

PROTOCOL

Title: School-based group interpersonal therapy for adolescents with depression in Nepal: protocol for a phase III realist cluster-randomised controlled trial

**Version 2.0
21st August 2026**

Table of Contents

Abbreviations.....	4
Key trial contacts	5
Funding.....	5
Roles	5
Overview	7
Background	7
Psychological interventions.....	7
Interpersonal psychotherapy.....	8
IPT in Nepal	8
Adaptation and feasibility	8
Realist randomised controlled trial.....	9
Aims and objectives	13
Methods	13
Design	13
Setting	14
Study population	14
Inclusion and exclusion criteria.....	14
Recruitment.....	15
Randomisation	16
Intervention and comparator	16
Intervention.....	16
Comparator	17
Trial outcomes and measurement	18
Assessment and follow up	18
Process evaluation data collection	19
Implementation.....	19
Mechanisms of change and contextual factors.....	20
Economic evaluation data collection.....	23
Power and sample size.....	24
Primary analysis.....	24
Mediation analysis for CMOCs	24
Moderation analysis for CMOCs	25
Moderated mediation analysis for CMOCs	25
Data management and analysis	25
Data management	25
Qualitative analysis	26
Quantitative analysis	26
Economic analysis.....	27

Minimising bias	28
Public and policy involvement	28
Ethical issues, safeguarding and SAEs.....	29
Dissemination	29
Availability of data and materials.....	29
<i>References</i>	30
<i>Tables</i>	34
<i>Figures</i>	45

Version control

Version	Date	Note of changes
1.0	1/5/25	N/A
2.0	21/8/26	• Clarification of assessment timepoints. • Addition of Ratnanagar municipality to study setting. Ratnanagar was added to the trial catchment due to <6 schools in Bharatpur large enough to reliably form separate girls' and boys' IPT groups. • Amendment to process evaluation sample size and removal of reference to case study schools. • Minimum viable IPT group size amended from 5-10 adolescents to 2-10 adolescents to avoid groups having to be dissolved on the rare occasions where we are unable to recruit 5 adolescents. • School eligibility and recruitment clarified - Protocol Version 1.0 specifies inclusion of schools with 100 or more adolescents across grades 7–9 (excluding one abnormally large school of over 850 students). In practice, the team has found it more logistically manageable to work through one municipality at a time, selecting the largest and most accessible eligible schools first before moving to the next municipality. • Visit windows are clarified: midline +/- 1 week; endline +/- 2 weeks; 32 week follow up +/- 2 weeks; follow up beyond 32 weeks +/- 4 weeks.

Abbreviations

95% CI – 95% confidence interval

CMOCs – Context-mechanism-outcome contingencies

(c)RCT - (Cluster) randomised trial

CSSRS – Columbia Suicide Severity Rating Scale

DMC – Data Monitoring Committee

DSRS – Depression Self-Rating Scale

EQ-5D - EuroQol-5 Dimension 5 levels

GAD-7 – Generalised Anxiety Disorder-7

IPT – Interpersonal psychotherapy

IPT-A – Interpersonal psychotherapy modified for adolescents

ITT – Intention to treat

LMICs – Low- and middle- income countries

mhGAP - WHO mental health GAP Action training package

MRC – Medical Research Council

PCL-5 – Post-Traumatic Stress Disorder Checklist for DSM-5

PHQ-A – Patient Health Questionnaire modified for adolescents

PRIME - PRogramme for Improving Mental health care

PTSD – Post-traumatic stress disorder

SEE – School Education Exam
SES – Standardised effect size
SOP – Standard operating procedure
TPO Nepal – Transcultural Psychosocial Organization Nepal
TSC – Trial steering committee
UCL – University College London
UKRI – UK Research and Innovation
UNICEF – United Nations Children’s Fund
US – United States
WHO – World Health Organization

Key trial contacts

Principal investigator	Kelly Rose-Clarke	kelly.rose-clarke@ucl.ac.uk
Nepal principal investigator	Nagendra Luitel	luiteln@gmail.com
Post-doctoral research fellow	Ceren Sonmez	c.sonmez@ucl.ac.uk
Trial statistician	John Hodsoll	john.hodsoll@kcl.ac.uk
Project coordinator	Sujan Shrestha	sshrestha@tponepal.org.np
Research supervisor	Bishnu Lamichhane	blamichhane@tponepal.org.np
Trial funder representative	Caroline Harris	Caroline.Harris@mrc.ukri.org
Trial Steering Committee chair	Helen Baker-Henningham	h.henningham@bangor.ac.uk
Data Monitoring Committee chair	Melissa Neuman	Melissa.Neuman@lshtm.ac.uk

Funding

Funder	Support
UKRI	£1,848,353

Roles

Sponsor

Trial Steering Committee (TSC)

TSC role

Based on MRC Guidelines for Management of Global Health Trials (2017) and the MRC Clinical Trials Unit template TSC Charter (extension of DAMOCLES IDMC template), the TSC has the following specific roles:

- provide expert oversight of the trial
- maintain confidentiality of all trial information that is not already in the public domain
- make decisions as to the future continuation (or otherwise) of the trial
- monitor recruitment rates and encourage the TMG to develop strategies to deal with any recruitment problems
- approve the protocol
- review regular reports of the trial from the Trial Management Group (TMG))
- receive letters of feedback from the IDMC and consider their recommendations

- assess the impact and relevance of any accumulating external evidence
- monitor follow-up rates and review strategies from TMG to deal with problems
- approve any amendments to the protocol, where appropriate
- approve any proposals by the TMG concerning any change to the design of the trial, including additional substudies
- oversee the timely reporting of trial results
- approve the statistical analysis plan

TSC members

Professor Helen Baker-Henningham, Professor in Psychology, Bangor University – Chair

Dr Rishav Koirala, Consultant psychiatrist and researcher, Nepal – Independent member

Dr Suraj Shakya, Clinical psychologist, Tribhuvan University, Nepal – Independent member

Professor Helen Weiss, Professor of epidemiology, London School of Hygiene and Tropical Medicine – Non-voting advisor

Data Monitoring Committee (DMC)

DMC role

The DMC will assume roles proposed by the DAMOCLES Study Group [1]:

- monitor evidence for treatment differences in the main efficacy outcome measures
- monitor evidence for treatment harm (e.g. serious adverse events, deaths)
- assess the impact and relevance of external evidence
- decide whether to recommend that the trial continues to recruit participants or whether recruitment should be terminated either for everyone or for some treatment groups and/or some participant subgroups
- decide whether trial follow-up should be stopped earlier
- assess data quality, including completeness (and by so doing encourage collection of high quality data)
- maintain confidentiality of all trial information that is not in the public domain
- monitor recruitment figures and losses to follow-up
- monitor compliance with the protocol by participants and investigators
- consider the ethical implications of any recommendations made by the DMC
- monitor planned sample size assumptions, preferably with regards to
 - (i) a priori assumptions about the control arm outcome and/or
 - (ii) emerging differences in clinically relevant subgroups, rather than on emerging, unblinded differences between treatment groups, overall
- suggest additional data analyses if necessary
- advise on protocol modifications proposed by investigators or sponsors (e.g. to inclusion criteria, trial endpoints, or sample size)
- monitor continuing appropriateness of patient information
- monitor compliance with previous DMC recommendations

DMC members

Dr Melissa Neuman, Associate Professor of Epidemiology, London School of Hygiene and Tropical Medicine – Chair

Professor Obi Ukoumunne, Professor of Medical Statistics, University of Exeter – Independent member

Dr Bibhav Acharya, Associate Professor of Psychiatry, University of California, San Francisco – Independent member

Overview

The protocol is for a phase III realist cluster randomised controlled trial (cRCT) of group interpersonal psychotherapy (IPT) to treat depression among adolescents in Nepal. The protocol builds on an earlier feasibility study and pilot cRCT of IPT in Nepal. In the feasibility study, we culturally and developmentally adapted the World Health Organization Group (WHO) group IPT manual for delivery in Nepali government secondary schools by lay workers [2]. We conducted the pilot cRCT in eight schools with 161 adolescents to assess the acceptability of IPT and cRCT procedures ahead of a phase III trial. We pre-specified criteria for progression to a phase III trial and all of these were met. We also integrated realist evaluation methods to generate and explore hypotheses about mechanisms of IPT and how they are affected by contextual factors. The phase III cRCT seeks to test these hypotheses by collecting qualitative and quantitative data, and in parallel will estimate the effectiveness and cost-effectiveness of IPT at a population level. This will be the first application of a realist cRCT design to evaluate a psychological intervention and will generate much-needed information about intervention processes and mechanisms, and how these can be modified and shaped by context.

Background

Depression affects over 264 million people worldwide, is a leading cause of disability and has profound negative consequences for the global economy [3]. When depression presents in adolescence it can have devastating and long-lasting effects on health and development [4]. Access to effective treatment is limited, especially in low- and middle-income countries (LMICs) where 90% of the world's adolescents live [5]. Researchers have tried to expand access to treatment in these settings by culturally and contextually adapting existing psychological interventions. Whilst this approach is promising, results from evaluations have been mixed; interventions shown to work in one setting are not always successful elsewhere [6]. The major reason for failures in intervention transfer is that we don't know enough about how interventions work and how this is affected by context in terms of setting and population characteristics. This is a major challenge to the provision of effective mental health care across the world.

Psychological interventions

Psychological interventions are treatments that seek to change behaviours, cognitions and feelings in order to improve mental health. Meta-analyses report medium-to-large effect sizes associated with psychological interventions for depression among children and adolescents but limited evidence from LMICs [7, 8]. We have little knowledge about how psychological interventions work. Correlational analyses from high-income settings suggest there may be common therapeutic factors, i.e. general elements common to all types of psychological intervention including communication skills, empathy and collaboration, as well as specific treatment factors related to individual therapies [9]. Few studies have examined how these factors bring about change. Even fewer have tried to understand how participant characteristics and other contextual factors affect treatment outcomes. This is important because effect sizes of treatments for depression have not increased in recent decades, individual responses are variable and uptake is low. Knowing how psychological interventions work could inform the development of interventions focused on mechanisms that are

more effective, efficient and acceptable, as well as the adaptation of interventions across different settings and populations [9].

Interpersonal psychotherapy

IPT is a psychological intervention that focuses on four common problems that trigger depression: grief, disputes, role transitions and social isolation [10]. Using techniques and strategies such as linking mood to event and event to mood, role play and skill-building, IPT encourages the individual to analyse and improve their interpersonal relationships. IPT was developed in the US to treat depressed adults but has also been used to treat other mental disorders and age groups. A meta-analysis of 17 evaluations of IPT for adolescents (IPT-A) reported large reductions in depressive symptoms following therapy ($d = 1.48$, $p < .0001$) [11]. IPT has also been adapted for use in LMICs by the WHO through the development of a group IPT manual for non-specialist providers [2]. However, results from evaluations of IPT in these countries are inconsistent. For example, in a study with HIV-positive adults in South Africa, group IPT significantly reduced depressive symptoms compared to treatment as usual, but 21/41 participants randomised to receive the intervention did not take it up and three dropped out [12]. In Uganda, group IPT for adolescents significantly improved depressive symptoms among girls but not boys [13]. These findings suggest that while there may be benefits of IPT in multiple contexts, research is needed to understand how, for whom and in what circumstances.

IPT in Nepal

Adaptation and feasibility

We adapted IPT for adolescents with depression in Nepal and tested it in a feasibility study (2018-20) [14]. We culturally and developmentally adapted the WHO group IPT manual through a participatory, multistage process involving: translation and review by Nepali mental health practitioners; literature reviews of existing adaptations of IPT and concepts of mental health and interpersonal problems in Nepal; qualitative research with adolescents, caregivers, teachers and community health workers [15, 16]; consultation with a youth advisory board; and piloting with IPT trainers and facilitators. Key adaptations were: integrating therapy into secondary schools for delivery by lay workers; adding components to promote parental engagement, including a pre-group session with the adolescent and parent; using locally acceptable terms for mental illness such as *man ko samasya* (heart-mind problem); framing the intervention as a training programme to de-stigmatise treatment; and including activities to strengthen relationships between group members. We did not adapt the therapeutic goals of IPT and conserved IPT-specific strategies and techniques, making edits only to the way these were described in the manual.

We undertook an uncontrolled feasibility study of the intervention with 62 boys and girls aged 13-19 in the mountainous district of Sindhupalchowk [17]. Adolescents attended 82.3% (standard deviation 18.9%) of group sessions. Depression and functional impairment improved between baseline and follow-up at 8-10 weeks post IPT: the Depression Self Rating Scale (DSRS) score decreased from 17.2 (95% confidence interval CI: 16.5 - 18.0) to 9.5 (95% CI 8.5 – 10.6) and functional impairment decreased from 12.6 (95% CI 11.4 – 13.7) to 3.2 (95% CI 2.5 – 3.9) [17]. The estimated intervention unit cost was USD 96.9 with facilitators operating at capacity. The adaptation and feasibility study were funded through a programme grant from the UK Medical Research Council (MRC), the National Institute for Health Research and the Department for International Development (grant reference: MR/R020434/1).

Realist randomised controlled trial

In 2022, we received funding from the UK Research and Innovation (UKRI) Future Leaders Fellowship scheme to undertake a 5-year programme of work to evaluate IPT through a realist randomised controlled trial (realist RCT). This is a relatively new approach integrating realist evaluation into a traditional RCT design [18]. Realist evaluation seeks to understand the mechanisms through which social programmes work and why these work better in some contexts than others [19]. To do this, realist evaluators build programme theory from which they derive hypotheses about what intervention mechanisms operate, the contexts in which they operate and the outcomes that are generated. These so-called context-mechanism-outcome configurations (CMOCs) are then empirically tested through analyses of qualitative and quantitative data collected during programme implementation.

The advantage of realist RCTs over traditional RCTs is that they can provide the least biased estimate of intervention impact at population level and at the same time open the intervention “black box” to understand mechanisms of action, and how they are influenced by context in terms of place and person characteristics [20, 21]. Local decision-makers can then use this information to decide whether an intervention tested in one place might also be useful and effective somewhere else.

Combining realist and RCT approaches is not without criticism. Some realist evaluators argue the two approaches are epistemologically and ontologically incompatible, and that RCTs cannot generate the kind of data required to answer realist questions. However, these concerns are addressed in various publications, and an increasing number of real-world examples demonstrate the philosophical congruity, feasibility and usefulness of realist RCTs [21, 22].

Bonell and colleagues describe a staged methodological framework for conducting realist RCTs [20]. Stage 1 involves elaborating a theory of change and developing CMOCs. Stage 2 involves collecting process data during the trial and using this to refine the CMOCs developed in Stage 1. In Stage 3, outcome data are used to estimate the effect of the intervention at the population level, whilst quantitative process and outcome data are integrated to test the CMOCs via various statistical and other analyses. The end products are an evaluation of the intervention in context and an evidence-based theory of change. We drew on this framework in the design of our realist RCT, expanding stage 1 to include formative research and a pilot cRCT to build a preliminary theory of change and CMOCs. Stage 1 is complete and this protocol refers to stages 2 and 3.

Formative research and pilot cRCT

Between 2022 and 2023, we undertook a realist review of adolescent psychological interventions in LMICs, and a systematic review of moderators and mediators of IPT-A to synthesise evidence on intervention mechanisms and contextual factors. In parallel, we held workshops with IPT practitioners across the globe to build a transcultural IPT theory of change for adolescents with depression. We also conducted a qualitative analysis of interview transcripts with participants and facilitators from the feasibility study in Nepal [23].

Through a pilot cRCT in 2023-24, we aimed to test the feasibility and acceptability of delivering and evaluating IPT in a trial in Chitwan district in the plains region of Nepal. We sought to test trial procedures (randomisation, masking, data collection, control conditions), explore initial outcome trends and generate data to refine our theory of change and CMOCs. The pilot cRCT was registered on the ISRCTN registry (ISRCTN14652885) and the protocol is available [online](#).

The design was a parallel two-arm pilot cluster-randomised controlled trial with schools as the unit of clustering. Baseline data were collected before randomisation. The control condition was

enhanced usual care. We assessed participants in intervention and control arms at baseline, after the second group session (midline 1), after the sixth group session (midline 2), at endline (within two weeks of the final group session) and at follow-up (12 weeks after the final group session). We included adolescent boys and girls aged 13-19 from grades 8, 9 and 11, with depression defined as a score of 11 or more on the PHQ-A and 4 or more on the functional impairment tool. Mental health outcomes were depression, anxiety and post-traumatic stress disorder (PTSD). The sample comprised 161 adolescents (81 intervention, 80 control) from eight schools.

We set and met *a priori* progression criteria related to implementation, acceptability, fidelity, serious adverse events (SAEs), eligibility, missing data and participant retention (Table 1). Depression scores at endline and follow-up were significantly lower in the intervention arm compared to control (endline standardised effect size SES -1.23 95% confidence interval CI -1.99, -0.47; follow-up SES -1.24 95% CI -2.03, -0.50). We also found significant improvements in functional impairment (follow-up SES -0.55 95% CI -1.10, -0.02), anxiety (endline SES -0.76 95% CI -1.29, -0.26; follow up SES -0.78 95% CI -1.30, -0.25) and PTSD (endline SES -0.43 95% CI -0.81, -0.08; follow-up SES -0.49, 95% CI -0.87, -0.14). We observed significant reductions in interpersonal conflict in the intervention arm compared to control at midline 2, endline and follow-up, emotion regulation at midline 1 and follow-up and social support at midline 2. We did not find significant differences for self-efficacy, hope or interpersonal psychology skills.

Implications of the pilot cRCT findings for the phase III trial protocol

The phase III trial will in general follow the same protocol as the pilot cRCT but with important changes to inclusion and exclusion criteria, data collection and implementation.

Changes to inclusion and exclusion criteria

In the pilot cRCT, adolescents with depression were only included if they also reported functional impairment. This was in line with previous studies in Nepal which have combined depression and functional impairment measures to optimise the sensitivity of screening procedures for depression [17, 24]. However, there has been a recent, rigorous validation of the PHQ-A in Nepal reporting sufficient sensitivity and selectivity for a cut off score of 11 or more on the PHQ-A [25]. These results negate the need for a measure of functional impairment at screening though we will still include functional impairment as a trial outcome. Therefore in the phase III trial, we will remove functional impairment from our inclusion criteria.

In the pilot cRCT, we included adolescents from grades 8, 9 and 11, but in the phase III trial we will include grades 7, 8, and 9. Previously we did not include adolescents in grade 7 because they were perceived as insufficiently mature to benefit from IPT. However, because adolescents frequently have to repeat grades when they fail exams, many older adolescents are studying in grade 7. Therefore, it is more appropriate to exclude adolescents on the basis of age rather than grade, so we will extend screening to grade 7. In the phase III trial, we will not include adolescents in grade 11 because the number of those adolescents is comparatively low, making it difficult to form IPT groups with sufficient numbers. Moreover, it was difficult to combine adolescents from grades 8, 9 and 11 in groups because grade 11 adolescents had different timetables and did not appear to bond with younger group members.

Adolescents were excluded from the pilot trial if they screened positive for *chhopne* (local term for symptoms in line with conversion disorder [26]) in the past three months because of concerns about whether IPT would be an appropriate treatment. Based on our experience in the pilot cRCT, we would prefer to include these adolescents and provide facilitator training and guidance in the manual on how to manage these symptoms. Therefore, we will drop this exclusion criteria for the phase III trial.

In the pilot cRCT, we assessed suicidality using the Columbia Suicide Severity Rating Scale (CSSRS) and excluded adolescents with suicide intention or plan in the last two weeks or attempt within the last three months [27]. These adolescents were categorised as high risk and deemed to require more intense care. In line with our standard operating procedure (SOP), high risk adolescents were referred to our counsellor for assessment, treatment and potentially further referral. According to the CSSRS, adolescents who endorse suicide method or lifetime suicide attempt are categorised as “moderate risk”. We referred these adolescents to the counsellor but also included them in the trial. However, during the pilot cRCT, the team raised concerns about the specificity of the CSSRS, specifically that we were excluding adolescents who were not at high risk and could have benefitted from IPT. The team was also concerned that the SOP was overly complex, that counsellors were overwhelmed and that it would not be feasible to scale up the existing SOP in the phase III trial. Consequently, we have undertaken extensive formative work to revise the CSSRS translation and our risk definitions, including interviews with adolescents and Nepali clinical experts. Findings from this work informed the revision of our exclusion criteria and SOP. We have removed the moderate-risk category and we will exclude high-risk adolescents defined as those with (i) suicide plan in the past 2 weeks, ii) suicide attempt in the past three months or iii) lifetime suicide attempt with suicidal ideation in the past two weeks. Moreover, we will only screen adolescents with the CSSRS if they endorse PHQ-A item 9 (*thoughts that you would be better off dead, or of hurting yourself in some way*) or if they are PHQ-A positive. This is because schools and parents indicated that it was not acceptable to assess all adolescents with the CSSRS in the absence of indicators of risk.

Changes to data collection

We timed midline, endline and follow-up surveys in the pilot cRCT according to how far the intervention arm had progressed through the intervention (i.e. based on group session). This was difficult to operationalise because IPT groups progressed at different speeds due to school holidays, exams and absences, and it was not obvious when to conduct surveys with adolescents in the control arm. In the phase III trial, all surveys will be conducted based on time since randomisation (designated week 0).

We will not conduct a second midline survey in the phase III trial because we do not think the scientific insight from the survey justifies the resources required and burden imposed on adolescents and schools. Furthermore, the purpose of the midline survey is to identify the mechanism through which IPT improves depression (i.e., indirect effects). Data from the pilot trial revealed that the indirect effects were similar across the two midline surveys (week 3 and week 6).

This protocol differs from the pilot cRCT in terms of the measure of school academic performance used. Although we collected data on adolescent end-of-year exam scores in the pilot cRCT, these were not standardised and we were therefore unable to analyse them. In longer term follow up of participants in the phase III trial, in order to undertake exploratory analyses of IPT’s impact on academic performance we will collect data on the Basic Education Examination (BEE), an examination set by the municipality for students in grade 8, and the School Education Exam (SEE), a national standardised exam for adolescents in grade 10. We will collect SEE results from adolescents when they reach grade 10. This will mean one year post IPT participation for those who are in grade 9 when they enroll in our trial, two years post IPT for those who are in grade 8 and three years post IPT for those who are in grade 7. For adolescents in grades 7, we will also collect their BEE results one year post IPT. In our analyses we will need to account for missing data due to school dropout, and variation in the duration between participation in IPT and undertaking the SEE/BEE. We will also collect data on academic self-efficacy using questions from the *Young Lives* non-cognitive instrument, translated and tested in Nepali [28]. We will collect academic self-efficacy data in baseline and

endline surveys as this is plausibly on the pathway between improved mental health and academic performance [29].

In the pilot cRCT we assessed the feasibility of conducting surveys with a subgroup of adolescents' mothers. We were able to collect data from 73% of mothers at baseline (40/55) including data on socioeconomic status and mental health. This suggests that it is feasible to conduct surveys with mothers in the phase III trial, but to improve the participation rate and reduce missing data we will conduct surveys with adolescents' mother or father – whomever is available to participate. This will enable us to test hypotheses concerning the influence of parents' socioeconomic status and mental health on adolescent trial outcomes.

Changes to implementation

In the pilot cRCT protocol, we stated IPT groups will comprise between 6-8 adolescents. For the phase III trial, we will relax these numbers to allow groups of 5-8 adolescents. This is to accommodate potential variation in the number of eligible adolescents across schools.

We narrowly met progression criteria for attendance at IPT sessions in the pilot cRCT: 54% of adolescents in the intervention arm attended >70% of IPT sessions. This included two adolescents whom we could not allocate to any IPT groups, and six who dropped out of the intervention before the sessions started because of stigma from teachers. Among the remaining adolescents, the main reasons for low attendance were because they had to study for exams, travelled away from home during the school holidays or were ill. To improve attendance in the phase III trial, we will introduce the following measures:

- Optimising the timing of the intervention – we now have a detailed calendar of school holidays, closures and exam periods and therefore optimised the timing of the intervention to avoid these as much as possible.
- Excluding grade 11 – in the pilot cRCT we combined grades 8/9 and 11 in some groups. Because of differences in class timetables for these grades we had few viable timeslots for IPT sessions. Timetables for grades 7, 8 and 9 have similar free periods and more mutually convenient timeslots, which will improve attendance.
- Guidance on managing absenteeism in the IPT manual – this includes scripted conversation and motivational-interviewing techniques.
- Showing a film to staff and students in intervention schools as part of an orientation session – we made a short film about IPT involving participants from the pilot trial to psychoeducate and destigmatise participation in IPT groups.

Theory of change and CMOCs

We developed a theory of change for IPT and CMOCs informed by the formative research and refined by the pilot trial findings [30]. Figure 1 presents the most recent version of the theory of change which includes roles for facilitators (delivering therapy and managing groups), school staff and caregivers (supporting adolescents), and adolescents (attending and participating in group sessions). Intervention mechanisms incorporate: social learning theory (adolescents acquire communication skills of sharing, listening, giving feedback, and negotiating); attachment theory (adolescents identify and evaluate key interpersonal relationships); mental health literacy (adolescents understand depression is common and can be treated); and problem-solving [31, 32]. Through these mechanisms, adolescents perceive the possibility of an improvement in their interpersonal circumstances and mood, internalise their membership of the IPT group (social identity theory), and improve their ability to empathise and communicate their interpersonal needs (emotion validation) [33]. These mechanisms generate intermediate impacts (relationship initiation, improved social support, conflict reduction, perception of hope, self-efficacy and emotion regulation), which in turn generate a reduction in depressive symptoms and improvements in general functioning. Qualitative

research from the pilot cRCT suggested a further causal pathway based on self-compassion theory, through which adolescents learn that they are not alone in suffering and difficult experiences are a normal part of life [34].

Prior to the pilot cRCT, we formulated five CMOCs derived from our theory of change. Table 2 presents these CMOCs and related qualitative and quantitative analyses from the pilot cRCT. Based on the analyses, we refined the CMOCs and these are outlined below in the section on the process evaluation.

Aims and objectives

We aim to evaluate the effectiveness and cost-utility of IPT for adolescents with depression in Nepal and to understand how IPT works, for whom and in what circumstances.

Objectives

1. Using a cluster randomised controlled trial design, examine the effectiveness of IPT on depression at 17 weeks post randomisation (the primary outcome) and on secondary outcomes
2. Using qualitative and quantitative process data, test hypotheses informed by the intervention theory of change about mechanisms and outcomes, and how they vary by contextual characteristics of place and person.
3. Conduct an economic evaluation combining cost-utility and benefit-cost analyses to assess costs and benefits accrued during intervention and follow-up, employing provider and modified societal perspectives.
4. Explore the implementation of IPT in terms of feasibility, fidelity, reach, dose and acceptability through a mixed methods process evaluation.
5. Conduct secondary analyses of qualitative and quantitative data collected 12, 18 and 30 months post-randomisation to examine how participation in IPT affects longer term outcomes.

Research under objectives 1-4 is funded by the UKRI fellowship scheme and will take place between April 2025 and June 2027. Research related to objective 5 will be funded under the fellowship renewal scheme (~£650k, 2027-30) and therefore a subsequent protocol will provide more detailed information related to this. The main trial analyses relate to objectives 1-4 and we plan to publish these findings prior to completing data collection related to objective 5.

Methods

Design

The trial is being implemented by Transcultural Psychosocial Organization Nepal (TPO Nepal) in partnership with UCL.

We will conduct a superiority phase 3 cRCT to test whether IPT is more effective at reducing symptoms of adolescent depression compared to enhanced usual care. Schools are the unit of allocation and the trial will comprise 48 schools (24 intervention, 24 control). Due to periods of school closure during holiday and exam periods, the trial will be implemented in eight waves over two years (Figure 2). Each wave will comprise six schools. Schools will be allocated to each wave based on their location so that schools in the same area are allocated to the same wave. This will make it easier for the team to travel from school to school to provide supervision and oversight.

Within each wave, we will assess primary and secondary outcomes at baseline and the following five time-points: weeks 17 (endline, post randomisation, primary analysis) and 32 weeks post randomisation (follow up 1)). During an extended follow up period we plan to survey participants at 12, 18 and 30 months post randomisation. We will also conduct a midline survey to assess intermediate outcomes at week 5 post randomisation. The primary analysis will be a cross-sectional comparison of mean depressive symptom scores at endline (week 17 post randomisation) across trial arms, adjusted for baseline scores and clustering.

Setting

The trial will be conducted in schools in Chitwan and Nawalparasi districts in the lowland region of Nepal. Both districts are mainly rural, though Bharatpur metropolitan city in Chitwan is the third most populous city in Nepal (~369k residents). In the 2021 national census, Chitwan had a population of 719,859 and 17.9% were adolescents aged 10-19 [35]. Nawalparasi had a population of 378,079 and 18.8% were adolescents. The percentage of the population aged 5 and over who could read and write was 83.7% in Chitwan and 82.4% in Nawalparasi, higher than the national average of 76.2%. The most populous ethnic group in Chitwan is Tharu and in Nawalparasi is Magar [35].

IPT will be delivered in public secondary schools. Most adolescents in Nepal attend public schools. Around 84% of school children are in public schools at basic education level (grades 1-8), 81% at secondary level (grades 9-10) and 71% at higher secondary level (grades 11-12) [36]. Public education in Nepal is free of charge until grade 10, though schools may charge registration fees and require students to purchase uniform and stationary [37].

In terms of mental health care in the study setting, there are several non-governmental organisations offering psychosocial counselling. Hospitals in urban areas run psychiatric outpatient clinics and some also have provision for counselling. From 2014-17, PRIME (PRogramme for Improving Mental health carE), led by TPO Nepal, provided mental health training to primary care providers in 10 facilities across the district but it is unclear how many of these providers are still in post. There has been limited research on adolescent mental health in the study setting. One study in Chitwan reported 27% prevalence of depression (using PHQ-9 cut off score of ≥ 10) among 371 students aged 15-19 [38].

Study population

The study population is adolescents with depression aged 13-19 years attending a school selected for the trial. Adolescents commonly repeat school years due to failing their end of year exams. Because of this, the age range in each grade is wide. We will focus recruitment on grades 8 and 9 but if there are insufficient adolescents with depression in these two grades (i.e. less than five boys or girls), we will recruit from grade 7. This age group is targeted because: depression often presents first in adolescence, with a peak age of onset of 20.5 [39]; data from the pilot trial and feasibility study in Nepal indicate that IPT is feasible and acceptable for this age group; it is feasible to work with adolescents in grades 7 to 9 because they have more free time and attend school regularly compared to students in higher grades; and evidence from other settings suggests IPT reduces depression for this age group [13].

Inclusion and exclusion criteria

Eligible schools are secondary schools within the public education system in Chitwan (Bharatpur and Ratnanagar municipalities) and Nawalparasi (Gaidakot and Devchuli municipalities). Because it is

logistically easier, the trial team will work in one municipality at a time (i.e. one municipality per wave). Among eligible schools in each municipality we will prioritise larger schools and those that are easiest to access from the road.

We will exclude private schools and single sex schools, as well as any schools in which we have previously run IPT groups as part of facilitator training. We will also exclude small schools defined as those with fewer than 100 students in grades 7, 8 and 9 combined, based on the most recent enrolment data available. This is because, based on an estimated 70% of adolescents returning signed consent forms and a 15% eligibility rate with a minimum of five adolescents per group, we are unlikely to be able to identify a sufficient number of adolescents to form boys and girls IPT groups in these schools. We will also exclude one exceptionally large school (>850 students) due to logistical challenges of screening such a large number of adolescents.

Within eligible schools, eligible participants will be:

- adolescent boys and girls aged 13-19;
- attending a participating secondary school;
- enrolled in Class 7, 8 or 9; and
- with depression (i.e. scoring 11 or more on the PHQ-A).

Adolescents will be excluded if they are:

- in Class 10 or above because their schedules are incompatible with the intervention; and
- at high risk of suicide defined as (i) suicide plan in the past 2 weeks, ii) suicide attempt in the past three months or iii) lifetime suicide attempt with suicidal ideation in the past 2 weeks.

Recruitment

We will approach schools, initially by phone call and arrange to meet with the principal to explain the study and obtain their approval. The research team will then run an orientation session with students.

Within participating schools, research assistants will visit classrooms to present the study to adolescents. All adolescents in the classrooms will be given a consent form and instructed to bring the consent form back (signed by themselves and a caregiver) by a certain date (approximately 3-4 days later) if they are interested in being screened and potentially participating in the trial. Research assistants will explain that signing the consent form will not guarantee their participation in the research as we will only recruit a certain number of students who meet the eligibility criteria. Each signed consent form will be assigned an ID number. ID numbers will be entered into an Excel spreadsheet and randomly ordered using an online platform (e.g., <https://www.random.org/lists/>). The research assistants will invite adolescents for screening according to this random order and screen them for eligibility until the target number is reached.

Based on our experience in the feasibility study and pilot cRCT, the optimum number of adolescents in an IPT group is eight, with a minimum acceptable number being five. Based on depression prevalence estimates from our pilot cRCT, recruiting eight boys and eight girls from grades 8th and 9th would be feasible in large schools. However, in small schools we expect that even five might be challenging and in that case we may need to screen grade 7 students who fall within our targeted age range (13-19). Our experience suggests it is preferable to have fewer same-grade adolescents in a group (with a minimum of five) than to recruit across more than two year-groups because of difficulties scheduling IPT sessions across multiple class timetables and the risk of having less cohesive groups due to differences in grades. Therefore, we would like to allow for screening grade 7

whenever it is absolutely necessary to form a group. To allow for this, we will implement the following recruitment procedure:

In all schools we will initiate screening in parallel in grades 8 and 9 and continue until either:

- (i) we recruit eight boys and eight girls (plausible in larger schools); or
- (ii) we have screened all adolescents in grades 8 and 9.

If (ii) occurs and we have recruited fewer than five boys and/or five girls, we will initiate screening in grade 7 and continue until we have recruited eight boys and/or eight girls in total across the three grades, or until we have screened all adolescents in grade 7.

Where recruitment does not yield sufficient numbers to form a group of 5, groups will not be dissolved; they will be retained and delivered as long as they comprise more than one adolescent (i.e., as small as two or three participants).

Those adolescents meeting the eligibility criteria will complete the baseline survey and excluded participants will be informed at the end of the screening by the research assistant. Those adolescents who are excluded due to suicidality (with a current plan, history of suicide attempt in the past three months or lifetime attempt and current ideation) will be escalated following the High Risk Protocol (See Full Trial Suicide and Abuse SOP).

Randomisation

Randomisation will be at the school level. Schools will be divided into eight groups or *waves* (six schools per wave). Schools will be grouped into waves based on their geographic location because combining neighbouring schools in one wave is more convenient for the trial team. Advantages of grouping schools into eight waves include: (i) the duration between identifying that an adolescent is depressed at screening and the adolescent starting the intervention is minimised; (ii) a staggered start to the surveys - meaning fewer research assistants and IPT facilitators are needed over the course of the trial; (iii) randomisation can be conducted after the baseline survey with minimal delay.

Randomisation will be conducted for each wave (i.e. on eight separate occasions) by two independent statisticians linked to the UCL MRC Clinical Trials Unit. In each wave, once screening is complete one statistician will be asked to submit a randomised list of numbers 1-6 allocated in a 1:1 ratio across intervention and control arms. The second statistician will be sent a list of the six school names and will randomly order these. The PI will then apply the school names in the order received from the second statistician to the allocations provided by the first statistician. The PI will then send the final allocation to the two statisticians for their approval.

Intervention and comparator

Intervention

We will evaluate the IPT intervention that we implemented in the feasibility study and pilot cRCT. The intervention involves two pre-group sessions and 10 group sessions. Sessions are held weekly unless there is a need to alter the schedule due to e.g. school closures or participant availability. In the first pre-group session, the facilitator meets the participant one to one to identify the most relevant IPT problem area, help the participant link their depressive symptoms to the problem area, and gather information about the participant's key relationships and history of depression. In the second pre-group session, the facilitator meets the participant and their caregiver together, ideally at home, to mobilise support and build rapport with the participant's family. Where a participant's depression is

related to their family circumstances, the facilitator will discuss with the participant in advance about what they are comfortable to share with their caregiver.

IPT groups are gender specific. Each group is allocated two facilitators: one designated as lead, the other as assistant. Group sessions take place in a quiet, private space in the school (such as an empty classroom or the library). In the initial group session, the facilitator focuses on encouraging participants to review and share their problems, and instilling hope for recovery. In the middle sessions (2–9), participants practice interpersonal skills and offer and receive support from group members to resolve their problems. In the last session, participants review and celebrate progress and make plans to tackle future problems.

In each pre-group and group session, participants review their depressive symptoms with the facilitator using a seven-item symptom checklist developed for the study. This review process helps participants link changes in symptoms to events in their daily lives, and enables facilitators to identify deterioration and suicidality. We will implement a standard operating procedure (SOP) to manage adolescents reporting suicidal thoughts, including risk assessment, consultation with an IPT supervisor, communication with parents and, where appropriate, one-to-one intervention for the adolescent with a psychosocial counsellor in parallel with the group sessions.

We will recruit lay people to train as IPT facilitators. An overview of facilitator recruitment criteria and training is presented in Figure 2 and described in detail in the trial clinical protocol. Project clinical supervisors who are trained in IPT will conduct the training. Supervisors will be supervised by an IPT master trainer. The training programme comprises three modules: (i) WHO's *Foundational Helping Skills* – a 10-day module to build basic psychosocial skills (ii) a one-day module focused on group management; and (iii) an IPT module involving a five-day didactic workshop focused on theory, structure, techniques and strategies, followed by supervision of a minimum of three practice cases including one individual and one practice group of four or five adolescents. Individual and group practice participants will be recruited from secondary schools that are not included in the trial. Practice group participants will be screened positive for, or with subthreshold symptoms of, depression.

At each stage of the training, we will assess facilitators' competency. Foundational helping skills will be assessed during standardised role-plays pre- and post-training using the Enhancing Assessment of Common Therapeutic factors (ENACT) rating scale [13]. We will assess group management skills in role plays and practice groups using the GroupACT tool [40]. Facilitators' understanding of the IPT model will be assessed after the didactic workshop using a paper-based knowledge test. During practice groups, supervisors will assess IPT skills using a standardised rating scale of activities carried out in each session. Based on their competency and availability, we will select ten from the trained facilitators to facilitate IPT in the trial.

Comparator

Participants attending schools in the control arm will receive enhanced usual care. This is justified in terms of balancing the generalisability of our findings and our ethical obligations as researchers. In intervention and control arms, we will train health workers in health posts and primary care centres using the WHO mental health GAP Action training package (mhGAP). Participants in the control clusters will receive a handout with information about the location of these trained health workers and how they can access treatment.

Adolescents in the control and intervention arms reporting a suicide plan in the past two weeks, suicide attempt in the past three months or lifetime suicide attempt and ideation in the past two

weeks will be assessed by a psychosocial counsellor employed through the project and referred to mental health services as per need. These participants will not be withdrawn from the outcome assessment and will continue with the intervention if they wish to and the clinical team deems it appropriate.

Prior to the baseline survey and randomisation, facilitators will provide an orientation to teachers in control and intervention schools. The orientation will focus on building support for adolescents participating in the trial. To incentivise participation among schools, in each wave after baseline we will invite one or two teachers per school in intervention and control arms to participate in a psychoeducation workshop.

Trial outcomes and measurement

Table 3 details all the trial outcomes and tools to measure them. The primary outcome is depression symptoms at baseline (week 17 post randomisation) measured as a continuous outcome. We will measure this using the Patient Health Questionnaire 9 item modified for adolescents (PHQ-A) [41]. This tool is widely used across cultures and contexts, has been adapted for and validated among adolescents in Nepal and is sensitive to change as demonstrated in our pilot cRCT [25].

We will assess other potential intervention effects by measuring the following secondary outcomes: depression symptoms at week 32, anxiety symptoms measured with the Generalised Anxiety Disorder-7 (GAD-7) validated in Nepal [25, 42]; functional impairment using a locally developed tool based on methods outlined in Bolton et al [43]; post-traumatic stress disorder (PTSD) symptoms with the 8-item version of the PTSD Checklist for DSM-5 (PCL-5) [44] based on longer versions previously used in Nepal [25, 45-47]; and quality of life with the EuroQol-5 Dimension 5 levels (EQ-5D) [48]. Other secondary outcomes are school attendance and academic self-efficacy. We will assess school attendance as the number of days attended in 24 days prior to the survey (excluding closures and Saturdays). We will also measure cognitive and non-cognitive academic skills. We will measure academic self-efficacy using questions from the Young Lives non-cognitive instrument [28], translated and tested in Nepali. Academic self-efficacy is a construct associated with goal setting and academic strategy use and conceivably lies on the pathway to improved academic performance [29]. In follow up surveys (12, 18 and 30 months post randomisation), we will conduct secondary analyses to explore potential long-term effects of IPT on school drop-out and performance in the BEE and SEE.

Informed by our theory of change, we will assess intervention effects at midline, baseline and follow-up on intermediate outcomes and whether these outcomes mediate the effect of the intervention on depression. Intermediate outcomes are: hope measured with the Herth Hope Index [49]; emotion regulation with the Difficulties in Emotion Regulation Scale (DERS) [50]; self-efficacy with the Generalised Self-Efficacy Scale [51]; social support with the Multidimensional Scale of Perceived Social Support [52]; and interpersonal conflict reduction using select items from the Network of Relationships Questionnaire [53]. In the intervention arm, we will measure group cohesion using a tool developed to assess cohesiveness in a trial of Problem Management Plus in Nepal [24], which we adapted for IPT and used in the pilot trial.

Assessment and follow up

We will conduct baseline surveys with adolescents before randomisation and collect data on socio-demographic characteristics, primary and secondary outcomes (mental health, functioning, quality of life) and intermediate outcomes (social-emotional skills and resources) as well as health service use, adversity and perceptions of school climate and gender norms.

We will resurvey adolescents at week 5 (midline), 17 (endline) and 32.(follow up 1), as well as at 12, 18, and 30 months post randomisation. Based on the pilot cRCT, we expect a 92% response rate at endline.

We will also conduct surveys at baseline with adolescents' parents to collect data on potential moderators of intervention effects. We will survey either the participants' mother or father, depending on who is available to participate in the survey. If the adolescent does not have a parent (e.g. they are orphaned) or the adolescent does not usually live with either of their parents, this will be recorded and linked to the adolescent's data. We will measure parents' mental health using the General Health Questionnaire 12-item [54], parenting behaviour with the Alabama Parenting Behaviour Scale [55] and adolescent behaviour with the Disruptive Behaviour International Scale – parent report [56]. We will survey parents again at endline to reassess moderators and collect data for the economic analysis (see below).

All surveys will be conducted by trained research assistants using tablets programmed with data collection software. Research assistants will read the survey questions to the adolescents/parents and enter their responses on the tablet. We will conduct as many surveys as possible in person. Where we cannot meet in person, we will conduct the interview on the phone. In-person surveys will be conducted in a private place at the participant's school (adolescents) or home (adolescents and parents) where participants cannot be overheard. For phone surveys, we will arrange a time for the survey in advance. Research assistants but not participants will be masked to allocation.

Process evaluation data collection

The process evaluation design is informed by frameworks and theory including the general theory of implementation and Medical Research Council guidance [57, 58]. Process data are described below and will be quantitative and qualitative. We will collect core process evaluation data (including fidelity and acceptability) across all sites and participants. We will purposively sample participants across different schools and waves. The process evaluation will assess intervention implementation, mechanisms of change and contextual factors.

Implementation

Under implementation we will assess fidelity, reach and acceptability across all sites and participants.

We will examine fidelity to explore whether IPT was delivered as per the intervention manual and by competent facilitators. Across IPT groups, we aim to assess 20-25% of IPT sessions delivered. During these sessions supervisors will observe the group in person or listen to audio recordings of the group and use checklists developed for each phase of IPT to assess the extent to which standardised components of IPT are being implemented and at what level (satisfactory, superior). Through focus group discussions with facilitators and supervisors we will explore deviations from the manual and whether they were intentional or otherwise, and whether they deviated or adhered to the theory of change. Within IPT, adolescents are encouraged to generate and select solutions to their interpersonal problems with the support of group members and facilitators. Through qualitative interviews with adolescents in case study schools, focus group discussions with facilitators and supervisors, and analysis of IPT session recordings, we will explore whether these solutions were consistent with our theory of change. We will explore how fidelity varies across schools, IPT groups and over the course of the trial, and potential contextual factors that affect this.

Reach is defined as the extent to which the target population comes into contact with the intervention [57]. We will assess the percentage of schools and individual adolescents approached that agree to participate.

To assess acceptability we will collect quantitative and qualitative data. In terms of quantitative data, across intervention schools we will assess the percentage of IPT sessions attended by adolescents using information from facilitator attendance records (overall and by gender, caste/ethnicity and socioeconomic status). If feasible, we will compare with school attendance records to understand whether an absence at a session coincided with an absence at school or whether adolescents were absent at the session only. In all intervention schools, we will administer treatment satisfaction surveys to IPT participants at endline. In terms of qualitative data on acceptability, we will conduct interviews and focus group discussions with adolescents, caregivers and teachers as well as facilitators and supervisors, using topic guides informed by the General Theory of Implementation [58]. We will explore: whether IPT is perceived as needed and valuable (coherence) and if this changes over the course of the intervention; how willing adolescents, families and schools were to invest in and commit to IPT (cognitive participation) and in what ways they demonstrated this commitment; how well IPT fitted into existing school and household routines (collective action); and how people felt IPT went (reflexive monitoring). We will also explore: intervention capability (whether IPT was workable within the education system); participant potential (individual and shared readiness for the intervention); institutional capacity (whether existing material and cognitive resources were sufficient, and how social norms and roles served to promote or undermine intervention activities); and contextual factors that affect intervention acceptability, for example whether implementation was better in schools that perceived a greater need and demonstrated more commitment to IPT.

To describe enhanced usual care in the control arm, we will examine quantitative data on adolescent mental health service use in surveys, review case notes for participants referred to the counsellor (number of participants assessed, number of counselling sessions conducted, number and type of referrals made) and conduct focus group discussions with the counsellors. Contamination across arms is extremely unlikely and only possible if for example, a participant in a control school and participant from the intervention arm were in the same household or friendship group, or a participant from the intervention arm moves to a school in the control arm. Through surveys with parents, we will identify if any participants live in the same household. From IPT facilitator notes, we will be informed if any intervention participants move schools.

Table 4 details qualitative and quantitative data sources for the process evaluation.

Mechanisms of change and contextual factors

Informed by realist evaluation principles and work by Bonell and colleagues on realist randomised trial methodology [18, 19, 59], the process evaluation seeks to refine our theory of change about how IPT works, for whom and in what circumstances. To do this we have generated hypotheses in the form of CMOs about how intervention mechanisms and contextual factors interact to generate outcomes. We define mechanisms as changes in the reasoning and behaviour of individuals or the way in which institutions operate in response to local use of intervention resources [19]. Whilst mechanisms cannot be directly observed they can be inferred for example through quantitative data on mediator variables or qualitative data collected from intervention participants. We define context as aspects of the place or people within which an intervention is implemented that might affect how the mechanisms operate. Context can relate to material or non-material features (social, psychological, organisational, political and economic), and operate at all levels of the social ecological framework including individual, family, school, community and government.

We also propose ‘dark logic’ hypotheses related to mechanisms of potential harm, where mechanisms of harm might generate ‘paradoxical effects’ on our primary or secondary outcomes or generate ‘harmful externalities’ in other domains [60].

We will draw on findings from analyses of the qualitative process data to refine (including removing or adding to) our hypotheses. This ensures we remain open to pathways that are not already specified in our process evaluation. We will then test these refined hypotheses using analyses drawing on quantitative trial data to check these data align with our theoretical predictions [59]. We will use mediation analyses to examine underlying mechanisms that explain the effect of IPT on adolescent depression. Moderation analysis will be used to assess whether a measure of contextual factor moderates the effect of IPT on depression. We will also use moderated mediation analysis to test whether mediation is contingent on a contextual factor.

We use categories proposed by Bonell and colleagues in a previous realist RCT to group our hypotheses [61]. These categories are intervention mediators, contextual factors affecting implementation, contextual factors modifying intervention effects and potential harms. In Table 5 we present data and analyses we will use to test each hypothesis.

Hypotheses about intervention mediators

In our theory of change we outline hypothetical pathways to reduced depression related to emotion regulation, conflict reduction, social support, self compassion, hope and self-efficacy. With the exception of self compassion, we generated hypotheses related to these pathways that could potentially be refined through qualitative analysis and tested quantitatively using trial data. For self-compassion we identified only one suitable tool to measure this construct that was developed in the US [62]. We were unable to establish the cultural equivalence of this tool in a Nepali context before the start of the trial, hence it was not included in the survey battery. Therefore, although we include a hypothesis related to self-compassion (H6) we are unable to test this in quantitative analyses.

H1: Adolescents participating in IPT will report increased social support at midline, endline and follow up due to bonding between group members, initiation of new relationships outside groups and/or improvement in the quality of existing relationships. This generates reductions in depression

H2: Adolescents participating in IPT will report reduced interpersonal conflict at midline, endline and follow-up due to improved communication of interpersonal needs and strategies to better manage anger. This generates reductions in depression.

H3: Adolescents participating in IPT will experience improved emotion regulation at midline, endline and follow up because they learn how to name and express their emotions. This generates reductions in depression.

H4: Adolescents participating in IPT will report increased hope at midline as participants perceive the possibility of change in mood and relationships. This generates reductions in depression

H5: Adolescents participating in IPT will report improved self-efficacy at midline, endline and follow-up because they experience generating solutions to and solving their interpersonal problems. This generates reductions in depression.

H6: Adolescents participating in IPT will report improved self-compassion at midline, endline and follow up because in IPT groups they normalise and validate their experience of depression. This generates reductions in depression.

Hypotheses about contextual factors affecting implementation

H7: Adolescents participating in IPT who are from the poorest backgrounds whose basic needs are not met will attend fewer IPT sessions because they have to undertake paid work and/or have household responsibilities to fulfil.

H8: Groups involving adolescents across two grades (8/9) rather than three (7/8/9) will be more cohesive and attendance among participants will be higher because participants will be closer in age and maturity and sessions will be easier to schedule across fewer timetables.

H9: In schools that have experienced a suicide, an attempted suicide or mass conversion disorder (chopnne), IPT will be more acceptable and implemented with more fidelity because staff will be more motivated to support the programme.

Hypotheses about contextual factors influencing intervention effects

We present hypotheses about contextual factors influencing intervention effects as CMOCs in line with our realist evaluation approach.

H10: Adolescents participating in IPT who attend schools with a climate characterised by a lack of understanding and supportive teacher and student relationships (context), will build strong peer to peer social support, a group identity and a sense of belonging with IPT group members (mechanism) that will offset the negative effect of the school climate on adolescents' depression (outcome).

H11: Adolescents participating in IPT from a lower socioeconomic background (poorer household, context) will be more engaged and motivated to address their problems (affordance) and have more problems to solve (vulnerability) compared to their peers from wealthier backgrounds. These adolescents will share their feelings and problems and participate in discussions through which they learn strategies to regulate their emotions, reduce conflict, build social support and improve self-efficacy (mechanism) and therefore experience more benefit from IPT compared to peers from a higher socioeconomic background (outcome).

H12: For girls, who are subject to gender norms which discourage them from expressing their feelings (context), group IPT will serve as a platform to become aware of and share feelings and learn emotion regulation skills (mechanism) which reduce their depression. On the other hand, boys who are exposed to gender norms which promote aggression and anger (context), through IPT learn how to reduce interpersonal conflict (mechanism) and therefore experience reductions in depression (outcome).

H13: For adolescents with a parent who has a mental health problem (context), through IPT they will increase their social support, normalise and validate their feelings, generate and implement solutions to their problems and reduce conflict at home (mechanism). This will offset the effect of the parent's mental health on the adolescent's depression (outcome).

Hypotheses about potential harms

It is very unlikely that IPT will cause harm but we will nonetheless examine the possibility. In addition to the hypotheses listed below, we will explore unanticipated harms in the qualitative process data.

H14: For adolescents with a background of trauma (context), who recount their traumatic experiences during group sessions and to whose experiences other group members cannot relate,

IPT will exacerbate their distress due to re-experiencing the stressful event (mechanism), which will result in a worsening of depression (outcome).

H15: Adolescents experiencing stigma from school and community members because of their participation in IPT would attend fewer sessions and the stigma would cause them distress and their depression worsens.

H16: In groups adolescents are encouraged to generate and select solutions to solve their interpersonal problems. If the group is poorly facilitated these solutions may run counter to our theory of change and could cause adolescents harm for example by inciting violence in the context of a dispute. This could exacerbate adolescents' depressive symptoms.

H17: Group members or facilitators may breach an adolescent's confidentiality by disclosing sensitive information about them to individuals outside the group. This would be distressing for the adolescent and could have harmful consequences resulting in the worsening of their depression.

Economic evaluation data collection

To estimate the costs of the intervention and control conditions, monthly project budget actuals will be entered in an Excel tool by TPO Nepal staff. We will divide costs into start-up or implementation costs and allocate them to different cost centres and intervention activities by treatment arm. Major expected cost categories for both intervention and control groups include *personnel, equipment, travel, supervision and monitoring, and finance and administration*. TPO staff members will allocate staff salaries between treatment arms; identify relevant costs incurred during the pilot which also contribute to the full trial cost; and document use of facilities and supplies in both Chitwan and Kathmandu office locations.

To estimate indirect societal costs, we will document productivity loss and informal care costs by recording adolescents' time spent on intervention-related activities (including travel and participation) if relevant, as well as time and costs associated with outside healthcare utilisation. At endline and follow-up, we will collect data on adolescent health care visits since randomisation, including the time and out-of-pocket expense for each visit (across a range of non-school based formal and informal healthcare types available locally). For each type of healthcare service received, we will establish whether its cost is fully covered by out-of-pocket expenditures, or whether it was partially or fully subsidised by the provider. We will record the estimated subsidised cost of services using publicly available information; for any subsidised costs we cannot confirm directly, we will contact service providers to establish an appropriate estimate or range. Finally, at endline and follow-up we will ask about how much time parents spent attending to adolescents due to their depression (e.g., staying home from work or accompanying them to appointments); whether adolescents engaged in any paid work in the previous period; adolescents' hourly wages; and the number of days adolescents missed work due to depression. Endline and follow up health cost surveys will be administered to caregivers because we anticipate they will have more insight into actual costs compared to adolescents. However, if adolescents are living away from their caregiver they will complete the surveys themselves.

The primary health outcome of interest for the economic evaluation analyses is the quality-adjusted life year (QALY). QALYs will be estimated by administering the EQ-5D-5L with adolescents at each planned timepoint for data collection. Using this data, we will derive QALY weights using existing value sets from India, which represent the closest cultural context for which EQ-5D-5L value sets are available [63].

Power and sample size

Primary analysis

The aim of the primary analysis is to test the overall impact of IPT on adolescent depressive symptoms compared to enhanced usual care through a two-armed cluster RCT with secondary schools as the unit of randomisation. To fulfil the study objectives, we plan to recruit 48 schools, with 24 schools per arm with 10 to 16 adolescents per school taking part, giving a range of 480 to 768 adolescent individuals taking part. As the sample size is fixed due to logistics, we will take the approach of estimating the effect which can be detected by our cluster randomised design. For the purpose of sample size calculation, we will need the intraclass correlation coefficient (ICC, the proportion of variation due to variation between clusters), the standard deviation of the depression measure, the arithmetic mean of the cluster size, the standard deviation of cluster size variation in order to calculate the coefficient of variation for cluster size (cluster SD / cluster mean), the number of clusters, desired power, alpha level, direction of test (here two-way) and the rate of dropout of schools and adolescents. This information will allow us to use `clustersampsi`, a stata module for cluster randomised trial power analysis.

We use the baseline standard deviation of included adolescents from the pilot trial PHQ-A, which is 3.05. Assuming the distribution of clusters is uniform across the 10 to 16 range, with an arithmetic mean of 13 and standard deviation of 1.97, results in a coefficient of variation for cluster size of 0.151. However, in the pilot study 12 / 161 (8%) of adolescents dropped out by endline (the primary endpoint), which would approximate one adolescent per cluster. Here we'll assume both 8% and a more conservative drop-out of two per cluster giving an additional missingness rate at individual level of 15%. No schools dropped out in the pilot study, but we will also assume one school in each group will be lost to follow-up. This leaves us with 46 schools with an average of 12 or 11 per cluster and coefficients of variation of 0.16 and 0.181 respectively. We use both a conservative ICC of 0.07 for the PHQ-A based on data collected on symptoms of depression using the Depression Self Rating Scale during the IPT feasibility study [17], and a more liberal value of 0.03 as ranges from 0.05 to 0.01 are more typical in community interventions. Finally, we assume a two-way test with 90% power and alpha set to 0.05. Given these parameters the smallest detectable effect size at the primary outcome timepoint at 8% missingness and ICC = 0.07 would be 0.29 (0.3 at 15% missingness) and the design effect (the effective reduction in sample size) due to clustering 1.79 (1.73 at 15% missingness). For an ICC of 0.03, the smallest detectable effect size would be 0.25 at 8% missingness and 0.26 at 15% missingness (design effects of 1.34 and 1.31 respectively). Hence, the study is well powered to detect small to moderate effects. These standardized effect sizes correspond to group difference 0.9 or less on the PHQ-A scale.

Mediation analysis for CMOCs

Whilst well-powered to detect modest effect sizes of 0.29, the sample size also means that we should have power to assess the CMOCs with a selection of multi-level moderation, mediation and mediated moderation models. Nevertheless, more complex models and effects such as the mediated direct effect mean sensitivities of the tests will be lower and detectable effects will necessarily be larger. Estimates of power for mediation and moderation analyses will depend on the CMOCs selected for testing but here we give some guidelines. For mediation, taking a simplistic approach using Table 3 from Fritz and McKinnon 2007 [64], a sample size (assuming the loss to follow up rates above) of 46 schools of cluster size 12 or 11 and ICC of 0.07 gives effective sample sizes of 308 and 292 (or 412 and 386 for an ICC of 0.03) and will provide 80% power to detect small to moderate mediation effects with a and b path coefficients of > 0.26. The minimum sample size for this permutation of effects (minimum 148) allows a substantial buffer (50%) wherein the true sample size

needed is larger due to the effects of clustering on the variance of the indirect effect being underestimated.

Moderation analysis for CMOCs

To evaluate power for a moderation analysis we ran simulations at different effect sizes for our fixed sample size and design effects of 1.79, 1.73, 1.34 and 1.31 (depending on the ICC and coefficient of variation of group size at missingness rates of 8% and 15%) to address clustering and assuming $\alpha = 0.05$. We used a simple model where treatment group interacted with a categorical variable such as gender dividing the treatment groups into equal sub-groups. The interaction effect size then represents the difference in the effect of treatment between the sub-groups and in the simulations ranged from 0.25 to 0.75 in increments of 0.01. As can be seen in Figure 4, assuming a drop out rate of one school per group and 8% or 15% of participants in remaining schools there will be power to detect a moderate to large ES of 0.635 (ICC = 0.07, missingness = 8%, DE = 1.79), 0.655 (ICC = 0.07, missingness = 15%, DE = 1.73), 0.555 (ICC = 0.03, missingness = 8%, DE = 1.34) or 0.575 (ICC = 0.03, missingness = 23%, DE = 1.28). Given the baseline SD of PHQ-A from the pilot trial (SD = 3.05) that gives a range of differences between groups of 1.7 to 2.0 on the PHQ-A scale.

Moderated mediation analysis for CMOCs

Finally, we consider moderated mediation models. In a moderated mediation there are four variables: the independent variable; the outcome; the mediator; and a fourth variable which moderates the indirect path (either path a, path b or both). For illustration purposes we use the case of a model in which the moderation happens to the a pathway only and let us assume gender moderates the relationship between treatment condition and the hope outcome. We use the above result for moderation as an estimate of the power for the moderated a pathway. Similarly, we can use a further simulation to find power of the b path, a model consisting of two predictors representing paths from the mediator and independent variable to the outcome. Assuming that the effects are independent, the power of the mediation or joint model is the power of the a path * power of the b path. If the value of this product is 0.8 or greater, then the a * b pathway will have 80% power. For example, this might occur if the power of the a path was 0.81 and the power of the b path 0.99 corresponding to a and b paths of at DE = 1.79 (the more conservative ICC estimate assuming missingness of 8% and 1 cluster per group). Other scenarios are shown in Table 6 along with the power curves for the b path from which the coefficients corresponding to power levels can be read in Figure 5. By Fritz and Mckinnon's schema we would expect the moderated a path effect size to be large and the b path to be medium, ranging small-medium to medium-large.

Data management and analysis

Data management

Pseudonymised survey data will be collected using tablets via an online platform. Data downloaded will be saved on TPO Nepal and UCL secure servers. Information linking participant IDs to their personal information and any data that contains participant information (e.g., audio recordings, transcripts) will be kept on TPO Nepal's server or a password-protected computer. Hard copies that contain participant information (e.g., consent forms) will kept in locked cabinets. The UCL research team members will not have access to participant information.

Audio recordings of interview and focus group discussions will be transcribed in full by the research team and securely transferred to TPO Nepal approved contractors for translation. The translations will be anonymised, stored on password protected devices and transferred securely to researchers

involved in the analysis. All reporting will be anonymised. To facilitate any secondary analyses, we will retain anonymised data for 20 years after the study as per MRC guidance.

Qualitative analysis

To analyse qualitative data pertaining to mechanisms of change and contextual factors, we will use dimensional analysis. This is a variant of grounded theory useful for refining CMOCs because of its focus on identifying and organising dimensions of experience — meanings, processes, conditions and consequences [65]. For analyses of implementation data, we will use the thematic analysis. We will develop a coding framework with deductive codes informed by topic guides and also allow codes to emerge from the data (inductive coding). Where possible, the researcher who conducted the interviews/focus group discussions will be involved in the transcription and analysis, and we will conduct member checks to validate the findings.

Quantitative analysis

A detailed, pre-specified statistical analysis plan will be finalised and approved by the DMC and the statistician of the TSC before the database lock. This plan will be available from the authors upon request. Reporting of results will adhere to the relevant Consolidated Standards of Reporting Trials (CONSORT) guidelines.

We define the target treatment effect as follows:

- **Population:** 13 to 19 years old adolescents from a randomised cluster (public mixed sex secondary school) who meet the trial inclusion / exclusion criteria and are identified as depressed by the PHQ-A (scoring 11 or more).
- **Outcome:** Depression scores at baseline, 17 weeks post randomisation.
- **Intervention:** IPT compared to enhanced treatment as usual irrespective of treatment being discontinued (Intention to Treat, ITT).
- **Population level summary measure:** Individual mean PHQ-A score at 17 weeks post baseline assessment for each trial arm.
- **Comparison:** Mean difference between the trial arms at the participant-level at baseline, 17 weeks post baseline assessment

Following the principle of ITT means adherence or compliance to allocated treatment is ignored and participants are followed up and analysed as part of their allocated groups. In this approach data is collected after treatment discontinuation and missing data is included as if it was observed. The pattern of withdrawals in each treatment arm will be assessed and any necessary remedial action will be approved by the TSC.

A descriptive summary of demographic and clinical variables at baseline and primary and secondary outcomes at baseline and post-randomisation (midline, endline and follow-up timepoints) will be reported by intervention group and overall. Mean, SD, medians and IQR will be used to summarise continuous variables and frequencies and percentages categorical variables respectively. Variables will be summarised at the cluster and individual level as appropriate.

The primary outcome of interest, the mean PHQ-A difference between the arms and 95% confidence intervals, will be estimated using mixed effects linear regression to take account of both the clustering due to the unit of randomisation being school and repeated measures by individual. The midline, endline and follow-up measures of the outcome measure are the dependent variable with a random intercept for cluster/school and an unstructured covariance matrix at the level of participant residuals to allow for repeated measures. This mmrm approach is preferred as it avoids model misspecification

and is unbiased when data is missing at either random or completely at random [66]. The model will include the baseline measure of PHQ-A (ANCOVA approach), a dummy variable for trial arm, for timepoint and a trial arm by timepoint interaction to allow the treatment effect to vary across time. As randomisation is stratified by wave, this will always be included as a covariate in the analysis model. The model will be estimated using restricted maximum likelihood (REML) allowing the inclusion of all participants, with missing data as ignorable under the assumption of missing at random. Continuous secondary outcomes will be analysed using the same methods. Standardised effect sizes will be calculated by dividing the treatment arm difference estimate by the baseline standard deviation. Model assumptions will be verified, linear mixed effects models assume residuals are normally distributed at the individual and cluster level and these will be checked appropriate data transformations will be applied if necessary and / or bootstrap procedures appropriate for multilevel models.

To assess and report missing data, we will adhere to CONSORT guidelines. This includes presenting a table comparing baseline characteristics of patients with and without follow-up data, stratified by treatment arm, to identify potential bias. We do not expect any missing data at baseline, but if there is substantive missing data we will consider using multiple imputation by chained equations to address this. Missing data in the outcomes is accounted for by the designated mixed model estimated using restricted maximum likelihood estimation, which is valid under the 'missing at random' (MAR) assumption. This allows us to include all participants with at least one follow-up measurement to satisfy the ITT principle. To substantiate the MAR assumption, we will examine various baseline factors (e.g., demographics, clinical characteristics) as potential predictors of missingness and include significant predictors in the analysis model to improve its accuracy. As a sensitivity analysis we will use multiple imputation with a substantive model which is consistent with the primary analysis model described above. Additionally, with multiple imputation we can consider using plausible ranges of patterns of missing data which allow departure from the MAR assumption based on the protocol of White et al, 2011 [67].

There are several planned additional secondary analyses of the primary outcome. To assess the effect of potential chance imbalance at baseline, effects of the size of grade 8- or 9-year groups per school will be controlled for and included as an additional covariate in the analysis models. Any changes in treatment effect will be noted. We will address the impact of treatment adherence using per protocol and Complier Average Causal Effects. In further secondary analysis we will evaluate CMOCs using planned multilevel mediation, moderation and moderated mediation analysis, noting that moderators (and potentially mediators) can occur at the cluster or individual level. These analyses are listed in Table 5.

Economic analysis

We will conduct a cost-utility analysis (CUA) to produce an incremental cost-utility ratio (ICUR) reflecting the incremental cost per additional QALY gained by participating in the intervention, with and without discounting. Probabilistic sensitivity analyses will vary cost inputs in 10,000 Monte Carlo simulations and will assess the probability of IPT being cost-effective at various willingness-to-pay thresholds. The time horizon for accumulating both costs and outcomes will be 30 months (the length of the full trial as well as follow-up) using 1-month time steps, capturing the longest possible timeframe for the intervention effects for which data collection is feasible. An annual discount rate of 3% will be applied to both costs and effects.

Subsequently we will conduct a benefit-cost analysis (BCA). The BCA will use two approaches to valuing outcomes: 1) a human capital valuation approach using Nepal's national minimum wage, and 2) a valuation of statistical life year approach (VSLY) for Nepal. Each approach will compare the net

present value (NPV) of the QALYs gained in each treatment arm. Benefit-cost ratios and net benefits will be calculated for each approach. The probability of intervention vs. control conditions' having positive NPV under each valuation approach will then be calculated via probabilistic sensitivity analysis with 10,000 replications.

Minimising bias

To mitigate selection bias we will conduct randomisation after the baseline survey. Randomisation of schools rather than individual adolescents will mitigate contamination. The trial team will have no way of predicting the allocation sequence. We will reduce detection bias (systematic differences in outcome assessments by arm) by masking research assistants. Clinical and research teams will be managed by different individuals in separate offices. The research assistants will not be informed about the nature of the intervention. To minimise attrition bias we will maximise school retention by building strong relationships with school staff, especially school principals at an early stage in the research process through orientation meetings and frequent communication. In intervention and control arms in each wave, we will also offer one or two teachers in each school the opportunity to participate in a psychoeducation workshop after endline. We will optimise adolescent retention at follow up by maintaining multiple lines of communication with adolescents and their family members (address, phone, social media ID), checking in with adolescents between follow-up surveys, incentivising participation with non-monetary items and phone credit, having flexible office hours so research assistants can conduct interviews with adolescents in the early morning or evening where necessary, and enabling research assistants to travel outside the study area (within Nepal) where adolescents have moved away.

To minimise reporting bias, we will publish a statistical analysis plan specifying primary, secondary and exploratory analyses. The trial statistician will be masked to allocation and the primary analyses will be on an intention-to-treat basis. Analyses will be overseen by the DMC. Exploratory analyses will focus on testing the CMOCs and hypotheses related to intervention fidelity, reach and acceptability (see process evaluation section).

Public and policy involvement

Our protocol is informed by several stakeholder groups. We adapted IPT for adolescents in Nepal through formative research with adolescents, parents, teachers and health workers in Nepal. We assessed the feasibility and acceptability of the intervention through qualitative interviews with the same stakeholder groups.

In the pilot cRCT we piloted methods for this phase III cRCT and sought feedback from stakeholders on their acceptability. The interim qualitative and quantitative findings from the trial were presented to our Trial Steering Committee (TSC) consisting of local Nepali and international experts. These findings informed the study procedures of the current protocol (e.g., screening procedures, inclusion and exclusion criteria, safety protocols, survey timing, survey battery, mechanism hypotheses). The same TSC will continue to meet throughout the Phase III trial. We will also convene a Data Safety and Monitoring Board to monitor the progress of the trial.

To input on the protocol and advise on trial issues, we have convened a youth advisory board in the study setting comprising ~8 local adolescent boys and girls who participated in the pilot cRCT. We will also convene an advisory board of a similar size for parents and teachers.

Ethical issues, safeguarding and SAEs

We have ethical approval from the Nepal Health Research Council (2144) and University College London Research Ethics Committee (0351).

At the school level, we will seek permission to conduct the study from the school principal. Eligible adolescents who wish to participate in the trial will be asked to take home a written information sheet and consent form. We will obtain written consent from adolescents which will be collected before the baseline survey is conducted. For adolescents aged 17 and younger we will take consent from their caregiver and adolescent assent. In the study setting some adolescents do not live with their parents due to parental migration or death, because the adolescents are living in a hostel close to their school or because they are married and live with their in-laws. For these adolescents we will accept consent from an adult responsible for their care, e.g. a grandparent, aunt or uncle, sibling or spouse aged 18 or over, parent-in-law or hostel warden. We will obtain consent for all adult participants from whom we collect data (e.g. teacher, IPT facilitator, parent).

We anticipate disclosures of suicidality throughout the trial. In our pilot trial 7.6% (85/1108) of adolescents screened were deemed to be at high risk of suicide and excluded from the trial. We have therefore developed an SOP to manage these cases. This involves the adolescent being assessed by a psychosocial counsellor and potentially referred to local mental health services. For surveys conducted by phone, the research assistant will schedule the survey for a time when the adolescent is at a specific location where the address is known, and is nearby an adult family member who could be reached in the event of a risk situation. If these conditions are not met, the interview will be rescheduled.

There is a separate SOP for disclosures of abuse which includes a safety assessment and planning and potential referral to the women and children government officer attached to the local municipality, or to a hospital where physical and mental health problems can be further assessed and managed.

Dissemination

We will proactively share findings with schools and advisory boards by developing simple research summaries, supported by workshops where requested. We will write an article for a Nepali newspaper about the research. In Chitwan and Kathmandu we will hold workshops to disseminate findings to policy makers and relevant organisations.

We will disseminate findings to academic audiences through publication in international peer-reviewed journals, and presentation at conferences. We will also hold a webinar for global policy makers and academics inviting participants from UNICEF and WHO.

To disseminate to audiences more broadly we will commission a film on adolescent depression, understandable to Nepali and English speakers. This will be disseminated through online platforms such as Twitter, The Conversation and Mental Elf.

Availability of data and materials

We will make data available after the main trial analyses have been conducted. Researcher requests for data should be directed to the principal investigator with a clear protocol. Approval for use of the data will be contingent on researchers obtaining the requisite ethical approval.

References

1. DAMOCLES Study Group and NHS Health Technology Assessment Programme, *A proposed charter for clinical trial data monitoring committees: helping them to do their job well*. Lancet, 2005. **365**(9460): p. 711-22.
2. World Health Organization and Columbia University, *Group interpersonal therapy (IPT) for depression. (WHO generic field-trial version 1.0)*. 2016, WHO: Geneva.
3. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators, *Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017*. The Lancet, 2018. **392**(10159): p. 1789-1858.
4. Thapar, A., et al., *Depression in adolescence*. Lancet, 2012. **379**(9820): p. 1056-67.
5. United Nations, *World Population Prospects 2017*, UN Department for Economic and Social Affairs (UN DESA) and Population Division, Editors. 2017: New York.
6. Singla, D.R., et al., *Psychological treatments for the World: lessons from low- and middle-income countries*. Annual Review of Clinical Psychology, 2017. **13**(149-181).
7. Barbui, C., et al., *Efficacy of psychosocial interventions for mental health outcomes in low-income and middle-income countries: an umbrella review*. Lancet Psychiatry, 2020. **7**(2): p. 162-172.
8. Zhou, X., et al., *Comparative efficacy and acceptability of psychotherapies for depression in children and adolescents: A systematic review and network meta-analysis*. World Psychiatry, 2015. **14**(2): p. 207-22.
9. Cuijpers, P., M. Reijnders, and M.J.H. Huibers, *The Role of Common Factors in Psychotherapy Outcomes*. Annual Review of Clinical Psychology, 2019. **15**(1): p. 207-231.
10. Markowitz, J.C. and M. Weissman, *Interpersonal psychotherapy: principles and applications*. World Psychiatry, 2004. **3**(3): p. 136-139.
11. Duffy, F., H. Sharpe, and M. Schwannauer, *The effectiveness of interpersonal psychotherapy for adolescents with depression - a systematic review and meta-analysis*. Child and Adolescent Mental Health, 2019. **24**(4): p. 307-317.
12. Petersen, I., et al., *A group-based counselling intervention for depression comorbid with HIV/AIDS using a task shifting approach in South Africa: A randomized controlled pilot study*. Journal of Affective Disorders, 2014. **158**: p. 78-84.
13. Bolton, P., et al., *Interventions for depression symptoms among adolescent survivors of war and displacement in northern Uganda: a randomized controlled trial*. JAMA, 2007. **298**(5): p. 519-27.
14. Rose-Clarke, K., et al., *Culturally and developmentally adapting group interpersonal therapy for adolescents with depression in rural Nepal*. BMC Psychology, 2020. **8**.
15. Hassan, E., et al., *Community perspectives on the implementation of a group psychological intervention for adolescents with depression: A qualitative study in rural Nepal*. Frontiers in Psychiatry, 2022. **13**.
16. Rose-Clarke, K., et al., *A cross-cultural interpersonal model of adolescent depression: a qualitative study in rural Nepal*. Social Science & Medicine, 2021. **270**: p. 113623.
17. Rose-Clarke, K., et al., *School-based group interpersonal therapy for adolescents with depression in rural Nepal: a mixed methods study exploring feasibility, acceptability, and cost*. Global Mental Health, 2022. **9**: p. 416-428.
18. Bonell, C., et al., *Realist randomized controlled trials: A new approach to evaluating complex public health interventions*. Social Science & Medicine, 2012. **75**: p. 2299–2306.
19. Pawson, R. and N. Tilley, *Realistic Evaluation*. 1997, London: SAGE.
20. Jamal, F., et al., *The three stages of building and testing mid-level theories in a realist RCT: a theoretical and methodological case-example*. Trials, 2015. **16**(1): p. 466.

21. Nielsen, S.B., S.Ø. Jaspers, and S. Lemire, *The curious case of the realist trial: Methodological oxymoron or unicorn?* Evaluation, 2024. **30**(1): p. 120-137.
22. Warren, E.A., G.J. Melendez-Torres, and C. Bonell, *Are realist randomised controlled trials possible? A reflection on the INCLUSIVE evaluation of a whole-school, bullying-prevention intervention.* Trials, 2022. **23**(1): p. 82.
23. Atmore, K., et al., *Exploring context, mechanisms and outcomes in group interpersonal therapy for adolescents with depression in Nepal: a qualitative realist analysis.* Cambridge Prisms: Global Mental Health [in press].
24. Jordans, M.J.D., et al., *Effectiveness of Group Problem Management Plus, a brief psychological intervention for adults affected by humanitarian disasters in Nepal: A cluster randomized controlled trial.* PLoS Med, 2021. **18**(6): p. e1003621.
25. Luitel, N.P., et al., *Translation, cultural adaptation and validation of Patient Health Questionnaire and generalized anxiety disorder among adolescents in Nepal.* Child Adolesc Psychiatry Ment Health, 2024. **18**(1): p. 74.
26. Sapkota, R.P., A. Brunet, and L.J. Kirmayer, *Characteristics of Adolescents Affected by Mass Psychogenic Illness Outbreaks in Schools in Nepal: A Case-Control Study.* Front Psychiatry, 2020. **11**: p. 493094.
27. Posner, K., et al., *The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults.* Am J Psychiatry, 2011. **168**(12): p. 1266-77.
28. Little, A.W. and O.B. Azubuike, *Young Lives School Surveys 2016–17: The Development of Non-Cognitive Instruments in Ethiopia, India and Vietnam*, Young Lives, Editor. 2017, University of Oxford: Oxford.
29. Marsh, H.W., et al., *OECD's brief self-report measure of educational psychology's most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries.* International Journal of Testing, 2006. **6**(4): p. 311- 360.
30. De Silva, M.J., et al., *Theory of Change: a theory-driven approach to enhance the Medical Research Council's framework for complex interventions.* Trials, 2014. **15**: p. 267.
31. Lipsitz, J.D., & Markowitz, J. C., *Mechanisms of change in interpersonal therapy (IPT).* Clinical Psychology Review, 2013. **33**(8): p. 1134–1147.
32. Ravitz, P., R. Maunder, and C. McBride, *Attachment, Contemporary Interpersonal Theory and IPT: An Integration of Theoretical, Clinical, and Empirical Perspectives.* Journal of Contemporary Psychotherapy, 2008. **38**(1): p. 11-21.
33. Tajfel, H., et al., *An integrative theory of intergroup conflict*, in *The Social Psychology of Intergroup Relations*, W.G. Austin and S. Worchel, Editors. 1979, Brooks/Cole: Monterey, CA. p. 33-47.
34. Neff, K.D., *Self-Compassion: Theory, Method, Research, and Intervention.* Annual Review of Psychology, 2023. **74**(Volume 74, 2023): p. 193-218.
35. Nepal Statistics Office, *National Population and Housing Census 2021.* 2021, Government of Nepal, Thapathali, Kathmandu.
36. Bhatta, P. and T. Pherali, *Nepal: Patterns of privatisation in education. A case study of low-fee private schools and private chain schools.* 2017, Education International Research.
37. Pradhan, U. and K. Valentin, *Free education? Blurred public-private boundaries in state-run schooling in Nepal.* Studies in Nepali History and Society, 2020. **25**(2): p. 277-298.
38. Gautam, P., et al., *Depression among Adolescents of Rural Nepal: A Community-Based Study.* Depression Research and Treatment, 2021. **2021**(1): p. 7495141.
39. Solmi, M., et al., *Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies.* Molecular Psychiatry, 2022. **27**(1): p. 281-295.
40. Kohrt, B.A., et al., *Therapist competence in global mental health: Development of the ENhancing Assessment of Common Therapeutic factors (ENACT) rating scale.* Behaviour research and therapy, 2015. **69**: p. 11-21.

41. Johnson, J.G., et al., *The patient health questionnaire for adolescents: Validation of an instrument for the assessment of mental disorders among adolescent primary care patients.* Journal of Adolescent Health, 2002. **30**(3): p. 196-204.
42. Spitzer, R.L., et al., *A brief measure for assessing generalized anxiety disorder: the GAD-7.* Archives of Internal Medicine, 2006. **166**(10): p. 1092-7.
43. Bolton, P. and A.M. Tang, *An alternative approach to cross-cultural function assessment.* Social Psychiatry and Psychiatric Epidemiology, 2002. **37**(11): p. 537-543.
44. Price, M., et al., *Investigation of abbreviated 4 and 8 item versions of the PTSD Checklist 5.* Psychiatry Res, 2016. **239**: p. 124-30.
45. Kane, J.C., et al., *Mental health and psychosocial problems in the aftermath of the Nepal earthquakes: findings from a representative cluster sample survey.* Epidemiology and Psychiatric Sciences, 2017: p. 1-10.
46. Kohrt, B.A., et al., *Cross-cultural gene- environment interactions in depression, post-traumatic stress disorder, and the cortisol awakening response: FKBP5 polymorphisms and childhood trauma in South Asia.* Int Rev Psychiatry, 2015. **27**(3): p. 180-96.
47. Thapa, S.B. and E. Hauff, *Psychological distress among displaced persons during an armed conflict in Nepal.* Soc Psychiatry Psychiatr Epidemiol, 2005. **40**(8): p. 672-9.
48. The EuroQol Group, *EuroQol-a new facility for the measurement of health-related quality of life.* Health Policy 1990. **16**(3): p. 199-208.
49. Herth, P., *Measuring hope: Development of the Herth Hope Index.* Journal of Nursing Scholarship, 2000. **32**: p. 309-315.
50. Gratz, K.L. and L. Roemer, *Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the difficulties in emotion regulation scale.* J. Psychopathol. Behav. Assess., 2004. **26**: p. 41-54.
51. Schwarzer, R. and M. Jerusalem, *Generalized Self-Efficacy scale.* , in *Measures in health psychology: A user's portfolio. Causal and control beliefs.*, J. Weinman, S. Wright, and M. Johnston, Editors. 1995, NFER-NELSON: Windsor, England. p. 35-37.
52. Zimet, G.D., et al., *The Multidimensional Scale of Perceived Social Support.* Journal of Personality Assessment, 1988. **52**(1): p. 30-41.
53. Buhrmester, D. and W. Furman, *The Network of Relationships Inventory: Relationship Qualities Version. Unpublished measure.* 2008: University of Texas at Dallas.
54. Goldberg, D.P., et al., *The validity of two versions of the GHQ in the WHO study of mental illness in general health care.* Psychol Med, 1997. **27**(1): p. 191-7.
55. Frick, P.J., *Alabama parenting questionnaire.* Journal of Clinical Child Psychology, 1991.
56. Burkey, M.D., et al., *Validation of a cross-cultural instrument for child behavior problems: the Disruptive Behavior International Scale - Nepal version.* BMC psychology, 2018. **6**(1): p. 51-51.
57. Moore, G.F., et al., *Process evaluation of complex interventions: Medical Research Council guidance.* BMJ, 2015. **350**: p. h1258.
58. May, C., *Towards a general theory of implementation.* Implementation Science, 2013. **8**(1): p. 18.
59. Bonell, C., G.J. Melendez-Torres, and E. Warren, *Realist Trials and Systematic Reviews: Rigorous, Useful Evidence to Inform Health Policy.* 2024, Cambridge: Cambridge University Press.
60. Bonell, C., et al., *'Dark logic': theorising the harmful consequences of public health interventions.* Journal of Epidemiology and Community Health, 2015. **69**(1): p. 95.
61. Bonell, C., et al., *Effects of the Learning Together intervention on bullying and aggression in English secondary schools (INCLUSIVE): a cluster randomised controlled trial.* Lancet, 2018. **392**(10163): p. 2452-2464.
62. Neff, K.D., *The self-compassion scale is a valid and theoretically coherent measure of self-compassion.* . Mindfulness, 2016. **7**: p. 264-274.

63. Jyani, G., et al., *Development of an EQ-5D Value Set for India Using an Extended Design (DEVINE) Study: The Indian 5-Level Version EQ-5D Value Set*. Value in Health, 2022. **7**: p. 1218–1226.
64. Fritz, M.S. and D.P. Mackinnon, *Required sample size to detect the mediated effect*. Psychol Sci, 2007. **18**(3): p. 233-9.
65. Bonell, C., E. Warren, and G. Melendez-Torres, *Methodological reflections on using qualitative research to explore the causal mechanisms of complex health interventions*. Evaluation, 2022. **28**(2): p. 166-181.
66. Bell, M.L. and B.A. Rabe, *The mixed model for repeated measures for cluster randomized trials: a simulation study investigating bias and type I error with missing continuous data*. Trials, 2020. **21**(148).
67. White, I.R., *Strategies for handling missing data in randomised trials*. Trials, 2011. **12**: p. A59.

Tables

Table 1: Progression criteria for the pilot cluster-randomised controlled trial

Research question	Criterion	Indicator	Green	Amber	Red	Outcome
Is it feasible to deliver IPT in secondary schools in Chitwan?	Intervention implementation	Percentage of planned IPT sessions delivered	>70%	40-70	<40	83%
Is the intervention acceptable to participants?	Participant treatment satisfaction	Percentage of participants rating IPT as 'quite helpful' or 'very helpful';	>67	30-66	<30	94%
	Treatment adherence	Percentage of participants who attend more than 70% of IPT group sessions*	>50	20-50	<20	54%
Is it feasible to train local lay people and school nurses to deliver IPT?	Fidelity to IPT	Percentage of session components rated superior or satisfactory, averaged across observed sessions	>60	30-59	<30	84.9%
	Serious adverse events	Difference in serious adverse events in the intervention arm compared to the control arm	<4	4	5	-1
Are trial procedures (randomisation and data collection) acceptable to participants and schools?	Eligible adolescents agree to participate	Percentage of eligible adolescents with informed consent at baseline**	>80%	50-80	<50	~ 90%
	Eligible schools agree to participate	Percentage of schools approached that agree to participate	>60%	30-60	<30	100%
	Missing data	Percentage of missing items on primary and secondary outcome	<15	15-50	>50	0%
	Participant retention	Percentage of participants completing the endline survey	>70%	30-70	<30	93%
*We included all IPT sessions (pre-group and group sessions). Anyone who attended equal or more than 9 sessions in total were included in this calculation (70% of 12 sessions is 8.4 sessions). **NOTE: Current study protocol does not have a separate consent procedure after randomization. Adolescents and parents give consent for both screening and the full trial. TBD to be determined.						

Table 2: Context-mechanism-outcome contingencies from the pilot trial and related findings

CMOC	Qualitative findings	Quantitative findings
CMOC 1 Context: For adolescents of higher socioeconomic status (SES) not experiencing structural vulnerability or intractable adversity and having more opportunities to implement new strategies. Mechanism: IPT enables participants to participate in discussions through which they learn and implement strategies to develop hope, reduce conflict, build relationships and improve self-efficacy. Outcome: This generates reductions in depression	• Participants described separate pathways through which poverty and adversity affected adolescents' experience of IPT. • Participants perceived lower SES groups benefit most because they are "more engaged and motivated to solve problems" (affordance), they "share more" and they are "more mature". • For lower SES groups to benefit they may need one supportive family member. • Participants described how adolescents with more adversity might benefit less than those with no adversity. Adolescents who live in a house where they have a supportive relationship with at least one adult benefit more. • Acceptance came up as a pathway through which adolescents with adversity/trauma can improve their depression.	In moderator analyses, SES was assessed via caste and food sufficiency and Bayesian models were used. • Caste: Some evidence that depression might reduce more in middle and upper (e.g, Brahman) than lower (e.g., Dalit) caste groups in IPT. • Food sufficiency: No evidence depression reduces more when food sufficiency is less than 3 months. • Measure of SES inadequate in pilot evidenced by discrepancies between mother and adolescent in terms of household assets.
CMOC 2 Context: For adolescents in schools with school climates characterised by strong adolescent-staff and adolescent-adolescent relationships and norms of mutual respect and social support and thereby having more opportunities to implement new strategies. Mechanism: supportive for implementation or amenable for adolescents to implement ipt skills	• Participants perceived that the school climate might have affected adolescents' attendance or decision to participate in IPT	In moderator analyses, school climate was assessed via the Beyond Blue School Climate selected items, and Bayesian models were used. • Some evidence for an interaction effect between school climate and intervention group. • When school climate was less positive, IPT's effectiveness over control group was more pronounced, when school climate was more positive, IPT's effectiveness over control group was less pronounced. • In other words, IPT's effectiveness was stable across different levels of school climate whereas, control

Outcome: This generates reductions in depression.		group seemed to have worse endline PHQ scores for less positive school climate levels.
CMOC 3 Context: For boys who do not experience cultural norms or structural violence which impede opportunities to implement new strategies. Mechanism: IPT enables participants to participate in discussions through which they learn strategies to develop hope, reduce conflict, build relationships and improve self-efficacy. Outcome: This generates reductions in depression.	• Participants thought girls had more insight into problems and awareness of their feelings, which allowed them to benefit more from IPT. • Participants thought younger boys and older girls benefit more. For older boys, they were more rebellious, less “compliant” and less willing to share, but older girls were more mature (affordance) and understand the difficulties due to gender roles, this was due to greater societal expectations from girls and less room to speak about their feelings (vulnerability) therefore they share more and are more engaged.	In moderator analyses gender as a moderator was explored using Bayesian models. • No moderating effects of gender observed.
CMOC 4 Context: For older adolescents more cognitively able to consider others’ perspectives, learn negotiation skills, develop solutions to interpersonal problems, understand links between events and mood, manage anger and name and express emotions. Mechanism: IPT enables participants to participate in discussions through which they learn strategies to develop hope, reduce conflict, build relationships and improve self-efficacy. Outcome: This generates reductions in depression.	• Participants perceived that older adolescents were more cognitively able to consider others’ perspectives (affordance), learn negotiation skills, develop solutions to interpersonal problems, understand links between events and mood, manage anger and name and express emotions. This was truer for girls. For boys, as they age, they became less compliant and willing to open up.	In moderator analyses age as a moderator was explored using Bayesian models. • No moderating effects observed.

CMOC 5 Context: For adolescents participating in initial sessions and participating in discussions which validate participants' experiences and instil hope. Mechanism: IPT enables participants to develop hope and motivation to engage in the intervention. Outcome: This generates immediate reductions in depression.	- Participants felt IPT helped mostly when the adolescent decided or had the opportunity to open up and share their problem. - Towards the middle sessions, when the problem area was discussed within the group then we saw gains.	Based on the longitudinal plot, changes in PHQ scores mainly took place before the first midline (around IPT session 3).
---	--	--

Table 3: Trial outcomes and tools

Outcome type	Outcome	Tool/measure	Assessment timepoint
Primary	Depression	PHQ-A	Baseline, midline, endline, follow-up
Secondary	Functional impairment	Locally developed tool	Baseline, midline, endline, follow-up
Secondary	Anxiety	Generalised Anxiety Disorder Assessment (GAD-7)	Baseline, midline, endline, follow-up
Secondary	Post-traumatic stress disorder	PTSD Checklist for DSM-5 (PCL) 8-item	Baseline, midline, endline, follow-up
Secondary	School attendance	No. of days attended in 24 days prior to baseline (excluding closures and Saturdays)	Baseline, midline, endline, follow-up
Secondary	Quality of Life	EuroQol-5 Dimension 5 levels (EQ-5D)	Baseline, midline, endline, follow-up
Secondary	Academic self efficacy	Young Lives Non-Cognitive Scales	Baseline, midline, endline, follow-up
Secondary (long term)	Academic performance	Basic Education Examination School Education Examination grade	Follow-up

Secondary (long term)	School drop out	Self report and school register	Endline and follow-up
Intermediate	Hope	Herth Hope Index	Baseline, midline, endline, follow-up
Intermediate	Emotion regulation	Difficulties in Emotion Regulation Scale (DERS-SF) 18 items	Baseline, midline, endline, follow-up
Intermediate	Self-efficacy	General Self-Efficacy Scale	Baseline, midline, endline, follow-up
Intermediate	Social support	Multidimensional Scale of Perceived Social Support (MSPSS)	Baseline, midline, endline, follow-up
Intermediate	Group cohesiveness	PM+ Group Cohesiveness scale adapted for IPT	(Intervention arm only) midline, endline
Intermediate	Conflict reduction	Network of Relationships Inventory	Baseline, midline, endline, follow-up
Moderator	School climate	Beyond Blues School Climate Scale	
Moderator	Gender norms	Johns Hopkins Global Early Adolescent Study Measure	Baseline, follow up
Moderator	Adversity	Johns Hopkins Global Early Adolescent Study Measure	Baseline
Moderator	Socioeconomic status	Social and Economic Measure for Parents	Baseline
Moderator	Parent mental health	General Health Questionnaire 12 items	Baseline, endline
Moderator	Parenting behaviour	Alabama Parenting Questionnaire 19 items	Baseline, endline
Moderator	Disruptive behaviour – parent report	Disruptive Behavior International Scale – Nepal version (DBIS-N) 10 items	Baseline, endline

Table 4: Qualitative and quantitative data sources for the process evaluation.

Data source	Timing	No. of participants	No. of schools (intervention/control)	Evaluation domain addressed
Quantitative data				
Baseline and follow-up surveys	Baseline, midline, endline, 32 week follow-up, 12 month follow-up, 18 month follow-up, 30 month follow-up	768 (16 per school)	N=48 (Int: N=24/Cont: N=24)	Outcomes Contamination Mechanisms Context
Parent surveys	Baseline, endline	768 (16 per school)	N=48 (Int: N=24/Cont: N=24)	Contamination Context
Facilitator competency assessments	Pre-training, post training, for each cycle and for each facilitator ENACT in pre-group sessions, GroupACT in 4 group sessions	All facilitators	N/A	Implementation
Session observations using IPT adherence checklist	Waves 1-8	All facilitators (~25% of total sessions)	N=24 (I)	Implementation
School attendance records	Baseline, endline	768 (16 per school)	N=48 (Int: N=24/Cont: N=24)	Outcomes, implementation
IPT session attendance	Per each IPT session	384 (16 per school)	N=24 (I)	Implementation
IPT treatment satisfaction survey	Endline	384 (16 per school)	N=24 (I)	Implementation
Group cohesion survey	Endline	384 (16 per school)	N=24 (I)	Implementation, mechanisms
Facilitator session and supervision records	Per each IPT session	All facilitators	N/A	Implementation, context
Counsellor notes from high risk and abuse-related cases	Per each case referred to the counsellors	All counsellors	N/A	Implementation

Qualitative data				
Audio recordings of sessions	(Supervisor records sessions observed) For each cycle: for each facilitator, rating two individual sessions and 4 group sessions	N/A	N=24 (I)	Implementation, mechanisms
Interviews with caregivers of adolescents participating in IPT	Endline	16 (2 per school)	N=6 (Intervention case study schools)	Implementation, context, mechanisms
Interviews with adolescents participating in IPT	Endline	32 (4 per school)	N=6 (Intervention case study schools)	Implementation, context, mechanisms
Interviews with adolescents in the control arm	Endline in case study schools	32 (4 per control case study school)	N=8 control case study schools	
Focus group discussions with facilitators and supervisors	Endline (Wave 4 and 8)	2-3 FGDs in Wave 4; 2-3 in Wave 8, 4-6 facilitators/supervisors per FGD	N/A	Implementation, context, mechanisms
Focus group discussions with counsellors and supervisors managing high-risk cases	Endline (Wave 2 and 6)	All counsellors and supervisors, 4-6 per FGD	N/A	Implementation
Teacher one to one interviews	Endline in case study schools	Principal (N=16 I=8/C=8), Teachers (N=16 I=8/C=8)	N=16 case study schools (Int: N=8/Cont: N=8)	Implementation, context
Interviews with caregivers of adolescents in the control arm	Endline in case study schools	8 (1 per control case study school)	N/A	Implementation

Table 5: Process evaluation hypotheses and associated data and analyses

Hypotheses	Qualitative data sources/analyses	Quantitative data sources/analyses
Hypotheses about intervention mediators		
H1: Adolescents participating in IPT will report increased social support at midline, endline and follow up due to bonding between group members, initiation of new relationships outside groups and/or improvement in the quality of existing relationships. This generates reductions in depression.	Analysis of SSI transcripts with adolescents and parents and FGD transcripts with facilitators from different groups (direct and indirect probing on hypothesis)	Mediation analysis to assess if IPT affects PHQ-A outcomes at endline through social support (MSPSS score at midline)
H2: Adolescents participating in IPT will report reduced interpersonal conflict at midline, endline and follow-up due to improved communication of interpersonal needs and strategies to better manage anger. This generates reductions in depression.	Analysis of SSI and FGD transcripts with adolescents and facilitators from different groups (direct and indirect probing on hypothesis)	Mediation analysis to assess if IPT affects PHQ-A outcomes at endline through interpersonal conflict (Social Adjustment Scale Self Report score at midline)
H3: Adolescents participating in IPT will experience improved emotion regulation at midline, endline and follow up because they learn how to name and express their emotions. This generates reductions in depression.	Analysis of SSI transcripts with adolescents and parents and FGD transcripts with facilitators from different groups (direct and indirect probing on hypothesis)	Mediation analysis of adolescent survey data (DERS-SF and PHQ-A outcomes)
H4: Adolescents participating in IPT will report increased hope at midline as participants perceive the possibility of change in mood and relationships. This generates reductions in depression.	Analysis of SSI transcripts with adolescents	Mediation analysis of adolescent survey data (Herth Hope Index and PHQ-A data)
H5: Adolescents participating in IPT will report improved self efficacy at midline, endline and follow-up because they experience generating solutions to and solving their interpersonal problems. This generates reductions in depression.	Analysis of SSI transcripts with adolescents and FGD transcripts with facilitators from different groups	Mediation analysis of adolescent survey data to assess if self-efficacy mediates PHQ-A.
H6: Adolescents participating in IPT will report improved self-compassion at midline, endline and follow up because in IPT groups they normalise and validate their experience of depression. This generates reductions in depression.	Analysis of SSI transcripts with adolescents and FGD transcripts with facilitators from different groups	We will not have quantitative data to test this hypothesis.

Hypotheses about contextual factors affecting implementation		
H7: Adolescents participating in IPT who are from the poorest backgrounds whose basic needs are not met will attend fewer IPT sessions because they have to undertake paid work and/or have household responsibilities to fulfil.	Analysis of SSI transcripts with adolescent and parents and FGD transcripts with facilitators from different groups (direct and indirect probing on hypothesis)	Analysis of adolescent IPT session attendance data and socio-demographic data from baseline survey
H8: Groups involving adolescents across two grades (8/9) rather than three (7/8/9) will be more cohesive and attendance among participants will be higher because participants will be closer in age and maturity and sessions will be easier to schedule across fewer timetables.	Analysis of FGD transcripts with facilitators and SSIs transcripts with teachers (direct and indirect probing on hypothesis)	Analysis of group cohesion and adolescent IPT session attendance data.
H9: In schools that have experienced a suicide, an attempted suicide or mass conversion disorder (chopnne), IPT will be more acceptable and implemented with more fidelity because staff will be more motivated to support the programme.	Analysis of SSI transcripts with teachers and principals, FGD transcripts with facilitators	Teacher attendance at IPT orientation
Hypotheses about contextual factors influencing intervention effects (CMOCs)		
H10: Adolescents participating in IPT who attend schools with a climate characterised by a lack of understanding and supportive teacher and student relationships (context), will build strong peer to peer social support, a group identity and a sense of belonging with IPT group members (mechanism) that will offset the negative effect of the school climate on adolescents' depression (outcome).	Analysis of SSI transcripts with adolescents sampled to represent different school climates (baseline Beyond Blue score), teachers and principals, and FGDs with facilitators, probing on hypothesis	Comparison of group cohesion data across schools Moderator analysis of adolescent survey data (Beyond Blues School Climate Scale as moderator variable, PHQ-A outcome)
H11: Adolescents participating in IPT from a lower socioeconomic background (poorer household, context) will be more engaged and motivated to address their problems (affordance) and have more problems to solve (vulnerability) compared to their peers from wealthier backgrounds. These adolescents will share their feelings and problems and participate in discussions through which they learn strategies to regulate their emotions, reduce conflict, build social support and improve self-efficacy (mechanism) and therefore	Analysis of SSI transcripts with adolescents from different socioeconomic backgrounds and facilitator FGD transcripts (direct and indirect probing on hypothesis)	Moderator analysis of adolescent survey data (socioeconomic moderator variable and PHQ-A outcome)

experience more benefit from IPT compared to peers from a higher socioeconomic background (outcome).		
H12: For girls, who are subject to gender norms which discourage them from expressing their feelings (context), group IPT will serve as a platform to become aware of and share feelings and learn emotion regulation skills (mechanism) which reduce their depression. On the other hand, boys who are exposed to gender norms which promote aggression and anger (context), through IPT learn how to reduce interpersonal conflict (mechanism) and therefore experience reductions in depression (outcome).	Analysis of SSI transcripts with boys and girls, and facilitator FGD transcripts (direct and indirect probing on hypothesis)	Moderated mediation analysis of adolescent survey data (gender moderator, emotion regulation and conflict mediator variables, PHQ-A outcome)
H13: For adolescents with a parent who has a mental health problem (context), through IPT they will increase their social support, normalise and validate their feelings, generate and implement solutions to their problems and reduce conflict at home (mechanism). This will offset the effect of the parent's mental health on the adolescent's depression (outcome).	Analysis of SSI transcripts with adolescent and parents with and without mental health problems (GHQ-12 score at baseline) and FGD transcripts with facilitators from different groups (direct and indirect probing on hypothesis)	Moderator analysis to assess if parent GHQ-12 score at baseline moderates effect of IPT on PHQ-A score at endline
Hypotheses about potential harms		
H14: For adolescents with a background of trauma (context), who recount their traumatic experiences during group sessions and to whose experiences other group members cannot relate, IPT will exacerbate their distress due to re-experiencing the stressful event (mechanism), which will result in a worsening of depression (outcome).	Analysis of SSI transcripts with adolescents with and without trauma (baseline PCL score) and FGD transcripts with facilitators from different groups (direct and indirect probing on hypothesis)	Analysis of adolescent survey data (PCL, PHQ-A outcomes)
H15: (Linked to H8 but related specifically to stigma due to adolescents' participation in IPT) Adolescents experiencing stigma from school and community members because of their participation in IPT (context) would attend fewer sessions and the stigma would cause them distress and their depression worsens.	Analysis of SSI transcripts with adolescents who have experienced stigma and facilitator FGD transcripts	Cross-sectional analysis of adolescent survey data (endline, questions about experience of stigma due to participation in IPT, PHQ-A)
H16: In groups adolescents are encouraged to generate and select solutions to solve their interpersonal problems. If the group is poorly facilitated these solutions may run counter to	Analysis of SSI transcripts with adolescents and facilitator FGD transcripts probing on the	We will not have quantitative data to test this hypothesis.

our theory of change and could cause adolescents harm for example by inciting violence in the context of a dispute. This could exacerbate adolescents' depressive symptoms.	nature of solutions generated within groups and their consequences Analysis of information from facilitator session and supervision notes on solutions and consequences	
H17: Group members or facilitators may breach an adolescent's confidentiality by disclosing sensitive information about them to individuals outside the group. This would be distressing for the adolescent and could have harmful consequences resulting in the worsening of their depression.	Analysis of SSI transcripts with adolescents and facilitator FGD transcripts probing for evidence of a breach in confidentiality	We will not have quantitative data to test this hypothesis.

Table 6: Power scenarios for moderated mediation

N Clusters	N individuals	Design effect	Missing	Power a	Power b	Overall power	a	b
46	552	1.79	8%	0.81	0.99	0.80	0.65	0.49
46	552	1.79	8%	0.85	0.94	0.80	0.68	0.40
46	552	1.79	8%	0.9	0.89	0.80	0.74	0.36
46	504	1.73	15%	0.81	0.99	0.80	0.66	0.50
46	504	1.73	15%	0.85	0.94	0.80	0.70	0.41
46	504	1.73	15%	0.9	0.89	0.80	0.75	0.34
46	552	1.34	8%	0.81	0.99	0.80	0.56	0.43
46	552	1.34	8%	0.85	0.94	0.80	0.59	0.35
46	552	1.34	8%	0.9	0.89	0.80	0.65	0.31
46	504	1.31	15%	0.81	0.99	0.80	0.58	0.44
46	504	1.31	15%	0.85	0.94	0.80	0.62	0.36
46	504	1.31	15%	0.9	0.89	0.80	0.67	0.33

Figures

Figure 1: Theory of change for group IPT for adolescents with depression in Nepal

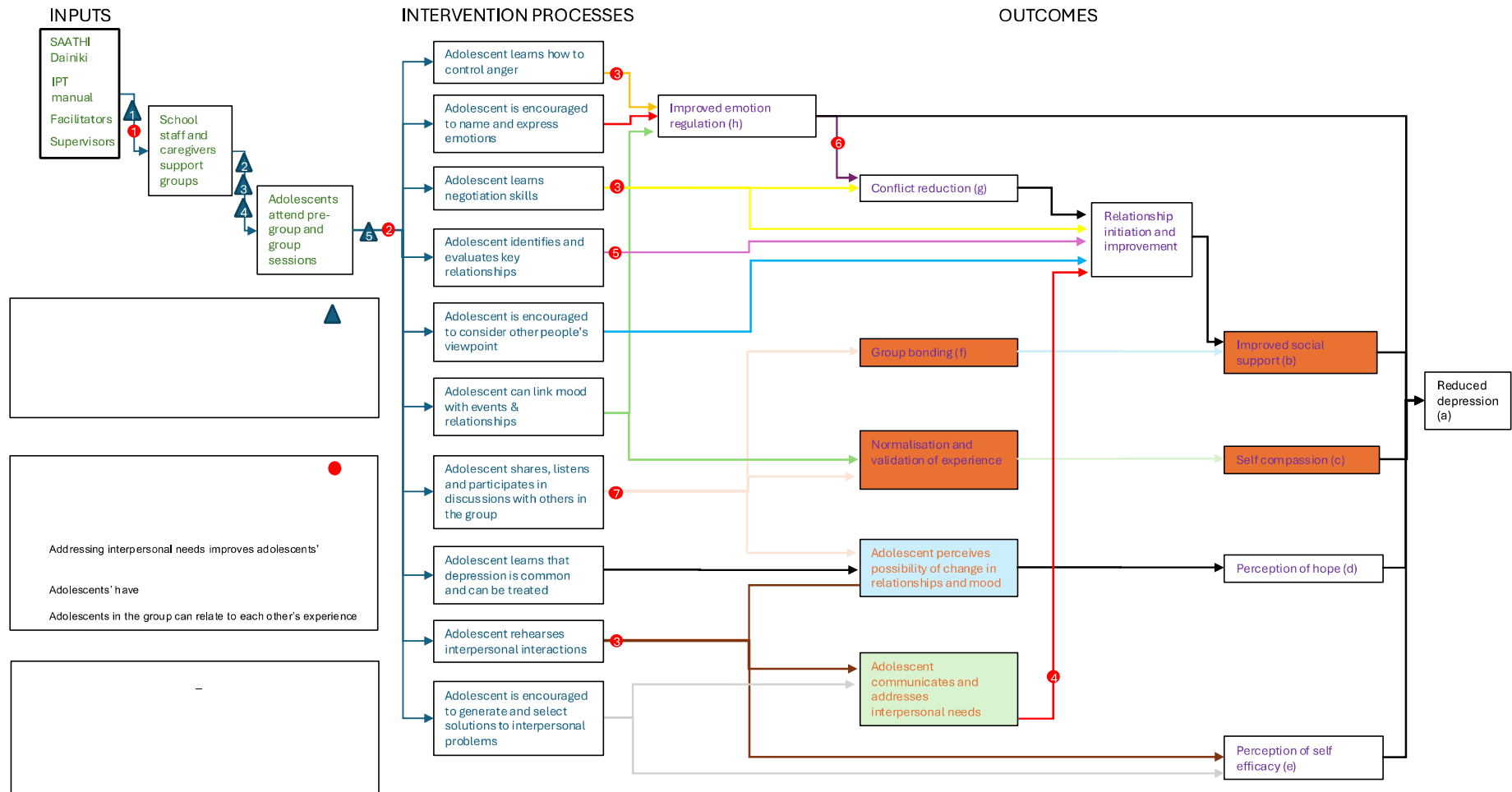


Figure 2: Phase III cRCT timeline

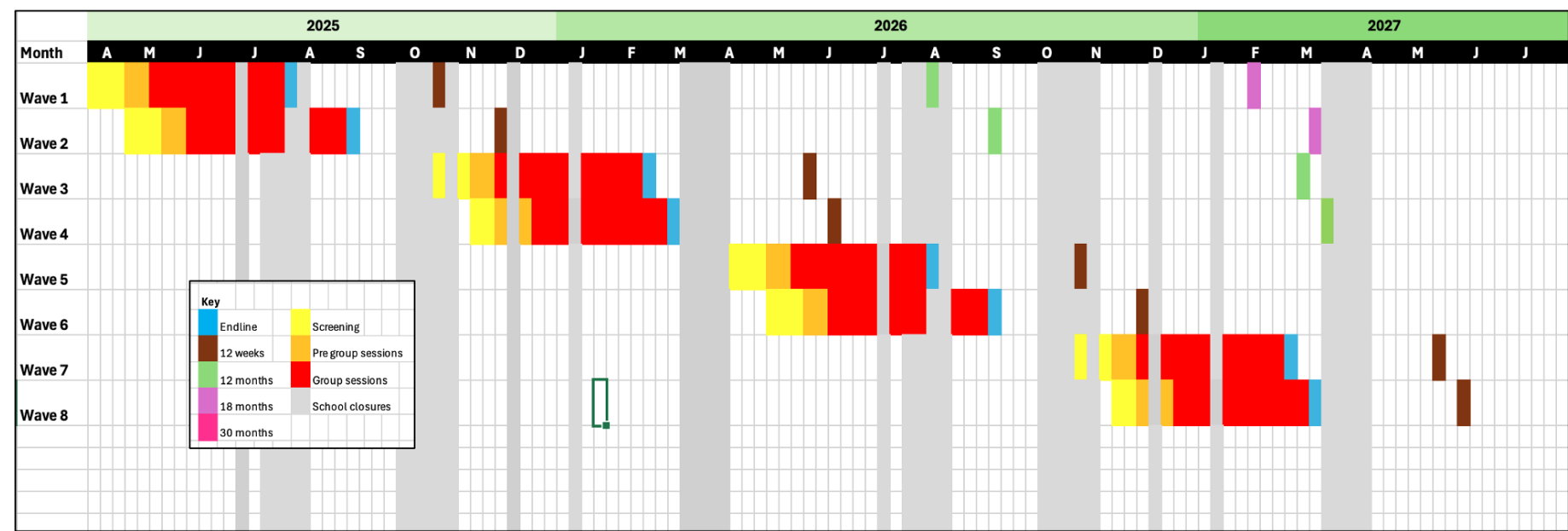


Figure 3: Overview of IPT facilitator training and competency assessment

TRAINING CONTENT	RECRUITMENT CRITERIA Essential • Education 10+2. • Aged 20-25 years • Nepali citizenship • Able to reside in Chitwan during the trial Desirable • Experience working with adolescents • Experience working in an NGO	MODULE 1: FOUNDATIONAL HELPING SKILLS (10 days) • Nonverbal communication and active listening • Verbal communication skills • Confidentiality • Rapport building and self-disclosure • Exploration, interpretation and normalisation of feelings • Empathy, warmth and genuineness • Assessment of harm • Connecting to social functioning and impact on life • Causal and explanatory models • Involvement of family • Goal setting • Promotion of hope • Coping mechanisms • Psychoeducation • Eliciting feedback	MODULE 2: GROUP MANAGEMENT SKILLS (1 day?) • Collaboratively develop and regularly review group ground rules and guidelines • Participation of all group members • Fostering empathy between group members • Guiding collaborative problem—solving • Mitigating barriers to attendance • Confidentiality among group members • Time management and pacing	MODULE 3: INTERPERSONAL THERAPY Didactic workshop (5 days) • Principles and theory • Initial, middle and termination phases • Safety planning • Treatment techniques Supervision of practice cases • Minimum of 2 individual cases and one group
ASSESSMENT TOOL	INTERVIEW	ENACT	GROUP-ACT	IPT KNOWLEDGE TEST IPT CHECKLIST

Figure 4: Power curves for moderation analysis

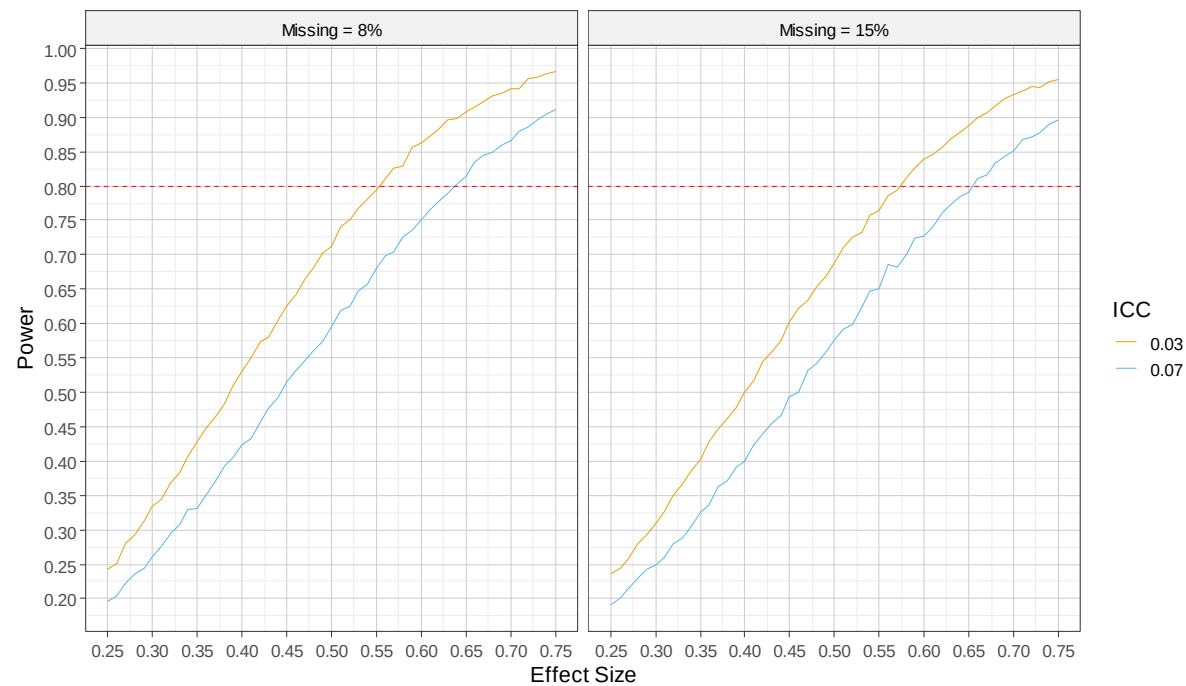


Figure 5: Power curves for the b path

