will look like the following:

$$Y_{ijt} = \beta_0 + b_{0,ij} + b_{1,i} + \beta_1 t + \beta_2 Group_{i,t} + \beta_3 Group_{it} * t + \beta_4 X_{ijt} + \epsilon_{ijt}$$

of which $Group_{i,t}$ denotes the treatment assignment of the i-th site at time t, X_{ijt}, additional covariates, $b_{0,ij}$ and $b_{1,i}$, random effects at individual and site levels respectively, ϵ_{ijt}, unexplained random error, and ***β's***, regression coefficients. For sensitivity analyses, we will explore the use of the piecewise model to model the time trend differently. We will adopt the intention-to-treat approach in general and use multiple imputation methods as the primary missing data strategy. In our pilot study, we have collected ~85% of expected longitudinal data points on the MFQ (primary clinical outcome), adjusting for attrition. Sensitivity analysis will be conducted to evaluate the impact of non-random missing and robust regression method will be used instead when the impact is high. SAS 9.4 will be our go-to software package for this project.

9.2 Power Calculation

Our power calculation for sites 1-3 is as follows; with our previously proposed six-step interrupted time series design that contains six sites each of 50 youth recruited (25 under ICP and 25 control), we expected to have sufficient power (.80) to detect a small to moderate effect size of 0.40, which is in line with the anticipated effect size learned from literature. To calculate the power, we conducted a Monte Carlo study with 50,000 replications to simulate the unique design that we have here. We also conservatively used .119 as the site level intra-class correlation and a 20% attrition rate. The power calculation was focused on the primary outcome with a pre-post analysis.

9.3 Implementation outcomes analysis:
Implementation outcomes will mainly use descriptive data analysis, including distribution of results from youth participant and clinician satisfaction questionnaires, the youth participant engagement checklist, and fidelity assessments (Hypotheses D, E, F, G).

Economic evaluation: The economic evaluations will take both the health care system and societal perspectives; these will include a cost-effectiveness analysis (CEA) and a cost-utility analysis (CUA) (Hypothesis H). The time horizon will be the youth's lifetime. For the CEA, we will account for the majority of direct costs to the health care system and to the patient/caregivers, as well as indirect costs. Costs will be adjusted for inflation to 2025 Canadian dollars using the Bank of Canada Consumer Price Index [Bank of Canada, 2000]. All health outcomes and costs will be discounted at 1.5% per year in line with Canadian Agency for Drugs and Technologies in Health guidelines. We will account for all costs associated with delivering the proposed intervention. This will include costs with personnel, supplies and services, equipment and program resources, among others. We will make use of a decision model in our CEA to help incorporate the benefits and costs beyond the time horizon of the existing data, and to help evaluate hypothetical scenarios. Long-term health-related costs for the TAU cohort will be determined through linkage to population-based administrative health care databases at ICES using encrypted unique patient identifiers (Jacobs & Yim, 2009). For each patient, we will determine the total utilisation of health resources and costs during the observation window.

The incremental cost-effectiveness ratio (ICER), the primary outcome of the CEA, will be calculated as the difference in discounted mean costs between the CARIBOU and TAU groups divided by the difference in the score obtained from Mood and Feelings Questionnaire. A 95% confidence interval around this estimate will be estimated non-parametrically using 1,000 bootstrap replications.

One-way deterministic sensitivity analyses will be performed to evaluate the robustness of our results. The ranges for the sensitivity analysis will be obtained from 95% confidence intervals. We will also perform a probabilistic sensitivity analysis using a second-order Monte Carlo simulation with 1,000 iterations. Using a net-benefit framework (Jacobs & Yim, 2009), a cost-effectiveness acceptability curve (CEAC) will be produced using varying willingness-to-pay thresholds. A CEAC provides the probability that an intervention is cost-effective, compared with the alternative, for a range of maximum monetary values that a decision-maker may be willing to pay for a particular unit change in the outcome analyzed. We will derive the CEAC from the joint distribution of incremental costs and incremental effects; we will use non-parametric bootstrapping of the observed data to estimate these joint distributions.

For the CUA, we will measure the change in utility as obtained from the Youth Quality of Life Scale, which will allow us to calculate quality adjusted life years (QALYs). The advantage of the CUA approach is that QALYs capture general well-being/disease burden directly from study participants. In addition, because the QALY is a widely utilized measure in economic evaluations, the results generated from this study will be easily compared to others.

As a subsequent analysis to the economic evaluation, we also propose to examine the real-world budget impact of implementing the CARIBOU integrated care pathway to hospitals that treat adolescents across the participating provinces. One of the main objectives of this analysis will be to determine how much the Ontario Ministry of Health would need to spend to implement this model of care either across the entire province or in a given jurisdiction. Another outcome of interest will be an estimate of the potential savings achieved by the system by implementing this intervention.

10 - PROJECT TIMELINE

Our data collection will end in December 2027. And end-of-grant knowledge translation will take place in 2028. The trial will be completed over the course of eight years. Figure 4 illustrates the CARIBOU-2 study timeline.

Figure 4. Project Timeline

	Sep 20-Feb 21	Mar 21-Aug 21	Sep 21-Feb 22	Mar 22-Aug 22	Sep 22-Feb 23	Mar 23-Aug 23	Sep 23-Feb 24	Mar 24-Aug 24	Sep 24-Feb 25	Mar 25-Aug 25	Sep 25-Feb 26	Mar 26-Aug 26	Sep 26-Feb 27	Mar 27-Aug 27	Sep 27-Feb 28	Mar 28-Aug 28	Sep 28-Feb 29
REB approvals and amendments																	
Site selection																	
Set up site training plan																	
Legal contracts																	
Participant recruitment																	
Data collection																	
Sites transition from TAU to ICP																	
Internal pilot assessment: feasibility outcomes																	
Formal data-check with DSMB																	
Data analysis																	
Post-study KT activities																	

11- RESEARCH TEAM:

11.1 Steering Committee:

The steering committee consists of: Dr. Darren Courtney, Dr. Melanie Barwick, Dr. Joanna Henderson, Dr. Stephanie Ameis and Dr. Peter Szatmari. Dr. Darren Courtney (15 hrs/wk contribution) is the nominated co-principal investigator and will oversee the project. He is the PI of the CARIBOU-1 pilot study. Dr. Melanie Barwick (2hrs/wk) is the other co-PI guiding the implementation science aspect of the study. She is an internationally renowned expert in implementation science and has had substantial experience with implementation of mental health initiatives for young people. Dr. Jo Henderson (2hrs/wk) is the Executive Director of Youth Wellness Hubs Ontario. Through this role, she has an intricate understanding of the mental health agency system in the community as well as other initiatives that are underway (eg. Ontario Structured Psychotherapy). She also has experience with conducting multi-centre RCTs in service delivery for youth with mental health concerns in community mental health agencies. Dr. Stephanie Ameis is an experienced clinician-scientist who will contribute her knowledge to the overall direction of the project. Dr. Peter Szatmari (2hrs/wk) is the senior investigator of the study guiding the aspects relating to conducting a clinical trial.

11.2 Others Co -investigators and Collaborators:

Dr. Brendan Andrade (<1hr/wk) is advising based on his experience studying multimodal mental health interventions in children and their caregivers, including large clinical trials. Dr. Kathryn Bennett (<1hr/wk) is an epidemiologist and methods specialist from McMaster University is advising on the implementation of clinical practice guidelines for child and adolescent mental health disorders. Dr. Wei Wang (<1hr/wk) is a statistician who will carry out the statistical analysis of the clinical outcomes. Dr. Lisa D. Hawke (1hr/wk) will advise on logistical organization of the study incorporation of the youth engagement initiative. She has played a lead role in the execution of the YouthCan study and will bring this experience to advise us on the conduct of the CARIBOU-2 study. Dr.Priya Watson (<1hr/wk) is a psychiatrist with expertise in “Negative Experiences in Psychotherapy” and will provide input on

how best to engage participants in treatment. Dr. Madison Aitken (1hr/wk) is a psychologist and will provide input on how to provide the caregiver intervention. Dr. Kristin Cleverley is a nurse researcher with expertise on care delivery in transitional-aged youth; she will provide input on how best to coordinate follow-up care for youth exiting the pathway. Dr Karolin Krause is a methodologist and will advise on outcome assessment methodologies (1hr/wk). Dr. Claire de Oliveira (<1hr/wk) is an expert in Health Economics and is advising on the Cost-Effectiveness and Cost-Utility Analyses. From 2021 to 2025, Dr. Bahar Amani serves as a Research Scientific Associate on the project, contributing substantial time across multiple phases of the study, including data collection, data management, analysis, and knowledge translation. As of November 2025, Dr. Amani continues to support the project in a reduced capacity (<1 hour/week) as a Research Methods Specialist with the Cundill Centre for Child and Youth Depression at CAMH. Dr. Amani's expertise in research methods, along with her in-depth knowledge of the CARIBOU-2 clinical trial design, remains integral to the project. Dr. Vasudha Gidugu (1 hour/week) is a Research Scientist at the SickKids Centre for Community Mental Health, one of the participating sites in the study, and provides input on the real-world implementation of the intervention. Dr. Carly Albaum (1 hour/week) is a psychologist with expertise in measuring psychotherapeutic treatment fidelity and provides key insights into the assessment of fidelity for the CBT component of the intervention. Dr. Albaum also supports the training of clinicians in the CARIBOU-2 intervention. Dr. Kiefer Cowie (1 hour/week) is a psychologist with extensive clinical experience in child and youth mental health and the delivery of psychotherapeutic interventions, including CBT, and contributes to fidelity measurement. Dr. Cowie also supports the training of clinicians in the CARIBOU-2 intervention. Dr. Victoria Iris Rae Anderson (1 hour/week) is a resident physician specializing in child and youth psychiatry and a co-developer of the CARIBOU-2 curriculum. She supports fidelity measurement to the CARIBOU-2 curriculum and the training of interventionists. Dr. Jia Hong "David" Dai (1 hour/week) is a psychiatry resident specializing in child and youth mental health, with extensive experience in delivering psychotherapy, including CBT, and contributes to the measurement of CBT fidelity.

11.3 Knowledge users:
Knowledge users are as follows: Mardi Daley and Gray Liddell are the project's main youth collaborators and bridge between the research team and the Youth Engagement Initiative at CAMH (see KT plan below).

11.4 International expert guidance:
Guidance will also be provided by the scientists on the International Advisory Board of the Cundill Centre for Child and Youth Depression. The board consists of international, leading experts in the field of youth depression: Drs. Allan Young (UK), Bonnie Zima (USA), and Sarah Hetrick (New Zealand).

12- ETHICAL, CONFIDENTIALITY AND SAFETY CONSIDERATIONS

12.1 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). Interviews and clinician meetings will be used to capture patients', families' and clinicians' impressions of TAU and the ICP. Adverse events including psychiatric hospitalizations, suicide attempts or deaths by any reason will be documented according to the protocols put in place at each site. The project's DSMB will also be notified. Any adverse event determinations made at the site level will be reported to the CAMH research team by local site clinical staff. Updates and outcomes will be documented accordingly using an Adverse Event Form, and once the event is resolved, the document will be timestamped.

12.2 Ethics, REB Approvals and Legal Contracts

We are studying how the intervention works externally. Youth participants will not be recruited from CAMH or the Hospital for Sick Children. CAMH is the research project lead site and we are seeking REB approval from CAMH as a result. The Hospital for Sick Children is providing support for the implementation of the intervention. Sites will have a variety of processes for ethics approval. Each site's research ethics approval process will be followed once sites are selected. Legal contracts with each site will also be completed after site selection is complete.

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). We are seeking approvals from The Research Ethics Board (REB) at CAMH and from The Hospital for Sick Children. Once participating mental health agencies are confirmed we will undergo the required research ethics procedures of each site. Any protocol modifications will be communicated to all relevant parties (co-investigators, REBs, ClinicalTrials.gov), as they are made.

12.3 Data Management and Sharing

Data Requestors (i.e., researchers) can only request de-identified or aggregated data, from this study, for their own research purposes. The Research Coordinator will be responsible for ensuring that all required steps are taken prior to sharing study data with data requestors. Please see Appendix H for the CAMH policy.

The Research Coordinator will be responsible overseeing the coordination of study visits. There will be multiple RAs on the research team. Data will be managed by the Research Coordinator, using the REDCap system. All electronic data files will be encrypted and password-protected. Any paper files will be kept in a locked file cabinet. All audio and/or video WebEX recordings will be securely transferred from the researcher's CAMH WebEX account to the project's folder on the X-Drive, which is located on secure CAMH servers. The transfer will occur once the recording files become available for download. Files saved on the CAMH WebEX account will be deleted within 30 calendar days. CAMH, and The Hospital for Sick Children are internationally recognized centers for delivering impactful and relevant research. The breadth and depth of resources available through these institutions enable such high quality research (e.g., YouthCAN (Henderson et al., 2017). The Cundill Centre for Child and Youth Depression, and other research centres at CAMH work collaboratively together to pool resources and infrastructure further solidifying our ability to provide robust scientific outputs.

Participant qualitative interviews will be audio and video recorded. The recording will be transcribed through WebEX and then verified by an RA for accuracy and to remove or change any information that could directly identify the participant. All WebEX recordings and transcriptions will be securely transferred from the researcher's CAMH WebEX account to the project's folder on the X-Drive, which is located on secure CAMH servers. The transfer will occur once the recording and transcription files become available for download. Files saved on the CAMH WebEX account will be deleted within 30 calendar days.

12.4 Data Safety Monitoring Board

The Data Safety Monitoring Board will consist of three RCT and data monitoring experts that are not investigators on this study. These experts will be responsible for:

1. Safeguarding the interests and safety of current and future study participants
2. Preserving integrity and credibility of the trial in order that future patients are treated optimally
3. Identifying Severe Adverse Events (i.e., death by any cause, suicide attempt, hospitalization for mental health reasons and potential link to intervention)

The following people have agreed to be on the DSMB for this project, with the intention to have our first meeting in March 2021 to review the charter and protocol:

1) Dr. Lehana Thabane, PhD, Professor and Associate Chair, Health Research Methods, Evidence, and Impact, McMaster University.
2) Dr. Suneeta Monga, MD, Associate Psychiatrist-in-Chief, Department of Psychiatry at the Hospital for Sick Children. Associate Professor, Department of Psychiatry at the University of Toronto.
3) Dr. Khrista Boylan, PhD, Associate Professor, University of McMaster

13- KNOWLEDGE TRANSLATION PLAN

13.1 Integrated Knowledge Translation:
Throughout the CARIBOU-1 pilot study, the Youth Engagement Initiative (YEI, Heffernan et al 2017) has been involved in the study design, selection of materials, selection of outcome measures and interpretation of the results. The YEI will continue to be involved to include the voice of youth partners (many of whom have lived experience) in the design of this study. Mardi Daley is our main contact with the YEI. We will also be including input from caregivers in the CARIBOU-2 study through a similar process. Connie Putterman is our main contact for this initiative and is just getting started. Moreover, our lead implementation team also consists of implementation specialists that work in various hospitals and mental health agencies in the community. As such, there is a heavy emphasis on integrated knowledge translation in this project.

13.2 Post-study Knowledge Translation:
After study completion, we will share our results with patients and families, including those who participated in the study, using tailored dissemination deliverables intended to share knowledge and inform decision-making. These deliverables will be co-developed by the investigators, a communications specialist, and youth and caregiver partners.

To inform researchers in the field of child and youth mental health and implementation science, we will present the findings to other researchers at national and international conferences meeting in the coming years and will publish the findings in peer-reviewed journals for the purpose of informing research and practice.

To inform policy and decision-makers, we provide feedback to the NICE Guideline Development Group on the process of utilizing the recommendations, which may be used for further refinement of the guideline-development process. Co-investigators on this project, Drs. Szatmari and Bennett, are in the process of collaborating with policy and decision-makers at the Canadian Academy of Child and Adolescent Psychiatry and the Academic Professors of Child and Adolescent Psychiatry of Canada to develop and implement a national strategy for implementation of high-quality guideline recommendations in academic clinical centres across the Canadian context.

To inform health practitioners, we will share our psychoeducation materials and updated CBT treatment manual developed with intensive partnership from youth collaborators and clinicians. These materials will be disseminated broadly and will be freely available to practice partners. Interest in accessing these materials from practitioners outside of CAMH has already been widely expressed.

14- RELEVANCE AND IMPACT

14.1 Service Development
Should our findings show that the CARIBOU-2 intervention is effective in reducing depressive symptoms and can be effectively implemented with fidelity; efforts to scale up the intervention to other centres would be indicated. Should our results not show differences between TAU and the CARIBOU-2 intervention, whether these relate to failures of the intervention and/or its implementation, then

further adaptation to the intervention and/or the implementation approach will be required, along with further effectiveness testing.

14.2 Informing Future trials:
The hybrid effectiveness-implementation design is a relatively novel approach to research in child and youth mental health. Through the process of conducting this trial, our findings will inform implementation science and future trials.

15- STUDY MONITORING, AUDITING, AND INSPECTING

15.1 Critical to Quality (CtQ) Factors
Critical to quality (CtQ) factors are attributes of a clinical research study that have a meaningful impact on participants' rights, safety and well-being, and reliability of study data/results. The Risk Mitigation Plan table below lists the CtQ factors and the risk mitigation plan for each factor. Refer to the Critical to Quality (CtQ) Factors Identification Worksheet for the documented assessment of all risk factors for this study and their scores based on their likelihood, detectability, and impact.

Risk Mitigation Plan Table

CtQ Factor	Risks to CtQ Factor	Required Activities	Risk Mitigation Plan
Informed Consent	Not obtaining informed consent from participant(s) prior to performing study activities	Obtain informed consent before any study specific procedures	• Use a Participant Consent Tracking Log to track and verify that the correct, REB-approved version of the consent form has been signed by all participants prior to completing any study-related tasks. • Train delegated site personnel on the informed consent process and proper documentation.
Delegation and Training	Individual performed study task without being trained and delegated	Complete research training and study-specific training prior to delegating study tasks to individuals on the task delegation log	• Verify completion of required training prior to delegation of an individual or delegation of new tasks to an individual on the Task Delegation Log. • Study personnel to set calendar reminders to prompt re-training 30 days prior to expiration.
Regulatory	Lapsing REB renewals	Complete REB renewal 30 days before expiry date.	• Use REB-Legal master tracking to review expiry dates and plan ahead of time.
Regulatory	Study data not submitted to Data and Safety Monitoring Board (DSMB) for review	Report to DSMB at least once a year with information as outlines in the DSMB charter.	• Plan a yearly meeting with DSMB members. • Provide a yearly report, as outlined in the charter, to DSMB. • Discuss with PI ongoingly whether reporting to the project's DSMB is needed.
Deviations	Deviations or deviation trends that affect participant safety and/or data integrity are not reported or escalated accordingly AND Corrective and Preventative Action Plans (CAPAs) for	Identify and document any deviations or deviation trends. Discuss deviations with PI (and Co-Is as applicable).	• Maintain clear and organized deviation records. • Ongoingly check in on status of deviations and verify whether there is a pattern that needs additional review and consider CAPAS.

	important protocol deviations or deviation trends not in place or are ineffective		
Source Records	Study events cannot be easily reconstructed (i.e. lack of documentation or disorganized documentation)	Save, store and document research and other research-related materials as stated in internal guides and protocol.	• Document research activities as instructed. • Maintain documents organized in the X-Drive as per Regulatory Study Binder guidelines and recommendations.
Data Quality and Management	Critical data points not collected	Keep baseline and follow up master log up to date	• Update baseline and follow up master logs regularly. • Verify REDCap completion.

15.2 Study Monitoring Plan

Site monitoring is conducted to ensure that the rights and well-being of research participants are protected, the reported trial data are accurate, complete, and verifiable, and the conduct of the clinical trial is in compliance with the currently approved protocol/amendment(s), ICH GCP, and applicable regulatory requirement(s). Reference the study monitoring plan for specific monitoring information.

Site monitoring is conducted to ensure that the rights and well-being of research participants are protected, the reported trial data are accurate, complete, and verifiable, and the conduct of the clinical trial is in compliance with the currently approved protocol/amendment(s), ICH GCP, and applicable regulatory requirement(s).

On-site internal peer-monitoring will be conducted by the central staff at the Cundill Centre for Child & Youth Mental Health throughout the course of the study, occurring at least once a year or as needed. Random review of certain data will take place at the monitoring visits.

15.3 Auditing and Inspecting

The PI will permit study-related audits, and inspections by the REB, CAMH, sponsor, and applicable granting agencies or regulatory bodies, including access to all study-related documents (e.g. source records, regulatory documents, data collection instruments, study data, etc.). The PI will ensure the capability for audits/inspections of applicable study-related facilities (e.g. research pharmacy, clinical laboratory, imaging facility, etc.).

16- CLINICAL TRIAL FINANCES

This study is funded through grants from the Canadian Institutes of Health Research and the Cundill Centre for Child & Youth Mental Health.

17- REFERENCES

Achenbach, T. M., & Edelbrock, C. (1991). Child behavior checklist. Burlington (Vt), 7, 371-392.

Asarnow, J. R., Emslie, G., Clarke, G., Wagner, K. D., Spirito, A., Vitiello, B., … Brent, D. (2009). Treatment of selective serotonin reuptake inhibitor-resistant depression in adolescents: Predictors and moderators of treatment response. Journal of the American Academy of Child and Adolescent Psychiatry, 48(3), 330–339. https://doi.org/10.1097/CHI.0b013e3181977476

Asarnow, J. R., Jaycox, L. H., Duan, N., Laborde, A. P., Rea, M. M., Murray, P., . . . Wells, K. B. (2005). Effectiveness of a Quality Improvement Intervention for Adolescent Depression in Primary Care Clinics. Jama, 293(3), 311. doi:10.1001/jama.293.3.311

Barwick, M., Barac, R., Kimber, M., Akrong, L., Johnson, S. N., Cunningham, C. E., ... & Godden, T. (2020). Advancing implementation frameworks with a mixed methods case study in child behavioral health. Translational behavioral medicine, 10(3), 685-704.

Bank of Canada. Consumer Price Index, 2000 to Present. Available at: https://www.bankofcanada.ca/rates/price-indexes/cpi/. Accessed January 7, 2021Brent, 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.

Brent, D. A., Holder, D., Kolko, D., Birmaher, B., Baugher, M., Roth, C., ... & Johnson, B. A. (1997). A clinical psychotherapy trial for adolescent depression comparing cognitive, family, and supportive therapy. Archives of general psychiatry, 54(9), 877-885.

Browne G. B., Arpin K., Corey P., Fitch M., Gafni A. Individual correlates of health service utilization and the cost of poor adjustment to chronic illness. Medical care. 1990 Jan 1:43-58.

Browne G, Roberts J, Weir R, Gafni A, Watt S, Byrne C. The cost of poor adjustment to chronic illness: lessons from three studies. Health & Social Care in the Community. 1994 Mar;2(2):85-93.

Busner, J., & Targum, S. D. (2007). The clinical global impressions scale: applying a research tool in clinical practice. Psychiatry (Edgmont), 4(7), 28–37.

Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean?(or it takes at least two to tango). Social science & medicine. 1997 Mar 1;44(5):681-92.

Chorpita, B. F., Daleiden, E. L., Park, A. L., Ward, A. M., Levy, M. C., Cromley, T., ... & Krull, J. L. (2017). Child STEPs in California: A cluster randomized effectiveness trial comparing modular treatment with community implemented treatment for youth with anxiety, depression, conduct problems, or traumatic stress. Journal of consulting and clinical psychology, 85(1), 13.

Costello, E. J., & Angold, A. (1988). Scales to assess child and adolescent depression: checklists, screens, and nets. Journal of the American Academy of Child & Adolescent Psychiatry, 27(6), 726-737.

Courtney, D. B., Cheung, A., Henderson, J., Bennett, K., Battaglia, M., Strauss, J., ... & Szatmari, P. (2019). Effectiveness of an integrated care pathway for adolescents with depression: a pilot clinical trial protocol. Journal of the Canadian Academy of Child and Adolescent Psychiatry, 28(3), 115.

Courtney DB, Watson P, Chan BW, Bennett K, Krause KR, Offringa M, Butcher NJ, Monga S, Neprily K, Zentner T, Rodak T. Forks in the road: Definitions of response, remission, recovery, and other dichotomized outcomes in randomized controlled trials for adolescent depression. A scoping review. Depression and Anxiety. 2021 Jul 26.

Curran Geoffrey, M., Mark, B., Brian, M., Pyne Jeffrey, M., & Cheryl, S. (2012). Effectiveness-implementation Hybrid Designs. Medical Care, 50(3), 217-226.

Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implementation science, 4(1), 1-15

Ebesutani, C., Reise, S. P., Chorpita, B. F., Ale, C., Regan, J., Young, J., Charmaine, H. M. M., & Weisz, J. R. (2012). The Revised Child Anxiety and Depression Scale-Short Version: Scale reduction via exploratory bifactor modeling of the broad anxiety factor. Psychological Assessment, 24(4), 833–845.

Edwards, T. C., Huebner, C. E., Connell, F. A., & Patrick, D. L. (2002). Adolescent quality of life, part I: conceptual and measurement model. Journal of adolescence, 25(3), 275-286.

Elwyn G, Barr PJ, Grande SW, Thompson R, Walsh T, & Ozanne E. (2013, June). Developing CollaboRATE: A fast and frugal patient-reported measure of the shared decision making process in the clinical encounter. Poster presented at the 7th International Shared Decision-Making Conference. Lima, Peru.

Ewing (2018) Ewing SW, Chang L, Cottler LB, Tapert SF, Dowling GJ, Brown SA. Approaching retention within the ABCD study. Developmental cognitive neuroscience. Aug 1;32:130-7.

Freedland, K. E., King, A. C., Ambrosius, W. T., Mayo-Wilson, E., Mohr, D. C., Czajkowski, S. M., ... & Riley, W. T. (2019). The selection of comparators for randomized controlled trials of health-related behavioral interventions: recommendations of an NIH expert panel. Journal of clinical epidemiology, 110, 74-81.

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.

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. The British Journal of Psychiatry, 210(1), 24-30.

Goodyer, I. M., Reynolds, S., Barrett, B., Byford, S., Dubicka, B., Hill, J., ... & Fonagy, P. (2017). Cognitive behavioural therapy and short-term psychoanalytical psychotherapy versus a brief psychosocial intervention in adolescents with unipolar major depressive disorder (IMPACT): a multicentre, pragmatic, observer-blind, randomised controlled superiority trial. The Lancet Psychiatry, 4(2), 109-119.

Gunlicks-Stoessel (2019) Gunlicks---stoessel M, Mufson L, Bernstein G, et al. (2019). Critical Decision Points for Augmenting Interpersonal Sequential Multiple Assignment Randomized Trial. J Am Acad Child Adolesc Psychiatry; 58(1), 80---91. doi:10.1016/j.jaac.2018.06.032

Henderson, J. L., Cheung, A., Cleverley, K., Chaim, G., Moretti, M. E., de Oliveira, C., ... & Szatmari, P. (2017). Integrated collaborative care teams to enhance service delivery to youth with mental health and substance use challenges: protocol for a pragmatic randomised controlled trial. BMJ open, 7(2), e014080.

Heffernan, O. S., Herzog, T. M., Schiralli, J. E., Hawke, L. D., Chaim, G., & Henderson, J. L. (2017). Implementation of a youth-adult partnership model in youth mental health systems research: Challenges and successes. Health Expectations, 20(6), 1183-1188.

Hughes, C. W., Emslie, G. J., Crismon, M. L., Posner, K., Birmaher, B., Ryan, N., ... & DEPRESSIVE, M. T. O.
C. M. (2007). Texas children's medication algorithm project: update from Texas consensus conference panel on medication treatment of childhood major depressive disorder. Journal of the American Academy of Child & Adolescent Psychiatry, 46(6), 667-686.

Joosten E. A., DeFuentes-Merillas L., De Weert G. H., Sensky T., Van Der Staak C. P., de Jong CA. Systematic review of the effects of shared decision-making on patient satisfaction, treatment adherence and health status. Psychotherapy and psychosomatics. 2008;77(4):219-26.

Kane JM, Robinson DG, Schooler NR, et al. (2016) Comprehensive versus usual community care for first---episode psychosis: 2---Year outcomes from the NIMH RAISE early treatment program. Am J Psychiatry. 2016;173(4):362---372. doi:10.1176/appi.ajp.2015.15050632

Kaufman, J., Birmaher, B., Brent, D., Rao, U. M. A., Flynn, C., Moreci, P., ... & Ryan, N. (1997). Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. Journal of the American Academy of Child & Adolescent Psychiatry, 36(7), 980-988.

Krause, K. R., Chung, S., Adewuya, A. O., Albano, A. M., Babins-Wagner, R., Birkinshaw, L., ... & Wolpert, M. (2021). International consensus on a standard set of outcome measures for child and youth anxiety, depression, obsessive-compulsive disorder, and post-traumatic stress disorder. The Lancet Psychiatry, 8(1), 76-86.

Law. D. & Jacob. J. (2015). Goals and goal based outcomes (GBOs): Some useful information (third edition) CAMHS Press: London.

Lewinsohn & Clarke (1984) Peter M Lewinsohn, Gregory N Clarke, Group treatment of depressed individuals: The ‘coping with depression’ course, Advances in Behaviour Research and Therapy, Volume 6, Issue 2, 1984, Pages 99-114, https://doi.org/10.1016/0146-6402(84)90005-5.

Lyneham, H. J., Sburlati, E. S., Abbott, M. J., Rapee, R. M., Hudson, J. L., Tolin, D. F., & Carlson, S. E. (2013). Psychometric properties of the child anxiety life interference scale (CALIS). Journal of Anxiety Disorders, 27(7), 711-719.

March J, Silva S, Petrycki S, et al. (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

Meyers, D. C., Durlak, J. A., & Wandersman, A. (2012). The quality implementation framework: a synthesis of critical steps in the implementation process. American journal of community psychology, 50(3-4), 462-480.

Mufson, L., Dorta, K. P., Wickramaratne, P., Nomura, Y., Olfson, M., & Weissman, M. M. (2004). A randomized effectiveness trial of interpersonal psychotherapy fordepressed adolescents. Archives of general psychiatry, 61(6), 577-584.

Nock, M. K., Holmberg, E. B., Photos, V. I., & Michel, B. D. (2007). Self-Injurious Thoughts and Behaviors Interview: development, reliability, and validity in an adolescent sample.

Patrick, D. L., Edwards, T. C., & Topolski, T. D. (2002). Adolescent quality of life, part II: initial validation of a new instrument. Journal of adolescence, 25(3), 287-300.

Posner, K., Brown, G. K., Stanley, B., Brent, D. A., Yershova, K. V., Oquendo, M. A., ... & Mann, J. J. (2011). The Columbia–Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. American journal of psychiatry, 168(12), 1266-1277.

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., ... & Hensley, M. (2011). Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Administration and Policy in Mental Health and Mental Health Services Research, 38(2), 65-76.

Richardson, L. P., Ludman, E., McCauley, E., Lindenbaum, J., Larison, C., Zhou, C., ... & Katon, W. (2014). Collaborative care for adolescents with depression in primary care: a randomized clinical trial. Jama, 312(8), 809-816.

Sackett, D. L., Rosenberg, W. M., Gray, J. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't.

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.

Tam, H. E., & Ronan, K. (2017). The application of a feedback-informed approach in psychological service with youth: Systematic review and meta-analysis. Clinical Psychology Review, 55, 41-55.

Weersing, V. R., Jeffreys, M., Do, M.-C. T., Schwartz, K. T. G., & Bolano, C. (2016). Evidence Base Update of Psychosocial Treatments for Child and Adolescent Depression. Journal of Clinical Child & Adolescent Psychology, 46(1), 11–43. https://doi.org/10.1080/15374416.2016.1220310

Weisz, J. R., Chorpita, B. F., Palinkas, L. A., Schoenwald, S. K., Miranda, J., Bearman, S. K., ... & Research Network on Youth Mental Health. (2012). Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth: A randomized effectiveness trial. Archives of general psychiatry, 69(3), 274-282.

Young J., Shaw BF, Beck AT, & Budenz D (1981). Assessment of competence in cognitive therapy. Unpublished manuscript, University of Pennsylvania.