Statistical analysis plan

MIND PRO trial

Version date: 19-11-2026

Contents

1. Background 2

2. Study Objectives 2

3. Study Design 3

Study population 3

Sample size calculation 4

Randomisation procedures 5

Study outcomes 5

Primary outcomes 5

Secondary outcomes 6

4. Statistical Analysis 8

General 8

Primary outcomes 8

Sensitivity analyses ..... 9~~8~~

Predefined subgroup analyses ..... 9~~8~~

Secondary outcomes 9

References ..... 11~~10~~

1. Background

The Mobile Health (mHealth) Intervention for Dementia Prevention through lifestyle Optimisation (MIND-PRO) study addresses the increasing prevalence of dementia among populations with lower socioeconomic status (SES) and a migrantion background. The study aims to evaluate the effectiveness and implementation of an mHealth app designed for self-managed lifestyle modifications with remote coaching to reduce dementia risk factors.

This Analysis plan will be used as a work description for all the persons who are involved in the analyses of the MIND PRO trial (ISRCTN92928122).

2. Study Objectives

The overall aim is to investigate whether a coach-supported mHealth intervention for lifestyle improvement can reduce the risk of dementia in those with low SES and/or a migration background aged 50-75 years.

The specific objectives are to investigate:

- The effectiveness of a coach-supported mHealth intervention for lifestyle improvement to reduce the risk of dementia in those with low SES and/or a migration background aged 50-75 years.
- The implementation of this coach-supported mHealth intervention for dementia prevention, operationalised as the coverage, acceptability, adoption, appropriateness, feasibility, fidelity, costs, and sustainability.

3. Study Design

The study is a single-centre, investigator initiated, prospective, open-label blinded endpoint randomized controlled trial with 12 months intervention. The study will be conducted in the Netherlands. We will use a type 2 hybrid implementation-effectiveness design to show proof of concept for effectiveness and implementability using a composite of three objectively measurable dementia risk factors as a composite effectiveness outcome.

Study population

People aged between 50 and 75, of low SES and/or migration background in the Netherlands, with one or more dementia risk factors and in possession of a smartphone are eligible for participation.

Inclusion criteria

To be eligible to participate in this study, a subject must meet all of the following criteria:

- Age ≥ 50 years ≤ 75 years;
- Basic level of literacy in Dutch;
- Possession of a smartphone;
- Migration background; OR Dutch background with low SES, operationalised using self-reported educational attainment in the Netherlands or in the country of origin, specifically defined as i) none, or only primary education', (ii) 'lower vocational or lower secondary education', (iii) 'intermediate vocational or intermediate or higher secondary education' as previously collected in HELIUS.
- Manifest cardiovascular disease, as diagnosed by specialist or general practitioner OR ≥ one dementia risk factors defined as:
    - Hypertension, defined by any of the following:
        - Diagnosis by specialist or general practitioner.
        - Currently on anti-hypertensive drugs.
        - Baseline blood pressure: ≥ 140/90 mmHg;
    - Dyslipidaemia, defined by any of the following:
        - Diagnosis by specialist or general practitioner
        - Use of lipid-lowering drugs
        - Baseline total cholesterol ≥5.0 mmol/L
    - Diabetes mellitus, defined by any of the following:

    - Diagnosis by specialist or general practitioner
    - Use of any blood glucose-lowering medication
- Active smoking (use of any sort of tobacco in any quantity)
- Overweight, defined by any of the following:
    - Body mass index (BMI) ≥ 30
    - Waist circumference men ≥ 102 cm, women ≥ 88 cm
- Lack of physical exercise, defined as below the World Health Organization (WHO) norm (five times a week 30 minutes or a total of 150 minutes per week of intermediate exercise)
- Depression
    - Currently on anti-depressive medication or receiving psychotherapy for depression
    - History of treatment (i.e. drug therapy or psychotherapy) for depression

Exclusion criteria

- Previously diagnosed with dementia by a specialist or general practitioner
- A score below the cut-off score of 21 on the Rowland Universal Dementia Assessment Scale (RUDAS) [1, 2], a validated dementia screening method specifically developed to be less susceptible to cultural, linguistic, and educational biases [3]. The RUDAS is available in multiple languages, including Dutch and Turkish.
- Any condition expected to limit 12 months compliance and follow-up, including metastasised malignancy or other terminal illness
- Any impairment interfering with operation of a smartphone
- Participating in another RCT on lifestyle behavioural change
- Present alcohol or illicit drug abuse impairing study participation.

Sample size calculation

We calculated the sample size based on the expected effect of the intervention on our primary effectiveness outcome (a composite z-score of the objectively measurable risk factors systolic blood pressure, BMI and non-HDL cholesterol) in those with low education as a proxy for low SES. In participants in the HATICE trial with a low educational level, the mean difference between intervention and control participants in change from baseline was 0.107 (pooled SD 0.454) for the composite z-score after 1.5 years. In the preDIVA trial, the difference in composite z-score between people with low SES developing dementia and those who did not during 6-8 years of follow up was 0.096 . We therefore consider a mean difference in change of 0.107 as potentially clinically relevant with respect to the risk of dementia. For specific migrant groups few data are available. With 277

participants per treatment arm, we will have 80% power (with alpha set at 0.05) to detect a 0.107 (pooled SD 0.454) difference on our primary outcome (z-score of systolic blood pressure, BMI and non-HDL cholesterol). To adjust for an anticipated drop-out of 20% in this specific target population we will recruit 692 participants.

Randomisation procedures

Randomization will take place in the app using a computer algorithm in a 1:1 manner. Participants will be informed that they are randomized to one of two different smartphone apps, that can provide support in changing their lifestyle to reduce the risk of dementia. Partners and close family members who also participate in MIND-PRO will automatically be allocated to the same treatment arm to limit contamination. Outcome assessment will be performed by an assessor who is blinded to treatment allocation. The coaches who support participants in behaviour change are, due to the nature of the intervention, not blinded. A certain level of unblinding during outcome assessment is conceivable, since a participant could express details about participation of the MIND-PRO app specifically to the outcome assessor. Participants are therefore instructed to not discuss the intervention with the assessors. If any unblinding occurs, its impact is likely limited since our primary outcome consists solely of objectively measured parameters.

More information about the intervention and the measurements can be found in the trial protocol.

Study outcomes

Primary outcomes

This is a type 2 hybrid implementation-effectiveness randomised controlled trial. There is a primary effectiveness outcome and there are several primary implementation outcomes.

1. Effectiveness: composite score of systolic blood pressure, non-HDL cholesterol, and BMI. We will use the z-score of the difference between baseline and 12 months follow-up values to be analysed as continuous variables, as previously done in the HATICE trial [4].

   Rationale for this outcome: Due to the long time lag between exposure (risk factors) and dementia onset, incident dementia cannot be used as outcome [5]. A dementia risk score can be used as proxy for proof of concept of effectiveness on dementia risk. To avoid reporting bias, we use a composite outcome including three objectively measurable variables from dementia risk scores appropriate to capture the potential effect of our multidomain intervention.

2. Implementation of the mHealth intervention: We will measure multiple aspects of implementation using mixed methods:
    - Coverage: Comparison of characteristics participants with eligible/source population
    - Acceptability: Satisfaction with the application e.g. user-friendliness, credibility, content, complexity (qualitative & quantitative research methods).
    - Adoption: intention, initial decision, or action to try or employ the mHealth intervention (quantitative analysis of the utilisation, usage and uptake of the mHealth intervention)
    - Appropriateness: qualitative analysis of the perceived fit or relevance of the mHealth intervention in the target population.
    - Feasibility: qualitative analysis to what extent the mHealth application can be carried out in a low socio-economic setting and in a population with a migration background.
    - Fidelity: qualitative evaluation of the degree to which the mHealth application is implemented as intended, compared to the original design.
    - Costs: analysis of the implementation costs related to the app and coaching time will be part of a health economic analysis.
    - Sustainability: quantitative evaluation of the extent to which the mHealth application is being used and incorporated during the 12 months of the implementation trial.

Secondary outcomes

The questionnaires used to assess the secondary outcomes are presented in **table 1 below**.

- Change in CAIDE dementia risk score between baseline and follow-up[6] (see Appendix 2)
- Change in individual modifiable components of the CAIDE risk score between baseline and follow-up (i.e. blood pressure, BMI, total cholesterol, physical activity[7])
- Disability[8]
- Depressive symptoms[9]
- Intervention costs
- Cost-effectiveness
- Cognitive functioning, assessed with culturally-sensitive cognitive screening tests, the RUDAS (The Rowland Universal Dementia Assessment Scale) and the Box Task[10].

- Digital measures for social daily functioning measured by BeHapp (remote behavioural monitoring app). Only in a subgroup of participants in the Dutch, Turkish and Surinamese Hindustan groups willing to install the BeHapp app; separate consent will be asked on the informed consent form.

Domain	**Questionnaire(s)**
1. Physical exercise	Self- administered short International physical activity questionnaire (IPAQ-SF)
2. Disability	WHO Disability Assessment Schedule 2.0 (WHODAS 2.0, 12-item)
3. Depressive symptoms	Geriatric Depression Scale 15-item (GDS-15)

Table 1. Self-assessment questionnaires

4. Statistical Analysis

General

Prior to analysis, all data will be checked for missing values and miscoding, and univariate analyses will be performed to compare the distribution of variables and to identify abnormalities/outliers.

Primary and secondary outcomes will be analysed according to the *intention to treat* principle for all participants who underwent baseline assessment and subsequent randomization and with available outcome data.

Adaptation dd 5-11-2026: We modified the linear mixed-effects models by including a random effect to account for clustering among partners or first degree relatives, and by replacing the random effect for ethnicity with a fixed effect.

Rationale for including a random effect for partners or first degree relatives
Partners and first-degree relatives were randomised to the same intervention group to avoid contamination, resulting in small clusters of participants. As observations within these clusters may be correlated, ignoring this clustering could lead to underestimated standard errors and overly narrow confidence intervals. Therefore, a random effect at the cluster level will be included in the linear mixed-effects models, in accordance with the principles outlined in the CONSORT extension for cluster randomised trials [14].

Rationale for removing the random effect for ethnicity while incorporating ethnicity as a fixed effect
We previously intended to use a linear mixed-effects models with individual observations nested within ethnic groups, operationalized by a random intercept and/or slope (dependent on model fit). However, given that the study includes only three ethnic groups, estimating random effects may lead to unstable variance estimates and singularity. Therefore, we decided to include ethnicity as a fixed rather than a random effect in the models, in line with recommendations for analyses involving very few groups (CASRAI guidelines) (https://casrai.org/guides/mixed-effects-models-random-vs-fixed-effects). We have changed the analyses plan below accordingly using track changes for transparency.

Primary outcomes

The primary effectiveness outcome is the difference between the composite Z score after 12 months and the baseline composite Z score ($(Z_{SBP} + Z_{non\text{-}HDL\text{-}cholesterol} + Z_{BMI})/3$). The baseline mean and SD will be used to calculate composite Z scores at both time points to be able to detect a change. For the

primary analyses we will use a linear mixed effects model with the difference in Z score as outcome, ~~and~~ adjusted for baseline Z score and ethnic group as fixed effect, and partner pairs as random intercept. Participants who are not part of a cluster of two participants will each be assigned a unique group identifier, so that they are treated as independent observations (clusters of size one), while the partner pairs contribute to estimation of the intraclass correlation~~, with individual observations nested within ethnic groups operationalized by a random intercept and/or slope (dependent on best model fit for primary outcome).~~

Implementation of the intervention will be analysed with graphical portray of the Likert scales for the questionnaires and we will perform deductive thematic analyses for the in-depth interviews.

Sensitivity analyses

Sensitivity analyses will only be done for effect of the intervention on the composite Z score of SBP, non-HDL cholesterol and BMI.
If needed, we will explore the impact of any baseline imbalances by including imbalanced factors as covariate in a sensitivity analysis. A per protocol analysis will be done, based on uptake within the first month and continued use during the whole intervention period. Finally, we will conduct a sensitivity analysis using multiple imputation.

Predefined subgroup analyses

Exploratory subgroup analyses will only be done for effect of the intervention on the composite Z score of SBP, non-HDL cholesterol and BMI.
Exploratory separate analyses on the effect of the intervention on the primary outcome will be performed for ethnic group, sex, age group, history of CVD and diabetes. Interaction terms will be included to test for between-subgroup differences in intervention effects. A p value of 0.10 is regarded significant interaction.

Secondary outcomes

The effect on the difference between baseline and 12M in the CAIDE dementia risk score and its individual components (i.e. blood pressure, BMI, non-HDL cholesterol, physical activity) will be analysed using the same model as for the primary analysis*.

In addition, as an exploratory analysis in the Dutch, Hindustan and Turkish groups, the effect on daily changes in general movement will be assessed based on the data collected with the BeHapp application[11-13].
Scales for disability, depressive symptoms and cognitive functioning, mostly ordinal, will also be analyzed as linear scales for the difference between baseline and 12M* if the data characteristics allow. Poisson regression or zero-inflated models may be applied to distributions resembling count or zero-inflated data.

A description of the cost-effectiveness analysis will be given in a separate analysis plan.

* The same model as used for the primary analysis will be employed for linear secondary outcomes: a linear mixed effects model with the difference between M12 and baseline as outcome, ~~and~~ adjusted for baseline value of the outcome and ethnic group as fixed effect, and partner pairs as random intercept. Participants who are not part of a partner group will each be assigned a unique group identifier, so that they are treated as independent observations (clusters of size one), while the partner pairs contribute to estimation of the intraclass correlation.
~~, with individual observations nested within ethnic groups operationalized by a random intercept and/or slope (dependent on best model fit for primary outcome).~~

References

1. Storey, J.E., et al., *The Rowland Universal Dementia Assessment Scale (RUDAS): a multicultural cognitive assessment scale.* Int Psychogeriatr, 2004. **16**(1): p. 13-31.
2. Goudsmit, M., et al., *One Size Does Not Fit All: Comparative Diagnostic Accuracy of the Rowland Universal Dementia Assessment Scale and the Mini Mental State Examination in a Memory Clinic Population with Very Low Education.* Dement Geriatr Cogn Dis Extra, 2018. **8**(2): p. 290-305.
3. Nielsen, T.R. and K. Jørgensen, *Cross-cultural dementia screening using the Rowland Universal Dementia Assessment Scale: a systematic review and meta-analysis.* Int Psychogeriatr, 2020. **32**(9): p. 1031-1044.
4. Richard, E., et al., *Healthy ageing through internet counselling in the elderly (HATICE): a multinational, randomised controlled trial.* Lancet Digit Health, 2019. **1**(8): p. e424-e434.
5. Montero-Odasso, M., et al., *Multidomain trials to prevent dementia: addressing methodological challenges.* Alzheimers Res Ther, 2022. **14**(1): p. 94.
6. Kivipelto, M., et al., *Risk score for the prediction of dementia risk in 20 years among middle aged people: a longitudinal, population-based study.* Lancet Neurol, 2006. **5**(9): p. 735-41.
7. Meh, K., et al., *Validity and reliability of IPAQ-SF and GPAQ for assessing sedentary behaviour in adults in the European Union: A systematic review and meta-analysis.* International journal of environmental research and public health, 2021. **18**(9): p. 4602.
8. Üstün, T.B., *Measuring health and disability: Manual for WHO disability assessment schedule WHODAS 2.0*. 2010: World Health Organization.
9. Sheikh, J.I. and J.A. Yesavage, *Geriatric Depression Scale (GDS): recent evidence and development of a shorter version*, in *Clinical gerontology*. 2014, Routledge. p. 165-173.
10. Kessels, R.P. and A. Postma, *The Box Task: A tool to design experiments for assessing visuospatial working memory.* Behavior Research Methods, 2018. **50**(5): p. 1981-1987.
11. Jagesar, R.R., J.A. Vorstman, and M.J. Kas, *Requirements and Operational Guidelines for Secure and Sustainable Digital Phenotyping: Design and Development Study.* J Med Internet Res, 2021. **23**(4): p. e20996.
12. Jongs, N., et al., *A framework for assessing neuropsychiatric phenotypes by using smartphone-based location data.* Transl Psychiatry, 2020. **10**(1): p. 211.
13. Eskes, P., et al., *The sociability score: App-based social profiling from a healthcare perspective.* Computers in Human Behavior, 2016. **59**: p. 39-48.
14. Campbell MK, Piaggio G, Elbourne DR, Altman DG. *Consort 2010 statement: extension to cluster randomised trials.* BMJ 2012;345:e5661.