A pragmatic randomized controlled cost effectiveness study of Well Parent Japan (WPJ) for Japanese mothers of children with Attention Deficit Hyperactivity Disorder (ADHD): The TRaining And Nurturing Support FOR Mothers*

Statistical Analysis Plan

Final version 1.0
(31 May 2022)

Based on Published Protocol [3] JMIR Res Protocol 2022 (Accepted 6 Feb 2022, Published19 Apr 2022)

Trial registration [1]: ISRCTN66978270

* The title of the published protocol paper was shortened to meet the journal requirements.

The following people have reviewed the Statistical Analysis Plan and are in agreement with the contents[6]

Name	Job title	Trial Role	Signature	Date
Shizuka Shimabukuro	Staff Scientist	Project Leader & Co-investigator		31.05.2022
Boliang Guo	RSS Cstat Asso. professor of medical statistics	Trial Statistician		May-30-2022
David Daley	Professor	Co-Investigator		30.05.2022
Gail Tripp	Professor	Chief Investigator		31.05.2022

Abbreviations

Abbreviation	Description
DMC	Data monitoring committee
SAP	Statistical Analysis Plan
TMG	Trial Management Group
TSC	Trial Steering Committee

Changes from protocol

The table below details changes to the planned analyses in the SAP compared to the protocol which after discussion with the TMG are not considered to require a protocol amendment.

Protocol version and section	Protocol text	SAP version and section	SAP text	Justification
Published version methods and table 1 which includes a tick for the health economic interview at T2	Both arms will be tested at baseline, immediately after the intervention arm completes WPJ at week 14, and at 3 months follow-up; that is, week 26. The study is being carried out through Japanese child and adolescent mental health care services at three different hospital sites in Japan (Fukui, Fukuoka, and Okinawa). Table 1 lists the schedule of parent self-report, teacher, and objective assessments at each time point	25/05/2020 3.2 b 5	Cost-effectiveness interview data at T2 was not collected for participants in the first wave of the study. The cost-effectiveness interview data collection was originally planned for T1 and T3, this was subsequently modified to include all three data points to improve recall/tracking of service utilization. Data from later waves cannot be used to model the missing data from wave 1 due to the impact of COVID-19 on service utilization. Analyses will be carried out for T1(service use in the three months prior to baseline) and T3 (service use in the three months following intervention).	The data was not collected at T2 for some of the participants and it was impossible to model. Focusing on data at T1 and T3 still provides a robust exploration of health economics for this trial.

Amendments to versions(4)

Version	Date	Change/comment	Statistician

Additional contributors to the SAP (non-signatory) (5)

Use this section to acknowledge other individuals who have made a significant contribution to the SAP.

Name	Trial role	Job Title	Affiliation

Contents

1. Introduction

This document details the rules proposed and the presentation that will be followed, as closely as possible, when analysing and reporting the main results from the study titled "The effectiveness and cost-effectiveness of Well Parent Japan for Japanese mothers of children with ADHD: Protocol for a randomized controlled trial". These analyses will assess the efficacy and safety of Well Parent Japan (WPJ), a new hybrid intervention in comparison with the treatment as usual (TAU) within routine Japanese mental health services and will be included in the clinical study report.

The purpose of the plan is to:

- Ensure that the analysis is appropriate for the aims of the trial, reflects good statistical practice, and that interpretation of a priori and post hoc analyses respectively is appropriate.
- Explain in detail how the data will be handled and analysed to enable others to perform or replicate these analyses.

Additional exploratory or auxiliary analyses of data not specified in the protocol may be included in this analysis plan. This analysis plan will be made available if required by journal editors or referees when the main papers are submitted for publication. Additional analyses suggested by reviewers or editors will be performed if considered appropriate. This should be documented in a file note.

Amendments to the statistical analysis plan will be described and justified in the final report of the trial and where appropriate in publications arising from the analysis. *Health economic and qualitative analysis plans are beyond the scope of this document.*

This document was developed with reference on Nottingham CTU SAP template (2018) and Cambridge CTU SAP template (2018), JAMA suggestion [1] on SAP and EMA guidelines for analysing randomised clinical trials [2-8].

1.1 Background and rationale (7)

In Japan, the value of parent training is recognized by physicians and allied health professions with wider acceptance and increased implementation. However, the availability of psychosocial interventions specifically designed for ADHD remains limited. In response we developed Well Parent Japan (WPJ), a hybrid program combining a culturally adapted version of the New Forest Parent Programme (NFPP), an empirically supported parenting program designed for families of children with ADHD, with Parent Stress Management Training (psychological support) for mothers. The hybrid program was evaluated against a waitlist control group in a small-scale RCT [9]. Participation in WPJ led to significant reductions in parenting stress, increased parenting self-esteem, implementation of more effective parenting strategies, and reduced maternal negativity. Mothers also reported reduced aggression, internalizing problems and inattention in their children. Well Parent Japan appears to be an efficacious psychosocial intervention for ADHD in Japan, with the group format and the session content well-tolerated.

The current study aims to extend the results from this pilot RCT [9] to a larger multi-center pragmatic trial to assess the effectiveness and cost-effectiveness of WPJ compared with treatment as usual (TAU) in child and adolescent mental health care services in Japanese hospitals.

1.2 Objectives (8)

This study aims to extend the results from the pilot RCT [9] to a larger multicenter pragmatic trial to assess the effectiveness and cost-effectiveness of WPJ compared with treatment as usual (TAU) in child and adolescent mental health care services in Japanese hospitals. Well Parent Japan will be compared against TAU at the end of the intervention and again after 3 months (short-term follow-up).

The primary objective is to compare the effectiveness of WPJ against TAU in reducing maternal parent domain parenting stress, that is perceived stress in the parent-child dyad arising from characteristics of the parent, measured with the Parent Stress Index [10, 11]. Parent domain stress was selected as the primary outcome measure given WPJ's strong focus on the emotional well-being of participating mothers.

Secondary objectives include comparison of the effectiveness of WPJ against TAU in improving child behaviour, parental well-being, and parenting practices and to explore the cost-effectiveness of WPJ.

Well Parent Japan is expected to reduce parent domain parent stress immediately post intervention and at 3 months follow-up compared with TAU. It is also predicted to improve key secondary outcomes such as child behaviour, parental well-being and parenting practices.

The null hypothesis is that there will be no difference in the parent domain stress scores for the WPJ group and the TAU group, controlling for baseline scores, immediately post-treatment and at 3 months follow-up. The alternative hypothesis is that there will be a significant reduction in parent domain stress scores for the WPJ compared to TAU groups at the end of treatment and at 3 months follow-up controlling for baseline scores.

For the secondary outcomes the null hypothesis is that there will be no significant difference in the secondary outcome scores between the WPJ and TAU groups at the end of treatment and at 3 months follow-up, controlling for baseline scores. The alternative hypotheses are that the WPJ group will show improvements in the secondary outcomes post treatment and at 3 months follow-up, controlling for baseline scores.

2. Study methods

2.1 Trial design (9)

The trial is a pragmatic multi-centre, randomised controlled trial (RCT) comparing the effectiveness and cost-effectiveness of WPJ (treatment arm) with TAU (control arm). Both arms will be tested at baseline, immediately after the intervention arm completes WPJ at week 14, and at 3 months follow-up (week 26). The study is being carried out through Japanese child and adolescent mental health care services at three different hospital sites in Japan (Fukui, Fukuoka, and Okinawa). Treatment allocation is a 1:1 ratio across sites.

Well Parent Japan (WPJ) is a new hybrid intervention created by combining culturally adapted versions of Parent Stress Management for ADHD (PSM) and the New Forest Parenting Programme (NFPP), a behavioural parent training intervention specific to ADHD.

Well Parent Japan is a 13-week group-based intervention that aims to reduce maternal stress and enhance maternal well-being, so that mothers can effectively implement NFPP strategies, modify their ADHD child's environment and reduce the expression of ADHD symptoms. The intervention has been evaluated in a pilot RCT and demonstrated considerable advantages over a wait list control group for mothers of children with ADHD.

2.2 Randomisation & Blinding (10)

The randomisation process is described in full within the clinical trial protocol [12].

Physicians at each site confirm participant eligibility prior to them receiving a detailed explanation of study procedures, including randomization to WPJ or TAU. Participants at each site are assigned an ID number, and only this number, but not participants identity, is shared with the primary research site (Okinawa Institute of Science and Technology, OIST) for data base entry and analyses. Once sufficient participants are recruited at each site (minimum of 12 for each of 3 waves of data recruitment and intervention) they complete baseline measures before being randomized into WPJ or TAU (block randomization) using a computer random number generator operated by a member of the research team (DD) who has no role in participant recruitment, treatment implementation, or data collection. Details of the randomisation method are held securely within the statistics master file.

Researchers, statistician, site clinical and administrative staff, and participants will be non-blind to participant allocation arm. Staff responsible for coding the parent-child interaction (Pasta task) and the five-minute speech sample (Expressed Emotion) will be blind to participant allocation arm. Teachers will be probably blind as they will not be informed of treatment allocation by the research team.

2.3 Sample size (11)

Parent domain parenting stress scores at week 14 is the primary outcome measure and informed our sample size calculation. Based on the results of our pilot RCT [9], to detect a 0.5 standardized effect size at week 14 using 80% power at a 2-tailed .005 significance level, assuming the correlation between the baseline and follow-up measures is 0.35, 112 participants were required. After adjusting for a 15% attrition rate, the target sample size was inflated to 132. The Stata `sampsi` command was used for power analysis.

2.4 Framework(12)

In line with the study objectives, specified in trial protocol and reiterated here in section 1.2, both primary and secondary outcomes are testing for superiority of the treatment effectiveness of WPJ over TAU.

2.5 Statistical interim analyses and stopping guidance (13)

There is no formal interim analysis planned.

2.6 Timing of final analysis (14)

The final analysis will be performed once the last recruited participant's last follow-up outcome data are available.

2.7 Timing of outcome assessments (15)

The schedule of data collection for all outcome measures is presented in Table 1. Data collection for primary and secondary outcome measures was planned for baseline (prior to randomization), week 14 and week 26 (3-month follow-up for the treatment group). In person data collection, for mother and child, (pasta task and five-minute speech sample) was scheduled for the baseline and week 14 assessments.

Baseline questionnaires were sent to mothers and teachers 3 weeks prior to the beginning of the WPJ program. Mothers and teachers were requested to send these back within 10 days of receipt. If the questionnaires were not returned within this time period mothers/teachers were contacted and reminded to send back the completed questionnaires. Baseline in-person assessments were carried out over the two weekends prior to the beginning of WPJ (5-12 days prior to WPJ). Mothers were informed of groups assignment after all parents completed the in-person assessment, i.e., 5-7 days prior to WPJ beginning.

Week 14 questionnaires were handed to the mothers at the end of the last session of the treatment group and also sent to mothers of the TAU group and all teachers via mail on the second day post WPJ completion by the treatment group. Mothers and teachers were requested to send these back within 10 days of receipt. If not received by that time mothers/teachers were contacted and reminded to send back the completed questionnaires. Week 14 in-person assessments were carried out over the two weekends following completion of WPJ by the treatment group.

Week 26 questionnaires were sent to mothers and teachers 12 weeks after completion of WPJ by the treatment group mothers. Mothers and teachers were requested to return these within 10 days of receipt. If completed questionnaires were not returned within this time frame mothers/teachers were contacted and reminded to return their completed questionnaires.

The timing of the week 14 and week 26 assessments were disrupted during the third wave of WPJ groups for the Kurume (Fukuoka) and Okinawa sites. COVID-19 closures resulted in suspension of the WPJ program at both sites. At the Kurume site program delivery was stopped for 13 weeks and at the Okinawa site for 16 weeks. Week 14 and week 26 data collection was therefore delayed by 13 weeks for participants at Kurume and by 16 weeks for participants in Okinawa.

3 Statistical Principals

3.1 Confidence intervals and P values (16-18)

All applicable statistical tests will be 2-sided and will be performed using a 5% significance level; No planned adjustment for multiplicity as the study has only one primary outcome [13]. All confidence intervals presented will be 95% and two-sided.

3.2 Adherence and protocol deviations (19)

Intervention adherence is recorded as (a) the fidelity of WPJ delivery and (b) number of missed interventions sessions. (i.e., number of treatment sessions not attended, and number of treatment sessions not attended and not followed by a catch-up session).

(a) Fidelity of WPJ delivery is assessed through review of approximately 20% (26/117 sessions) of the intervention sessions, across the three sites, selected at random, with the following constraints: inclusion of two recordings for each of the 13 WPJ sessions, at least eight recordings from each intervention site and each wave of data collection. Recordings of these 26 sessions were reviewed by a research assistant, blind to study site, against the Group Leaders manual for major and minor session content. Inclusion of major and minor content (percentage), across sites, will be reported in the trial paper; (b) Non-attendance will be recorded as the total number of group sessions missed and the total number of group sessions missed and not caught up. These will be summed for groups, sites, data collection waves and overall. These will be recorded as percentages with the number of available sessions as the denominator. The overall percentage of missed sessions (caught up and not caught up) will be reported in the trial paper.

(b)Protocol deviations

1. The primary protocol variation is the period of time over which the 13-week WPJ program was delivered to the third wave of participants at the Kurume and Okinawa sites. This violation resulted from the shut-down of hospital and community sites where the intervention was provided, due to community spread of COVID-19. The Fukui site was not impacted. Specifically, at the Kurume site program delivery was stopped for 13 weeks and at the Okinawa site for 16 weeks. Week 14 and week 26 data collection was therefore delayed by 13 weeks for participants at Kurume and by 16 weeks for participants in Okinawa.

2. In response to the extended program suspensions, the affected treatment groups were provided with one additional group session. This additional session was designed to refresh participants memories of the programme content covered prior to the suspension of the groups, before new content was introduced.

3. The cost-effectiveness interview for the 3rd treatment group at the Kurume site was carried out via telephone, rather than in person, to reduce the risk of COVID infection. This was considered necessary as data collection was being undertaken at a developmental support center for children with disabilities.

4. One participant was identified as the paternal grandmother, rather than the mother, of a child with ADHD. The grandmother was responsible for the child's care, i.e., was the child's primary caregiver.

5. Cost-effectiveness interview data at T2 was not collected for participants in the first wave of the study. The cost-effectiveness interview data collection was originally planned for T1 and T3, this was subsequently modified to include all three data points to improve recall/tracking of service utilization. Data from later waves cannot be used to model the missing data from wave 1 due to the impact of COVID-19 on service utilization. Analyses will be carried out for T1(service use in the three months prior to baseline) and T3 (service use in the three months following intervention).

6. In the protocol paper, T2 data collection was described as taking place 14 weeks after baseline data collection. In practice, data collection took place between weeks 14 and 16, depending on the timing of questionnaire return and scheduling of the in-person assessments. Time 3 data collection is described in the protocol as taking place 26 weeks after baseline data collection. In practice, data collection took place between the 26th and 28th week post baseline data collection. Any deviations beyond these data collection periods will be reported in the trial paper. T2 and T3 data collection for the third treatment group at the Kurume and Okinawa sites did not follow this timeline due to the suspension of program delivery as described under point 1 above.

7. Four study participants (mothers of children with ADHD) self-identified their interest in participating the RCT rather than being referred to the study by one of the site physicians. All participants received a detailed explanation of the study from the site research assistants who checked for study inclusion and exclusion criteria.

All protocol variations will be reported in the main Trial paper.

3.3 Analysis populations (20)

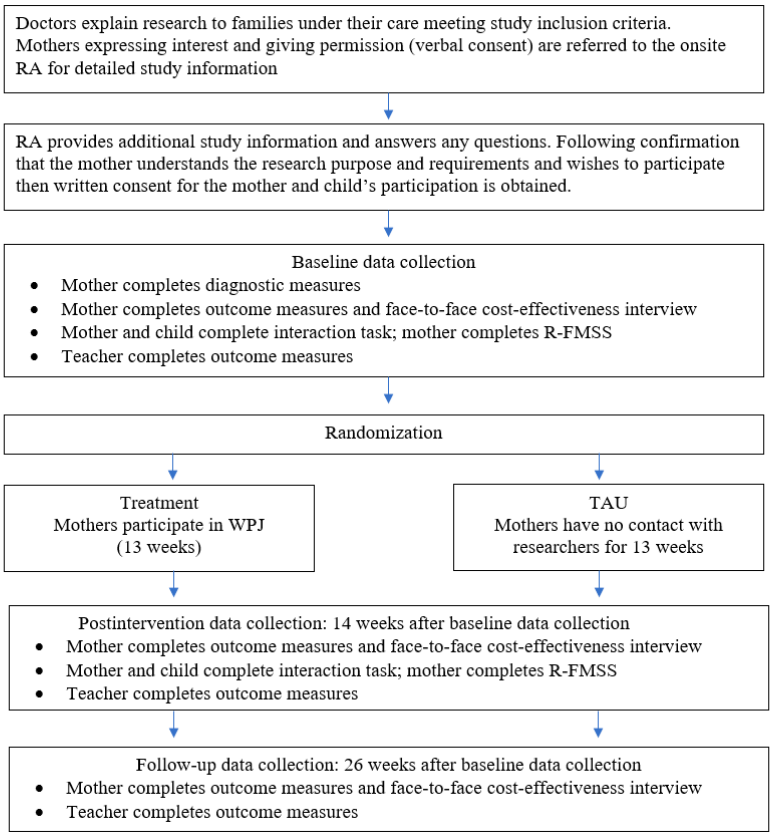
The intention-to-treat population will include all randomised participants, regardless of their eligibility, according to the treatment they were randomised to receive.

4 Trial population

4.1 Screening data (21)

The number of participants screened will be presented in CONSORT flow diagrams (figure 1).

Figure 1. Consort flow



R-FMSS: revised Five-Minute Speech Sample; RA: research administrator; TAU: treatment as usual; WPJ: Well Parent Japan.

4.2 Eligibility (22)

Inclusion criteria

Fluency in Japanese (reading and writing); parenting a child aged 6-12 years diagnosed with ADHD and attending elementary school, and for whom participation in a group-based behavioral intervention for the mother is not contraindicated. Mothers of children diagnosed with ADHD and autism spectrum disorder were eligible to participate. As a pragmatic trial, referring doctors were asked to exercise clinical judgment regarding mothers' ability to understand the program content and their suitability to participate in a group program.

Exclusion criteria

Limited pragmatic speech or a functional intellectual disability in the child; current or recent (ie, within 2 months of the starting date of WPJ) participation in another parenting program. Mothers of children receiving medication for the management of their ADHD symptoms were asked to keep their child's medication constant throughout the study. Medication status changes will be recorded but will not result in the family being removed from the trial.

As randomization takes place after confirmation of eligibility, no ineligible participants are expected to be randomized. Should this occur, the number of ineligible participants

randomised will be reported, with reasons for ineligibility, and presented in CONSORT diagrams.

4.3 Recruitment (23)

A CONSORT flow diagram (figure 1) will be used to summarise the number of participants who were:

- assessed for eligibility at screening (number referred to site RAs for detailed explanation)
    - eligible at screening
    - ineligible at screening*
- eligible and randomised
- eligible but not randomised*
- received the randomised allocation
- did not receive the randomised allocation*
- lost to follow-up*
- discontinued the intervention*
- randomised and included in the primary analysis
- randomised and excluded from the primary analysis*

*Reason will be being provided

4.4 Withdrawn/follow-up (24)

The level of consent withdrawal will be tabulated (classified as "consent to continue follow-up and data collection" "consent to continue data collection only", "complete – no further follow-up or data collection")." This will be presented in CONSORT diagram format rather than as a table, with numbers and reasons for withdrawal and/or exclusion from analysis given at each stage.

Figure 1: Example of skeleton CONSORT flow diagram

4.5 Baseline participant characteristics (25)

Participants will be described with respect to child's age, gender, diagnosis and medication use, mother's age, education, occupational status, total family income, and marital status, at baseline, both overall and separately for the two randomised groups. The details of descriptive statistics are reported in 5.2.1. Tests of statistical significance will not be undertaken for baseline characteristics; rather the clinical importance of any imbalance will be noted [3].

5 Analysis

5.1 Outcome definitions (26)

Brief information regarding the primary and secondary outcomes is summarised in table 1.

Table 1: Summary of the clinical efficacy outcome measures

Outcome measures	Scale, description and source	Derivation of scores	Time point (weeks)		
Primary outcome			**0**	**14**	**26**
Change from baseline in Parent Stress Index (PSI) at 14 weeks	Parenting Stress Index (PSI), 78 items, response range is 1 to 5	PSI Parent Domain Stress	✓	✓	✓
Secondary outcomes					
Change from baseline in Parent Stress Index (PSI) at 26 weeks	Parenting Stress Index (PSI), 78 items, response range is 1 to 5	PSI Parent Domain Stress	✓	✓	✓
Change from baseline in measures at 14 and 26 weeks.	Beck Depression Index (BDI), 21 items, response range 0 to 3	Total score	✓	✓	✓
Change from baseline in measures at 14 and 26 weeks.	Parenting Scale (PS), 30 items, response range 1 to 7	• Japanese score Laxness • Japanese score Overreactivity	✓	✓	✓
Change from baseline in measures at 14 and 26 weeks.	Parent sense of competence (PSOC), 17 items, response range 1 to 6	• Satisfaction • Efficacy	✓	✓	✓
Change from baseline in measures at 14 and 26 weeks.	Parent locus of control (PLOC), 37 items, response range 1 to 5	• Efficacy • Responsibility • Parent control • Child control	✓	✓	✓
Change from baseline in measures at 14 and 26 weeks.	Mother rated SNAP, 26 items, response range 0 to 3	• Inattention • Hyperactivity/impulsivity • ODD	✓	✓	✓
Change from baseline in measures at 14 and 26 weeks.	Parent rated Vanderbilt Assessment Scale, 8 items, response range 1 to 5	• Academic performance • Social performance	✓	✓	✓
Change from baseline in measures at 14 and 26 weeks.	Parent Impairment Rating Scale, 8 items, response range 0 to 6	• Peer relationship • Sibling relationship • Parent relationship • Academics • Self-Esteem • Family Functioning • Global	✓	✓	✓
Change from baseline in measures at 14 and 26 weeks.	Teacher rated SNAP, 26 items, response range 0 to 3	• Inattention • Hyperactivity/impulsivity • ODD	✓	✓	✓
Change from baseline in measures at 14 and 26 weeks.	Teacher rated Vanderbilt Assessment Scale, 8 items, response range 1 to 5	• Academic performance • Behavioural performance	✓	✓	✓
Change from baseline in measures at 14 and 26 weeks.	Family Strain Index, 6 items, response range 0 to 4	Total score			
Change from baseline in measures at 14 and 26 weeks.	Mother-child interaction observation	• Positive parenting • Negative parenting • Negative child affect	✓	✓	✓

Outcome measures	Scale, description and source	Derivation of scores	Time point (weeks)		
Change from baseline in measures at 14 and 26 weeks.	Expressed emotion	• Relationship • Warmth • Positive comment • Negative comment	✓	✓	✓

5.2 Analysis methods (27)

All analyses will be conducted on an Intention-To-Treat (ITT) basis [2]. Exploratory analysis will be conducted first for all measures; All participant demographic and outcome measures will be summarised by arm across follow-up times if repeatedly measured, with n (non-missing sample size), mean, standard deviation, median, maximum and minimum for continuous variables, the frequency and percentages (based on the non-missing sample size) of observed levels for all categorical measures.

5.2.1 Summary of primary and secondary outcomes analysis

Treatment effect estimate and its precision (95% confidence interval, 95%CI) on all repeated outcomes measure will be quantified by multilevel modelling (MLM) with baseline measures, minimization factors included as covariates and participants as higher-level analytical units [3, 13, 14]. Assumption for all linear regression will be assessed by visually and numerically exploring the residual values of each MLM modelling. No formal adjustment for multiple significance testing will be applied, in that there is only one primary outcome and secondary outcomes will be considered supportive to the primary analysis. The significance of all parameters will be tested using 2-tailed 0.05 significance level. Skewed outcome variables will be transformed for linear regression modelling if needed with reference on data exploratory results.

5.2.2 Analysis of primary outcome

The primary outcome is parenting stress of PSI at 14 weeks post randomisation so the parameter (95%CI) reflecting treatment effects will be group difference of mean change from baseline to 14 weeks. The parameter (95%CI) will be derived using ANCOVA approach by means of MLM with covariate including baseline measure, binary group status, follow up time, interaction term of group and time, participant will be the level 2 analytical unit [3]. The proposed MLM model equation could be written as

$$y_{ij} = \beta_{0j} + BX_{ij} + e_{ij}$$

$$\beta_{0j} = \beta_0 + \mu_{0j}$$
$$\mu_{0j} \sim N(0, \sigma^2_{\mu_0})$$
$$e_{ij} \sim N(0, \sigma^2_e)$$

With y_{ij} is the PSI change from baseline score, i= follow up time,

j= individual participant indicator, X are covariates vector including baseline measure, arm status, dummy coding for discrete time, interaction term of arm × time, B is a vector of fixed effect regression coefficients, β_{0j} is regression intercept parameter which was set random at participant level (if any significant variability show from model exploring), μ_{0j} is the departure of jth participant

change score from overall mean change and distributed normally with mean 0 and variance $\sigma^2_{\mu_0}$, e_{ij} is the regression residual term which follows a normal distribution with mean 0 and variance σ^2_e. Residual Q-Q plot will be examined to check the normality assumption for regression modelling. As participants will be recruited from various centres, centre will be included as a higher-level analytical unit if data exploratory analysis showed there is great variability at centre level.

5.2.3 Sensitivity analyses of primary outcome

- Model shown in section 5.2.2 will be performed on observed data only (safety dataset).
- Model shown in section 5.2.2 will be performed with parameters' standard errors adjusted for centre cluster effect.

5.2.4 Secondary analysis of primary outcome

- Model shown in section 5.2.2 will be performed on per protocol population.

5.2.5 Planned subgroup analyses of primary outcome

There is no planned subgroup analysis.

5.2.6 Analysis of secondary outcomes

All the secondary outcome will be analysed with same analytical modelling as for primary outcome shown in section 5.2.2.

5.3 Missing data[28]

Missing values in all outcomes will be checked and reported across treatment group and follow up time. As the outcome will be repeatedly measured, a two level logistic regression with participant as a level 2 unit will be performed to test the influence of treatment status and baseline measures on outcome missingness. The missing value patterns and the results from multilevel logistic regression modelling will be used to inform missing value imputation under Missing at Random (MAR) assumption [15]. Although multilevel modelling for repeated measures could be automatically taken into account missing outcomes under MAR assumption may be used to give sensible results [16], to make sure all randomised participants will be included in the analysis, the missing values will be imputed using analytical multilevel modelling to quantify the treatment effect estimates [17]. Results of modelling on observed data will be used as sensitivity analysis to check the robustness of results sensitive to missing value[18]. The then latest version Stata and REALCOM-IMPUTE software will be used to perform multiple imputations via analytical model by means of Markov chain Monte Carlo (MCMC) approach for multilevel data [17]. MCMC procedure setting include one chain, burn-in length=1000, chain length=5000, a thinning of 10 and non-informative priors for all parameters included in the model. The random seed number for MCMC stage will be the one automatically generated by REALCOM-IMPUTE software at each MCMC run. Twenty imputed datasets will be generated initially with possible imputing number changing after checking imputation performance [19]. Results from imputed dataset will be combined using Rubin's imputation rules to produce a pooled treatment effect estimate (95% CI) and a pooled p-value for the test of null hypothesis of no treatment effect [15].

5.4 Additional analyses/exploratory analysis (29)

There is no planned exploratory analysis.

5.5 Harms & Adverse events (30)

All AEs, SAEs, ADEs and SADEs will be documented in the participant's records and CRF, and will have to be followed until resolution, or for at least 30 days after discontinuation in use of the device, whichever comes first. Those AEs will be reported to REC within the statutory timeframes and in the annual reports to REC. The number (%) of participants experiencing each AE/SAE will be presented for each treatment arm categorised by severity (across follow-up time) [8]. For each participant, only the maximum severity experienced of each type of AE will be displayed. The number (%) of occurrences of each AE/SAE will also be presented for each treatment arm. The denominator used to work out the percentage of each type of harm at each follow up time will be the numbers of ITT population. No formal statistical testing will be undertaken.

5.6 Statistical software (31)

The analysis will be carried out using Stata, REALCOM and possibly MLwiN. All the software will be the then latest version available in University of Nottingham (UoN) when study data is ready for analysis. All the data will be stored in UoN secure server and analysed in UoN computers. All the data and analytic code will be archived as per instruction from study Trial Statistician Boliang Guo who will be the data custodian for this study.

6 References (32)

1. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, Williamson PR, Altman DG, Montgomery A, Lim P *et al*: **Guidelines for the content of statistical analysis plans in clinical trials**. *JAMA* 2017, **318**(23):2337-2343.
2. EMA: **Statistical principles for clinical trials (E9): ICH tripartite guideline**: European Medicines Agency; 1998.
3. EMA: **Guideline on adjustment for baseline covariates**. London, UK: European Medicines Agency; 2013.
4. EMA: **Guideline on the investigation of subgroups in confirmatory clinical trials**. 7 Westferry Circus, Canary Whard, London E14 4HB: European Medicines Agency; 2014.
5. EMA: **General principles for planning and design of multi-regional clinical trials**: European Medicines Agency; 2016.
6. EMA: **Guideline on multiplicity issues in clinical trials**. 30 Churchill Place, Canary Wharf London E14 5EU, United Kingdom: European Medicines Agency; 2016.
7. EMA: **Guideline on Missing Data in Confirmatory Clinical Trials**. Westferry Circus, Canary Wharf, London E14 4HB,United Kingdom: European Medicines Agency; 2011.
8. Ioannidis JPA, Evans SJ, Gøtzsche PC, O'Neill RT, Altman DG, Schulz K, Moher D: **Better reporting of harms in randomized trials: an extension of the CONSORT statement**. *Annals of Internal Medicine* 2004, **141**.
9. Shimabukuro S, Daley D, Thompson M, Laver-Bradbury C, Lovern K, Tripp G: **Supporting Japanese mothers of children at risk for Attention Deficit**

Hyperactivity Disorder (ADHD): A small scale randomized control trial of Well Parent Japan. *Journal of Child and Family Studies* 2020, **29**(6):1604-1616.

10. Abidin R: **Parenting Stress Index-Manual**. Charlottesville: Pediatric Psychology Press; 1983.
11. Narama M, Kanematsu Y, Araki A, Maru M, Nakamura N, Takeda J: **Nihonban parent stress index (PSI) no shinraisei, datousei no kentou [Review of the reliability and validity of the Japanese version of the parenting stress index (PSI)]**. Shouni Hoken Kenkyu. 1999;58:610-16. Japanese.
12. Shimabukuro S, Daley D, Endo T, Harada S, Tomoda A, Yamashita Y, Oshio T, Guo B, Ishii A, Izumi m, Nahahara Y, Yamamoto k, Yao A, Tripp G: **The effectiveness and cost-effectiveness of Well Parent Japan for Japanese mothers of children with ADHD: Protocol for a randomized controlled trial**. JMIR Research Protocols 2022, 11(4): 1-14.
13. Kahan BC, Morris TP: **Improper analysis of trials randomised using stratified blocks or minimisation**. *Statistics in Medicine* 2012, **31**(4):328-340.
14. Kahan BC, Morris TP: **Analysis of multicentre trials with continuous outcomes: when and how should we account for centre effects?** *Statistics in Medicine* 2013, **32**(7):1136-1149.
15. Carpenter J, Kenward MG: **Multiple Imputation and its Application**. The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom: John Wiley & Sons, Ltd; 2013.
16. Goldstein H: **Multilevel Statistical Models**, 4th edn. The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom: John Wiley & Sons Ltd; 2011.
17. Carpenter J, Goldstein H, Kenward MG: **REALCOM-IMPUTE software for multilevel multiple imputation with mixed response types**. *Journal of Statistical Software* 2011, **5**:1-14.
18. White IR, Horton NJ, Carpenter J, Pocock SJ: **Strategy for intention to treat analysis in randomised trials with missing outcome data**. *BMJ* 2011, **342**.
19. White IR, Royston P, Wood AM: **Multiple imputation using chained equations: Issues and guidance for practice**. *Statistics in Medicine* 2011, **30**(4):377-399.