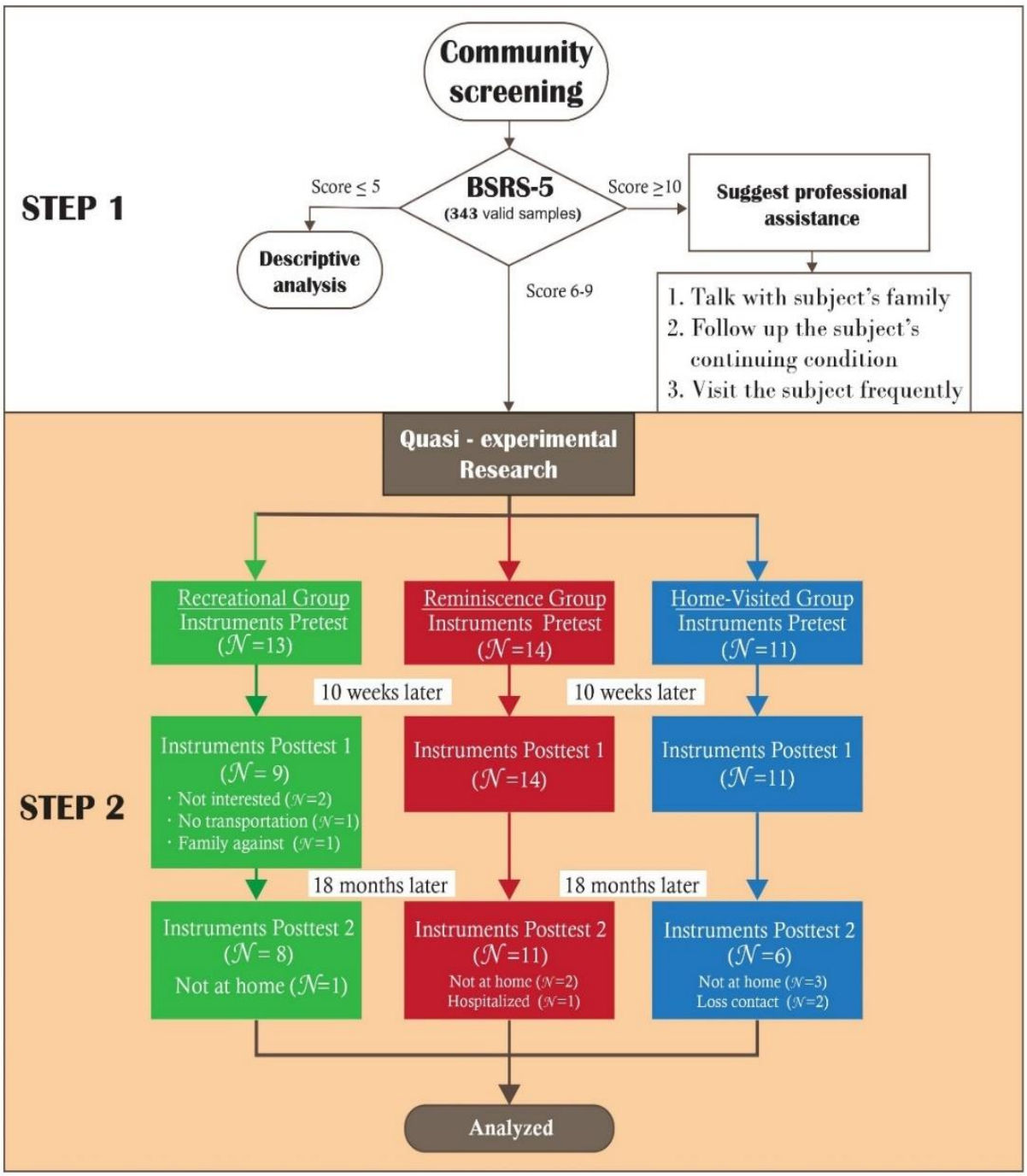
Except cluster randomization was used firstly, the snowball sampling was also introduced, included, 1. most depressed participants would not join unless be accompanied with a friend or relative; 2. they and their family had concerns and rejected when receiving the invitation via telephone due to the common telephone fraud in Taiwan. Therefore, the study was finally a quasi-experimental and analyzed in pre - protocol (PP) approach. The process is shown as Figure 1.

Figure 1. Framework and participants at each stage of this study



Materials and Methods

The SPMSQ was reported as valid for detecting moderate to severe cognitive impairment as efficacious for screening older people (Pfeiffer, 1975). It was a screening test for cognition function of older people by face-to-face interviews. Cronbach's α for the SPMSQ was 0.92, and test-retest reliability was 0.83 in this study.

The 5-item Brief System Rating Scale (BSRS-5) Taiwan version consists of six items developed by Lee, Lee, Yen, Lin, & Lue (1990) which can be completed in five minutes. It has a Cronbach's α of 0.77- 0.90 and a test-retest reliability of 0.82 (Tsai, Chen, Chen, Chan, & Huang, 2021). Participants were asked about the mental conditions in the past week including anxiety, depression, hostility, interpersonal sensitivity, and additional symptoms. The scores of each item ranges from zero to four (0, not at all; 4, extremely). A total score of BSRS-5 is ranging from 0 to 24. The scores above six indicates different level of depression: 6 to 9 is mild; 10 to 14 is moderate; and ≥ 15 is severe. The Cronbach's α in this study was 0.73.

Geriatric Depression Scale-Short Form (GDS-SF) and Social Support Inventory (SSI) were used to evaluate the alternations of depression and social support at baseline (pretest), on that day or no less than 5 days after intervention closing ceremony (posttest 1) and 18 months later (posttest 2). Participants completed the study instruments by themselves in a private area. The feedbacks about their feelings were also recorded as qualitative data at posttest 2.

Interventions

The interventions were designed as a ten-week program composed of physical activity, theme activities (reminiscence, recreational and home-visit), and footbath with massage, as mix models in this study. One researcher led the physical activity, Five-Element Gymnastics (FEG). It integrates Chinese traditional martial arts: China aura, Qigong, and artistic gymnastics (Five-Element Health Promotion Center, 2018). The only difference between Reminiscence and Recreational groups was theme activities. We arranged different sets of eight-theme-activities weekly in Reminiscence group followed as a small team discussion and Recreational group as a handcraft workshop. Discussion in Reminiscence group with team members led by two leaders to recall members' memory; the themes were summer foods and activities, our childhood, old movie watching, my hobbies, my marriage ceremony, hand-made children's toys, my favorite paintings, local religion festival.; while in Recreational group, one teacher taught team members how to play watercolor paintings, singing, old movie watching, hand-made children's toys, traditional Chinese painting, poker (card game), cross-stitch weaving, and paper-cutting (Supplementary Table 2).

After the 2-hour interventions every week, all participants in Reminiscence and Recreational groups were given a 10-minute warm water footbath and a 10-minute foot massage focused on the meridian acupuncture points including spleen meridian, stomach meridian, and foot reflection areas (Supplementary Table 3). We also used incentive of prize and bonus to encourage the full attendance for ten weeks. For Home-Visit group, we provided

in person visit by nursing students accompanied with researchers or alone every week. Blood pressure, finger stick blood sugar checking were routines in Home-Visit schedule, extra nursing instructions were offered responding to participants' health needs.

Data Analysis

The *t* test or one-way analysis of variance (ANOVA) was used for comparing continuous variables between or among subgroups. A one-way analysis of covariance (ANCOVA) was also used to compare GDS-SF and SSI among groups post intervention adjusting for inhabitancy or / and SSI, respectively, because they were significantly different at baseline.

We used Wilcoxon's signed-rank test to compare the differences of outcomes within group, i.e., the GDS-SF and SSI at two time points after intervention with the baseline, respectively. Significant (α) level was set at 0.05. SAS 9.4 and IBM SPSS 19.0 were used for all analyses.

In-depth and semi-structured interviews with participants at posttest 1and 2, followed by a content analysis.