ISSN: 2755-9351

Journal of Alzheimer's Disease & Reports

SCIENTIFIC Research and Community

Research Article

Open Access

Results of a Longitudinal Physical and Cognitive Training of People with Dementia, from Alzheimer Athens

Tsapanou A*[1,2], Zoi P[1] and Sakka P[1]

[1]Alzheimer Athens, Greece

[2]Department of Neurology, National and Kapodistrian University of Athens, Greece

ABSTRACT

Background: Physical and cognitive training programs have shown promise in improving functional and cognitive outcomes in individuals with dementia. However, large-scale, longitudinal evidence remains limited, particularly within real-world clinical settings.

Objective: This study investigates the longitudinal effects of a structured physical and cognitive training program implemented by Alzheimer Athens, focusing on functional and cognitive changes in people with dementia, in a 10-year investigation.

Methods: We analyzed data from 956 patients enrolled in a multi-component intervention program, followed for up to ten-time points. Primary outcomes included the Mini Mental State Examination (MMSE), a verbal fluency task, and the Instrumental Activities of Daily Living (IADL) questionnaire. Mixed-effects models with REML estimation were used to evaluate the impact of time, cognitive status, gender, age, and educational level on outcome measures. Post- hoc comparisons were adjusted using Tukey's method.

Results: Over time, overall cognition and verbal fluency performance remained stable through the fourth to seventh follow-up, and declined thereafter, similarly to the daily functionality. Results were mostly driven by early-stage demented people. Higher cognitive status and educational level were associated with better performance across measures. Age was a significant predictor, with older individuals generally scoring lower on MMSE and verbal fluency, and having more affected functionality.

Conclusions: Despite disease progression, participation in the routine multimodal care program was associated with relatively stable functional abilities over several follow-up periods and slower decline in some cognitive domains. However, causal effects cannot be inferred from this observational study design. The findings support the integration of multimodal training in dementia care pathways.

***Corresponding author**

Angeliki Tsapanou, MSc, PhD Assistant Professor of Clinical Psychology-Neuropsychology National and Kapodistrian University of Athens, Greece.

Received: October 09, 2026; **Accepted:** October 16, 2026; **Published:** October 25, 2026

Keywords: Intervention, Cognitive Training, Physical Training, Dementia

Introduction

Dementia, especially Alzheimer's disease type (AD) encompasses a group of progressive neurodegenerative conditions characterized by cognitive decline, functional impairment, and loss of independence [1]. Non-pharmacological approaches—particularly physical exercise and cognitive training—have garnered increased attention for their potential to slow cognitive and functional deterioration in older adults with dementia or mild cognitive impairment (MCI) [2,3]. Physical activity has been shown to improve brain structure and function, enhancing hippocampal volume, cerebral blood flow, and neurotrophic factor expression [4]. These changes are associated with improvements in attention, executive function, and overall cognition [5]. In a group of older Greek adults, a 36-week exercise training had significant improving results in both physical and cognitive function [6].

Similarly, cognitive stimulation programs, including memory exercises, problem-solving tasks, and language-based activities, have demonstrated beneficial effects on cognitive reserve and delay in clinical progression [7]. Meta-analyses have confirmed the efficacy of cognitive training in enhancing global cognition and specific domains such as working memory and processing speed in individuals with MCI [8]. Regarding the Greek population, a pilot study of people with MCI showed that a training in programming tasks improved mental and cognitive health of the participants, while a different pilot randomized trial showed that a positive effect of a multi-domain multimodal-based cognitive training program in elderly with MCI, with broader benefits on cognition by inducing more cooperative transfer effects over different domains [9,10].

More recently, a growing body of evidence supports the synergistic benefits of combined physical and cognitive interventions. Multidomain programs have shown larger effects on cognition than either intervention alone. A systematic review by Zhu et al. found

that such combined approaches yielded significant improvements in global cognitive functioning among older adults [11]. Similarly, Gavelin et al. reported moderate-to-large effect sizes for working memory, processing speed, and executive function in MCI populations undergoing combined interventions [12].

Large prevention trials such as the Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER) demonstrated that a multidomain lifestyle approach—including diet, physical activity, cognitive training, and vascular risk management—resulted in significant cognitive benefits over two years [13]. Comparable findings emerged from the Australian Maintain Your Brain project, indicating improved cognitive and mental health outcomes in at-risk older adults [14].

Despite promising evidence from controlled trials, fewer studies have examined long-term outcomes of combined physical and cognitive training when delivered as routine care in real-world clinical settings, particularly using multi-year longitudinal follow-up. In the present study, "naturalistic" refers to observational data derived from routine clinical practice, where participants received standard program services without randomization, a control group, or experimental manipulation, and follow-up assessments occurred as part of usual care. Furthermore, many interventions lack extended follow-up or consistent cognitive assessments beyond baseline, limiting our understanding of cognitive trajectories under non-pharmacological care.

The present study seeks to address these gaps by analyzing longitudinal data (over an approximately nine-year, ten-assessment follow-up) from Alzheimer Athens' comprehensive training program for individuals with dementia. This real-world program integrates structured physical activity and cognitive training, with assessments spanning up to ten time points. We focus on three core outcomes, covering multiple cognitive domains: Mini-Mental State Examination (MMSE), verbal fluency task (letter), and the Instrumental Activities of Daily Living (IADL). Our aim is to evaluate longitudinal changes and the influence of demographic and clinical variables on patient trajectories.

Methods

The sample consisted by the participants from all the day call day of Alzheimer's Athens, which is a non-profit organization dedicated to improving the quality of life for people with dementia and their families (Alzheimer Athens - Alzheimer Athens). This study was approved by the Institutional Review Board of the Alzheimer Athens, # 02/09/2024 – 1. This study is based on a naturalistic, real-world clinical cohort drawn from Alzheimer Athens day care centers. This study did not include a control or comparison group. All participants included in the analyses were enrolled in the Alzheimer Athens day care center program and received the structured physical and cognitive training activities that constitute routine care within the organization. No group of eligible non-participants, untreated individuals, or usual-care-only patients was available for comparison, and no contemporaneous external comparison cohort was included. Therefore, the study was designed to describe longitudinal cognitive and functional trajectories among program participants rather than to evaluate treatment efficacy relative to an alternative intervention or standard care.

Participants were recruited through routine clinical intake procedures. Specifically, all individuals who enrolled and participated in the Alzheimer Athens day care center program during the study period were eligible for inclusion. Cognitive status was defined as the baseline clinical severity stage assigned at intake by the Alzheimer Athens multidisciplinary team, using the center's standardized classification system (A1–A5). This variable reflects increasing severity of cognitive impairment, with A1 representing the mildest stage and A5 representing the most advanced stage. Classification was based on routine clinical evaluation including cognitive screening and functional status. Cognitive status (A1–A5) was reassessed during follow-up visits as part of routine clinical care. However, the present analyses used baseline cognitive status to characterize participants according to initial disease severity and to examine whether longitudinal trajectories differed according to severity at study entry.

Because cognitive status classification was not consistently reassessed at later follow-ups, cognitive status was treated as a baseline clinical covariate rather than a time-varying variable. Longitudinal change was evaluated using repeated measurements of the primary outcomes (MMSE, verbal fluency, and IADL) across follow-ups. The MMSE was selected as a primary outcome because it is one of the most widely used measures of global cognitive functioning in dementia research and clinical practice, providing a standardized assessment of overall cognitive status and facilitating comparison with previous studies. The IADL questionnaire was selected because it assesses functional independence in everyday activities, a clinically meaningful outcome closely related to disease progression, autonomy, and quality of life in individuals with dementia. For the IADL measure used in this study, higher scores indicate greater functional impairment, whereas lower scores indicate better functional functioning. Sensitivity analyses indicated that excluding baseline cognitive status from the models resulted in only minor changes in estimates, suggesting that the main longitudinal patterns were not driven by this baseline classification. All cognitive, functional, and psychological measures were administered as part of the routine clinical evaluation and follow-up monitoring conducted at Alzheimer Athens day care centers. The assessment battery is implemented in standard care in order to support accurate clinical characterization, document disease progression, and guide individualized intervention planning. Given the multidimensional nature of dementia, measures were selected to capture complementary domains relevant to diagnosis and clinical monitoring, including global cognition, executive functioning, mood and depressive symptomatology, and everyday functional independence.

Verbal fluency (letter) was included as a sensitive measure of language and executive retrieval processes, which are commonly affected in dementia and are directly targeted by cognitive stimulation activities. It also provides domain-specific information that may not be captured by global screening measures such as the MMSE. Additional outcomes (e.g., ACE, FAB, Trail Making Test, mood scales) were part of the routine multidomain clinical assessment to support diagnostic characterization and longitudinal monitoring of cognitive, executive, and psychological status.

Statistical Analysis

Means and standard deviations were used to express the values of all outcome variables, while categorical data were presented as frequencies and percentages. We estimated the mixed-effects models using restricted maximum likelihood (REML) since it provides less biased estimates of variance components compared with maximum likelihood (ML), especially in moderate sample sizes that are observed as time lapses. Furthermore, all reported model comparisons involved the same fixed-effects structure,

for which REML-based inference is valid [15]. Mixed models with REML estimation were applied for all outcome variables to account for the effect of time, expressed through subsequent follow ups and the effects of cognitive status, gender, age and educational status. The normality of residuals as well as the visual inspection of the related QQ plots was assessed to examine assumptions of each model. All reported post hoc comparisons were based on the Tukey's criterion. Statistical significance was set as 0.05 and the analysis was carried out using Jamovi 2.6.13.0.

A linear mixed-effects model was fitted to examine factors associated with each outcome score. The dependent variable was continuous, modeled using a Gaussian error distribution with an identity link. The model included fixed effects for time Measurement, Cognitive status, Education, Sex, and Age, as well as the following interaction terms:
Measurement × Cognitive Status,Measurement × Education, Measurement × Sex, Education × Sex, and the three-way interaction Measurement × Education × Sex.
A random intercept for participant (id) was included to account for within-subject correlation across repeated measurements. All covariates were mean-centered prior to model estimation. Model parameters were estimated using restricted maximum likelihood with the bobyqa optimizer. Degrees of freedom for fixed effects were calculated using the Satterthwaite approximation. Wald confidence intervals were used for parameter estimation.

Missing data were handled using linear mixed-effects models estimated with restricted maximum likelihood (REML). This approach uses all available observations and does not require complete data across all follow-up assessments. Parameter estimates are valid under the assumption that data are Missing at Random (MAR), meaning that the probability of missingness may depend on observed variables but not on unobserved outcome values after conditioning on model covariates. No imputation procedures were applied.

Because the fully specified model with a ten-level categorical Measurement factor and its higher-order interactions was not estimable at the sparse later assessments (the design matrix becomes rank-deficient owing to empty stage-by-timepoint cells), the primary inferential analysis modelled time as a continuous predictor (linear and quadratic terms), with time-by-covariate interactions and a random intercept per participant. This model is estimable, summarizes the overall trajectory through interpretable slopes with confidence intervals, and is reported with omnibus Type-III Wald tests for every main effect and interaction and with marginal and conditional R^2 (see Supplementary Table 3). Categorical, timepoint-by-timepoint post hoc comparisons were restricted to the first four follow-up assessments, where retention exceeded approximately 20% of the baseline sample (see Supplementary Table 5).

To examine whether attrition was selective, baseline characteristics were compared between participants who returned for at least one follow-up and those assessed only at baseline (see Supplementary Table 2). Differences were tested using Welch t-tests for continuous variables and chi-square tests for categorical variables, with Cohen's d as the effect-size index.

Pairwise post hoc comparisons were conducted only for factors demonstrating statistically significant effects in the mixed-effects models. All contrasts were estimated directly from the fitted mixed models using estimated marginal means and were adjusted for multiple comparisons using Tukey's method.

Results

The analysis involves the changes recorded in a longitudinal follow up of ten measurements for the observed patients. The outcome indices that were recorded were, Ace, BDI, FAB, GDS, GLV7, GLV8, Verbal fluency (letter, animals), IADL, MMSE, Trail Making Test a and b. The demographic characteristics of the participants appear in Table 1. Table 2 shows the changes observed in the cognitive status across the 5 different time points where assessments were made. Despite the fact that many patients were less frequently reported after the first cognitive assessment, (956 pts to 380 at the 2nd assessment, 164 at the 3rd, 26 at the 4th and only 1 at the 5th), what is apparent is that the vast majority, exceeding 75% in each status, of cognitive assessments does not change during the study period.

Table 1: Participant Flow: Observations Available Per Primary Outcome by Follow-up

Baseline characteristic	Retained (>=2 visits, n=633)	Lost after baseline (n=323)	Test	p	Effect size
Age, M (SD)	79.1 (7.2)	77.8 (8.0)	t = 2.53	.012	d = 0.18
MMSE, M (SD)	23.8 (5.7)	19.4 (7.6)	t = 9.25	<.001	d = 0.66
IADL, M (SD)	8.7 (7.0)	12.3 (10.4)	t = -5.42	<.001	d = -0.41
Letter fluency, M (SD)	9.2 (5.4)	6.6 (4.6)	t = 7.17	<.001	d = 0.51
Female, %	73.8%	68.4%	chi2 = 2.78	.096	-
Education (4 cat.)	-	-	chi2 = 10.83	.013	-
Baseline stage A5, %	28.8%	44.9%	chi2 = 47.84	<.001	-
FU8 (~7 yr)	61	52	59	6%	
FU9 (~8 yr)	28	23	27	3%	
FU10 (~8.7 yr)	15	11	14	2%	

Median inter-visit interval = 11.8 months. Percentage relative to the baseline MMSE sample (n = 951).

Table 2: Baseline Comparison of Retained Participants vs Those Lost After Baseline

Baseline characteristic	Retained (>=2 visits, n=633)	Lost after baseline (n=323)	Test	p	Effect size
Age, M (SD)	79.1 (7.2)	77.8 (8.0)	t = 2.53	.012	d = 0.18
MMSE, M (SD)	23.8 (5.7)	19.4 (7.6)	t = 9.25	<.001	d = 0.66
IADL, M (SD)	8.7 (7.0)	12.3 (10.4)	t = -5.42	<.001	d = -0.41
Letter fluency, M (SD)	9.2 (5.4)	6.6 (4.6)	t = 7.17	<.001	d = 0.51
Female, %	73.8%	68.4%	chi2 = 2.78	.096	-
Education (4 cat.)	-	-	chi2 = 10.83	.013	-
Baseline stage A5, %	28.8%	44.9%	chi2 = 47.84	<.001	-

Higher IADL = greater impairment. Drop-outs were more impaired at baseline on every measure, indicating non-random (not MCAR) attrition.

Importantly, these figures (956, 380, 164, 26 and 1) refer to the number of participants whose cognitive stage (A1–A5) was re-evaluated at successive clinical staging assessments (Table 2), and not to the number of participants contributing the primary outcome measures. Retention for the analysed outcomes was substantially higher: for the MMSE, for example, 951, 630, 362 and 242 observations were available at the first four follow-ups, respectively (see Table 3 and Supplementary Table 1).

This research paper focuses on the changes observed in the MMSE, the Verbal fluency and the IADL. The observed values across time appear in Table 3.

IADL outcomes

The overall change in IADL across the 10 follow-up measurements indicates a consistent change in functionality, though not always statistically significant. Specifically, as shown in Figure 1, baseline values of IADL scale are significantly lower comparing to any other timepoint (lower score indicated better functionality). The 2nd, 3rd and 4th follow up appear no significant changes, while the 5th and on, till the 9th follow up, show a statistically significant difference to the IADL values of the 2nd timepoint. The 10th follow up is not significantly higher in certain comparisons as only 15 patients were recorded at that timepoint. The mean differences along with the post hoc tests are reported in Table 4. Examining the changes in time by cognitive level indicates that cognitive levels A4 and A5 record consistently higher values of IADL comparing to A1, A2 or A3 that do not differ between them, across all time points as shown in Figure 2. The results of the post hoc differences appear in Table A1 of the Appendix. A statistically significant effect was also recorded regarding age, where higher IADL values were expected for older patients ($p<0.001$) (Figure 3). Follow up #10 was omitted due to the small sample of values. Functional change over time was not uniform across baseline severity. In stage-stratified mixed-model estimates (see Supplementary Table 4), IADL was statistically unchanged over time only in the mildest baseline stage (A1; slope 0.13 points per follow-up, 95% CI [-0.24, 0.49]), whereas more impaired groups showed progressively steeper functional worsening (A5 slope 1.30 [1.09, 1.50]). Apparent functional stability was therefore confined to the least impaired participants and does not support a general claim of preserved functionality.

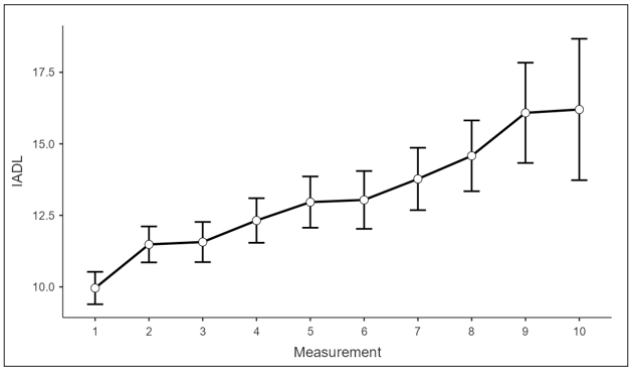


Figure 1: IADL Scores Across Time

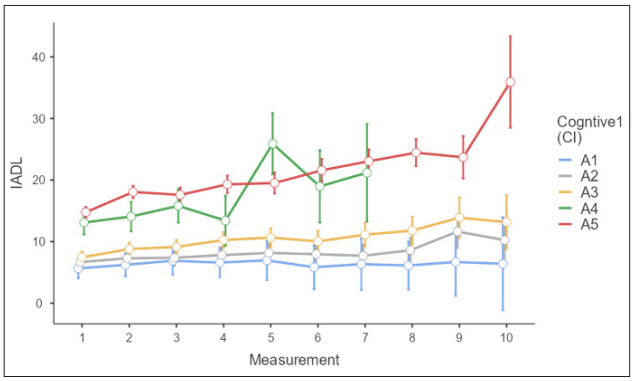


Figure 2: IADL Scores Across Time by Baseline Cognitive Status

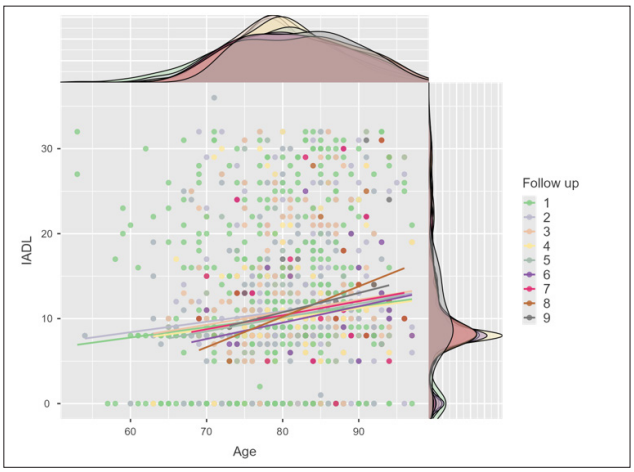


Figure 3: IADL Scores by Age Across Time along with Probability Densities

Restricted to FU1-4 (retention >= ~20-25% of baseline). The earlier apparent late-visit stability does not appear once depleted later visits are excluded.

The mixed-effects model identified a significant main effect of Measurement ($p < 0.001$) and Cognitive Status ($p < 0.001$) on IADL scores. Therefore, Tukey-adjusted post hoc contrasts were performed to characterize differences between follow-up assessments and cognitive status groups.

Table 3: Continuous-Time Mixed-Model Effects (omnibus Type-III Wald tests)

Effect	MMSE chi2 (df), p	IADL chi2 (df), p	Verbal fluency chi2 (df), p
Time (linear)	0.49 (1), .484	0.07 (1), .785	1.15 (1), .284
Time^2	1.41 (1), .235	0.95 (1), .329	31.80 (1), <.001
Cognitive status	1511.7 (4), <.001	346.8 (4), <.001	282.5 (4), <.001
Education	26.5 (3), <.001	15.8 (3), .001	5.4 (3), .142
Sex	0.25 (1), .615	13.1 (1), <.001	0.78 (1), .379
Age	5.17 (1), .023	0.45 (1), .501	4.52 (1), .034
Time x Cognitive status	332.2 (4), <.001	73.7 (4), <.001	21.96 (4), <.001
Time x Education	30.0 (3), <.001	8.73 (3), .033	10.35 (3), .016
Time x Sex	3.27 (1), .071	7.79 (1), .005	8.95 (1), .003
Marginal / Conditional R2	.62 / .89	.28 / .84	.35 / .64

N = 2,639/950 (MMSE), 2,228/930 (IADL), 2,379/911 (fluency) observations/participants.

Table 4: Stage-Specific Change Per Follow-Up (model-based slopes, 95% CI)

Baseline stage	MMSE (points/visit)	IADL (points/visit)	Verbal fluency (points/visit)
A1	-0.07 [-0.29, 0.15] ns	0.13 [-0.24, 0.49] ns	0.16 [-0.14, 0.46] ns
A2	-0.12 [-0.22, -0.01]	0.32 [0.14, 0.49]	-0.27 [-0.42, -0.12]
A3	-0.35 [-0.46, -0.23]	0.66 [0.47, 0.85]	-0.16 [-0.33, 0.00]
A4	-0.86 [-1.31, -0.41]	1.54 [0.82, 2.27]	-0.12 [-0.76, 0.52] ns
A5	-1.53 [-1.66, -1.39]	1.30 [1.09, 1.50]	-0.48 [-0.66, -0.30]

Higher MMSE/fluency = better; higher IADL = worse. Change is minimal at A1 and steepens with baseline severity. ns = 95% CI includes zero.

Table R5: Truncated Sensitivity Analysis (follow-ups 1-4; categorical time)

Effect	MMSE chi2 (df), p	IADL chi2 (df), p	Verbal fluency chi2 (df), p
Measurement (FU1-4)	1.90 (3), .593	1.89 (3), .595	16.39 (3), <.001
Cognitive status	940.7 (4), <.001	222.6 (4), <.001	255.1 (4), <.001
Education	38.8 (3), <.001	9.47 (3), .024	53.3 (3), <.001
Sex	0.85 (1), .358	5.08 (1), .024	6.24 (1), .012
Age	5.43 (1), .020	0.22 (1), .638	4.49 (1), .034
Measurement x Cog status	114.9 (12), <.001	39.5 (12), <.001	40.8 (12), <.001
Marginal / Conditional R2	.58 / .90	.25 / .84	.36 / .72

MMSE Outcomes

The overall change in MMSE across the 10 follow-up measurements indicates a consistent decrease, though not always statistically significant. Specifically, as shown in Figure 4, baseline values of MMSE are significantly higher comparing to any other timepoint. The 2nd, and 3rd follow up appear no significant changes, while the 4th decreases compared to the 2nd, the 5th compared to the 3rd and so on for almost all timepoints, till the 9th. The 10th follow up is not significantly lower in certain comparisons as only 15 patients were recorded at that timepoint. The exact differences appear in Table 5. Examining the changes in time by gender, in Figure 5, indicates that females record consistently lower values of MMSE comparing to males although tend to lessen across time. The exact differences appear in Table A2 of the Appendix. Examining the changes in time by cognitive level (see Figure 6) indicates that cognitive level A5 records consistently lower values of MMSE comparing to A1, A2 or A3 that do not differ between them, across all time points, while A4 shows no difference to the rest of the groups, mostly due to the large variation observed in the relatively small sample of A4 level patients in the study. The exact differences appear in Table A3 of the Appendix. A statistically significant effect was also recorded regarding age, where lower MMSE values were expected for older patients ($p<0.001$) (see Figure 7). Follow ups # 9 and #10 were omitted due to the small sample of values. Overall, MMSE declined significantly over the follow-up period (categorical effect of Measurement, $\chi^2(9) = 213.2$, $p < .001$), but the decline was concentrated in the more advanced baseline stages. In stage-stratified

estimates (see Supplementary Table 4), MMSE was essentially stable in the mildest stage (A1; slope -0.07 points per follow-up, 95% CI [-0.29, 0.15], non-significant) and declined progressively faster with increasing baseline severity (A5 slope -1.53 [-1.66, -1.39]). Overall cognition therefore declined, driven by more impaired participants, while milder stages remained comparatively stable over the earlier follow-ups.

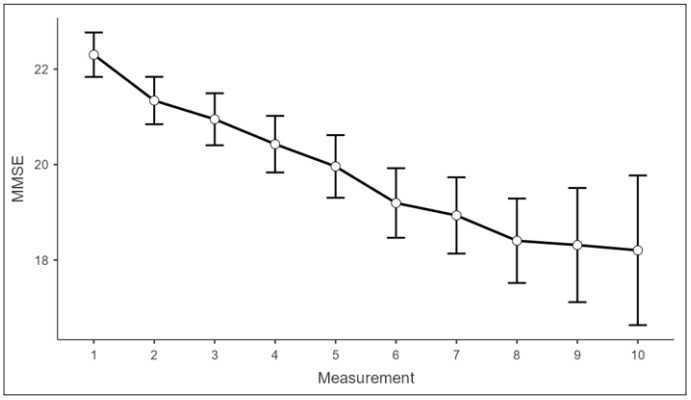


Figure 4: MMSE Scores Across Time

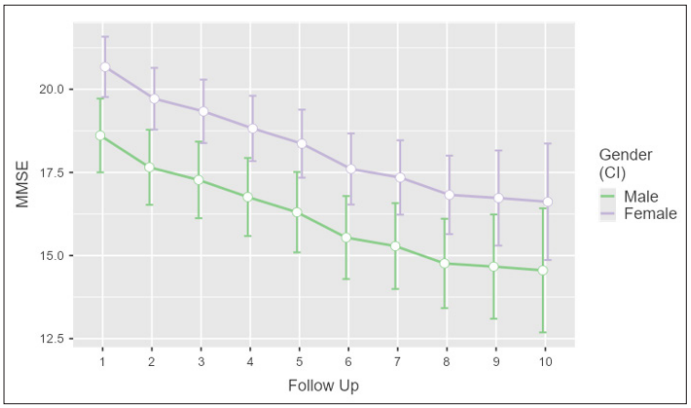


Figure 5: MMSE Scores Across Time by Gender

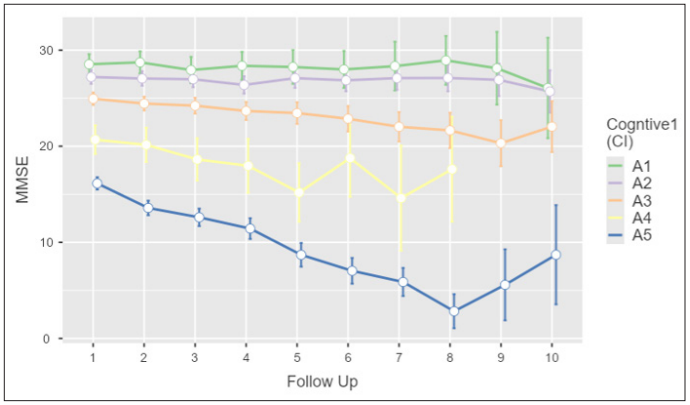


Figure 6: MMSE Scores Across Time by Cognitive Status

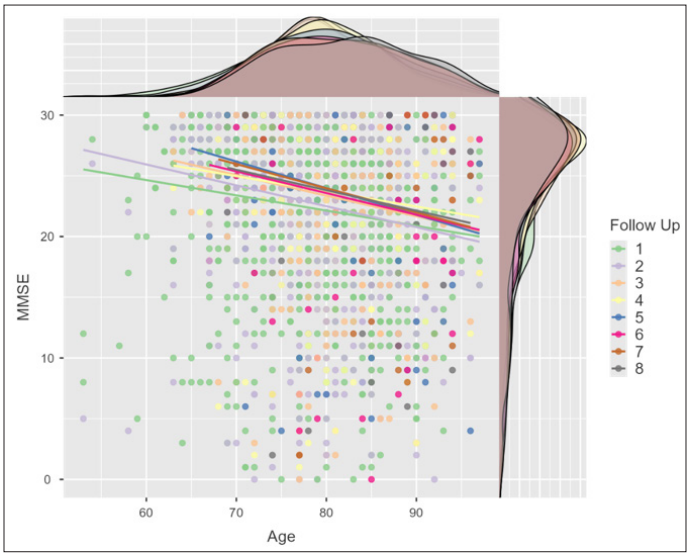


Figure 7: MMSE Scores by Age Across Time along with Probability Densities

Verbal Fluency Outcomes

The overall change in verbal fluency across the 10 follow up measurements indicates no changes till the7th follow up, where lower values appear, that decrease at the 8th follow up and even more at the 9th. The 10th follow- up values of verbal fluency are no different to the values of the 8th and 9th follow up. These appear in Figure 8 and the exact differences appear in Table 6. Examining the changes in time by cognitive level (see Figure 9) indicates that cognitive level A5 records consistently lower values of verbal fluency comparing to A1, A2 or A3 that do not differ between them, across all time points, while A4 shows no difference to the rest of the groups, mostly due to the large variation observed in the relatively small sample of A4 level patients in the study. The exact differences appear in Table A4 of the Appendix. Examining the changes in time by educational status (see Figure 10) indicates that university graduates score significantly higher than primary school or uneducated patients across time. The exact differences appear in Table A5 of the Appendix. A statistically significant effect was also recorded regarding age, where lower verbal fluency values were expected for older patients ($p<0.001$) (see Figure 11). Follow ups # 9 and #10 were omitted due to the small sample of values. Overall, results indicated that people who followed the physical and cognitive program showed stable verbal fluency the first years after their diagnosis, and before the uneventful decline.

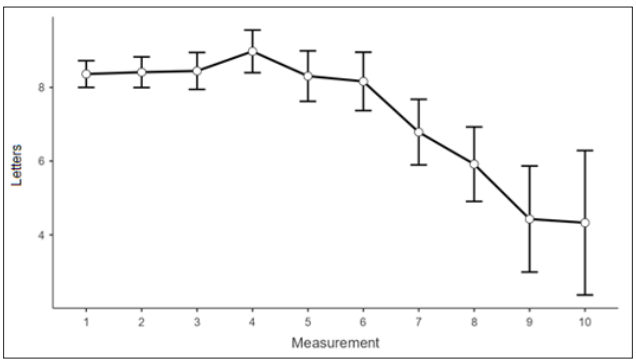


Figure 8: Verbal Fluency Scores Across Time

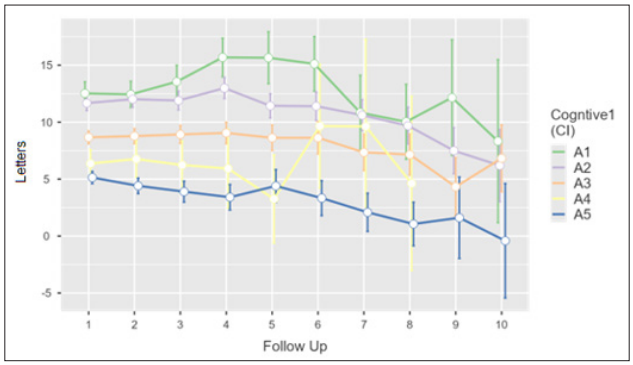


Figure 9: Verbal Fluency Scores across Time by Cognitive Status

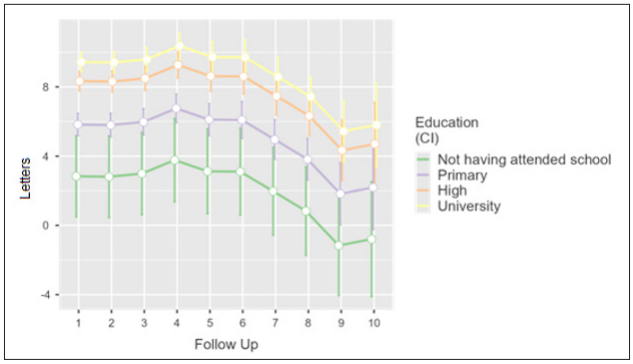


Figure 10: Verbal Fluency Scores Across Time by Educational Level

Dementia stages: Trajectories differed by baseline severity. Statistically significant decline over time was concentrated in the more advanced baseline stages (for verbal fluency, e.g., A5 slope -0.48 points per follow-up, 95% CI [-0.66, -0.30]), whereas the mildest stage showed no significant change (A1; [-0.14, 0.46]). In other words, the overall decline was driven by the more impaired groups rather than by the milder stages (see Figures 2, 6, 9 and Supplementary Table 4).

Discussion

This longitudinal analysis describes cognitive and functional trajectories among individuals with dementia who participated in a structured physical and cognitive training program delivered as part of routine clinical care of Alzheimer Athens. The primary findings suggest that functional abilities, either improved or remained quite stable the first years followed the diagnosis. In contrast, MMSE remained quite stable and declined progressively, while, also, more specifically, verbal fluency remained stable until later years after the initial diagnosis of dementia.

The stability of cognitive status in the majority of participants over time supports the assumption that baseline cognitive function is representative throughout most of the follow-up period. The relatively low percentage of individuals undergoing repeated cognitive assessments and the small sample sizes in later follow-ups limited the granularity of subgroup analysis but did not undermine the overall trends observed.

Importantly, demographic factors played a significant role in outcomes. Older age was associated with lower MMSE and Verbal fluency scores but higher IADL scores, which may reflect increased familiarity with daily tasks or differential disease trajectories. Higher education correlated with better Verbal fluency performance, consistent with the cognitive reserve hypothesis. Gender differences were evident in MMSE scores, with women recording lower values—a finding warranting further investigation. IADLs are clinically meaningful outcomes in dementia, as they reflect real-world autonomy and often decline early due to their reliance on executive and memory processes. However, IADL instruments also have important cross-cultural limitations [16]. Although the IADL scale used in this study provides valuable longitudinal information on functional trajectories, interpretation should consider that IADL performance may be influenced by sociocultural factors (e.g., gender roles, education, and lifestyle) in addition to cognitive decline. Future work should prioritize culturally validated functional measures and complementary ecological assessments to strengthen cross-cultural interpretability. The longer effect of the functionality, despite cognitive decline, underscores the value of functional interventions in dementia care. These findings align with prior literature supporting the role of physical and cognitive training in preserving autonomy and engagement in meaningful daily activities 2. Interventions that maintain functionality—even amid cognitive loss—can contribute to better quality of life, caregiver relief, and delay in institutionalization. This functional resilience is particularly relevant for prolonging independence and reducing caregiver burden7.

Our findings support the ongoing integration of structured, multimodal training into dementia care pathways. Future research should address current limitations, such as missing follow-up data and limited long-term cognitive reassessment. Incorporating biomarkers and neuroimaging may also enhance understanding of underlying mechanisms. Overall, the evidence suggests that continued participation in structured programs can contribute to better outcomes for individuals living with dementia. Beyond direct cognitive and physical effects, participation in structured group-based programs may reduce social isolation, a well-established but often neglected risk factor for poorer cognitive and physical health in older adults. Social isolation and loneliness have been linked to cognitive decline, increased dementia risk, and worse functional outcomes [17]. Although social engagement was not directly measured in this study, increased social contact through regular participation may have contributed to the observed functional stability, and should be examined in future work using validated measures.

The pattern of results indicates that, even in the context of a progressive neurodegenerative condition, structured multimodal training can yield measurable functional benefits and help maintain certain cognitive abilities for a substantial period 4. Improvements in IADL highlight the possibility that interventions may bolster compensatory mechanisms, enabling patients to better navigate daily life even when formal cognitive testing shows deterioration 5.

Strengths of the study include its large baseline sample size, naturalistic clinical setting, and extended follow-up period, which provide valuable insight into long-term intervention outcomes beyond the scope of most controlled trials [18]. The integration of both physical and cognitive training within a single program mirrors real-world clinical practice and enhances ecological validity [11]. However, limitations must be acknowledged: The substantial attrition across follow-up assessments, resulting in progressively smaller sample sizes over time, with the extent of missingness varying by outcome [12]. Consequently, estimates at later time points are less precise and should be interpreted cautiously; for this reason, interpretation of longitudinal patterns is primarily focused on results up to the fourth follow-up

assessment, where retention exceeded approximately 20% of the baseline sample. Substantial attrition was observed over time, which is expected in long-term dementia cohorts due to disease progression and related clinical factors. Although mixed-effects models accommodate incomplete follow-up under the Missing at Random (MAR) assumption, estimates from later assessments should be interpreted cautiously. A formal attrition analysis (see Supplementary Table 2) showed that drop-out was associated with observed baseline severity (e.g., lower baseline MMSE and more advanced baseline stage), indicating that the data are not Missing Completely At Random. Because the mixed models condition on the observed data, including baseline stage and earlier observed outcomes, this pattern is consistent with the MAR assumption under which the model-based estimates remain valid, and does not by itself indicate a MAR violation. The residual, untestable concern is Missing-Not-At-Random drop-out (i.e., drop-out driven by the concurrent, unobserved outcome). Nevertheless, naive descriptive statistics at later visits are computed on a positively selected (survivor) subgroup and should be interpreted cautiously; the continuous-time and truncated analyses therefore serve as robustness checks against the influence of the sparse later assessments. Additionally, the absence of a control group prevents definitive conclusions regarding causality, and the reliance on a single baseline cognitive assessment for many participants may underestimate within-person variability. A limitation is that dementia stage (A1–A5) could not be modelled as a time-varying factor due to missing follow-up staging assessments; therefore, baseline cognitive status should be interpreted as a descriptive clinical indicator rather than a longitudinal measure of progression. Future research should address these issues through strategies such as more frequent cognitive evaluations, inclusion of comparison groups, and incorporation of neuroimaging or biomarker data to explore mechanisms of change [13,14]. Lastly, because all participants received the program and no contemporaneous comparison group was available, observed changes cannot be directly attributed to the intervention and should be interpreted as longitudinal outcomes occurring during participation in routine clinical care.

These findings reinforce the importance of multimodal, non-pharmacological interventions in the management of dementia. The observed improvements in daily functioning, even amidst cognitive deterioration, highlight the role of such programs in preserving autonomy and enhancing quality of life. Future efforts should aim to expand access to these interventions, ensure longer follow-up, and incorporate biomarkers or imaging to deepen our understanding of the mechanisms involved.

The present findings support further investigation of multimodal training approaches within dementia care pathways, ideally through controlled prospective studies capable of evaluating causal effects on cognitive and functional outcomes.

Conclusion

This study provides real-world evidence on the potential benefits of structured physical and cognitive training in individuals with dementia of different stages, based on longitudinal data from Alzheimer Athens. While cognition declines followed an expected trajectory over time, functional abilities are sustained for longer time. Additionally, verbal fluency appeared relatively preserved until later follow-ups, especially among those with higher educational level. The differential effects observed across cognitive subgroups further emphasize the importance of individualized approaches to intervention [2].

Data availability statement: Data available upon request to the corresponding author

Declaration of generative AI: AI tool was used for language editing

Table 1: Demographic Characteristics of the Participants

		N	%
Sex	Men	268	28.0%
	Women	688	72.0%
Educational Status	Not having attended school	18	1.9%
	Primary school	266	27.8%
	High school	364	38.1%
	University Degree	308	32.2%
Cognitive status at baseline	A1 (cognitively healthy)	91	9.5%
	A2 (mild dementia)	201	21.0%
	A3 (moderate dementia)	277	29.0%
	A4 (moderately severe dementia)	60	6.3%
	A5 (severe dementia)	327	34.2%

Table 2: Changes in the Cognitive Assessment at the 5 different Time Points

1st Cognitive Assessment (n=956)	2nd Cognitive Assessment (n=380)	3rd Cognitive Assessment (n=164)	4th Cognitive Assessment (n=26)	5st Cognitive Assessment (n=1)
91 A1	22 A1	7 A1	1 A1	
		1 A2		
	5 A2	2 A3		
	1 A4			
201 A2	1 A1			
		2 A1		
	79 A2	39 A2	3 A2	
			1 A3	
		5 A3		
		1 A5	1 A4	
		1 A2		
	19 A3	6 A3		
		3 A4		
		1 A5		
	1 A5			
277 A3	2 A2	1 A4		
	105 A3	40 A3	7 A3	
		6 A4	1 A5	
	20 A4	1 A3		
		5 A4	1 A4	
		2 A5		
	9 A5			
60 A4	12 A4	1 A4		
		2 A5		
	5 A5	1 A3		
		2 A5		
327 A5	2 A3			
	2 A4	1 A4		
	95 A5	1 A4		
		34 A5	1 A4	1 A5
			8 A5	

Table 3. Baseline and follow-Up Values for the Main Outcomes of the Study

	F-up	N	Min	Max	M	SD
MMSE	1st	951	0	30	22.3	6.70
	2nd	630	0	30	22.65	6.91
	3rd	362	0	30	23.39	6.30
	4th	242	1	30	23.67	6.39
	5th	159	4	30	23.82	6.53
	6th	110	0	30	23.34	6.94
	7th	82	2	30	23.72	6.94
	8th	61	2	30	23.43	7.44
	9th	28	11	30	24.46	5.62
	10th	15	4	30	24.27	7.09
		N	Min	Max	M	SD
	1st	843	0	36	10.02	8.56
	2nd	504	0	36	10.78	7.54

IADL	3rd	306	0	32	10.68	6.83
	4th	204	0	32	10.03	6.27
	5th	126	0	32	10.28	6.11
	6th	89	0	32	9.82	6.32
	7th	71	0	31	10.58	6.45
	8th	52	0	31	10.9	6.69
	9th	23	0	31	11.3	6.79
	10th	11	0	29	10.09	7.18
		N	Min	Max	M	SD
Verbal fluency	1st	832	0	45	8.29	5.24
	2nd	560	0	40	8.85	5.54
	3rd	331	0	48	9.22	5.82
	4th	227	0	44	10.43	6.63
	5th	147	0	37	10.27	6.20
	6th	104	0	32	10.28	5.49
	7th	79	0	20	9.1	4.91
	8th	59	0	21	8.8	5.56
	9th	27	1	17	8.22	4.89
	10th	14	1	15	9.14	4.45

Table 4: Statistical Assessment of the IADL differences Across Time

Measurement	vs	Measurement	Difference	SE	p_{tukey}
1	-	2	-1.526	0.235	<.001
1	-	3	-1.609	0.285	<.001
1	-	4	-2.362	0.333	<.001
1	-	5	-3.005	0.402	<.001
1	-	6	-3.081	0.468	<.001
1	-	7	-3.814	0.512	<.001
1	-	8	-4.622	0.594	<.001
1	-	9	-6.125	0.868	<.001
1	-	10	-6.242	1.243	<.001
2	-	3	-0.083	0.287	1.000
2	-	4	-0.836	0.336	0.274
2	-	5	-1.479	0.403	0.010
2	-	6	-1.555	0.469	0.032
2	-	7	-2.288	0.513	<.001
2	-	8	-3.096	0.594	<.001
2	-	9	-4.599	0.868	<.001
2	-	10	-4.716	1.242	0.006
3	-	4	-0.753	0.348	0.482
3	-	5	-1.396	0.415	0.027
3	-	6	-1.472	0.481	0.069
3	-	7	-2.205	0.524	0.001
3	-	8	-3.013	0.605	<.001
3	-	9	-4.516	0.875	<.001
3	-	10	-4.633	1.248	0.008
4	-	5	-0.643	0.431	0.896
4	-	6	-0.719	0.494	0.909
4	-	7	-1.452	0.536	0.171
4	-	8	-2.260	0.616	0.009

4	-	9	-3.763	0.883	<.001
4	-	10	-3.880	1.253	0.062
5	-	6	-0.076	0.525	1.000
5	-	7	-0.810	0.568	0.919
5	-	8	-1.617	0.642	0.260
5	-	9	-3.121	0.902	0.020
5	-	10	-3.237	1.267	0.240
6	-	7	-0.734	0.598	0.968
6	-	8	-1.541	0.670	0.390
6	-	9	-3.045	0.920	0.033
6	-	10	-3.161	1.280	0.286
7	-	8	-0.807	0.687	0.976
7	-	9	-2.311	0.938	0.289
7	-	10	-2.428	1.291	0.683
8	-	9	-1.504	0.958	0.863
8	-	10	-1.620	1.304	0.965
9	-	10	-0.116	1.384	1.000

Table 5: Statistical Assessment of the MMSE differences Across Time

Measurement	vs	Measurement	Mean Difference	SE	ptukey
1	-	2	0.961	0.155	<.001
1	-	3	1.352	0.193	<.001
1	-	4	1.874	0.227	<.001
1	-	5	2.341	0.269	<.001
1	-	6	3.107	0.313	<.001
1	-	7	3.367	0.355	<.001
1	-	8	3.898	0.404	<.001
1	-	9	3.988	0.576	<.001
1	-	10	4.098	0.774	<.001
2	-	3	0.391	0.195	<.001
2	-	4	0.913	0.228	0.595
2	-	5	1.380	0.270	0.003
2	-	6	2.147	0.314	<.001
2	-	7	2.407	0.356	<.001
2	-	8	2.937	0.405	<.001
2	-	9	3.027	0.577	<.001
2	-	10	3.137	0.774	<.001
3	-	4	0.522	0.239	0.002
3	-	5	0.989	0.279	0.466
3	-	6	1.755	0.322	0.015
3	-	7	2.015	0.363	<.001
3	-	8	2.546	0.411	<.001
3	-	9	2.636	0.581	<.001
3	-	10	2.746	0.778	<.001
4	-	5	0.467	0.292	0.015
4	-	6	1.233	0.333	0.848
4	-	7	1.494	0.373	0.008
4	-	8	2.024	0.420	0.003
4	-	9	2.114	0.588	<.001

4	-	10	2.224	0.783	0.012
5	-	6	0.766	0.352	0.124
5	-	7	1.026	0.390	0.472
5	-	8	1.557	0.435	0.202
5	-	9	1.647	0.598	0.013
5	-	10	1.757	0.790	0.153
6	-	7	0.260	0.411	0.441
6	-	8	0.791	0.454	1.000
6	-	9	0.881	0.612	0.772
6	-	10	0.990	0.801	0.915
7	-	8	0.531	0.476	0.967
7	-	9	0.620	0.629	0.983
7	-	10	0.730	0.814	0.993
8	-	9	0.090	0.651	0.997
8	-	10	0.200	0.831	1.000
9	-	10	0.110	0.906	1.000

Table 6: Statistical Assessment of the Verbal Fluency differences Across Time

Measurement	vs	Measurement	Difference	SE	ptukey
1	-	2	-0.049	0.201	1.000
1	-	3	-0.084	0.248	1.000
1	-	4	-0.614	0.290	0.516
1	-	5	0.057	0.346	1.000
1	-	6	0.201	0.402	1.000
1	-	7	1.577	0.453	0.018
1	-	8	2.446	0.515	<.001
1	-	9	3.936	0.734	<.001
1	-	10	4.036	1.000	0.002
2	-	3	-0.035	0.252	1.000
2	-	4	-0.565	0.294	0.652
2	-	5	0.106	0.349	1.000
2	-	6	0.250	0.404	1.000
2	-	7	1.626	0.454	0.013
2	-	8	2.495	0.515	<.001
2	-	9	3.985	0.735	<.001
2	-	10	4.085	1.000	0.002
3	-	4	-0.530	0.306	0.776
3	-	5	0.140	0.360	1.000
3	-	6	0.284	0.413	1.000
3	-	7	1.661	0.462	0.012
3	-	8	2.530	0.522	<.001
3	-	9	4.020	0.738	<.001
3	-	10	4.120	1.002	0.002
4	-	5	0.671	0.373	0.737
4	-	6	0.815	0.424	0.654
4	-	7	2.191	0.471	<.001
4	-	8	3.060	0.530	<.001
4	-	9	4.550	0.744	<.001
4	-	10	4.650	1.007	<.001
5	-	6	0.144	0.448	1.000

5	-	7	1.521	0.494	0.065
5	-	8	2.390	0.550	<.001
5	-	9	3.880	0.758	<.001
5	-	10	3.980	1.017	0.004
6	-	7	1.377	0.519	0.195
6	-	8	2.245	0.573	0.004
6	-	9	3.735	0.774	<.001
6	-	10	3.835	1.029	0.008
7	-	8	0.869	0.597	0.909
7	-	9	2.359	0.792	0.086
7	-	10	2.459	1.040	0.348
8	-	9	1.490	0.817	0.720
8	-	10	1.590	1.059	0.892
9	-	10	0.100	1.149	1.000

Supplementary Tables

The following quantitative results were added in response to the reviewers. All models were fitted on the original dataset; the per-timepoint sample sizes and baseline demographics reproduce Tables 1 and 3 exactly. Across all three primary outcomes the overall linear effect of time was non-significant once baseline cognitive status and its interaction with time were modelled; trajectories were governed by baseline severity (Table 3), with change minimal at the mildest stage and steepening with severity (Table 4). A formal attrition analysis (Table 2) showed that drop-out was related to observed baseline severity (e.g., baseline MMSE 19.4 vs 23.8 for those lost vs retained; Welch $t = 9.25$, $p < .001$, $d = 0.66$), i.e., the data are not Missing Completely At Random but consistent with the Missing-at-Random assumption under which the mixed models remain valid. Restricting categorical comparisons to follow-ups 1-4 (Table R5) removed the apparent late-visit stability, confirming it was an artefact of selective retention.

References

1. Wimo A, Guerchet M, Ali G-C (2017) The worldwide costs of dementia 2015 and comparisons with 2010. Alzheimer's & Dementia 13: 1–7.
2. Baker LD, Frank LL, Foster-Schubert K (2010) Effects of aerobic exercise on mild cognitive impairment: a controlled trial. Archives of Neurology 67: 71–79.
3. Barnes DE, Yaffe K, Belfor N (2009) Computer-based cognitive training for mild cognitive impairment: results from a pilot randomized, controlled trial. Alzheimer Disease & Associated Disorders 23: 205–210.
4. Erickson KI, Voss MW, Prakash RS (2011) Exercise training increases size of hippocampus and improves memory. Proceedings of the National Academy of Sciences of the United States of America 108: 3017–3022.
5. Colcombe S, Kramer AF (2003) Fitness effects on the cognitive function of older adults: a meta-analytic study. Psychological Science 14: 125–130.
6. Bardopoulou MS, Patsaki I, Chondronikola C (2024) Effect of a 36-week supervised exercise training program on physical and cognitive function in older patients with dementia. In Vivo 38: 286-294.
7. Woods B, Aguirre E, Spector AE, Orrell M (2012) Cognitive stimulation to improve cognitive functioning in people with dementia. Cochrane Database of Systematic Reviews 2012: CD005562.
8. Chan ATC, Ip RTF, Tran JYS, Chan JYC, Tsoi KKF (2024) Computerized cognitive training for memory functions in mild cognitive impairment or dementia: a systematic review and meta-analysis. npj Digital Medicine 7: 1.
9. Demetriadis S, GV, Sapounidis T (2015) Robot programming and tangible interfaces for cognitive training. In: Handbook of Research on Innovations in the Diagnosis and Treatment of Dementia 28.
10. Goumopoulos C, Skikos G, Frounta M (2023) Feasibility and effects of cognitive training with the COGNIPLAT game platform in elderly with mild cognitive impairment: pilot randomized controlled trial. Games for Health Journal 12: 414–425.
11. Zhu X, Yin S, Lang M, He R, Li J (2016) The more the better? A meta-analysis on effects of combined cognitive and physical intervention on cognition in healthy older adults. Ageing Research Reviews 31: 67–79.
12. Gavelin HM, Dong C, Minkov R (2021) Combined physical and cognitive training for older adults with and without cognitive impairment: a systematic review and network meta-analysis of randomized controlled trials. Ageing Research Reviews 66: 101232.
13. Ngandu T, Lehtisalo J, Solomon A (2015) A 2-year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. The Lancet 385: 2255–2263.
14. Heffernan M, Andrews G, Fiatarone Singh MA (2019) Maintain Your Brain: protocol of a 3-year randomized controlled trial of a personalized multi-modal digital health intervention to prevent cognitive decline among community dwelling 55 to 77 year olds. Journal of Alzheimer's Disease 70: S221–S237.
15. Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67: 1–48.
16. Calia C, O'Donald F, Franzen S (2025) Cross-cultural functional assessment for dementia: a commentary. The Clinical Neuropsychologist 1–26.
17. Tragantzopoulou P, Giannouli V (2021) Social isolation and loneliness in old age: exploring their role in mental and physical health. Psychiatriki 32: 59–66.

18. Li H, Li J, Li N, Li B, Wang P, et al. (2011) Cognitive intervention for persons with mild cognitive impairment: a meta-analysis. Ageing Research Reviews 10: 285–296.

Copyright: ©2026 Angeliki Tsapanou, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.