Statistical Analysis Plan

Caregiver-administered acupressure as long-term intervention to slow the progression of early/mild Alzheimer's disease with blood amyloid and tau biomarkers: a randomized controlled trial

Affiliated institution:

School of Chinese Medicine, LKS Faculty of Medicine, The University of Hong Kong

Principal Investigator
Zhang-Jin Zhang, MD, PhD
School of Chinese Medicine, LKS Faculty of Medicine, The University of Hong Kong
Tel no.: (+852)3917-6445
E-mail: zhangzj@hku.hk

Abbreviations

Abbreviations: AD, Alzheimer's disease; AEs, adverse events; Aβ, amyloid; CAE, Comfy Acupressure for the Elderly; DSF, Digit Span Forward; DSR, Digit Span Reverse; EQ-5D-5L, 5-level EQ-5D version; EQ VAS, EQ visual analogue scale; GDS-15, 15-item Geriatric Depression Scale; GFAP, glial fibrillary acidic protein; HK-MoCA, Montreal Cognitive Assessment, Hong Kong version; IADL-CV, Chinese Lawton Instrumental Activities of Daily Living Scale; LAC, least acupressure control; MBI-C, Chinese version of the Modified Barthel Index; NfL, neurofilament light chain protein; PSQI, Pittsburgh Sleep Quality Index; QALYs, quality-adjusted life years; QoL-AD, Quality of Life in Alzheimer's Disease; SD, standard deviation,

Last edited and approved: 23 October 2026

Introduction

This statistical analysis plan (SAP) describes the planned analyses to be performed on data collected in the clinical trial titled "Caregiver-administered acupressure as long-term intervention to slow the progression of early/mild Alzheimer's disease with blood amyloid and tau biomarkers: a randomized controlled trial". No statistical analyses will be performed until the final version of this SAP has been approved.

Study design

It is an assessor-blinded, randomized, LAC-controlled trial. A total of 188 eligible subjects will be recruited. They will be randomly assigned to either a "Comfy Acupressure for the Elderly (CAE)" group or a "least acupressure control (LAC)" control group, with 94 subjects per group for 12 months.

Hypothesis and aims

The hypotheses of this trial are:
(a) the 12-month CAE early long-term intervention could produce greater effects in slowing the progression of Alzheimer's Disease (AD) by improving cognitive deterioration, blood amyloid and tau pathological products, functional dependency, and comorbid symptoms in older adults living with early/mild AD compared to LAC;
(b) the CAE intervention is cost-effective compared with LAC.
The aims of this study are:
(a) to compare the effectiveness of CAE and LAC 12-month interventions in improving cognition, functional dependency, comorbid symptoms, blood amyloid and tau biomarkers in older adults living with early/mild AD;
(b) to examine whether the 12-month CAE intervention is cost-effective compared with LAC by conducting a nested economic evaluation using health-related quality of life measured by the 5-level EQ-5D version (EQ-5D-5L) to derive quality-adjusted life years (QALYs).

Patient population

Older adults aged 60-85 years older adults with early/mild AD are eligible for enrolment.

Inclusion criteria

Subjects will be eligible for this study if they:
(a) are aged 15-22 years;
(b) have at least one of the following abnormal plasma amyloid and tau levels: Aβ42/Aβ40 ratio ≤ 0.17, p-tau217 > 0.15 pg/ml, or/and p-tau181 > 3.3 pg/ml;[1-4]
(c) have mild neurocognitive disorder with the HK-MoCA in classifying mild neurocognitive disorder, in which mild neurocognitive disorder is defined as those whose HK-MoCA score falls between the 16th and 2nd percentile of his/her peers adjusted for age and education.[5-6]

Exclusion criteria

Subjects will be excluded if they:
(a) show moderate-to-severe dementia, as evidenced by a HK-MoCA score falling below the 2nd percentile of his/her peers adjusted for age and education;
(b) cannot proficiently communicate due to language function impairment;
(c) have severe skin lesions on acupressure areas;
(d) had a surgery on the head or neck;
(e) have a medical condition that is a contraindication for acupressure.

Randomization, central allocation and blindness

To minimize the imbalance of participant groups, stratified randomization will be conducted with age, care centre, and baseline HK-MoCA score. Random codes will be produced in advance using simple, complete, non-sequential random codes with a block of 4.

One independent research assistant will be in charge of randomization and central allocation. After confirmation of patients' eligibility and completion of the baseline assessment, participants will be randomly assigned to CAE or LAC groups in a ratio of 1:1. Randomization information will be sealed in sequentially numbered opaque envelopes.

Central allocation will be conducted, i.e., site investigators who are responsible for subjects' eligibility will obtain the opaque envelope from the central coordinator for an eligible subject who will be allocated to one of the two groups. Each envelope will be opened by acupuncturists after the participant completes the baseline assessment. Clinical assessors and data analysts will be blind to patients' treatment.

LAC is applied to the control group to control "placebo" effects of CAE and increase study adherence of participants of the control group, based on the following two principles: (1) the acupoints chosen are unrelated or less related to the treated syndromes according to TCM theory; and (2) the number of acupoints used, intensity, and duration of acupressure are kept to the least level at which the therapeutic effects produced are minimal, but participants are aware of receiving active acupressure.

Intervention

(a) CAE regimen: Subjects assigned to the CAE group will receive CAE intervention, 3 sessions per week for 12 consecutive months (156 weeks). Details of the intervention procedures can be found in the study protocol.
(b) LAC regimen: Subjects assigned to the LAC group will receive LAC intervention, 3 sessions per week for 12 consecutive months (156 weeks). Details of the intervention procedures can be found in the study protocol.

Study outcomes

Primary outcome

The primary outcome will be evaluated by the mean change of the Montreal Cognitive Assessment (MoCA) score between baseline and endpoint. The MoCA test assesses seven domains of cognitive function, including visuospatial, naming, attention, language, abstraction, recall, and orientation domains, for a total possible score of 30 points.[6-7] Assessments will be conducted at baseline and thereafter once bimonthly.

Secondary outcomes

Clinical assessments will be conducted at baseline and thereafter once bimonthly. Biochemical assessments be conducted at baseline, 6 and 12 months.
The secondary outcomes will include:
(a) the Digit span test for attentional function, short-term memory, and working memory, including the Digit Span Forward (DSF) and the Digit Span Reverse (DSR);[8]
(b) the Quality of Life in Alzheimer's Disease (QoL-AD) scale for quality of life;[9-11]
(c) the Chinese Lawton Instrumental Activities of Daily Living Scale (IADL-CV) for functional independence;[12-13]
(d) the Chinese version of the Modified Barthel Index (MBI-C) for functional independence;[14-15]
(e) the 15-item Geriatric Depression Scale (GDS-15) for depression;[16-17]
(f) the Pittsburgh Sleep Quality Index (PSQI) for sleep;[18]
(g) the 5-level EQ-5D version (EQ-5D-5L) for health-related quality of life;[19]
(h) plasma Aβ42/Aβ40 ratio;
(i) plasma p-tau181 level;
(j) plasma p-tau217 level;
(k) plasma neurofilament light chain protein (NfL) level; and
(l) plasma glial fibrillary acidic protein (GFAP) level.

Safety outcomes

The severity of adverse events (AEs) will be assessed according to the Common Terminology Criteria for Adverse Events (CTCAE) v5.0 criteria.[20] AEs will be systematically recorded, in which AEs that first appear in the study or worsen relative to the pre-study status will be recorded at each visit, including date and time of onset, duration, severity, relationship to intervention, and action taken accordingly. The causality between treatment procedures and AEs will be assessed.

Sample size determination

Clinical and biomarker parameters are used and compared in the estimation of sample size. The greater number will be applied in the study.

Clinical parameter: In our aforesaid trial, CAE treatment for 12 weeks produced a 1.67 ± 3.63 (standard deviation, SD) score greater improvement than the routine care control on HK-MoCA in subjects with mild cognitive impairment and mild dementia. We conservatively assume that the CAE 12-month intervention could yield at least 1.67 ± 3.63 (SD) greater improvement on HK-MoCA than LAC. The following formula is then used to calculate the sample size,

$$n = \frac{2}{(M/SD)^2} \times C_{p,power}$$

where n is the number of subjects required in each arm; M is the mean difference that is equal to 1.67. SD is standard deviation, which is 3.63. $C_{p,power}$ is equal to 7.9 when the two-tailed level of α and power (1-β) are set at 0.05 and 80%, respectively. It requires 75 subjects per group.

Biomarker parameter: To ensure that the power yielded from the sample size obtained in the aforesaid clinical parameter is sufficient to detect the Aβ/tau-associated mechanisms, we separately calculate and compare it with the sample size using the biomarker parameter. Due to the absence of studies on the effects of acupressure/manual therapy on AD-related biomarkers, we refer to data obtained from a placebo-controlled trial of aducanumab, an approved monoclonal antibody on plasma p-tau181 in patients with early AD.[21] For a conservative estimate, we use the results of low doses (3 or 6 mg/kg). The low dose of aducanumab once every 4 weeks for 76 weeks reduced the plasma p-tau181 level by an average of 1.02 pg/ml (SD, 0.54 pg/ml) versus placebo. At the most conservative point, we assume that one-year acupressure produced only 1/4 of the effects of the low-dose aducanumab on plasma p-tau181. M (mean difference) = 0.26 pg/ml and SD = 0.54 pg/ml are then obtained and applied to the above formula, with a two-tailed level of α = 0.05 and power (1-β) = 80%. 69 subjects per group are needed, confirming that the sample size required to detect the Aβ/tau mechanisms is less than that needed to detect the clinical difference. The greater number of 75 subjects per group is hence applied for the further estimation.
With a sample of 75 participants per group after accounting for anticipated dropout, α = 0.05 (two-tailed) and 80% power are adequately powered to detect a minimum between-group difference of 0.25 pg/ml in plasma p-tau181 levels (SD = 0.54 pg/ml), equivalent to approximately 25% of the effect size observed with low-dose aducanumab in the reference trial.[21] This effect size corresponds to Cohen's d $\approx$ 0.46, indicating a medium effect, suggesting the current sample size provides sufficient sensitivity for exploratory detection of biologically meaningful changes in AD-related plasma biomarkers. Furthermore, using Fisher's r-to-z transformation with two-tailed α = 0.05 and 80% statistical power, a sample size of 75 participants per group (150 total) is sufficient to detect a minimum correlation of r $\approx$ 0.23 between changes in blood biomarkers and changes in cognitive outcomes. This detection threshold is lower than the range of effect sizes reported in the existing literature for associations between biomarkers and cognitive function. Specifically, it was found that longitudinal changes in plasma p-tau181 were associated with cognitive decline (cognitively impaired: r = 0.34, P < .001),[22] and plasma levels of GFAP, p-tau181, and NfL were negatively correlated with MoCA/ Mini-Mental State Examination scores (r = 0.266 - 0.325, p $\leqslant$ 0.001).[23] Therefore, the current sample size is sufficient to examine the inter-correlation between blood biomarkers and changes in cognitive outcomes. As no studies have yet investigated the effects of acupressure or manual therapy on AD-related biomarkers, this biomarker analysis will also serve as an exploratory investigation.

Economic evaluation**:** The economic evaluation is a secondary, exploratory component of the trial; the sample size is determined by the primary clinical outcome and is not separately powered to detect differences in cost-effectiveness.

Discontinuation rate: Although no data are available for estimation of the discontinuation rate of long-term (one-year) acupressure intervention, we assume a discontinuation rate of 20% in

the proposed study based on the following elaboration. Our preliminary pilot study showed only a 10.2% (12/118) dropout rate for caregiver-administered acupressure for 12 weeks in elderly with moderate and severe dementia. One 6-month self-administered acupressure study on cognitive functions in older adults with mild cognitive impairment reported a 16% dropout rate.[240] Our three previous studies showed even lower dropout rates (7.0-9.7%) in self-administered acupressure for the treatment of stress disorders, insomnia, and chemotherapy-induced peripheral neuropathy for 8-16 weeks.[25-27] According to our experience in previous studies, compared to self-administered acupressure, caregiver-administered performed acupressure (such as CAE) was more acceptable and comfortable. Moreover, the CAE will be conducted in care centres rather than at home. This will largely increase adherence. Therefore, a 20% dropout rate over one year is realistic. A total of 188 subjects (n = 94 each group) are then recruited in the consideration of this dropout rate.

Sample size for stratified randomization: The 188 participants proposed for recruitment are distributed across 8 strata defined by 2 age groups (60–74 years; 75–85 years) × 2 care setting types (care and attention homes for the elderly and elderly activity centres) × 2 HK-MoCA score percentile groups (16th–9th; <9th–2nd), yielding approximately 23–24 participants per stratum, substantially exceeding the minimum of 4 required to complete at least one block per stratum (block size = 4). Stratified randomization improves statistical efficiency by ensuring balance across key prognostic factors (design effect $\leqslant 1$), such that the standard two-group formula already yields a conservative estimate and no upward adjustment is warranted. Although a downward adjustment is theoretically permissible, it seems not to be adopted in practice, as the required reduction depends on the correlation between stratification factors and the primary outcome, a parameter that cannot be reliably estimated prior to trial initiation, and any overestimation would risk insufficient statistical power. Therefore, 94 subjects per group (188 in total) remain valid and appropriate for the proposed stratified randomization design.

Data processing and analysis

General principles

One biostatistician who is blinded to interventions will be responsible for statistical analysis. The analysis will be carried out on the intention-to-treat (ITT) population. For all primary and secondary outcomes, two-tailed P-values will be reported in addition to confidence intervals. The nominal level of statistical significance (1-α) will be 5%. The analysis and reporting of the results will follow the Consolidated Standards of Reporting Trials (CONSORT) guidelines.

Analysis software

All analyses will be performed using R Studio version 4.0.0 or later.

Analysis Populations

All analyses will be carried out on the intention-to-treat (ITT) population, for which participants have been randomized, completed the baseline assessment regardless of protocol compliance.[28] The flow of patients throughout the trial will be shown using a CONSORT

diagram (Figure 1).

Demographic characteristics and baseline comparisons

Baseline comparisons will be tabulated for the following variables (Proposed Table 1):
(a) Age, years
(b) Gender (Categories: male / female)
(c) Marital status (Categories: married or living with partner / single, separated, divorced, or widowed)
(d) Education level (Categories: uneducated / primary or below / lower secondary / upper secondary / university or above)
(e) Past occupation (Categories: managers and administrators / professional and associate professional / skilled and semi-skilled worker / unskilled worker / housework / others)
(f) Diagnosis of AD (Categories: Yes / No)
(g) Duration of AD
(h) Past medical history (Categories: cardiometabolic disease / malignant neoplasm / stroke / Parkinson's disease / psychotic disorder / affective disorder)
(i) Current medication (Categories: cholinesterase inhibitors / glutamate regulators / antidepressants / antipsychotics / hypnotics)
(j) Baseline HK-MoCA score
(k) Baseline DSF score
(l) Baseline DSR score
(m) Baseline QoL-AD score
(n) Baseline IADL-CV score
(o) Baseline MBI-C score
(p) Baseline GDS-15 score
(q) Baseline PSQI score
(r) Baseline EQ-5D-5L score
(s) Baseline EQ visual analogue scale (EQ VAS) score
(t) Plasma Aβ42/Aβ40 ratio
(u) Plasma p-tau217 level, pg/mL
(v) Plasma p-tau181 level, pg/mL
(w) Plasma neurofilament light chain protein (NfL) level, pg/mL
(x) Plasma glial fibrillary acidic protein (GFAP) level, pg/mL

Continuous variables will be presented using mean with standard deviation (SD). Independent-samples t-test will be used to detect differences in continuous baseline variables.

Categorical variables will be presented by counts with percentages. Percentages will be calculated according to the number of patients for whom data are available. Categorical variables will be analyzed using the Chi-square test or Fisher's exact test.

Discontinuation will be summarized by counts with percentages. Chi-square test or Fisher's exact test will be used to detect differences between the two groups.

Primary outcome

A linear mixed-effect model will be applied to compare the primary outcome (changes in total HK-MoCA score from baseline), with time (baseline, month 2, month 4, month 6,

month 8, month 10, and month 12) and group (CAE and LAC) as categorical fixed factors and random intercepts within a scaled identity covariance matrix. Gender, age, baseline HK-MoCA, baseline amyloid and tau levels, medication, routine care, and study sites will serve as covariates. Two-tailed P-values will be reported in addition to confidence intervals (Proposed Tables 2).

Subgroup analysis will be further conducted to compare outcomes to detect whether these subgroup factors are associated with outcomes. The following subgroups include:
(a) Gender (female and male)
(b) Age (60–74 years and 75–85 years)
(c) Care setting types (care and attention homes for the elderly and elderly activity centres)
(d) Baseline HK-MoCA (16th–9th percentile group and <9th–2nd percentile group)

Effect modification will be assessed by the treatment-by-time-by-subgroup interaction, tested by an F test with Satterthwaite degrees of freedom; subgroup-specific estimates of the between-arm difference in HK-MoCA change from baseline to month 12 are obtained from the same interaction model, so that all strata share a common covariate set. Interaction P values are reported without adjustment for multiple comparisons and are therefore nominal, whereas post hoc pairwise interaction tests within the subgroups will be compared against a Bonferroni-adjusted threshold of 0.0125 for four comparisons.

Secondary outcomes

A linear mixed-effect model will be applied to compare the secondary outcomes, including changes in scores of the HK-MoCA domains, DSF, DSR, QoL-AD, EQ-5D-5L, MBI-C, GDS-15, and PSQI, with time (baseline, month 2, month 4, month 6, month 8, month 10, and month 12) and group (CAE and LAC) as categorical fixed factors and random intercepts within a scaled identity covariance matrix. Gender, age, baseline HK-MoCA, baseline amyloid and tau levels, medication, routine care, and study sites will serve as covariates. Two-tailed P-values will be reported in addition to confidence intervals (Proposed Tables 2).

Two-way repeated measures analysis of variance (ANOVA) will be used to detect significant differences in blood biomarkers (Aβ42/Aβ40 ratio, p-tau181, p-tau217, NfL and GFAP) between the two groups. Two-tailed P-values will be reported in addition to confidence intervals (Proposed Tables 3). Linear regression will be conducted to examine inter-correlations between clinical outcomes and biomarker levels.

Safety outcomes

All AEs for frequency threshold equal to or exceeding 5% that occurred throughout the trial will be tabulated by treatment group and AE category (mild, moderate, severe, life-threatening, or death). Chi-square test or Fisher's exact test will be used to detect differences between the two groups (Proposed Table 4).

Details of serious AEs and death cases will be summarized and listed in a separate table.

Economic evaluation

A within-trial cost-utility analysis will be conducted from the perspective of the healthcare

payer. QALYs will be estimated by the area-under-the-curve method from EQ-5D-5L utility values measured at the seven assessment time points, with baseline utility included as a covariate; no discounting will be applied given the 12-month time horizon. The cost of delivering CAE and LAC will be estimated by micro-costing, using the treatment logs to obtain, for each participant, the number of sessions completed, the duration of each session, and the category of staff delivering it, valued at the corresponding salary rates inclusive of on-costs; the costs of training workshops and supervision will be apportioned across participants. Participants' use of health and social care services in the preceding 6 months will additionally be recorded at baseline, month 6, and month 12 from the service records maintained by the participating centres, verified with the participant, their caregiver, or centre staff by the research assistant, and valued using published Hong Kong unit costs. Costs arising solely from research procedures, including blood sampling and biomarker assays, will be excluded, as they would not be incurred if the intervention were delivered in routine practice. The incremental cost-effectiveness ratio, expressed as the incremental cost per QALY gained, will be calculated, and uncertainty will be characterized by non-parametric bootstrapping, with results presented on the cost-effectiveness plane and as cost-effectiveness acceptability curves. Missing utility and cost data will be handled by multiple imputations.

Interim analysis

No interim outcome analyses were done prior to the completion of the study to avoid assessment bias.

Missing data

A linear mixed-effect model will be conducted for all analyses of primary and secondary outcomes without imputation of missing data.[29]

Sensitivity analysis

Three sensitivity analyses will be performed. First, multiple imputation is undertaken by fully conditional specification with predictive mean matching (five donors), generating 50 imputed datasets over 20 iterations; for each outcome the imputation model included treatment allocation, that outcome's scores at all five timepoints, all covariates from the primary analysis model, and the baseline characteristics associated with attrition; the primary model is refitted in each imputed dataset and estimates combined using Rubin's rules with the Barnard–Rubin small-sample correction to the degrees of freedom. Second, a delta-adjustment (pattern-mixture) analysis examines departures from the missing-at-random assumption, in which imputed follow-up values in the intervention group are shifted in the unfavourable direction in increments of 0.1 baseline standard deviations, with the tipping point defined as the smallest shift at which a between-group difference lost statistical significance. Third, the analyses are repeated in the per-protocol population, defined as participants who completed the allocated intervention and received at least 75% of the treatment sessions scheduled during their period of participation.

Proposed Tables

Table 1. Baseline characteristics

Characteristic	CAE (n=94)	LAC (n=94)	Total (n=188)	*P* value
Age, years [a]	XX ± XX	XX ± XX	XX ± XX	XX
Gender (Female) [b]	XX (XX)	XX (XX)	XX (XX)	XX
Marital status [b]				XX
Married or living with partner	XX (XX)	XX (XX)	XX (XX)	
Single, separated, divorced, or widowed	XX (XX)	XX (XX)	XX (XX)	
Education level [b]				XX
Uneducated	XX (XX)	XX (XX)	XX (XX)	
Primary or below	XX (XX)	XX (XX)	XX (XX)	
Lower Secondary	XX (XX)	XX (XX)	XX (XX)	
Upper Secondary	XX (XX)	XX (XX)	XX (XX)	
University or above	XX (XX)	XX (XX)	XX (XX)	
Past occupation [b]				XX
Managers and administrators	XX (XX)	XX (XX)	XX (XX)	
Professional and associate professional	XX (XX)	XX (XX)	XX (XX)	
Skilled and semi-skilled worker				
Unskilled worker	XX (XX)	XX (XX)	XX (XX)	
Housework	XX (XX)	XX (XX)	XX (XX)	
Others / no information	XX (XX)	XX (XX)	XX (XX)	
Have previous diagnosis of AD [b]	XX (XX)	XX (XX)	XX (XX)	XX
Duration of AD, months [a]	XX ± XX	XX ± XX	XX ± XX	XX
Past medical history [b]				XX
Cardiometabolic disease	XX (XX)	XX (XX)	XX (XX)	
Malignant neoplasm	XX (XX)	XX (XX)	XX (XX)	
Stroke				
Parkinson's disease	XX (XX)	XX (XX)	XX (XX)	
Psychotic disorder	XX (XX)	XX (XX)	XX (XX)	
Affective disorder	XX (XX)	XX (XX)	XX (XX)	
Past medical history [b]				XX
Cholinesterase inhibitors	XX (XX)	XX (XX)	XX (XX)	
Glutamate regulators	XX (XX)	XX (XX)	XX (XX)	
Antidepressants				
Antipsychotics	XX (XX)	XX (XX)	XX (XX)	
Hypnotics	XX (XX)	XX (XX)	XX (XX)	
HK-MoCA [a]	XX ± XX	XX ± XX	XX ± XX	XX
DSF [a]	XX ± XX	XX ± XX	XX ± XX	XX
DSR [a]	XX ± XX	XX ± XX	XX ± XX	XX
QoL-AD [a]	XX ± XX	XX ± XX	XX ± XX	XX
IADL-CV [a]	XX ± XX	XX ± XX	XX ± XX	XX

Characteristic	CAE (n=94)	LAC (n=94)	Total (n=188)	*P* value
MBI-C [a]	XX ± XX	XX ± XX	XX ± XX	XX
GDS-15 [a]	XX ± XX	XX ± XX	XX ± XX	XX
PSQI [a]	XX ± XX	XX ± XX	XX ± XX	XX
EQ-5D-5L [a]	XX ± XX	XX ± XX	XX ± XX	XX
EQ VAS [a]	XX ± XX	XX ± XX	XX ± XX	XX
Plasma Aβ42/Aβ40 ratio	XX ± XX	XX ± XX	XX ± XX	XX
Plasma p-tau217 level, pg/mL [a]	XX ± XX	XX ± XX	XX ± XX	XX
Plasma p-tau181 level, pg/mL [a]	XX ± XX	XX ± XX	XX ± XX	XX
Plasma NfL level, pg/mL [a]	XX ± XX	XX ± XX	XX ± XX	XX
Plasma GFAP level, pg/mL [a]	XX ± XX	XX ± XX	XX ± XX	XX

Abbreviations: AD, Alzheimer's Disease; Aβ, amyloid; CAE, Comfy Acupressure for the Elderly; DSF, Digit Span Forward; DSR, Digit Span Reverse; EQ-5D-5L, 5-level EQ-5D version; EQ VAS, EQ visual analogue scale; GDS-15, 15-item Geriatric Depression Scale; GFAP, glial fibrillary acidic protein; HK-MoCA, Montreal Cognitive Assessment, Hong Kong version; IADL-CV, Chinese Lawton Instrumental Activities of Daily Living Scale; LAC, least acupressure control; MBI-C, Chinese version of the Modified Barthel Index; NfL, neurofilament light chain protein; PSQI, Pittsburgh Sleep Quality Index; QoL-AD, Quality of Life in Alzheimer's Disease

[a] Continuous data are expressed as mean ± SD and were examined using independent samples t-test. Those which statistical differences $P<0.05$ reached the significance level are indicated in bold font.

[b] Categorical data are expressed as count (percentage) and were examined using Chi-square (χ^2) or Fisher Exact test. Those which statistical differences $P<0.05$ reached the significance level are indicated in bold font.

Table 2. Primary and secondary outcomes of the clinical assessment

	Change from baseline					
	CAE (n=94)		LAC (n=94)		Analysis (Between group)	
Outcomes	Mean (95%CI)	Within-group *P* value [c]	Mean (95%CI)	Within-group *P* value [c]	Difference Mean (95%CI)	*P* value [d]
Primary outcome						
HK-MoCA [a]						
Month-2	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-4	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-6	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-8	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-10	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-12	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Secondary outcomes						
DSF [a]						
Month-2	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-4	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-6	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-8	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-10	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-12	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
DSR [a]						
Month-2	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-4	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-6	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-8	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-10	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-12	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
QoL-AD [a]						
Month-2	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-4	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-6	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-8	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-10	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-12	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
IADL-CV [a]						
Month-2	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-4	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-6	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-8	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-10	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-12	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
MBI-C [a]						
Month-2	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-4	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-6	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-8	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-10	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-12	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
GDS-15 [b]						
Month-2	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-4	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-6	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-8	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-10	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-12	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX

	Change from baseline					
	CAE (n=94)		LAC (n=94)		Analysis (Between group)	
Outcomes	**Mean (95%CI)**	**Within-group *P* value [c]**	**Mean (95%CI)**	**Within-group *P* value [c]**	**Difference Mean (95%CI)**	***P* value [d]**
PSQI [b]						
Month-2	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-4	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-6	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-8	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-10	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-12	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
EQ-5D-5L [a]						
Month-2	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-4	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-6	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-8	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-10	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-12	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
EQ VAS [a]						
Month-2	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-4	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-6	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-8	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-10	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX
Month-12	XX (XX, XX)	XX	XX (XX, XX)	XX	XX (XX, XX)	XX

Abbreviations: CAE, Comfy Acupressure for the Elderly; DSF, Digit Span Forward; DSR, Digit Span Reverse; EQ-5D-5L, 5-level EQ-5D version; EQ VAS, EQ visual analogue scale; GDS-15, 15-item Geriatric Depression Scale; HK-MoCA, Montreal Cognitive Assessment, Hong Kong version; IADL-CV, Chinese Lawton Instrumental Activities of Daily Living Scale; LAC, least acupressure control; MBI-C, Chinese version of the Modified Barthel Index; PSQI, Pittsburgh Sleep Quality Index; QoL-AD, Quality of Life in Alzheimer's Disease.

[a] A greater positive value represents improvement in symptoms.

[b] A greater negative value represents improvement in symptoms.

[c] Data was represented as change in the scores from baseline. *P* value was calculated using a mixed-effects model with baseline adjustment to illustrate pre- and post-treatment within-group differences. Those which statistical differences $P<0.05$ reached significance level are indicated in bold font.

[d] *P* value was calculated using a mixed-effects model with baseline adjustment to illustrate between-group differences. Those which statistical differences $P<0.05$ reached significance level are indicated in bold font.

Table 3. Biochemical assays

	Change from baseline			
	CAE (n=94)	LAC (n=94)	Analysis (Between group)	
Outcomes	**Mean (95%CI)**	**Mean (95%CI)**	**Difference Mean (95%CI)**	***P* value [c]**
Aβ42/Aβ40 ratio [a]				
Month-6	XX (XX, XX)	XX (XX, XX)	XX (XX, XX)	XX
Month-12	XX (XX, XX)	XX (XX, XX)	XX (XX, XX)	XX
p-tau217, pg/mL [b]				
Month-6	XX (XX, XX)	XX (XX, XX)	XX (XX, XX)	XX
Month-12	XX (XX, XX)	XX (XX, XX)	XX (XX, XX)	XX
p-tau181, pg/mL [b]				
Month-6	XX (XX, XX)	XX (XX, XX)	XX (XX, XX)	XX
Month-12	XX (XX, XX)	XX (XX, XX)	XX (XX, XX)	XX
NfL, pg/mL [b]				
Month-6	XX (XX, XX)	XX (XX, XX)	XX (XX, XX)	XX
Month-12	XX (XX, XX)	XX (XX, XX)	XX (XX, XX)	XX
GFAP, pg/mL [b]				
Month-6	XX (XX, XX)	XX (XX, XX)	XX (XX, XX)	XX
Month-12	XX (XX, XX)	XX (XX, XX)	XX (XX, XX)	XX

Abbreviations: Aβ, amyloid; GFAP, glial fibrillary acidic protein; NfL, neurofilament light chain protein.

[a] A greater positive value represents improvement in symptoms.

[b] A greater negative value represents improvement in symptoms.

[d] *P* value was calculated using two-way repeated measures analysis of variance (ANOVA) to illustrate between-group differences. Those which statistical differences $P<0.05$ reached significance level are indicated in bold font.

Table 4. Reported adverse events in different categories, n (%) [a]

Adverse Event Category	CAE (n=94)	LAC (n=94)	P-value [a]
Any	XX (XX)	XX (XX)	XX
Mild	XX (XX)	XX (XX)	XX
Event 1	XX (XX)	XX (XX)	
Moderate	XX (XX)	XX (XX)	XX
Event 1	XX (XX)	XX (XX)	
Severe	XX (XX)	XX (XX)	XX
Event 1	XX (XX)	XX (XX)	
Life-threatening	XX (XX)	XX (XX)	XX
Event 1	XX (XX)	XX (XX)	
Death	XX (XX)	XX (XX)	XX
Event 1	XX (XX)	XX (XX)	

Abbreviations: CAE, Comfy Acupressure for the Elderly; LAC, least acupressure control.
[a] Data was expressed as number of patients (%). Those which statistical differences $p<0.05$ reached significance level are indicated in bold font.

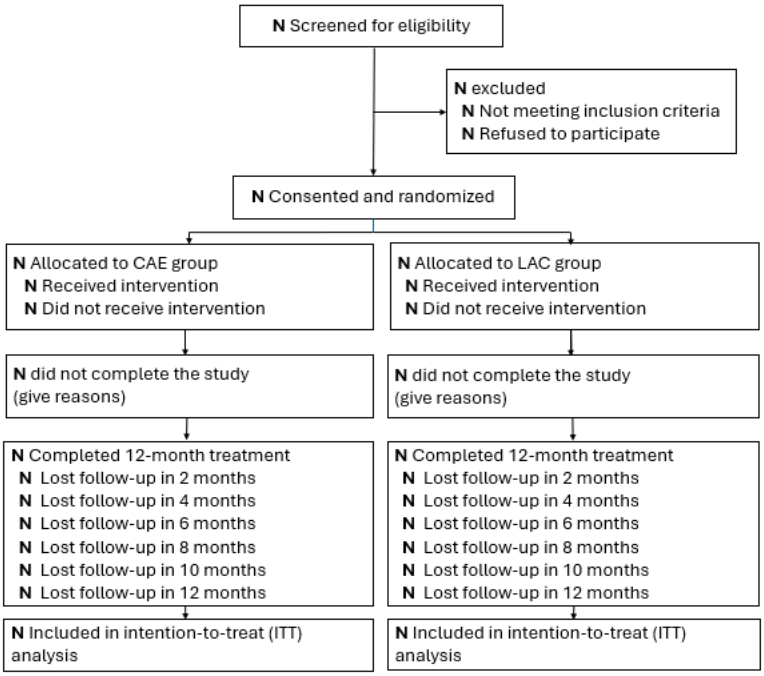


Figure 1. CONSORT Flowchart

References

1. Jack CR Jr, Andrews JS, Beach TG, Buracchio T, Dunn B, Graf A, Hansson O, Ho C, Jagust W, McDade E, Molinuevo JL, Okonkwo OC, Pani L, Rafii MS, Scheltens P, Siemers E, Snyder HM, Sperling R, Teunissen CE, Carrillo MC. Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. *Alzheimers Dement*. 2024 Aug;20(8):5143-5169.
2. Mielke MM, Fowler NR. Alzheimer disease blood biomarkers: considerations for population-level use. *Nat Rev Neurol*. 2024;20(8):495-504.
3. Palmqvist S, Tideman P, Mattsson-Carlgren N, Schindler SE, Smith R, Ossenkoppele R, Calling S, West T, Monane M, Verghese PB, Braunstein JB, Blennow K, Janelidze S, Stomrud E, Salvadó G, Hansson O. Blood Biomarkers to Detect Alzheimer Disease in Primary Care and Secondary Care. *JAMA*. 2024;332(15):1245-1257.
4. Wang J, Huang S, Lan G, Lai YJ, Wang QH, Chen Y, Xiao ZS, Chen X, Bu XL, Liu YH, Zeng F, Zhang L, Li A, Cai Y, Sun P, He Z, Doré V, Fripp J, Bourgeat P, Chen Q, Yu JT, Tang Y, Zetterberg H, Masters CL, Guo T, Wang YJ; Translational Biomarker Research of AgIng and Neurodegeneration (TBRAIN). Diagnostic accuracy of plasma p-tau217/Aβ42 for Alzheimer's disease in clinical and community cohorts. *Alzheimers Dement*. 2025;21(3):e70038.
5. Yeung PY, Wong LLL, Chan CC, Yung CY, Leung LMJ, Tam YY, Tang LN, Li HS, Lau ML. Montreal Cognitive Assessment - Single Cutoff Achieves Screening Purpose. *Neuropsychiatr Dis Treat*. 2020 Nov 6;16:2681-2687.
6. Yeung PY, Wong LL, Chan CC, Leung JL, Yung CY. A validation study of the Hong Kong version of Montreal Cognitive Assessment (HK-MoCA) in Chinese older adults in Hong Kong. *Hong Kong Med J*. 2014;20(6):504-10.
7. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005;53(4):695-9.
8. Jones G, Macken B. Questioning short-term memory and its measurement: Why digit span measures long-term associative learning. *Cognition*. 2015;144:1-13.
9. Logsdon R, Gibbons LE, McCurry SM, Teri L. Quality of Life in Alzheimer's Disease: patient and caregiver reports. *J Ment Health Aging*. 1999;5:21–32.
10. Logsdon RG, Gibbons LE, McCurry SM, Teri L. Assessing quality of life in older adults with cognitive impairment. *Psychosom Med*. 2002;64:510–9.
11. Torisson G, Stavenow L, Minthon L, Londos E. Reliability, validity and clinical correlates of the Quality of Life in Alzheimer's disease (QoL-AD) scale in medical inpatients. *Health Qual Life Outcomes*. 2016;14:90.
12. Lawton MP, Brody EM. Assessment of older people: selfmaintaining and instrumental activities of daily living. *Gerontologist*. 1969;9(3):179-86.
13. Tong AYC, Man DWK. The Validation of the Hong Kong Chinese Version of the Lawton Instrumental Activities of Daily Living Scale for Institutionalized Elderly Persons. *OTJR: Occupational Therapy Journal of Research*. 2002;22(4):132-142.
14. Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. *J Clin Epidemiol*. 1989;42(8):703-9.
15. Leung SO, Chan CC, Shah S. Development of a Chinese version of the Modified Barthel Index-- validity and reliability. *Clin Rehabil*. 2007;21(10):912-22.
16. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, Leirer VO. Development

and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res*. 1982-1983;17(1):37-49.
17. Chiu HF, Lee HC, Wing YK, Kwong PK, Leung CM, Chung DW. Reliability, validity and structure of the Chinese Geriatric Depression Scale in a Hong Kong context: a preliminary report. *Singapore Med J*. 1994;35(5):477-80.
18. Buysse DJ, Reynolds CF 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res*. 1989;28(2):193-213.
19. Wong ELY, Ramos-Goni JM, Cheung AWL, Wong AYK, Rivero-Arias O. Assessing the Use of a Feedback Module to Model EQ-5D-5L Health States Values in Hong Kong. *Patient*. 2018;11(2):235-247.
20. Shah S. Common terminology criteria for adverse events. USA: National Cancer Institute; 2022; 784:785.
21. Toda Y, Iwatsubo T, Nakamura Y, Matsuda N, Miyata M, Jin M, Chen T, Kuribayashi K, Tian Y, Hughes R, Yamamoto J, Muralidharan KK, Rubel C, Hutchison RM, Budd Haeberlein S. Japanese Subgroup Analyses from EMERGE and ENGAGE, Phase 3 Clinical Trials of Aducanumab in Patients with Early Alzheimer's Disease. *J Prev Alzheimers Dis*. 2024;11(5): 1260-1269.
22. Moscoso A, Grothe MJ, Ashton NJ, Karikari TK, Lantero Rodríguez J, Snellman A, Suárez-Calvet M, Blennow K, Zetterberg H, Schöll M; Alzheimer's Disease Neuroimaging Initiative. Longitudinal Associations of Blood Phosphorylated Tau181 and Neurofilament Light Chain With Neurodegeneration in Alzheimer Disease. *JAMA Neurol*. 2021;78(4):396-406.
23. Liang C, Yan X, Tian J, Yang Y, Xiao X, Huang Y, Wang T. Plasma GFAP, NfL, and p-tau181 levels as early biomarkers of dementia in Chinese adults: Shenzhen community cohort study. *Aging Clin Exp Res*. 2025;37(1):105.
24. Sun J, Zeng H, Pan L, Wang X, Liu M. Acupressure and Cognitive Training Can Improve Cognitive Functions of Older Adults With Mild Cognitive Impairment: A Randomized Controlled Trial. *Front Psychol*. 2021;12:726083.
25. Cheung DST, Tiwari A, Yeung WF, Yu DSF, So MKP, Chau PH, Wang XM, Lum TYS, Yuk Fung HYK, Ng BYM, Zhang ZJ, Lao L. Self-Administered Acupressure for Caregivers of Older Family Members: A Randomized Controlled Trial. *J Am Geriatr Soc*. 2020;68(6):1193-1201.
26. Cheung DST, Yeung WF, Chau PH, Chan SY, Chan WL, Yeh CH, Lao L, Zhang Z, Lin CC. Improving chemotherapy-induced peripheral neuropathy in cancer patients using a combined qigong and self-administered acupressure intervention: a randomized controlled trial. *ESMO Open*. 2025;10(9):105565.
27. Yeung WF, Yu BY, Chung KF, Zhang ZJ, Lao L, Ho FY, Suen LK, Ho LM. Self-administered acupressure for insomnia disorder: A randomized controlled trial. *Phytomedicine*. 2022;99:153993.
28. Gupta SK. Intention-to-treat concept: A review. Perspect Clin Res. 2011;2(3):109-12. doi: 10.4103/2229-3485.83221
29. Twisk J, de Boer M, de Vente W, Heymans M. Multiple imputation of missing values was not necessary before performing a longitudinal mixed-model analysis. J Clin Epidemiol. 2013 Sep;66(9):1022-8. doi: 10.1016/j.jclinepi.2013.03.017