Statistical Analysis Plan (SAP)

Official Title of the study:

Transdiagnostic, Cognitive and Behavioral Intervention for in School-aged Children with Emotional and Behavioral Disturbances (MindMyMind RCT)

NCT number:

NCT03535805

Date of the document:

8th august 2019

Section 1: Administrative information

Working Title

Transdiagnostic, Cognitive and Behavioral Intervention versus Treatment As Usual for School-Aged Children with Emotional and Behavioral Disturbances: Statistical Analysis Plan for The Mind My Mind Study - A Pragmatic Randomized Effectiveness Trial

Collaborators/Authors

Pia Jeppesen[1,2], Rasmus Trap Wolf[1,3], Sabrina M. Nielsen[4,5], Robin Christensen[4,5], Kerstin von Plessen[1,6], Niels Bilenberg[7,8], Per Hove Thomsen[9,10], Mikael Thastum[11], Simon-Peter Neumer[12,13], Louise Berg Puggaard[1], Mette Maria Agner Pedersen[1], Anne Katrine Pagsberg[1,2], Wendy Silverman[14], Christoph U. Correll[15,16,17,18].

[1]Child and Adolescent Mental Health Centre, Mental Health Services – Capital Region of Denmark, Copenhagen, Denmark

[2]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark

[3]Danish Centre for Health Economics, Department of Public Health, University of Southern Denmark, Denmark

[4]Musculoskeletal Statistics Unit, The Parker Institute, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark

[5]Research Unit of Rheumatology, Department of Clinical Research, University of Southern Denmark, Odense University Hospital, Denmark

[6]Division of Child and Adolescent Psychiatry, Department of Psychiatry, Lausanne University Hospital CHUV, Switzerland.

[7]Department for Child and Adolescent Psychiatry, Mental Health Services in the Region of Southern Denmark.

[8]University of Southern Denmark, Odense, Denmark.

[9]Research Center at Department for Child- and Adolescent Psychiatry, Psychiatry, Aarhus University Hospital, Skejby, Denmark.

[10]Institute of Clinical Medicine, Aarhus University, Aarhus, Denmark.

[11]Centre for the Psychological Treatment of Children and Adolescents, Department of Psychology and Behavioural Sciences, Aarhus BSS, Aarhus University, Denmark.

[12]Centre for Child and Adolescent Mental Health, Oslo, Norway.

[13]The Arctic University of Norway, Centre for Child and Youth Mental Health and Child Welfare, North Norway (RKBU North), Tromsø, Norway.

[14]Anxiety and Mood Disorders Program, Yale Child Study Center, Yale School of Medicine, New Haven, CT, USA.

[15]Donald and Barabara Zucker School of Medicine at Hofstra/Northwell, Department of Psychiatry and Molecular Medicine, Hempstead, New York, USA

[16]The Zucker Hillside Hospital, Department of Psychiatry, Glen Oaks, New York, USA.

[17]Feinstein Institute for Medical Research, Center for Psychiatric Neuroscience, Manhasset, New York, USA

[18]Charité Universitätsmedizin, Department of Child and Adolescent Psychiatry, Berlin, Germany

Trial Registration

Clinical Trials Identifier: NCT03535805

Danish Ethics Committee: Journal-nr.: 17011408, Date of approval: June 20th, 2017

SAP Version

Version 1.0 Date: August 5th, 2019

Protocol Version

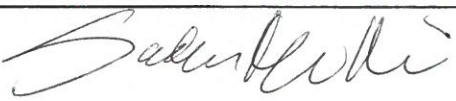
Version 1.2 Date: March 27th, 2019

Roles and Responsibility – non-signatory names and contribution

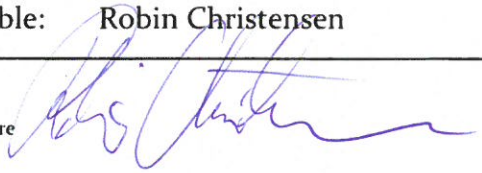
Project Leader:	Pia Jeppesen, PhD, Associate Clinical Professor
Advisory Board:	Anne Katrine Pagsberg, Professor
	Kerstin Plessen, Professor
	Niels Bilenberg, Professor
	Per Hove Thomsen, Professor
	Mikael Thastum, Professor
	Simon-Peter Neumer, Ph.D., Senior Researcher
	Wendy Silverman, Professor
	Christoph U. Correll, Professor
Senior Biostatistician:	Robin Christensen, MSc, PhD (Biostatistics), Professor
Statistical Analysis:	Sabrina M. Nielsen, MSc, PhD Fellow (Biostatistics)

Roles and Responsibility – signatures

Person performing the analyses: Sabrina M. Nielsen

Signature  7/8 -19 Date

Senior biostatistician responsible: Robin Christensen

Signature  Aug. 8, 2019 Date

Chief investigator/clinical lead: Pia Jeppesen

Signature  August 7th 2019 Date

Section 2: Introduction

Background and Rationale

Impairing emotional and behavioral problems are common in children and adolescents and mark a three-fold increased risk of mental disorder in young adulthood.[1-3] Evidence-based psychological interventions are recommended for indicated prevention and first-line treatment[4], but access to treatment is often limited.

A new, transdiagnostic and modular, cognitive and behavioral therapy program '*Mind My Mind*' (MMM) comprising evidence-based interventions for children with emotional and behavioral problems was designed to be delivered by educational psychologists in the Danish municipalities. A feasibility randomized controlled trial (RCT) (NCT03448809), demonstrated that the study design was acceptable among children, parents, and therapists, and it provided data to estimate the sample size needed for a pivotal RCT.

Aim and Objectives

Our aim is to investigate the effectiveness and cost-effectiveness of MMM compared with '*Treatment As Usual*' (TAU) for children and adolescents with impairing anxiety, depressive symptoms and/or behavioral problems. Both beneficial and harmful effects will be evaluated.

The primary objective is to compare the effectiveness of MMM, relative to TAU, on the impact of mental health problems reported by the parent using the impact scale of the Strengths and Difficulties Questionnaire (SDQ) at the end of treatment (week 18). The SDQ-impact scale evaluates the impact of mental health problems on the everyday life of the child, i.e., how much the difficulties upset the child and interfere with the child's home life, friendships, classroom learning, and leisure activities.

The hypothesis was that the parent-reported impact of problems will be significantly lower for children in the MMM group compared with children in the TAU group after the 18-week intervention period (primary hypothesis).

Key secondary objectives are to compare the effectiveness of MMM, relative to TAU, measured at week 18 on level of (*i*) Anxiety (Spence Children's Anxiety Scale [SCAS], parent-reported), (*ii*) Depressive symptoms (Mood and Feelings Questionnaire [MFQ], parent reported), (*iii*) Daily functioning of child (Weiss Functional Impairment Rating Scale [WFIRS], parent reported), (*iv*) School attendance (proportion of school-days within the last 4 weeks, where the child is present as recorded by the parent), (*v*) Top-problem scores (parent reported), Quality of life (KIDSCREEN-27) two subscales on (*vi*) physical well-being and (*vii*) psychological well-being (child reported), Behavioral problems (Eyberg Child Behavior Inventory [ECBI]), (*viii*) ECBI intensity score (parent reported), (ix) ECBI problem

score (parent reported), and (*x*) Emotional and behavioral problems (SDQ total difficulties scale, parent reported).

The hypotheses were that the severity of key secondary outcomes of psychopathology, daily functioning, school attendance and health-related quality of life would show significant group differences in favor of MMM compared with TAU at week 18 (secondary hypotheses).

Exploratory objectives are:
To compare the proportion of children with an SDQ impact score reduction ≥ 1 point from baseline to week 18, and the proportion of children scoring below the SDQ inclusion cut-off at week 18 (see inclusion criteria 2, page 11)

To compare the sustained effectiveness of MMM, relative to TAU from baseline to week 26 (i.e., 8 weeks after cessation of the MMM intervention) on the impact of mental health problems reported by the parent (using SDQ), as well as the level of (*i*) Anxiety (SCAS, parent-reported), (*ii*) Depressive symptoms (MFQ, parent reported), (*iii*) Daily functioning of child (WFIRS, parent reported), (*iv*) School attendance (proportion of school-days within the last four weeks where the child is present as recorded by the parent), (*v*) Top-problem scores (parent reported), Quality of life (KIDSCREEN-27) two subscales on (*vi*) physical well-being and (*vii*) psychological well-being, child reported), Behavioral problems (ECBI), (*viii*) ECBI intensity score (parent reported) (ix) ECBI problem score (parent reported), and (*x*) Emotional and behavioral problems (SDQ total difficulties scale, parent-reported). Furthermore, to compare proportion of children with SDQ impact score reduction ≥ 1 point from baseline to week 26, and proportion of children scoring below the SDQ inclusion cut-off at week 26.

To compare the effectiveness of MMM, relative to TAU, on the different measures of psychopathology, functioning, and quality of life that are not already part of the primary, secondary or exploratory objectives. These exploratory outcomes include responses from children at week 18 (i.e., at the end of MMM intervention), and at week 26 (i.e., 8 weeks after cessation of the MMM intervention). Some of the questionnaires could not be administered to children below the age of 8 years (SCAS and MFQ) or below the age of 11 years (SDQ), as they are not standardized for children this age. Hence, we compare the impact of mental health problems (SDQ-impact score) and the emotional and behavioral difficulties (SDQ-total difficulties score) as reported by children of age 11-16 years, and anxiety (SCAS) and depressive symptoms (MFQ) as reported by children of age 8-16 years in the randomized design. The Top-problem score and the Health-Related Quality of Life (KIDSCREEN-27) had no age-restrictions and was administered to all children.

Furthermore, we will explore the level of satisfaction measured by the parent- and child-reported Experience of Service Questionnaire (ESQ), the teacher-reported impact of mental health problems (SDQ-impact scale, range 0-6), the teacher-reported emotional and behavioral difficulties (SDQ-total difficulties scale), the parent-reported KIDSCREEN-27

scale scores (Physical well-being, Psychological well-being, Autonomy & Parents, Peers & Social Support and School Environment); and finally, the parent's response on the Parental Stress Scale (PSS) in MMM, compared to TAU at week 18, and at week 26.

Measures of potential harms

There are no well validated instruments available to monitor safety and potential hams of psychological interventions in children and adolescents. Instead of introducing specific questionnaires, we defined two binary composite scores. One for suicidality and negative cognitions (the primary measure of harms) based on selected items from MFQ for children aged 8-16 years, and another one for poor quality of family relationships, free time and friendships (the secondary measure of harms) based on selected items from KIDSCREEN-27 for all the children. The self-reported MFQ is prioritized because suicidality and negative cognitions are internalized symptoms that parents are often not or only insufficiently aware of.

If at least one out of the selected items (four items of suicidality, and four items of guilt, hopelessness and negative self-evaluation) is scored "always" (the most severe response) at follow-up, but not at entry, then a potential harm is counted as present (yes/no) in the domain of suicidality and/or negative cognitions. If at least one out of the selected items (five items of family relations and free time, and four items of friendship), is scored "always" (the most severe response) at follow-up, but not at entry, then a potential harm is counted as present (yes/no) in the domain of family relations/free time and/or friends.

We compare the effectiveness of MMM, relative to TAU, on the proportion of children with suicidality and/or negative cognitions at week 18, and at week 26. We compare the effectiveness of MMM, relative to TAU, on the proportion of children with poor quality of family relationships, free time and friendships at week 18, and at week 26.

The economic evaluation will comprise both a cost-effectiveness analysis (CEA) and a cost-utility analysis (CUA) comparing the effects of MMM versus TAU with the cost of MMM versus TAU, including health services, social services and parental cost from baseline to the end of the post-intervention follow-up period (week 26). The detailed listing of the cost to the different stakeholders (municipality, region, private households) will not be available for the present report and will be delivered at a later time point.

Section 3: Study Methods

Trial Design and Interventions

A total of 396 children were included and randomized. The trial was designed as a pragmatic, multi-site, randomized, parallel-group, controlled trial for children aged 6-16 years with emotional and behavioral difficulties. Treatment allocation was performed in a 1:1 ratio. Patients were randomized to either Mind My Mind (MMM) or Treatment as Usual (TAU) -control.

MMM: is a newly developed, transdiagnostic and modular cognitive and behavioral therapy (CBT) program comprising evidence-based interventions for children and adolescents with anxiety, depression or disruptive behavior. The evidence-based CBT methods were organized into modules, and the intervention was tailored to the individual child by the dosing and the sequencing of the modules. Parents were engaged in child's therapy and supported the child in doing the homework. Parent management training was offered for behavioral disturbances. The educational psychologists delivered the therapy after a one-week training in the manualized intervention, followed by weekly supervision. The treatment fidelity was monitored by video observation of therapy sessions. The full MMM training program consisted of 9-13 sessions plus one booster session, which each were completed within 17 weeks.

TAU: The parents in the TAU group were offered two sessions (at week 2 and week 17) to support them to seek help for the child in the municipality. This care coordination visit was provided by psychologists (or other local professionals) who hold records of the currently available treatment options in the municipality. TAU varies considerably from no intervention to counselling, talk therapy, pedagogical advice, network meetings, and/or individual support in the school setting. Some children are offered CBT interventions, but access to manualized treatment is almost inexistent.

Randomization and Allocation Concealment

Allocation concealment was done through centralized randomization, which was provided by the Data & Documentation department named DEFACTUM, Social and Health Services and Labour Market, Central Denmark Region. A computer-generated allocation sequence with variable and unknown block size was employed. Randomization was performed centrally, with stratification, and the sequence was generated by an independent statistician using a random number generator. The principal investigator, the project manager and all researchers and therapists were blinded to the sequence used. To optimize comparability between the two treatment groups, randomization was stratified across three factors:

1) Municipality (2 levels: Vordingborg & Næstved OR Holstebro & Helsingør)
2) Child's age (2 levels: 6–10 years OR 11–16 years)
3) The principal problem classified by the psychologist during the visitation (3 levels: the classification of the top-problem as 1=anxiety OR 2=depressive symptoms OR 3=behavioral problems; the classification of the principal problem into one of the three categories was mandatory, otherwise the child was not eligible).

For each stratum (with a total of 12 strata), computer-generated random allocation sequence lists were prepared by DEFACTUM. This central treatment allocation system was set up to ensure allocation concealment, so that the person enrolling participants did not know in advance which treatment the next child would get. Randomization was done after eligibility criteria and baseline assessments were electronically confirmed in the web-based case report form (see section 5). Following this approach, we masked knowledge of forthcoming assignments.

Power and Sample Size Considerations

The minimal relevant difference (MIREDIF) in impact of mental health problems between the two randomly allocated intervention groups (MMM versus TAU) measured by the SDQ-impact-scale at week 18 (and subsequently at week 26) was set at 1.0 units on the parent-reported SDQ-impact-scale (range 0-10), potentially corresponding to a change from severe to moderate, or from moderate to little or no impact in one important aspect of the child's life (distress, home-life, friendships, classroom learning or leisure activities).

We pre-specified the use of a two-sided type-1 error of 0.05 ($P<0.05$), with a risk of type-2 error of 0.10 (corresponding to a statistical power of 90%). To estimate the sample size needed, we used an expected SD of 2.7 (based on parameters from the feasibility trial) for the SDQ-impact-scale scores. Hence the required sample size necessary to detect or reject a difference of at least 1.0 point was estimated as 2 × 154 children =308 in total:

R software (version 3.4.3)code:

```
power.t.test(delta=1.0, sd=2.7, sig.level = 0.05, power=0.90)
```

Allowing for an attrition rate for up to 25%, it was decided to include and randomize 308/0.75 = 412 children in total. We randomized a total number of 396 children within the recruitment period and decided not to extend the recruitment period since attrition rates were lower than expected (approximately 15%, whereas power calculations allowed 25% attrition), meaning that the necessary number of children with follow-up data had been included.

Framework

The Mind-My-Mind trial protocol states that a secondary objective is to determine whether a potentially superior effect of MMM (over TAU) can be sustained even with a further 8-week extension added (assessed after 26 weeks). Therefore, the exploratory outcomes are tested for superiority to confirm a potential claim of sustained effectiveness related to MMM therapy for the primary outcome (SDQ-impact).

Statistical Interim Analyses and Stopping Guidance

Interim data analysis for effectiveness was not conducted due to the relatively small sample size and short duration of this clinical trial. Safety data were not assessed systematically in the municipalities; however, the steering committee held the responsibility to act as a 'Data and Safety Monitoring Board' (DSMB). The steering committee met regularly throughout the whole data collection period and monitored the safety based on reports of local experiences (without disclosing the group identity of cases). Thus, we applied no pre-specified guidelines for determining stopping rules due to a safety concern. The clinical

opinion from the steering committee deliberations did not reveal any *ad hoc* safety concerns.

Timing of the Final Analysis

The final analysis for the MMM vs. TAU comparison is planned to take place in two separate stages: (1) First, we will perform the statistical analyses based on all the repeated measures from baseline to 18 weeks; this will constitute the first main report of the trial to the advisory board and will be prepared for the MMM/TAU comparison when every randomized child has reached the 18-week assessment, and when the data for the primary endpoint should have been received and cleaned (anticipated to be August 6'st, 2019); (2) Longer-term (sustained effect) endpoints at 26 weeks for the MMM/TAU comparison will be analyzed when every child has reached the 26 weeks assessment and data for the primary endpoint have been received and cleaned (anticipated to be September 15'th, 2019). Data analysis for the cost-effectiveness outcomes will be performed at a later stage (estimated to take place in 2020) when the relevant information necessary for those analyses will be available from the Danish register: the Danish National Prescription Registry, the National Patient Register, and the Danish National Health Service Register.

Timing of Outcome Assessments

The schedule of study procedures is given in the table below. The expected visit dates and visit windows was defined as:

Variable View	x0	w0	w2	w4	w6	w8	w10	w12	w14	w18	w26
Sex	x										
Age	x										
Region	x										
Principal domain of problems	x										
Developmental delays	x										
School absenteeism	x										
DAWBA diagnosis	x										
Physical illness	x										
Living arrangement	x										
Parent registered as informant	x										
Mother's highest education	x										
Immigration history of parents	x										
Number of children in the household	x										
Mother's mental health problems	x										
Both parents had mental health problems	x										
Adherence 9-13 sessions MMM										x	
Strengths and Difficulties Questionnaire (SDQ impact score, parent reported)		x	x	x	x	x	x	x	x	x	x
*Spence Children's Anxiety Scale (SCAS, parent-reported)		x								x	x
*Mood and Feelings Questionnaire (MFQ. Parent-reported)		x								x	x
*Weiss Functional impairment rating Scale (WFIRS-P, parent reported)		x								x	x
*School attendance (percent school-days in last 4 weeks)		x								x	x
*Top-problem scores, parent-reported		x	x	x	x	x	x	x	x	x	x
*KIDSCREEN Physical Well-Being (self-reported, t-scores)		x								x	x
*KIDSCREEN Psychological Well-Being (self-reported, t-scores)		x								x	x
*Eyberg Child Behaviour Inventory (ECBI, intensity of problems)		x								x	x
*Eyberg Child Behaviour Inventory (ECBI, no of problems)		x								x	x
*Emotional and behavioral difficulties (SDQ total difficulties score, parent-reported)		x								x	x
Self-reported impact of mental health problems (SDQ impact score, range 0-10)		x	x	x	x	x	x	x	x	x	x
Self-reported emotional and behavioral difficulties (SDQ total difficulties score)		x								x	x
Self-reported anxiety reported (SCAS)		x								x	x
Self-reported depressive symptoms (MFQ)		x								x	x
Self-reported Top-problem scores		x	x	x	x	x	x	x	x	x	x
KIDSCREEN Autonomy & Parents (self-reported, t-scores)		x								x	x
KIDSCREEN Peers & Social Support (self-reported, t-scores)		x								x	x
KIDSCREEN School Environment (self-reported, t-scores)		x								x	x
Parent-reported Experience of Service Questionnaire (ESQ)										x	x
Self-reported Experience of Service Questionnaire (ESQ)										x	x
Teacher-reported impact of mental health problems (SDQ impact score, range 0-6)		x								x	x
Teacher-reported emotional and behavioral difficulties (SDQ total difficulties score)		x								x	x
KIDSCREEN Physical Well-Being (parent-reported, t-scores)		x								x	x
KIDSCREEN Psychological Well-Being (parent-reported, t-scores)		x								x	x
KIDSCREEN Autonomy & Parents (parent-reported, t-scores)		x								x	x
KIDSCREEN Peers & Social Support (parent-reported, t-scores)		x								x	x
KIDSCREEN School Environment (parent-reported, t-scores)		x								x	x
Parental Stress Scale (PSS)		x								x	x
Suicidality and negative cognitions binary composite score										x	x
Autonomy and social relationships binary composite score										x	x
Referral or contact to regional mental health services (parent-reported)										x	x

x: indicates that it was assessed/collected.

*: Confirmatory secondary outcomes presented in hierarchical order

Section 4: Statistical Principles

Confidence Intervals and *P* values

Level of statistical significance and use of confidence intervals: All applicable statistical tests will be 2-sided and will be performed using a 5% statistical significance level. All confidence intervals presented will be 95% and two-sided. We will not apply explicit adjustments for multiplicity, rather we will analyze the confirmatory secondary outcomes in a prioritized order (see "*gatekeeping procedure*" below). The analyses of the key secondary outcomes will be performed in sequence until one of the analyses fails to show the statistically significant difference, or until all analyses have been completed at a statistical significance level of 0.05. The key secondary statistical tests will be reported with *P* values for hypothesis tests and claims of statistical significance.

Adherence and Protocol Deviations

Adherence to the MMM intervention will be assessed based on the percent of participants who have completed 9-13 sessions of MMM ($n_{completed}$) out of all allocated to MMM ($n_{allocated}$). This outcome is defined as: % Adherence = ($n_{completed}/n_{allocated}$) *100%.

Other protocol deviations include active or passive refusal to continue to fill out the questionnaires (all participants were asked to continue to reply to questionnaires regardless of any type of drop out from treatment and/or referral to other mental health services).

Analysis Populations

The primary analyses will be based on the Intention to Treat (ITT) population, i.e., based on the Full Analysis Set. The ITT principle asserts the effect of a treatment policy (that is, the planned treatment regimen), rather than the actual treatment given (i.e., it is independent of treatment adherence). Accordingly, participants allocated to a treatment group (MMM or TAU, respectively) should be followed up, assessed and analyzed as members of that group, irrespective of their adherence to the planned course of treatment (i.e., independent of withdrawals and cross-over phenomena).

Section 5: Trial Population

Recruitment, Screening data, and Eligibility

Reporting of screening data will be used to describe the representativeness of the trial sample (Figure 1). Information to be included is depicted in the CONSORT flow diagram and the draft baseline characteristics table 1 below. The number of ineligible patients randomized, if any, will be reported, with reasons for ineligibility.

Help-seeking parents were invited to register their child in the local Pedagogical Psychological Services (in Danish '*Pædagogisk Psykologisk Rådgivning*', PPR) in their municipality.

A two-stage standardized screening for eligibility was implemented in the PPR:
Stage 1: The child and parent answered the strengths and difficulties questionnaire (SDQ), and if the SDQ-parent-scores were above the lower threshold for eligibility, the IT system automatically proceeded to administer the Spence Children's Anxiety Scale (SCAS), the Mood and Feelings Questionnaire (MFQ), and questions about child' and parents' characteristics (family constitution, health, social and school functioning). At this stage children were excluded, if problems were too mild for intervention. The lower threshold was based on parent-reported answers to the SDQ and an algorithm that combined scores of emotional and behavioral problems with functional impairments.
Stage 2: All children above the inclusion threshold were assessed with a clinical psychopathological interview by a trained psychologist in the PPR. The goal was to identify the mental health problems, formulate the Top-problem in the participants' own words, categorize the principal problem as 1) anxiety, 2) depressive symptoms or 3) behavioral problems - or detect a more severe mental health disorder in need of other treatment (see first exclusion criterion).

Finally, the PPR psychologist decided if the child was eligible according to the study inclusion criteria:

Inclusion criteria:

1) Aged 6–16 years and in 0-9th grade (excluding the second semester of the 9th grade).
2) SDQ scores reported by the parent are above the lower cutoff: a total difficulties score of ≥14 and/or emotional problems ≥5, and/or conduct score ≥3; combined with a functional impairment score of ≥1. Scores above this cutoff place the child's difficulties within the top 10 percent of mental health problems in the general age-matched population in Denmark.
3) The child and parents determine one top problem that has to fall within the domains of anxiety, depressive symptoms or behavioral problems, according to the classification by the PPR psychologist.
4) The child and at least one of the two parents understand and speak Danish sufficiently to participate in the treatment.

5) Written informed consent from the holders of the parental rights and responsibilities (usually both parents).

Exclusion criteria:

1) Indications that the child may have a severe mental disorder, like autism-spectrum disorder, attention-deficit/hyperactivity disorder (ADHD), schizophrenia-spectrum psychosis, eating disorder, severe obsessive-compulsive disorder, repeated self-harm, abuse or dependence of alcohol or psychoactive drugs, or other mental disorder requiring referral to a more intensive assessment or treatment within child and adolescent mental health services (after systematic assessment and according to usual recommendations and guidelines).

2) Indications of intellectual functional impairment, severe learning difficulties, or other special needs that would interfere negatively with the MMM training. The judgment is made as a best estimate by the PPR psychologist on the basis of the available information. A formal intelligence test is not required.

3) A prior diagnosis of any developmental or mental disorder after assessment by the regional child and adolescent psychiatrist, regardless of present status or treatment. A prior examination that did not result in a diagnosis of any specific mental health disorder will not exclude the child. The PPR psychologist must consult the study PI who decides whether there is sufficient information to exclude the child because of a significant prior psychiatric history.

4) Prior participation in the MMM pilot or current study.

5) The child and/or parents are unable to participate in weekly sessions throughout the next 13–18 weeks.

Enrolment procedures:

If the child was found eligible, the PPR psychologist asked the legal guardians (usually both parents) to give informed consent for their child's participation in the trial.

The PPR psychologist checked "yes" or "no" to each of the inclusion and exclusion criteria and uploaded the signed consent form to the database.

An automatic e-mail notified the research leader who reviewed the consent form and the answers to the inclusion and exclusion criteria before she approved the child for inclusion in the MMM study.

Once the child was enrolled, the database automatically sent an e-mail and SMS to the child and the parents with a link to the web-based baseline questionnaires.

The IT system automatically approved the child for randomization once the baseline data were collected. All standardized questionnaires had to be completed by each of the respondents.

Withdrawal/follow-up

Level of withdrawal, from intervention and/or from follow-up: *see draft Figure 1 below.* Timing of withdrawal/lost to follow-up data: *see draft Figure 1 below.* Reasons and details of how withdrawal/lost to follow-up data will be presented: see draft Figure 1 below.

Baseline Patient Characteristics

List of baseline characteristics to be summarized: *see draft Table 1 below.* Details on baseline characteristics will be summarized descriptively.

Section 6: Analysis

Outcome Definitions

The data assessment table (above), displaying the outcome measures collected at various time points, indicates how many data points will be part of the analysis. Also, Table 1 and Table 2 below define each outcome explicitly, clearly identifying primary and secondary variables. This display includes a clear specification of the hierarchical order (gatekeeping). Tables 1, 2, and 3 identify the specific measurement variable and its units (e.g., SDQ-impact scores) and provide descriptions and details of any data manipulations or derivations to be performed in the footnotes. If the calculation of a score is more complex, using a validated algorithm, then the algorithm will be provided in a footnote. Scoring, including handling of missing data items, will follow the procedure proposed by the instrument developers, which will be described and justified in the footnote. Sufficient detail will be provided in order for the reader to understand how the scores and results are to be calculated for each outcome.

Analysis Methods

Our primary analyses will be based on the ITT population, including all randomized participants with available data at baseline. Missing data will be handled indirectly and statistically modeled using repeated-measures linear mixed models (see below). These models will be valid if data are 'Missing at Random' (MAR): "*Any systematic difference between the missing values and the observed values can be explained by differences in observed data*".[5] Contrasts between groups will be estimated based on repeated-measures analysis of covariance applied in mixed linear models (i.e., at 18 weeks from baseline).

The primary statistical model will consist of fixed effects and random effects. Fixed effects define the expected values of the observations, and random effects define the variance and covariances of the observations. In this study, participants were randomly assigned to two treatment groups (MMM vs TAU), and observations were made at nine time points for the primary outcome measure (i.e., at baseline and 2, 4, 6, 8, 10, 12, 14, and 18 weeks from baseline [see data collection table above]) for each participant. Basically, there are two fixed-effect factors: group and time. Random effects result from variation between

and within participants. We anticipate that measures on the same patient at different times are correlated, with measures taken closely together in time being more highly correlated than measures taken more apart in time; observations on different participants will be assumed as being independent.

The objectives of a repeated measures designs are to make inferences about the expected values of the observations, that is, about the means of the populations from which participants are sampled. This objective is achieved by taking into account treatment and time effects in the model. Data will be analyzed using SAS and R, with the particular outcome variable at baseline level as a covariate, using a multilevel repeated measures random effects model with participants as the random effect factor based on a restricted maximum likelihood (REML) model.

For continuous outcomes (e.g., SDQ-impact score) the change from baseline will be the response (dependent) variable, and the baseline value (one for each participant), treatment group (two levels), and time (nine levels) will be included as covariates, as well as the interaction between treatment group and time; Patient ID will be handled as a random effect. As the study design was based on a stratified randomization technique, we will also adjust for the three stratification variables. This statistical model holds all between-group comparisons at all assessment points up to 18 weeks from baseline (including baseline) and allows for evaluation of the average effect, as well as the trajectory over time from baseline to 18-week follow-up.

Categorical outcomes for dichotomous endpoints (including responder status and harms) will be analyzed with logistic regression based on a Generalized Linear Mixed Model with the same fixed effects and covariates as the respective analysis of covariance. Since Odds Ratios (*OR*s) for outcomes of common incidence either over- or under estimate the corresponding risk estimate, we will convert all the calculated *OR* values and 95% confidence intervals into approximate Risk Ratios (*RR*) in the text. This approach will also allow us to calculate numbers-needed-to-treat (NNTs) and numbers-needed-to-harm (NNHs) for efficacy and safety outcomes by dividing 1 by the risk difference that can guide clinicians in their decision making.

Missing Data and Sensitivity Analyses

Robustness is a concept that refers to the sensitivity of the overall conclusions to various limitations of the data, assumptions, and analytic approaches to data analysis. Robustness implies that the treatment effect and primary conclusions of the trial are not substantially affected when analyses are carried out based on alternative assumptions or analytic approaches.

Loss to follow-up and missing data for various reasons is difficult to avoid in randomized trials and in particular in pragmatic trials like the Mind-My-Mind trial. We will apply the analysis framework suggested by White et al (2011) in which missing data related to the ITT approach depend on making plausible assumptions about the missingness of the data and including all participants in subsequent sensitivity analyses[6]:

1. Attempt to follow up all randomized participants, even if they withdraw from allocated treatment (i.e., contact all individuals unless they explicitly stated that they had withdrawn their consent)

2. Perform a main analysis of all observed data that are valid under a plausible assumption about the missingness of the data (i.e., Model-based: data as observed; using linear mixed models, assuming that data are 'Missing at Random' [MAR])

3. Perform sensitivity analyses to explore the effect of departures from the assumption made in the main analysis (i.e., a non-responder-imputation: using the value at baseline to replace missing data will correspond to a non-responder imputation; these models will potentially be valid even if data are 'Missing Not At Random' [MNAR])

4. Account for all randomized participants, at least in the sensitivity analyses (covered by #2 and #3 above, plus the corresponding analyses based on the per protocol population).

The interpretation of the corresponding statistical measures of uncertainty of the treatment effect and treatment comparisons will involve consideration of the potential contribution of bias to the *p*-value, 95% confidence interval, and of the inference in general.

#1+2: Our primary analysis population will be all participants with available data at baseline, statistically modelled using repeated-measures linear mixed models (see above). These models will be valid if data are 'MAR'.

#3+4 Sensitivity: We will analyze all variables, with missing data being handled by multiple imputation techniques of the baseline level.

When the full analysis set, the per protocol set (defined as participants with available data at baseline, week 18, and/or week 26), and the analyses on the two different enrolment periods lead to essentially the same conclusions, confidence in the trial results is increased.

Additional Analyses

For erroneous reasons this trial included participants without appropriate confirmation in www.clinicaltrials.gov; 244 of 396 (61.6%) participants were unintentionally included before final trial registration was approved (and publicly available) in ClinicalTrials.gov. However, no data were analyzed before the entire study was completed and the database was locked. As a consequence, we will perform further exploration of the sensitivity of conclusions to the choice of the set of participants analyzed. In the Mind-My-Mind Trial, we will perform sensitivity analyses on two different enrolment (time) periods, corresponding to before (244 participants) and after registration (152 participants); the first mentioned period corresponds to approximately the first seven months of active inclusion.

Harms

Participants may experience some temporary discomfort as they learn new behaviors. Furthermore, some discomfort may be associated with answering questionnaires on psychopathology. However, due to the lack of evidence on potential harms related to psychological interventions in children and adolescents, the suicidality, negative cognitions, and poor quality of family relationships, free time and friendships will be monitored (see Table 4 below).

The trial data will be linked with data from the national registries on outpatient services, inpatient services, and emergency room visits to measure group-differences in service use related to acute crisis, self-harm and suicide attempts during the study period. The advantage of using register-based data to capture these rare events is the complete follow-

up of all included and randomized children, independent of any attrition from the study and from answering questionnaires. However, the annual update of the registries takes a full calendar year, so data will not be available before autumn 2020, and are therefore not part of the present study report.

Statistical Software

Analyses will be performed using SAS (proc mixed or proc glmm) and R version 3.5.0 (or newer R Project for Statistical Computing) with the packages lme4, nlme and emmeans.

Section 7: Manuscript Outline

Appendix 1: All protocol versions (merged); pdf format
Appendix 2: Statistical Analysis Plan; pdf format
Figure 1: Trial Flow Diagram

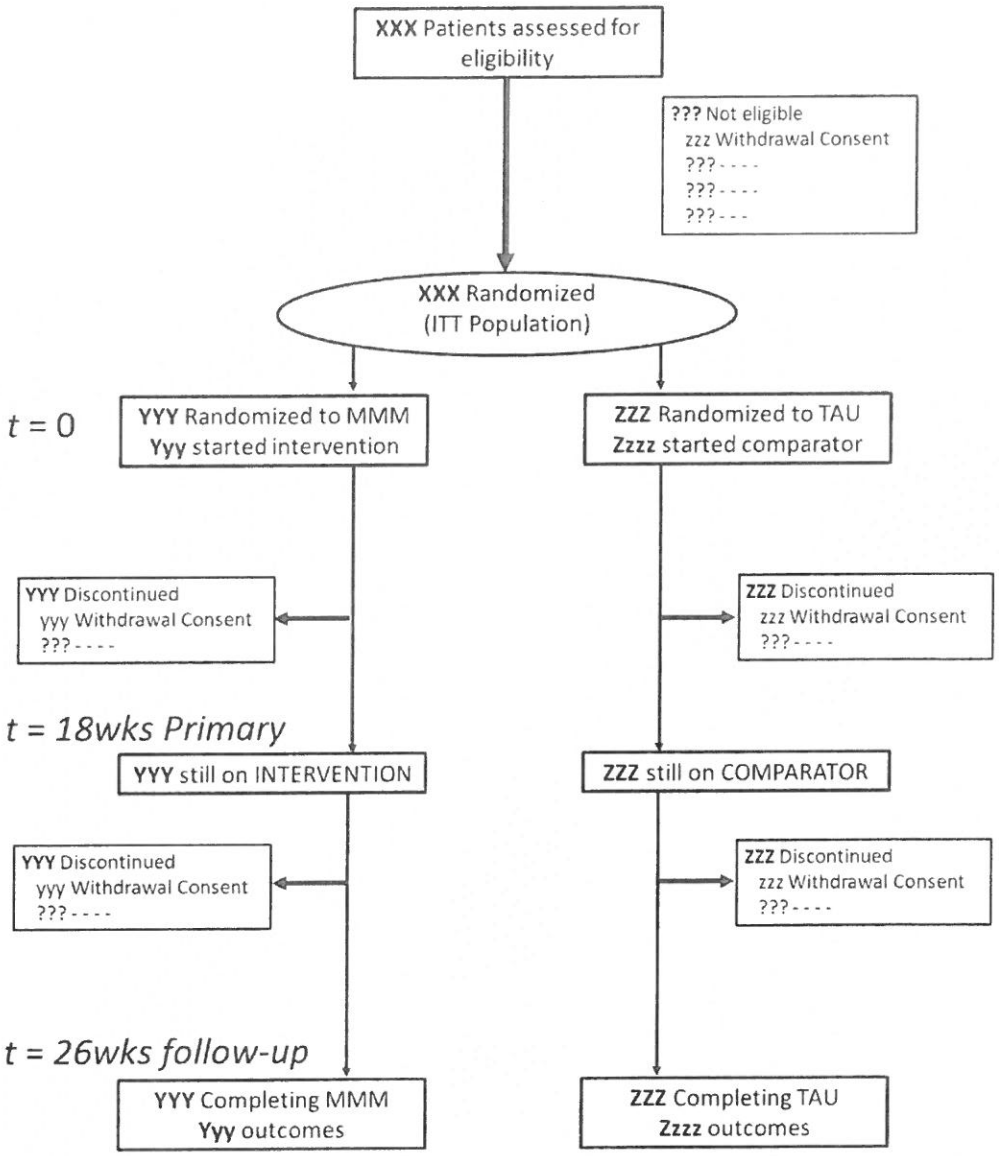


Table 1. Baseline Characteristics (The ITT Population)

Characteristic[1]	MMM group (n = ??)	TAU group (n = ??)	Total ($n = n_{MMM} + n_{TAU}$)
Demographic characteristics			
Girls, no. (%)			
Age, mean (SD), y			
Age-group: 6-10 y, no. (%) 11-16 y, no. (%)			
Region: Holstebro-Helsingør, no. (%) Vordingborg-Næstved, no. (%)			
Principal domain of problems: Anxiety, no. (%) Depressive symptoms, no. (%) Behavioral problems, no. (%)			
Developmental delays: Language, no. (%) Any other[2], no (%)			
School absenteeism: Above 4 weeks last y, no. (%)			
Number of DSM-IV/V Mental disorder based on DAWBA: Anxiety disorder, no. (%) Depressive disorder, no. (%) Behavioral disorder, no. (%) Neurodevelopmental disorder, no. (%) Any disorder, no. (%) Comorbidity, ≥ 2 disorders, no. (%)			
Physical illness (asthma, diabetes, eczema, epilepsy, other), no. (%)			
Living arrangement: Both parents, no. (%) Single parent, no. (%) Other/reconstituted family, no. (%)			
Parent registered as informant: Mother, no. (%) Father, no. (%)			
Mother's highest education: Elementary school, 9-10 y, no. (%) High school/ skilled, 11-14 y, no. (%) Bachelor and above, 15-17 y, no. (%) Higher education, ≥17, no (%)			
Immigration history of parents: Two parents born in Denmark, no. (%) One foreign born, no. (%) Two foreign born, no. (%)			
Number of children in the household: only the index child 2, no. (%) ≥3, no. (%)			
Mother's mental health problems: Anxiety, no. (%) Depression, no. (%), Other, no. (%)			
Both parents had mental health problems, no. (%)			
Outcome measures			
Primary outcome measure:			
SDQ Impact score (0-10)			
Key secondary outcome measures:			
Anxiety (SCAS, parent-reported; 0-114)			
Depressive symptoms (MFQ, parent reported; 0-68)			
Level of daily functioning of child (WFIRS, parent reported; 0-150)			
School attendance (percent school-days in the last 4 weeks; 0-100%)			
Top-problem score (parent reported; 1-10)			
KIDSCREEN-27, converted to T-values (SD) Physical well-being Psychological well-being			
Behavioral problems (ECBI, parent reported) Intensity score (7-252) Problem score (0-36)			

Emotional & behavioral problems (SDQ Total difficulties; 0-40)			

Abbreviations:

SDQ, The Strengths and Difficulties Questionnaire; SCAS, The Spence Children' s Anxiety Scale; MFQ, The Mood and Feelings Questionnaire, WFIRS-P, The Weiss Functional Impairment Rating Scale-Parent Reported; ECBI, The Eyberg Child Behaviuor Inventory; Kidscreen-27, a Health-Related Quality of Life (HRQOL) with five dimensions, of which we use the scales: Physical Well-Being (5 items), Psychological Well-Being (7 items). DAWBA: Development and Well Being Assessment, used for DSM-IV/V diagnoses.

[i] Data are presented as mean (SD) unless otherwise stated.

[ii] Any other developmental delays include motor, social communication and learning difficulties.

Figure 2: Trajectory Least Squares Means scores over time for the child's impact of mental health problems reported by the parent (SDQ) from baseline to week 18, plus the extended follow-up (26 weeks).

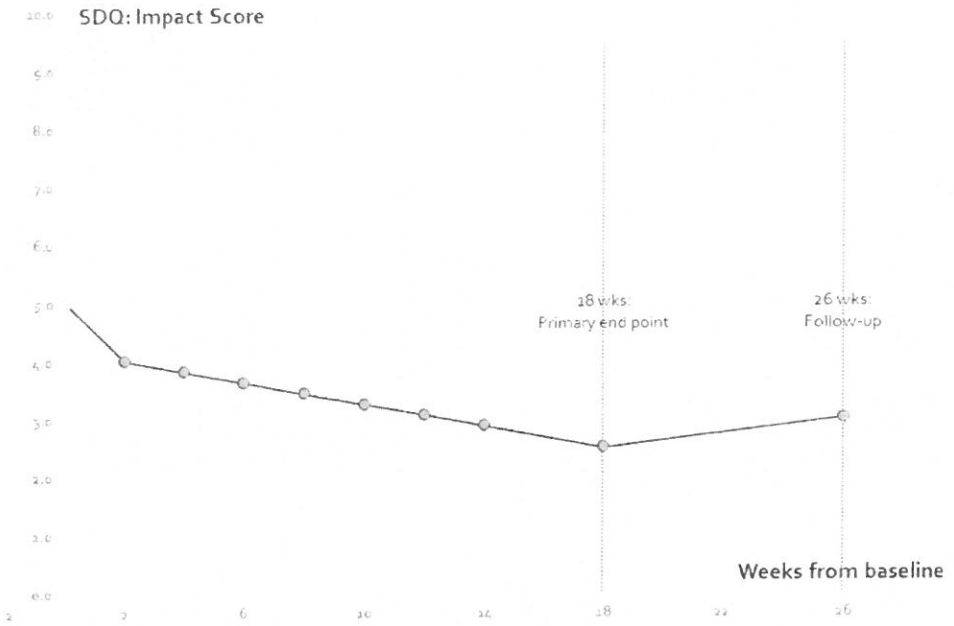


Table 2. Change from Baseline in Primary and Key Secondary Outcomes at 18 Weeks (The ITT Population)

End Point[1]	MMM group (n = XXX)	TAU group (n = XXX)	Difference Between Groups (95% CI)	*P* Value
Primary outcome measure:				
Parent reported SDQ Impact score (0-10)				
Secondary outcome measures:				
Anxiety (SCAS, parent-reported; 0-114)				
Depressive symptoms (MFQ, parent reported; 0-68)				
Level of daily functioning of child (WFIRS, parent reported; 0-150)				
School attendance (percent school-days in the last 4 weeks, parent reported; 0-100%)				
Top-problem score (parent reported; 1-10)				
Self-reported KIDSCREEN-27, converted to T-values (SD) Physical well-being Psychological well-being				
Behavioral problems (ECBI, parent reported) Intensity score (7-252) Problem score (0-36)				
Emotional & behavioral problems (SDQ Total difficulties, parent reported; 0-40)				
Responder indices:				
Parent reported SDQ Impact score ≥ 1 point reduction from baseline, no. (%)				
Scoring below the parent reported SDQ inclusion cut-off[2], no. (%)				

Abbreviations:

SDQ, The Strengths and Difficulties Questionnaire; SCAS, The Spence Children's Anxiety Scale; MFQ, The Mood and Feelings Questionnaire, WFIRS-P, The Weiss Functional Impairment Rating Scale-Parent Reported; ECBI, The Eyberg Child Behavior Inventory; Kidscreen-27, a Health-Related Quality of Life (HRQOL) with five dimensions, of which we use the scales: Physical Well-Being, Psychological Well-Being.

[1] Data are presented as mean (SD) unless otherwise stated.

[2] Inclusion cut-off on the parent-reported SDQ: a total difficulties score of ≥14 and/or emotional problems ≥5, and/or conduct score ≥3; combined with a functional impairment score of ≥1.

Table 3. Exploratory Outcomes: Change from Baseline at 26 Weeks (The ITT Population)

End Point[1]	MMM group (n = XXX)	TAU group (n = XXX)	Difference Between Groups (95% CI)	*P* Value
Primary outcome measure:				
Parent reported SDQ Impact score (0-10)				
Secondary outcome measures:				
Anxiety (SCAS, parent-reported; 0-114)				
Depressive symptoms (MFQ, parent reported; 0-68)				
Level of daily functioning of child (WFIRS, parent reported; 0-150)				
School attendance (percent school-days in the last 4 weeks, parent reported; 0-100%)				
Top-problem score (parent reported; 1-10)				
Self-reported KIDSCREEN-27, converted to T-values (SD) Physical well-being Psychological well-being				
Behavioral problems (ECBI, parent reported) Intensity score (7-252) Problem score (0-36)				
Emotional & behavioral problems (SDQ Total difficulties, parent reported; 0-40)				
Responder indices:				
Parent reported SDQ Impact score ≥ 1 point reduction from baseline, no. (%)				
Scoring below the parent reported SDQ inclusion cut-off[2], no. (%)				

Abbreviations:

SDQ, The Strengths and Difficulties Questionnaire; SCAS, The Spence Children' s Anxiety Scale; MFQ, The Mood and Feelings Questionnaire, WFIRS-P, The Weiss Functional Impairment Rating Scale-Parent Reported; ECBI, The Eyberg Child Behaviuor Inventory; Kidscreen-27, a Health-Related Quality of Life (HRQOL) with five dimensions, of which we use the scales: Physical Well-Being, Psychological Well-Being.

[1] Data are presented as mean (SD) unless otherwise stated.

[2] Inclusion cut-off on the parent-reported SDQ: a total difficulties score of ≥14 and/or emotional problems ≥5, and/or conduct score ≥3; combined with a functional impairment score of ≥1.

Table 4. Potential harms and negative outcomes at end of treatment (week 18), and at follow-up (week 26)

	MMM (n=XXX	TAU (n=XXX)	Difference Between Groups (95% CI)	*P* Value
Potential harms measures				
Primary measure of harms at week 18: Composite score of suicidality and/or negative cognition, no. (%)				
Primary measure of harms at week 26: Composite score of suicidality and/or negative cognition, no. (%)				
Secondary measure of harms at week 18: Composite score of poor quality of family relationships, free time and/or friendships, no. (%)				
Secondary measure of harms at week 26: Composite score of poor quality of family relationships, free time and/or friendships, no. (%)				
Referrals to CAMHS				
Parent-reported contact to CAMHS, from entry to week 18, no (%)				
Parent-reported contact to CAMHS, from entry to week 26, no (%)				

Abbreviation:
CAMHS, child and adolescent mental health services

Table 5. Other Exploratory Outcomes at 18 weeks and at 26 Weeks

End Point		*18 weeks*				*26 weeks*		
	N	MMM group (n = XX)	TAU group (n = XX)	Difference Between Groups (95% CI)	N	MMM group (n = XX)	TAU group (n = XX)	Difference Between Groups (95% CI)
Self-reported SDQ impact score (0-10)								
Self-reported SDQ total difficulties score (0-40)								
Self-reported anxiety (SCAS; 0-114)								
Self-reported depressive symptoms (MFQ; 0-66)								
Self-reported Top-problem score (1-10)								
KIDSCREEN Autonomy & Parents (self-reported, t-scores)								
KIDSCREEN Peers & Social Support (self-reported, t-scores)								
KIDSCREEN School Environment (self-reported, t-scores)								
Parent-reported Experience of Service Questionnaire (ESQ)								
Self-reported Experience of Service Questionnaire (ESQ)								
Teacher-reported SDQ impact score (0-6)								
Teacher-reported SDQ total difficulties score (0-40)								
KIDSCREEN Physical Well-Being (parent-reported, t-scores)								
KIDSCREEN Psychological Well-Being (parent-reported, t-scores)								
KIDSCREEN Autonomy & Parents (parent-reported, t-scores)								
KIDSCREEN Peers & Social Support (parent-reported, t-scores)								
KIDSCREEN School Environment (parent-reported, t-scores)								
Parental Stress in role functioning (PSS)								

Abbreviations:
SDQ, The Strengths and Difficulties Questionnaire; SCAS, The Spence Children's Anxiety Scale; MFQ, The Mood and Feelings Questionnaire, WFIRS-P, The Weiss Functional Impairment Rating Scale-Parent Reported; ECBI, The Eyberg Child Behaviour Inventory; Kidscreen-27, a Health-Related Quality of Life (HRQOL); PSS; Parental Stress Scale

Appendix Table 1. Change from Baseline in Primary and Key Secondary Outcomes at 18 Weeks (The ITT Population: missing data handled using multiple imputation)

End Point[1]	MMM group (n = XXX)	TAU group (n = XXX)	Difference Between Groups (95% CI)	*P* Value
Primary outcome measure:				
Parent reported SDQ Impact score (0-10)				
Secondary outcome measures:				
Anxiety (SCAS, parent-reported; 0-114)				
Depressive symptoms (MFQ, parent reported; 0-68)				
Level of daily functioning of child (WFIRS, parent reported; 0-150)				
School attendance (percent school-days in the last 4 weeks, parent reported; 0-100%)				
Top-problem score (parent reported; 1-10)				
Self-reported KIDSCREEN-27, converted to T-values (SD) Physical well-being Psychological well-being				
Behavioral problems (ECBI, parent reported) Intensity score (7-252) Problem score (0-36)				
Emotional & behavioral problems (SDQ Total difficulties, parent reported; 0-40)				
Responder indices:				
Parent reported SDQ Impact score ≥ 1 point reduction from baseline, no. (%)				
Scoring below the parent reported SDQ inclusion cut-off[2], no. (%)				

Abbreviations:
SDQ, The Strengths and Difficulties Questionnaire; SCAS, The Spence Children' s Anxiety Scale; MFQ, The Mood and Feelings Questionnaire, WFIRS-P, The Weiss Functional Impairment Rating Scale-Parent Reported; ECBI, The Eyberg Child Behaviuor Inventory; Kidscreen-27, a Health-Related Quality of Life (HRQOL) with five dimensions, of which we use the scales: Physical Well-Being, Psychological Well-Being.

[1] Data are presented as mean (SD) unless otherwise stated.

[2] Inclusion cut-off on the parent-reported SDQ: a total difficulties score of ≥14 and/or emotional problems ≥5, and/or conduct score ≥3; combined with a functional impairment score of ≥1.

Appendix Table 2. Exploratory Outcomes at 26 Weeks (The ITT Population: missing data handled using multiple imputation)

End Point[1]	MMM group (n = XXX)	TAU group (n = XXX)	Difference Between Groups (95% CI)	*P* Value
Primary outcome measure:				
Parent reported SDQ Impact score (0-10)				
Secondary outcome measures:				
Anxiety (SCAS, parent-reported; 0-114)				
Depressive symptoms (MFQ, parent reported; 0-68)				
Level of daily functioning of child (WFIRS, parent reported; 0-150)				
School attendance (percent school-days in the last 4 weeks, parent reported; 0-100%)				
Top-problem score (parent reported; 1-10)				
Self-reported KIDSCREEN-27, converted to T-values (SD) Physical well-being Psychological well-being				
Behavioral problems (ECBI, parent reported) Intensity score (7-252) Problem score (0-36)				
Emotional & behavioral problems (SDQ Total difficulties, parent reported; 0-40)				
Responder indices:				
Parent reported SDQ Impact score ≥ 1 point reduction from baseline, no. (%)				
Scoring below the parent reported SDQ inclusion cut-off[2], no. (%)				

*Non-Responder Imputation: Defined as Baseline Observation Carried Forward.

Abbreviations:

SDQ, The Strengths and Difficulties Questionnaire; SCAS, The Spence Children's Anxiety Scale; MFQ, The Mood and Feelings Questionnaire, WFIRS-P, The Weiss Functional Impairment Rating Scale-Parent Reported; ECBI, The Eyberg Child Behavior Inventory; Kidscreen-27, a Health-Related Quality of Life (HRQOL) with five dimensions, of which we use the scales: Physical Well-Being, Psychological Well-Being.

[1] Data are presented as mean (SD) unless otherwise stated.

[2] Inclusion cut-off on the parent-reported SDQ: a total difficulties score of ≥14 and/or emotional problems ≥5, and/or conduct score ≥3; combined with a functional impairment score of ≥1.

Appendix Table 3. Change from Baseline in the Primary and Secondary Outcomes at 18 Weeks: Per Protocol Population*

End Point[1]	MMM group (n = XXX)	TAU group (n = XXX)	Difference Between Groups (95% CI)	*P* Value
Primary outcome measure:				
Parent reported SDQ Impact score (0-10)				
Secondary outcome measures:				
Anxiety (SCAS, parent-reported; 0-114)				
Depressive symptoms (MFQ, parent reported; 0-68)				
Level of daily functioning of child (WFIRS, parent reported; 0-150)				
School attendance (percent school-days in the last 4 weeks, parent reported; 0-100%)				
Top-problem score (parent reported; 1-10)				
Self-reported KIDSCREEN-27, converted to T-values (SD) Physical well-being Psychological well-being				
Behavioral problems (ECBI, parent reported) Intensity score (7-252) Problem score (0-36)				
Emotional & behavioral problems (SDQ Total difficulties, parent reported; 0-40)				
Responder indices:				
Parent reported SDQ Impact score ≥ 1 point reduction from baseline, no. (%)				
Scoring below the parent reported SDQ inclusion cut-off[2], no. (%)				

*Only participants with data both at baseline and at week 18 are included in the analysis.

Abbreviations:

SDQ, The Strengths and Difficulties Questionnaire; SCAS, The Spence Children' s Anxiety Scale; MFQ, The Mood and Feelings Questionnaire, WFIRS-P, The Weiss Functional Impairment Rating Scale-Parent Reported; ECBI, The Eyberg Child Behavior Inventory; Kidscreen-27, a Health-Related Quality of Life (HRQOL) with five dimensions, of which we use the scales: Physical Well-Being, Psychological Well-Being.

[1] Data are presented as mean (SD) unless otherwise stated.

[2] Inclusion cut-off on the parent-reported SDQ: a total difficulties score of ≥14 and/or emotional problems ≥5, and/or conduct score ≥3; combined with a functional impairment score of ≥1.

References

1. Costello EJ, Copeland W, Angold A. Trends in psychopathology across the adolescent years: What changes when children become adolescents, and when adolescents become adults? J Child Psychol Psychiatry. 2011;52(10):1015-1025. doi:10.1111/j.1469-7610.2011.02446.x

2. Copeland WE, Adair CE, Smetanin P, et al. Diagnostic transitions from childhood to adolescence to early adulthood. J Child Psychol Psychiatry. 2013;54(7):791-799. doi:10.1111/jcpp.12062

3. Reef J, Van Meurs I, Verhulst FC, Van Der Ende J. Children's problems predict adults' DSM-IV disorders across 24 years. J Am Acad Child Adolesc Psychiatry. 2010;49(11):1117-1124. doi:10.1016/j.jaac.2010.08.002

4. Arango C, Díaz-Caneja CM, McGorry PD, et al. Preventive strategies for mental health. The Lancet Psychiatry. 2018;5(7):591-604. doi:10.1016/S2215-0366(18)30057-9

5. Detry MA, Ma Y. Analyzing Repeated Measurements Using Mixed Models. Jama. 2016;314(4):407-408. doi:10.1001/jama.2015

6. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. BMJ. 2011;342:d40. doi:10.1136/bmj.d40