Effects of a Community-Based Multicomponent Physical Exercise Program on Physical Function, Frailty, and Sarcopenia in Older Adults in the Municipality of Santoña: A Quasi-Experimental Study

The Santoña Municipal Multicomponent Exercise Program

Item	Detail
Protocol version	Version 1.0 (13 October 2026). Based on the 2025–2026 assessment database and the original assessment protocol (IFIS questionnaire and adapted Borg scale)
Version date	13 October 2026
Protocol status	Approved by the CEIH-UGR on 29 September 2026; recruiting (Year 1, 2026/2027)
Sponsor / promoting institutions	University of Granada (UGR), Faculty of Sport Sciences, Sport and Health University Research Institute (iMUDS); Ayuntamiento de Santoña (Cantabria, Spain)
Principal investigator (IP)	Andrés B. Fernández-Revelles (UGR), teaching and research staff of the University of Granada (Section 11.1).
Research team	Javier Fernández-Ortega (UGR); Andrés B. Fernández-Revelles (UGR); Juan A. Rangel-García (Ayuntamiento de Santoña); Javier Loza-Herbella (Ayuntamiento de Santoña)
Study setting	Municipal sports facilities ceded by the Excmo. Ayuntamiento de Santoña (Cantabria, Spain), including the Mini pabellón de las Instalaciones Deportivas Municipales (IDM)
Ethics committee	Comité de Ética en Investigación Humana, Universidad de Granada (CEIH-UGR), approved on 29 September 2026 (Granada)
Planned duration / sample	4 consecutive academic years (2026/2027–2029/2030); ~60 participants recruited per annual cohort; cumulative target N ≈ 240

Table of Contents

Protocol Development Note ..... **Error! Bookmark not defined.**
Table of Contents 2
1. Protocol Synopsis 5
2. Background and Rationale 6
2.1 Frailty and Falls in Older Adults 6
2.2 Multicomponent Exercise and the Vivifrail Framework 6
2.3 Rationale for a Formal, Multi-Year Protocol 6
2.4 Preliminary Data 7
3. Study Objectives and Hypotheses 8
3.1 General Objective (per the CEIH-UGR submission) 8
3.2 Secondary Objectives 8
3.3 Hypotheses 8
4. Study Design 10
4.1 Overview 10
4.2 Study Setting 10
4.3 Study Duration and Timeline 10
5. Study Population 11
5.1 Recruitment 11
5.2 Inclusion Criteria 11
5.3 Exclusion Criteria 11
5.4 Withdrawal and Discontinuation Criteria 12
5.5 Sample Size 12
6. Intervention: Multicomponent Exercise Program 14
6.1 Theoretical Framework 14
6.2 FITT-VP Structure 14
6.3 Session Structure 15
6.4 Personnel and Supervision 15
6.5 Individual Adaptations 15
6.6 Fidelity and Quality Control 15
7. Outcome Measures and Assessment Protocol 16
7.1 Assessment Timepoints 16
7.2 Assessment Battery Overview 16
7.3 Block 1: Demographics and Questionnaires 16
7.4 Block 2: Anthropometry 17
7.5 Block 3: Muscular Strength and Functional Performance 17
7.6 Block 4: Cardiorespiratory Health 17
7.7 Composite Fall-Risk Classification 18

7.8 Note on Gait/Walk Test Terminology 18
7.9 Cognitive Screening (MMSE) 18
8. Safety, Adverse Events, and Risk Management 19
8.1 Pre-Participation Medical Screening 19
8.2 Adverse Event Definitions and Classification 19
Causality 19
Severity 19
8.3 Adverse Event Reporting Procedures 19
8.4 Stopping Rules 20
9. Data Management 21
9.1 Data Collection and Entry 21
9.2 Data Storage and Security 21
9.3 Confidentiality 21
10. Ethical Considerations 22
10.1 Ethical Approval 22
10.2 Nature of the Ethical Review and Absence of Clinical Interference 22
10.3 Validated Tools and Team Experience 22
10.4 Compensation, Credits, and Institutional Custody 22
10.5 Informed Consent 22
10.6 Declaration of Helsinki and Data Protection 23
11. Study Organization 24
11.1 Research Team and Roles 24
11.2 Collaborating Institutions 24
11.3 Funding 24
12. Statistical Analysis Plan (SAP) 25
12.1 General Principles 25
12.2 Design Summary for Analysis 25
12.3 Analysis Populations 25
12.4 Outcome Definitions 25
12.5 Descriptive Statistics 26
12.6 Handling of Missing Data 26
12.7 Primary Outcome Analysis 26
12.8 Secondary Outcome Analyses 26
12.9 Pooled Multi-Year (4-Year) Analysis 27
12.10 Subgroup Analyses 27
12.11 Sensitivity Analyses 27
12.12 Multiplicity 28
12.13 Sample Size / Power Justification for the SAP 28
12.14 Software and Reporting Standards 28

12.15 Interim Analyses 28
13. Dissemination Plan 29
14. Study Timeline 29
15. Limitations 29
16. References 31
Supplementary Material 37
Appendix S1: Sample Training Session ("Sesión Tipo") 37
Warm-up (≈10 minutes, 8–10 repetitions per exercise) 37
Main Part: Balance and Gait (2 sets, with ≈1-minute rest between sets) 37
Main Part: Upper-Body Strength (2 sets × 10 reps) 37
Main Part: Lower-Body Strength (2 sets × 10 reps) 38
Cool-Down / Stretching (≈10 minutes; 2 sets × 3 reps, 10-s hold each) 38
Safety Observations (from the source session record) 38
Appendix S2: Borg RPE Scale (0 to 10), adapted version used by the program 39
Appendix S3: Adverse Event Severity Classification 39
Appendix S4: International Fitness Scale (IFIS), Spanish version used by the program 39

1. Protocol Synopsis

Field	Summary
Title	Effects of a Community-Based Multicomponent Physical Exercise Program on Physical Function, Frailty, and Sarcopenia in Older Adults in the Municipality of Santoña: A Quasi-Experimental Study
Design	Prospective, quasi-experimental, repeated-measures, single-arm design (no passive control group), conducted over 4 consecutive annual cohorts; each participant serves as their own control
Population	Community-dwelling, pre-frail/frail older adults ≥65 years referred via the municipal program, meeting the operational eligibility criteria in Section 5.2
Intervention	Vivifrail-based supervised multicomponent exercise program (strength, balance, gait, flexibility, and aerobic components) (Izquierdo, 2019), twice-weekly group sessions (≈66 sessions/cycle) across each ~9-month (≈36-week) annual program cycle, October–June
Primary outcome	Change in Short Physical Performance Battery (SPPB) total score from baseline to end of each annual program cycle
Key secondary outcomes	Handgrip strength (Jamar dynamometry); Timed Up and Go (TUG); 30-second Chair Stand Test; 6-Minute Walk Test (6MWT); anthropometrics (weight, height, BMI, waist circumference); self-perceived fitness (IFIS); cognitive status (MMSE, descriptive/exploratory); composite fall-risk classification; attendance/adherence
Duration	4 consecutive academic years: 2026/2027, 2027/2028, 2028/2029, and 2029/2030; baseline assessment in October and end-of-program assessment in June of each cycle
Sample size	~60 participants recruited per annual cohort; cumulative target N ≈ 240 across the 4 years
Primary analysis	Within-cohort paired comparison (Wilcoxon signed-rank / paired t-test per normality) per annual cycle, plus a pooled multi-year linear mixed-effects model across the 4 years

2. Background and Rationale

2.1 Frailty and Falls in Older Adults

Population ageing is a defining public-health challenge of the coming decades (Beard et al., 2016) and is associated with a rising prevalence of frailty (Fried et al., 2001; Rockwood & Mitnitski, 2007) and falls (Ganz & Latham, 2020), two interrelated geriatric syndromes characterized by reduced strength, impaired balance, and diminished mobility (Woo, 2019). Frailty carries a substantially elevated risk of adverse health outcomes, including hospitalization, disability, and mortality (Chu et al., 2021), and falls represent a considerable and largely preventable economic burden on health systems (Heinrich et al., 2010; Davis et al., 2010). Functional decline of this kind frequently precedes disability, offering a critical window for early, targeted intervention to preserve autonomy and reduce downstream healthcare costs (Laffon De Mazières et al., 2017).

Physical performance is a key marker of functional ageing. The Short Physical Performance Battery (SPPB) (Guralnik et al., 1994) is a validated, widely used tool for its assessment (Cabrero-García et al., 2012; Bergland & Strand, 2019; Kameniar et al., 2024), sensitive enough to detect a Minimal Clinically Important Difference (MCID) of approximately one point (Kwon et al., 2009; Vasunilashorn et al., 2009), and predictive of disability, institutionalization, and mortality (Gawel et al., 2012; De Fátima Ribeiro Silva et al., 2021; Cavanaugh et al., 2018; Máximo et al., 2025), which makes it well suited for monitoring the real-world effectiveness of exercise interventions in this population. Muscle strength, balance, and gait parameters are themselves closely interrelated components of the frailty and sarcopenia phenotype (Serra-Prat & Palomera, 2019), and gait variability in particular is an independent marker of incident mobility disability risk (Brach et al., 2007).

2.2 Multicomponent Exercise and the Vivifrail Framework

Multicomponent exercise programs, combining strength, balance, gait, and aerobic training, are currently regarded as one of the most effective non-pharmacological strategies for preventing and reversing frailty in older adults (Casas Herrero & Izquierdo, 2012; Yang et al., 2024; Sáez De Asteasu et al., 2019; Tou et al., 2021), including measurable benefits for gait speed specifically (Hortobágyi et al., 2015). The Vivifrail project (Izquierdo, 2019) provides an internationally validated, evidence-based framework for prescribing this type of exercise, tailoring content and progression to each participant's functional level; the volume and intensity of the strength component in particular should be individually dosed, since strength gains in adults over 55 show a clear dose–response relationship with training variables (Silva et al., 2014), and multimodal exercise dose is itself associated with health-related quality of life gains (Kaushal et al., 2019). While the efficacy of multicomponent training is well established in short-term, tightly controlled clinical trials (typically 8–12 weeks), evidence on its long-term implementation (≥24 weeks, and across multiple consecutive years) in real-world community and municipal settings remains comparatively limited. Most of the underlying evidence still derives from lab-based or highly controlled clinical-trial conditions, which, while offering high internal validity, often lack ecological validity, that is, they do not reflect the infrastructure, heterogeneous baseline functional levels, and biopsychosocial determinants of long-term adherence found in real municipal programs. Pragmatic, multi-year, real-world evaluations are therefore needed to determine whether efficacy demonstrated under laboratory conditions translates into sustained real-world effectiveness (Millan-Domingo et al., 2022), and to understand how functional gains are built and consolidated over time when a program is embedded, on an ongoing basis, within routine municipal health-promotion services.

2.3 Rationale for a Formal, Multi-Year Protocol

The municipal exercise service described in this protocol has already been delivered in Santoña for several consecutive years (with prior 32-week and 36-week cohorts evaluated in 2022 and 2023 (Fernández-Revelles et al., 2022, 2023), and a further 36-week cohort evaluated between October 2024 and June 2025, n = 86; Fernández-Revelles et al., 2026, under review; a fourth cohort followed in 2025–2026; Fernández-Ortega et al., 2026, under review). Those prior evaluations were conducted as routine program-evaluation activities rather than as a prospectively registered research protocol. The present document formalizes the ongoing collaboration between the Ayuntamiento de Santoña and the University of Granada into a single, prospective 4-year quasi-experimental protocol, with a harmonized and expanded assessment battery (see Section 7) and a Statistical Analysis Plan (Section 12), so that data collected across the next four annual cohorts (2026/2027–2029/2030) can be analyzed both individually (year by year, replicating the existing pipeline) and jointly (pooled across the 4-year period). Because the 2024–2025 and 2025–2026 cohorts (Fernández-Ortega et al., 2026; Fernández-Revelles et al., 2026, both under review) preceded this protocol, they are treated as pilot cohorts of the present protocol: they are not counted among the participants of the formal project, whose Year 1 is the 2026/2027 cycle, and their data are used for preliminary evidence, for the sample-size calculation and in pre-specified sensitivity analyses (Section 12.11).

2.4 Preliminary Data

In the 2024–2025 pilot cohort (36-week program, October 2024–June 2025, n = 86; 84% female; mean age 76.5 years; Fernández-Revelles et al., 2026, under review), the SPPB total score increased significantly from 10.20 ± 1.97 at baseline to 11.17 ± 1.76 post-intervention ($p < .001$); handgrip strength increased from 23.23 ± 6.77 kg to 24.44 ± 6.96 kg ($p < .001$); Timed Up and Go (TUG) time decreased from 11.12 ± 3.87 s to 10.58 ± 3.58 s ($p = .034$); and the prevalence of a composite "high fall-risk" classification fell from 41.9% to 19.8% (McNemar's $p = .001$). Mean adherence was 74.8%, and participants with high baseline fall risk attended significantly fewer sessions than low-risk participants (67.4% vs. 78.4%, $p = .021$). A second pilot cohort ran in 2025–2026 (baseline October 2025, end-of-program June 2026; Fernández-Ortega et al., 2026, under review) with the same core battery as the 2024–2025 cohort (SPPB, TUG, 6-meter gait speed, handgrip dynamometry, MMSE, fall history and attendance), and not yet the expanded battery described in Section 7. Among the 92 participants with paired SPPB data, the total score rose from 10.11 ± 1.90 to 11.43 ± 1.32 (mean change 1.33 points; SD of the individual change 1.38; $dz = 0.96$; Wilcoxon signed-rank $p < .001$); 67 participants (72.8%) improved, 21 (22.8%) did not change and 4 (4.3%) worsened, and 20 (21.7%) already started at the maximum score of 12. Baseline MMSE was 27.0 ± 3.1 (range 11 to 30). These findings, from the two pilot cohorts (Fernández-Ortega et al., 2026; Fernández-Revelles et al., 2026, both under review), motivate and inform the design, the sample-size calculation (Section 5.5) and the analysis plan of the present 4-year protocol.

3. Study Objectives and Hypotheses

3.1 General Objective

To evaluate the longitudinal (pre- and post-intervention) effects of a community-based multicomponent physical exercise program across 4 academic years in older adults residing in the municipality of Santoña (Cantabria), with physical performance, measured by the SPPB total score (Guralnik et al., 1994), as the primary evaluative endpoint.

3.2 Secondary Objectives

- To evaluate the effect of the program on upper-limb muscle strength (handgrip dynamometry), a key clinical marker of sarcopenia (Syddall et al., 2003).
- To evaluate the effect of the program on functional mobility (Timed Up and Go, TUG), given its established predictive value for future disability and mortality (Podsiadlo & Richardson, 1991; Son et al., 2022; Bergland et al., 2017).
- To evaluate the effect of the program on lower-limb functional strength/power (30-second Chair Stand Test), which may outperform SPPB in discriminating disability risk among higher-functioning older adults (Máximo et al., 2025).
- To evaluate the effect of the program on submaximal aerobic capacity/functional endurance (6-Minute Walk Test) (Lipkin et al., 1986).
- To evaluate the effect of the program on body composition indicators (BMI, waist circumference).
- To evaluate changes in self-perceived physical fitness (IFIS questionnaire) (Merellano-Navarro et al., 2017), and its concordance with performance-based measures (Mayhew et al., 2020).
- To describe global cognitive status (MMSE) at baseline and its change over the program, and to descriptively explore its association with exercise participation (Folstein et al., 1975; Davis et al., 2021).
- To evaluate the effect of the program on a composite fall-risk classification (Ambrose et al., 2013) and on frailty category (pre-frail vs. frail vs. minimal-limitation).
- To examine whether program effects are consistent across the four annual cohorts (Year × Time interaction), i.e., whether the intervention is reproducibly effective when delivered repeatedly, year after year, in the same municipal setting.
- To describe adherence to the program across the 4-year period and to examine baseline predictors of adherence (e.g., baseline fall-risk status, sex, age, cognitive function), consistent with previously identified predictors of exercise adherence in older adults (Aartolahti et al., 2015; Farrance et al., 2016; Liu & Miyawaki, 2020; Davis et al., 2021; McPhate et al., 2013; Fielding et al., 2007; Room et al., 2017; Wu et al., 2025; Lee et al., 2024).
- To monitor and describe the type, frequency, and severity of adverse events associated with program participation over the 4-year period.

3.3 Hypotheses

General hypothesis: continued participation in a 36-week multicomponent physical exercise program (based on the Vivifrail model) will generate statistically significant improvements in physical function, overall

muscular strength, and cardiorespiratory capacity, acting as an effective strategy to reverse indicators of frailty and sarcopenia in independent, community-dwelling older adults. The sub-hypotheses below operationalize this general hypothesis for the purposes of the Statistical Analysis Plan (Section 12).

- H1 (primary): Participation in the multicomponent exercise program will be associated with a statistically significant and clinically meaningful (≥1-point) improvement in SPPB total score from baseline to end of program (Kwon et al., 2009), replicated across each of the 4 annual cohorts.
- H2: Participation will be associated with significant improvements in handgrip strength, 30-second Chair Stand repetitions, and 6MWT distance.
- H3: The prevalence of the composite high fall-risk classification will decrease significantly from baseline to end of program.
- H4: Participants with high baseline fall risk will show lower session adherence than participants with low baseline fall risk, consistent with prior cohorts.
- H5: The magnitude of the pre–post improvement in the primary outcome will not differ significantly across the 4 program years (i.e., no significant Time × Year interaction), supporting the reproducibility of program effects over time.

4. Study Design

4.1 Overview

This is a prospective, quasi-experimental, repeated-measures design of a single group (experimental group only) (Casas Herrero & Izquierdo, 2012; Yang et al., 2024). A passive control group is deliberately not used: given the community-based, health-promotion nature of the program in a municipal setting, the objective of the municipal program is not to deny access to exercise and its benefits to older residents of the locality by withholding it from a comparison arm. Each participant therefore serves as their own control, comparing baseline and end-of-program records. The study is organized as four consecutive annual program cycles (academic years 2026/2027 through 2029/2030), each corresponding to one full delivery of the municipal multicomponent exercise program (baseline assessment in week 0 → ~36-week supervised intervention → end-of-program assessment in week 36). Because the municipal program is a pre-existing, continuously operating community service, some participants may enroll in more than one annual cycle (returning participants); this repeated-enrollment structure is explicitly addressed in the Statistical Analysis Plan (Section 12.9).

4.2 Study Setting

The intervention is delivered in a municipal sports facility ("Mini pabellón de las Instalaciones Deportivas Municipales", IDM) in Santoña, Cantabria, Spain, as part of a public physical-activity service for older adults organized by the Ayuntamiento de Santoña. Data collection and statistical analysis are coordinated by the Department of Physical Education and Sports, Faculty of Sport Sciences (iMUDS), University of Granada.

4.3 Study Duration and Timeline

The project spans 4 consecutive academic years: 2026/2027, 2027/2028, 2028/2029, and 2029/2030. Each cycle runs approximately 36 weeks (≈9 months), consistent with the October-to-June structure observed in the 2024–2025 and 2025–2026 pilot cohorts (Fernández-Ortega et al., 2026; Fernández-Revelles et al., 2026, both under review; baseline assessment in the first 1–2 weeks of October; end-of-program assessment in the final 1–2 weeks of June), with a summer recess outside the intervention period. Year 1 of the project is the 2026/2027 cycle; the 2024–2025 and 2025–2026 cycles are pilot cohorts (Section 2.4).

Assessment windows. Baseline (T0) is performed during the first two weeks of October and the end-of-program assessment (T1) during the last two weeks of June, so that the interval between T0 and T1 is about 36 weeks in every cycle. The exact dates are fixed each year in the municipal sports calendar and recorded in the annual report.

Follow-up. No post-program follow-up visit is planned. The summer recess (about July to September) is treated as a natural detraining interval: for participants who re-enroll, the difference between the end-of-program value of year n and the baseline value of year n + 1 estimates how much of the gain is retained over the break (Section 12.9).

5. Study Population

5.1 Recruitment

Participants are identified through a purposive referral pathway from the local primary health care center ("Centro de Salud"): older adults aged over 65 years with mobility limitations are referred for a frailty screening using a physical-function test battery. Screening and subsequent follow-up are performed by researchers and sport-science educators trained in the standardized timing of the physical tests and in the use of the Jamar hand dynamometer, following the protocol of the American Society of Hand Therapists (ASHT).

The operational eligibility criteria below are used in this study; frailty/risk status is defined operationally by the SPPB, TUG, and gait-speed thresholds below rather than by SHARE-FI screening.

5.2 Inclusion Criteria

- Persons of either sex, aged 65 years or older.
- No disability in basic activities of daily living: Barthel Index score ≥ 90 points (Mahoney & Barthel, 1965).
- Altered mobility: Timed Up and Go (TUG) ≥ 10 seconds, i.e., outside the normative "freely mobile" band for community-dwelling older adults (Podsiadlo & Richardson, 1991).
- Possible frailty, defined by an SPPB score < 12 points (maximum score 12) (Guralnik et al., 1994).
- Fall risk, defined by at least one of: 1 fall in the previous year requiring medical assistance, more than 1 fall without medical assistance, TUG > 20 seconds, or gait speed over 6 meters (VM) > 7.50 seconds (equivalent to < 0.8 m/s) (Ambrose et al., 2013; Studenski, 2011; Cesari et al., 2005).
- Reduced strength: handgrip strength < 27 kg (men) or < 16 kg (women), or more than 15 seconds to complete 5 chair stands, in line with the EWGSOP2 sarcopenia cut-off points (Cruz-Jentoft et al., 2019).
- Voluntary signature of the informed consent form.

Eligibility rule. All the criteria listed above must be met for entry into the program. The only exception concerns people who took part in a previous edition: they may continue in the program even if, after that edition, they no longer meet one or more of the criteria (for example, they now start with an SPPB of 12 and therefore no longer meet the criterion of possible frailty). In the 2025–2026 pilot cohort (Fernández-Ortega et al., 2026, under review), 20 of 92 participants (21.7%) started with an SPPB of 12, which is consistent with this exception. Returning status and the number of previous cycles completed are recorded for every participant, so that returning participants who no longer meet the entry criteria can be handled in the sensitivity analyses (Section 12.11).

5.3 Exclusion Criteria

- Terminal illness.
- Residence in a geriatric care facility (nursing home).
- Diagnosed with, or currently under evaluation for, moderate-to-severe cognitive impairment.
- Unstable cardiovascular disease.
- Acute infectious disease.

- Uncontrolled arterial hypertension (systolic > 180 mmHg or diastolic > 110 mmHg).
- Uncontrolled diabetes mellitus.
- Advanced organ failure.
- Disabling musculoskeletal pathology precluding safe participation.

Cognitive screening at baseline. No MMSE threshold is applied for entry. The exclusion criterion relies on a documented medical diagnosis, or an ongoing diagnostic work-up, of moderate-to-severe cognitive impairment. The MMSE (Folstein et al., 1975) is administered at baseline and at the end of the program as a descriptive and exploratory measure of global cognition, and it is not used to include or exclude participants. In the 2025–2026 pilot cohort (Fernández-Ortega et al., 2026, under review) the baseline MMSE was 27.0 ± 3.1 (range 11 to 30; one participant below 18), and low scores are reported descriptively and examined in a sensitivity analysis (Section 12.11).

5.4 Withdrawal and Discontinuation Criteria

Participation is voluntary and a participant may withdraw at any time, without giving a reason and without any consequence for access to the municipal program. A participant is also withdrawn from the study when (i) a physician advises it or a condition listed in Section 5.3 appears during the cycle, or (ii) a severe exercise-related adverse event requires it (Section 8.4). Stopping attendance for a period is not by itself a reason for withdrawal: as in previous editions, participants who interrupt their training are still invited to the end-of-program assessment, and their attendance is recorded and used in the adherence analyses (Sections 12.10 and 12.11). Data collected before withdrawal are kept and analyzed according to the intention-to-treat principle (Section 12.3), unless the participant asks for them to be deleted, in which case they are removed in line with data-protection law. The reason for withdrawal is recorded.

5.5 Sample Size

Annual recruitment of approximately 60 participants per cohort is anticipated, yielding a cumulative total sample of approximately 240 participants across the 4 years of the project. This figure is administratively set (based on referral capacity and facility capacity) rather than derived from a formal a priori power calculation; the corresponding achieved power is estimated below for context. The pilot cohorts of 2024–2025 and 2025–2026 (Fernández-Ortega et al., 2026; Fernández-Revelles et al., 2026, both under review) are not part of this figure.

Basis of the calculation. The 2025–2026 pilot cohort (Fernández-Ortega et al., 2026, under review) gives the variability needed for the primary outcome: in 92 participants with paired data the SPPB changed by 1.33 points on average, with a standard deviation of the individual change of 1.38 (Section 2.4).

Evaluable participants needed. For the minimal clinically important difference of one SPPB point (Kwon et al., 2009; Perera et al., 2006), the standardized effect is dz = 1 / 1.38 = 0.72. A two-sided paired comparison with $\alpha = 0.05$ needs about 17 evaluable participants for 80% power and 22 for 90% power with the paired t-test, and about 18 and 24 with the Wilcoxon signed-rank test (asymptotic relative efficiency 0.955).

Planned sample. In the previous editions there were no drop-outs: every participant who started the program completed the post-intervention assessment, although some stopped training for a period. A conservative planning assumption of 10% lost before the end-of-program assessment is nevertheless used. With about 60 participants enrolled per year (about 54 completers), one annual cohort has 80% power to detect a mean change of 0.54 SPPB points (dz = 0.39). The pooled sample of about 240 participants (about 216 completers) has 80% power to detect 0.26 points (dz = 0.19). With a pessimistic loss of 25% (about 45 and 180 completers)

the detectable changes are 0.59 and 0.29 points, so the planned sample is sufficient for the primary outcome under both scenarios. Only 1 to 5 participants per year (about 2% to 8% of the annual intake) are expected to come from previous editions, according to the knowledge of the team, so the correlation created by re-enrollment is small and the figures above are close to the power of the mixed model of Section 12.9; the number of returning participants is recorded each year and the power is recalculated with the observed figure.

6. Intervention: Multicomponent Exercise Program

6.1 Theoretical Framework

The exercise program adheres to the Vivifrail training model (Izquierdo, 2019), an internationally validated reference framework for the prevention of frailty, sarcopenia, and fall risk in older adults. The intervention is multicomponent in nature, integrating muscular strength, balance, coordination, flexibility, and cardiovascular endurance training within each session, consistent with current evidence on effective multicomponent exercise design for frail and pre-frail populations (Yang et al., 2024; Bohrer et al., 2019).

6.2 FITT-VP Structure

The intervention is prescribed and documented according to the FITT-VP framework (Frequency, Intensity, Time, Type, Volume, Progression), summarized in Table 6.1.

Component	Prescription
Frequency	Twice-weekly supervised group sessions, confirmed by the 2025–2026 attendance records (≈66 sessions delivered October–June; see Supplementary Material S1)
Intensity	Controlled using the Borg Rating of Perceived Exertion (RPE) scale, 0–10 (see Supplementary Material, Appendix S2)
Time	50–60 minutes per session, structured as warm-up + main part + cool-down (see Section 6.3)
Type	Muscular strength, aerobic endurance, flexibility, and balance training
Volume	Approximately 2–3 sets of 8–12 repetitions for strength exercises; 15–20 minutes of continuous walking for aerobic components (session-specific volumes detailed in Section 6.3 and Supplementary Material S1)
Progression	Individualized progression criteria (e.g., addition of weighted ankle cuffs, increased resistance-band tension, progression from seated to standing exercise variants)

The 2025–2026 attendance register (monthly grids, October through June, ~66 scheduled sessions) confirms a twice-weekly delivery schedule, consistent with the previously reported 2024–2025 cohort (Fernández-Revelles et al., 2026, under review) and resolving the earlier discrepancy against a generic "3 days/week" FITT-VP example found in an older internal template.

Progression criteria. Progression is based on performance, not on the calendar, and follows the Vivifrail model (Izquierdo, 2019) and the international exercise recommendations for older adults (Izquierdo et al., 2021). Instructors apply the following rules and record any individual adaptation on the session sheet:

- Strength: the target intensity is 5 to 7 on the adapted Borg scale (Appendix S2; Borg, 1982). When a participant completes 2 sets of 10 repetitions at a rating of 5 or less in two consecutive sessions, the load is raised one step (heavier ankle weights, a stiffer elastic band, or a move from seated to standing variants).
- Balance and gait: the exercise moves from a wide to a narrow base of support (feet together, semi-tandem, tandem, single leg) and from two hands to one hand to no hand support on the chair, when the participant completes the whole set without losing balance.
- Reduction: the load or the range of movement is reduced when the rating reaches 8 or more, or when pain, dizziness or breathlessness limits the execution.

6.3 Session Structure

Each session follows a standardized three-part structure, adapted from the Vivifrail multicomponent program guidelines. A fully worked example of a typical session ("sesión tipo"), as currently delivered, is reproduced in full in Supplementary Material, Appendix S1. In brief:

- Warm-up (≈10 minutes): general joint-mobility exercises for the neck, shoulders, elbows, wrists, spine, hips, knees, and ankles.
- Main part (≈40 minutes): balance and gait exercises (e.g., single-leg stance, heel-to-toe standing, weight transfers, tandem walking); upper-body strength exercises using elastic resistance bands and stress balls; lower-body strength exercises (sit-to-stand variants, hip abduction/extension, calf raises), with progression via weighted ankle cuffs.
- Cool-down / flexibility (≈10 minutes): static stretching of the neck, arms, shoulders, posterior thigh, anterior thigh, and ankles.

6.4 Personnel and Supervision

All sessions are conducted in group format and supervised by a trained exercise professional with experience delivering physical-activity programs for older adults. Physical assessments are performed by researchers and sport-science educators trained in standardized test administration, including dynamometry per the ASHT protocol.

6.5 Individual Adaptations

Individual adaptations to exercise selection, load, and range of motion are applied when necessary to ensure participant safety and continued participation, consistent with the Vivifrail model's emphasis on individualized functional-level-based prescription.

6.6 Fidelity and Quality Control

Fidelity. Each session is documented on the program’s session sheet (Appendix S1: date, group, number of participants, ages, exercises, time and observations, signed by the instructor) and attendance is recorded in the monthly attendance register. All assessors and instructors follow the same written methodological guide, and the same assessors perform T0 and T1 whenever possible. Before each baseline the team holds a standardization session (timing, dynamometer position, verbal cues of the 6MWT), and the principal investigator reviews a sample of session sheets each term against the prescribed content.

7. Outcome Measures and Assessment Protocol

7.1 Assessment Timepoints

All outcome measures are assessed at baseline (T0, at program entry, each annual cycle) and at end of program (T1, at completion of the ~36-week annual cycle), by the same trained personnel who performed the baseline measurements, to minimize inter-rater variability. Assessors are not blinded to participants' intervention status; a strict standardized protocol is followed for each test at both timepoints to ensure objective, comparable measurement.

No interim assessment is planned. Outcomes are measured at two time points per cohort, week 0 and week 36; attendance is recorded continuously during the cycle.

7.2 Assessment Battery Overview

The assessment protocol is organized into four functional blocks plus two cross-cutting measures (fall-risk screening and cognitive screening) carried over from the program's existing practice. Blocks and individual tests marked "(new)" represent an expansion of the battery beyond what was used in the previously reported 2024–2025 cohort (Fernández-Revelles et al., 2026, under review) and is still being used, as of the 2025–2026 cohort (Fernández-Ortega et al., 2026, under review), i.e. these additions are planned but do not yet appear in the most recent completed assessment database.

Block	Domain	Measure(s)
1	Demographics & self-report	Age, sex; International Fitness Scale (IFIS) questionnaire (Merellano-Navarro et al., 2017) (new)
2	Anthropometry	Height, body weight, BMI, waist circumference (new)
3	Muscular strength & functional performance	SPPB (Guralnik et al., 1994), primary outcome; Timed Up and Go (TUG) (Podsiadlo & Richardson, 1991), currently in use; handgrip dynamometry, currently in use; 30-second Chair Stand Test (Máximo et al., 2025) (new)
4	Cardiorespiratory health	6-Minute Walk Test, 6MWT (Lipkin et al., 1986) (new, distinct from the 6-meter gait-speed test in Section 7.7; see Section 7.8)
—	Cognitive status (descriptive)	Mini-Mental State Examination (MMSE), currently in use (Folstein et al., 1975)
—	Fall-risk screening	Fall history, TUG, 6-meter gait-speed test → composite classification (Section 7.7), currently in use

7.3 Block 1: Demographics and Questionnaires

Administered at the start of contact with the participant, in a quiet, seated setting. Age and sex are recorded to characterize the sample and allow stratification of results by sex and age group. The IFIS questionnaire assesses self-perceived physical fitness (aerobic capacity, strength, speed/agility, and flexibility) relative to same-age peers, and has demonstrated acceptable validity in older adult populations (Merellano-Navarro et al., 2017); it shows only moderate concordance with performance-based measures, which is itself informative when both are collected (Mayhew et al., 2020). The assessor reads each item verbatim, without influencing the participant's response.

IFIS items and scoring. The Spanish version used by the program (Appendix S4) has five items answered on a five-point scale (1 = very poor, 2 = poor, 3 = acceptable, 4 = good, 5 = very good), in which the participant compares himself or herself with people of the same age: (1) general fitness, (2) cardiorespiratory fitness, (3)

muscular strength, (4) speed and agility and (5) flexibility. Each item is analyzed separately and, also as the mean of the five items (range 1 to 5).

7.4 Block 2: Anthropometry

Measurements are taken with the participant in light clothing and barefoot. Height and body weight are used to calculate BMI, allowing classification of nutritional status (underweight/normal weight/overweight/obesity). Waist circumference is measured twice (non-consecutively) at the midpoint between the lower costal margin and the iliac crest, along the mid-axillary line (World Health Organization, 2011), with the participant's arms relaxed at their sides; the measurement is recorded after a full exhalation, to a precision of 0.1 cm, and repeated if the two measurements differ by more than 5 mm. Waist circumference serves as an anthropometric indicator of central adiposity and cardiometabolic/vascular risk.

7.5 Block 3: Muscular Strength and Functional Performance

The SPPB (primary outcome) (Guralnik et al., 1994) evaluates lower-extremity functional performance, mobility, and frailty/disability risk through three chained sub-tests: a standing balance test (feet together, semi-tandem, and tandem positions, held for 10 seconds each; unipedal stance duration is itself an established indicator of fall risk in older outpatients (Hurvitz et al., n.d.)), a 4-meter habitual gait-speed test, and a 5-repetition chair-stand test. The total score ranges from 0–12; scores of 0–6 indicate low performance/frailty, 7–9 intermediate performance/pre-frailty, and 10–12 high performance/minimal limitation (Cabrero-García et al., 2012; Bergland & Strand, 2019). A 1-point change is considered the MCID (Kwon et al., 2009).

The Timed Up and Go (TUG) test (Podsiadlo & Richardson, 1991) is administered alongside the SPPB and continues to be collected pre- and post-intervention, as confirmed by the 2025–2026 assessment database; it is not superseded by the newer battery. The participant rises from a standard armchair, walks 3 meters, turns, walks back, and sits down again, timed to the nearest hundredth of a second. TUG is used both as part of the eligibility/fall-risk criteria (Section 5.2) and as a standalone secondary outcome, given its established predictive value for future disability (Son et al., 2022) and all-cause mortality (Bergland et al., 2017).

Handgrip strength is measured using a Jamar hydraulic dynamometer, twice per hand, per the ASHT standardized procedure; the highest value obtained is used for analysis. It serves as a key clinical marker for sarcopenia screening (Syddall et al., 2003).

The 30-second Chair Stand Test evaluates lower-limb strength-endurance and functional power (primarily knee and hip extensors), and has been reported to outperform the full SPPB in discriminating disability risk among higher-functioning older adults (Máximo et al., 2025). Using the standard protocol with a chair (seat height 43–45 cm) braced against a wall, the participant begins seated with arms crossed over the chest and, on the signal "go", stands fully (complete hip and knee extension) and sits back down (until the back touches the backrest) as many times as possible in 30 seconds; only fully completed repetitions are counted.

7.6 Block 4: Cardiorespiratory Health

The 6-Minute Walk Test (6MWT) (Lipkin et al., 1986) assesses global aerobic functional capacity and submaximal exertional tolerance, measured as the maximum distance (m) walked along a flat 30-meter course, marked with cones, in 6 minutes. Participants may stop and rest if fatigued, but the stopwatch is never paused. Standardized verbal encouragement is given at each minute (e.g., "You're doing great, 4 minutes to go"); no other phrasing is used, to avoid biasing pacing. At the end of 6 minutes, the participant stops in place and the exact distance is calculated to 2 decimal places based on the number of completed 30-

meter laps plus any partial final distance. It is recommended to conduct Block 4 on a separate day, or at the end of the assessment session, so that fatigue does not affect the strength and balance tests.

7.7 Composite Fall-Risk Classification

Consistent with the 2025–2026 database ("riesgo caídas" sheet), participants are classified as "high fall risk" (CON RIESGO) if they meet at least one of the following criteria, otherwise "low/no risk" (SIN RIESGO): (i) 1 fall in the previous year requiring medical assistance, or more than 1 fall without medical assistance (Ambrose et al., 2013); (ii) TUG time > 20 seconds (Podsiadlo & Richardson, 1991); or (iii) gait speed over 6 meters > 7.50 seconds, equivalent to < 0.8 m/s (Studenski, 2011; Cesari et al., 2005). This is tracked pre- and post-intervention for every cohort, and its post-intervention prevalence is a secondary outcome (Section 12.4).

7.8 Note on Gait/Walk Test Terminology

Two distinct tests intentionally coexist in this protocol and should not be confused in the CRF, database, or analysis code. The 6-meter gait-speed test ("6 MT", reported in seconds, embedded in the fall-risk criteria above and currently in active use per the 2025–2026 database) is a short, timed-distance walk used to compute gait speed. The 6-Minute Walk Test (6MWT) (Lipkin et al., 1986; Block 4, Section 7.6) is an unrelated, submaximal endurance test reported in meters over a fixed 6-minute duration, and is part of the planned battery, though it does not yet appear in the 2025–2026 completed database. Both tests should be retained as distinct variables (e.g., gaitspeed_6m_sec vs. sixmwt_distance_m) in the database and SAP.

7.9 Cognitive Screening (MMSE)

The Mini-Mental State Examination (MMSE) (Folstein et al., 1975) is administered pre- and post-intervention, as confirmed by the 2025–2026 database, and serves two purposes: (i) describing the cognitive profile of the participants at baseline, since no MMSE threshold is applied for eligibility (Section 5.3), and (ii) providing a secondary, exploratory measure of change in global cognitive status over the program period, given the documented association between cognitive function and both physical function and exercise adherence in older adults (Davis et al., 2021).

8. Safety, Adverse Events, and Risk Management

8.1 Pre-Participation Medical Screening

All participants must undergo mandatory medical assessment by their primary care physician prior to enrollment. Primary care physicians are responsible for evaluating clinical suitability, confirming that safety criteria for physical exercise are met, and formally referring eligible patients to the program.

8.2 Adverse Event Definitions and Classification

An adverse event (AE) is defined as any unfavorable sign, symptom, injury, or illness experienced by a participant during the program period, regardless of whether it is judged related to the exercise program. AEs are classified both by causality and by severity.

Causality

- Exercise-related: any AE occurring during a session or as a direct consequence of prescribed exercises (e.g., a fall in the training room, a muscle strain during a strength exercise).
- Not exercise-related: health events occurring outside program hours/setting, but recorded because they may interrupt attendance (e.g., a fall at home, hospitalization for a pre-existing chronic condition).

Severity

Category	Description	Practical examples
Mild	Minor injuries or discomfort that do not require urgent medical attention and do not prolong-ly interrupt daily activities.	Severe delayed-onset muscle soreness persisting > 48 h; transient post-exertional muscle/joint discomfort interrupting session continuity; mild dizziness or momentary orthostatic hypotension on standing; minor footwear-related blisters/chafing during walking exercises.
Severe	Situations that compromise participant health, require medical attention or hospitalization, or cause temporary or permanent disability.	Accidental falls resulting in fracture, fissure, or head trauma; syncope during or after a session; hypertensive crisis or cardiovascular events (angina, myocardial infarction); severe musculoskeletal injury (muscle tears, severe sprains).

8.3 Adverse Event Reporting Procedures

Reporting pathway.

- Recording: the instructor present records every adverse event on the session sheet on the same day and completes an adverse-event form (participant code, date, description, relation to the exercise, severity as in Section 8.2, action taken and outcome).
- Severe events: emergency services (112) are called when needed, and the adverse-event officers and the participant's primary-care physician are informed within 24 hours.
- Return to exercise: the participant resumes the program only after medical clearance (Section 8.4).
- Review and reporting: the adverse-event officers review the adverse-event log every month; an annual safety summary (number, type, severity and relation to the exercise) is sent to the CEIH-UGR

and to the Ayuntamiento, and severe exercise-related events are communicated to the CEIH-UGR within 7 working days. Events not related to the exercise are recorded but not escalated.

8.4 Stopping Rules

Individual stopping rules. A session is interrupted immediately for a participant who reports chest pain, marked breathlessness, dizziness or near-fainting, palpitations or a new acute pain, and the participant does not continue until assessed. After a severe exercise-related adverse event the participant is suspended from the program until a physician gives written clearance. A participant who develops a condition listed in Section 5.3 (for example unstable cardiovascular disease or uncontrolled hypertension) is withdrawn as described in Section 5.4.

Program-level rule. The principal investigator, the adverse-event officers and the CEIH-UGR review the cycle if a death or a life-threatening event possibly related to the exercise occurs, or if two or more severe exercise-related adverse events occur in the same cycle. Data collection may be paused until the review is complete; the municipal service itself remains under the responsibility of the Ayuntamiento.

9. Data Management

9.1 Data Collection and Entry

In practice, the 2025–2026 cohort's assessment data was compiled in a structured spreadsheet workbook (one sheet per test: SPPB, fall risk, dynamometry, MMSE, attendance, ages), organized by participant ID, with pre/post columns and computed improvement scores. This confirms the current de facto data-collection and entry approach for the annual cohorts.

Data capture. Data are recorded in a structured Excel workbook (one sheet per test, one row per participant code, pre and post columns), with a locked template and a data dictionary shared by all cohorts, so that variable names and units are identical across years. A second member of the team checks the entries against the original paper records, and automated range and consistency checks are run in R before each analysis.

9.2 Data Storage and Security

Scientific direction, statistical treatment, and custody of anonymized data are managed centrally from the University of Granada.

Storage, backup and retention. The Excel workbook is stored on an institutional computer, with a backup copy on the institutional Drive. Data are kept for 10 years after the end of the study and are then deleted. Access to the identifiable dataset is restricted as described in Section 9.3.

9.3 Confidentiality

Participants' data will be pseudonymized/anonymized for research and public-health evaluation purposes, in accordance with the Declaration of Helsinki and applicable Spanish and EU data-protection law (Reglamento General de Protección de Datos, RGPD/GDPR; Ley Orgánica 3/2018, de Protección de Datos Personales y garantía de los derechos digitales, LOPDGDD). Only the coordinating research team will have access to the identifiable dataset linking participant identity to study codes; analyses will be conducted on de-identified data.

10. Ethical Considerations

10.1 Ethical Approval

The study was approved by the Comité de Ética en Investigación Humana of the University of Granada (CEIH-UGR) in Granada on 29 September 2026. The exercise program itself is delivered as a routine municipal health-promotion service by the Ayuntamiento de Santoña and operates independently of this research; the approval covers the collection, analysis, and research use of participant data under this protocol.

10.2 Nature of the Ethical Review and Absence of Clinical Interference

Participants are not clinical patients recruited in a hospital setting but members of the general community; the intervention therefore does not interfere with any necessary medical treatment, and, given the nature of an active physical-exercise, social-domain intervention, no placebo condition of any kind is used. The study does not involve the extraction or processing of any biological samples (blood, tissue, etc.); outcomes are evaluated exclusively through the functional and anthropometric variables described in Section 7, collected at two key timepoints per cohort: week 0 (pre) and week 36 (post).

10.3 Validated Tools and Team Experience

All tools, tests, and methodologies selected are widely validated by the international scientific community for use in older adult populations (see the citations throughout Section 7); scales, stadiometers, and measuring tapes will carry official calibration and CE marking, and the hand dynamometer is a Jamar-type clinical-standard device. The SPPB and 6MWT in particular have high test–retest reliability reported in the literature (Guralnik et al., 1994; Lipkin et al., 1986; Kameniar et al., 2024), and execution is standardized under an identical methodological guide across all evaluators. The Principal Investigator and the attached research team hold higher academic training in Physical Activity and Sport Sciences and have prior experience in functional assessment and in directing exercise sessions for older adults, having conducted these evaluations in previous years (Fernández-Revelles et al., 2022, 2023).

10.4 Compensation, Credits, and Institutional Custody

There is no financial compensation for participants; participation is entirely voluntary and disinterested, with the supervised training program itself being the direct, freely provided benefit. The project is not linked to the award of undergraduate or postgraduate academic credit to participants, so no separate Departmental Council approval is required on that basis. The data-collection and intervention site is the municipal sports facilities ceded by the Ayuntamiento de Santoña under an institutional collaboration agreement for health promotion; the Principal Investigator maintains an active contractual/statutory link with the University of Granada, and scientific direction, statistical treatment, and custody of anonymized data are managed centrally from the UGR.

10.5 Informed Consent

Participation is entirely voluntary. All participants will be informed that their data, collected during routine clinical and physical assessments associated with the municipal program, may be used for research and public-health evaluation purposes, and written informed consent will be obtained prior to data collection for research purposes, consistent with the ~9-month consent period specified for each annual cycle.

Consent procedure. Written informed consent is obtained at enrollment, with one information sheet (Annex 1.1) and one consent form (Annex 1.2) that cover the whole 4-year protocol; both are written in Spanish and are provided as separate documents. Consistent with the 9-month consent period in the selection criteria, consent is renewed at the start of each cycle for participants who re-enroll. Participants can withdraw their consent at any time without any consequence for their participation in the municipal program (Section 5.4).

10.6 Declaration of Helsinki and Data Protection

The study will be conducted in accordance with the ethical principles of the Declaration of Helsinki and will adhere to Spanish and EU data-protection regulations (RGPD/GDPR; LOPDGDD) to ensure participant confidentiality throughout data collection, storage, and analysis.

11. Study Organization

11.1 Research Team and Roles

Name	Affiliation	Role
Javier Fernández-Ortega	University of Granada (iMUDS)	Statistical analysis; manuscript preparation
Andrés B. Fernández-Revelles	University of Granada (iMUDS)	Co-investigator; statistical analysis; manuscript preparation
Juan A. Rangel-García	Ayuntamiento de Santoña	Co-investigator; intervention delivery and physical assessments
Javier Loza-Herbella	Ayuntamiento de Santoña	Co-investigator; intervention delivery and physical assessments

Principal investigator (IP). Andrés B. Fernández-Revelles, of the teaching and research staff of the University of Granada, who holds the scientific direction, the statistical treatment and the custody of the anonymized data.

Adverse-event oversight. Javier Loza-Herbella and Juan A. Rangel-García (Ayuntamiento de Santoña) monitor the adverse events, keep the adverse-event log and report to the principal investigator (Section 8).

Data quality. Javier Fernández-Ortega (UGR) supervises the annual data-quality checks described in Section 9.1.

11.2 Collaborating Institutions

University of Granada (UGR), Department of Physical Education and Sports, Faculty of Sport Sciences, Sport and Health University Research Institute (iMUDS); Ayuntamiento de Santoña (Cantabria, Spain), through its municipal sports/physical-activity service.

11.3 Funding

Funding. The exercise program is organized by the Ayuntamiento de Santoña as a municipal curricular program and is not part of a funded research project. The Ayuntamiento has arranged for the program to be delivered following the guidelines of the scientific literature and the available evidence. No specific research funding supports the data collection or analysis described in this protocol.

Conflicts of interest. The investigators declare no conflicts of interest.

12. Statistical Analysis Plan (SAP)

12.1 General Principles

This SAP is pre-specified to the extent possible given the source materials available and should be finalized and version-locked prior to the start of data collection for Year 1, and prior to database lock for each annual cohort thereafter. All statistical tests will be two-sided, with statistical significance set at $\alpha = 0.05$, unless otherwise specified. Analyses will be performed in R (version 4.5 or later, fixed at the start of each annual analysis). Package versions are locked with renv and the session information is archived with every annual analysis, so that all results can be reproduced. Mixed models use lme4 (Bates et al., 2015) with lmerTest, marginal means use emmeans, and effect sizes use the effectsize and rstatix packages.

12.2 Design Summary for Analysis

The design is a repeated annual-cohort, single-arm, pre–post quasi-experimental design. Each participant is assessed at baseline (T0) and end-of-program (T1) within each annual cycle in which they participate (Year 1, 2, 3, and/or 4, where Year 1 is the 2026/2027 cycle). The 2024–2025 and 2025–2026 pilot cohorts (Fernández-Ortega et al., 2026; Fernández-Revelles et al., 2026, both under review) are analyzed separately as a sensitivity analysis (Section 12.11). Some participants may be assessed in more than one annual cycle (returning participants). Two complementary analytic strategies are therefore specified: (a) a within-year analysis, replicating the approach used in the previously published cohort, conducted separately for each of the 4 annual cohorts; and (b) a pooled, multi-year analysis across all 4 years, using mixed-effects models that account for repeated participation (Section 12.9).

12.3 Analysis Populations

Population	Definition
Intention-to-treat (ITT)	All enrolled, eligible participants with at least one post-baseline assessment for the outcome in question (available-case analysis), consistent with the approach used in the prior published cohort
High-adherence / sensitivity population	Subset of the ITT population who attended > 70% of prescribed sessions in that annual cycle, replicating the prior published sensitivity analysis
Safety population	All enrolled participants who attended at least one session, used for adverse-event summaries

12.4 Outcome Definitions

Outcome	Type	Definition / derivation
SPPB total score	Primary	Sum of balance, gait-speed, and chair-stand sub-scores (0–12); change = T1 − T0
TUG (seconds)	Secondary	Best of standardized trials; change = T1 − T0 (lower is better)
Handgrip strength (kg)	Secondary	Maximum of two trials per hand; change = T1 − T0
30-second Chair Stand (reps)	Secondary	Completed repetitions in 30 s; change = T1 − T0
6MWT distance (m)	Secondary	Total distance walked in 6 minutes; change = T1 − T0
BMI (kg/m²), waist circumference (cm)	Secondary / descriptive	Derived from height, weight, and waist measurements; change = T1 − T0

Outcome	Type	Definition / derivation
IFIS score	Exploratory / descriptive	Self-perceived fitness domains; summarized descriptively, change T1 − T0 where applicable
MMSE score	Exploratory / descriptive	Global cognitive status; change T1 − T0, and described at baseline (no entry threshold)
Composite fall-risk category	Secondary (categorical)	"High risk" vs. "Low risk", as defined in Section 7.7
Frailty category	Secondary (categorical)	Low performance/frailty (SPPB 0–6), intermediate/pre-frailty (7–9), high performance/minimal limitation (10–12)
Adherence (%)	Secondary / covariate	Attended sessions ÷ total scheduled sessions × 100, per annual cycle

12.5 Descriptive Statistics

Continuous variables will be summarized as mean ± SD (or median [IQR] where distributions are non-normal); categorical variables as n (%). A baseline characteristics table will be produced for each annual cohort (replicating the prior published Table 1 format, stratified by fall-risk category), and a pooled baseline table across all 4 years, stratified by year, will also be produced to examine cohort comparability over time.

12.6 Handling of Missing Data

Missing data. The primary analysis follows the intention-to-treat principle in its available-case form: every participant with at least one value of an outcome contributes, and the mixed model of Section 12.9 uses all available observations under a missing-at-random assumption. In previous editions every participant who started completed the post-intervention assessment, so missing end-of-program data are expected to be rare, and the following procedures are specified for completeness. The reason for any missing end-of-program data is recorded (withdrawal, illness, absence at the assessment, other) and the baseline characteristics of completers and non-completers are compared. As sensitivity analyses for the primary outcome, multiple imputation by chained equations (50 imputed data sets; predictors: baseline value, age, sex, adherence and baseline fall risk) is run (Sterne et al., 2009), together with a tipping-point analysis in which non-completers are assumed not to have improved.

12.7 Primary Outcome Analysis

Within each annual cohort, normality of the SPPB pre–post difference will be assessed using the Shapiro-Wilk test. Where normality is not rejected, a paired-samples t-test will be used; where normality is rejected (as observed in the prior cohort for all functional variables), the Wilcoxon signed-rank test will be used, consistent with the previously published analysis. The mean (or median) pre–post difference will be reported with a 95% confidence interval and an appropriate effect size (Cohen's dz for the paired t-test, or matched-pairs rank-biserial correlation for the Wilcoxon test). The proportion of individual participants who improved, remained unchanged, or worsened on the SPPB will also be reported, as in the prior cohort.

12.8 Secondary Outcome Analyses

Continuous secondary outcomes (TUG, handgrip strength, 30-second Chair Stand repetitions, 6MWT distance, BMI, waist circumference, MMSE) will be analyzed using the same paired-testing approach as the primary outcome (Section 12.7), selected per outcome-specific normality testing. Categorical outcomes

(composite fall-risk category, frailty category) will be analyzed using McNemar's test for paired proportions (Mcnemar, 1952), as in the prior published cohort.

12.9 Pooled Multi-Year (4-Year) Analysis

To formally evaluate whether intervention effects are consistent (reproducible) across the four program years, and to make efficient use of repeated observations from returning participants, a pooled analysis across all 4 annual cohorts will be conducted using linear mixed-effects models (LMM), fitted separately for each continuous outcome. The proposed general model structure is:

Outcome ~ Time × Year + Age + Sex + Baseline severity + (1 | Participant ID)

where Time is a two-level within-participant factor (baseline vs. end-of-program), Year is a categorical fixed effect for the annual cohort (Year 1–4), the Time × Year interaction tests whether the magnitude of pre–post change differs across program years, and a random intercept per participant accounts for the non-independence of repeated observations from participants who enroll in more than one annual cycle. Models will be fitted using restricted maximum likelihood (REML), with fixed-effect estimates reported alongside 95% confidence intervals. For categorical outcomes (fall-risk category, frailty category), an equivalent generalized linear mixed model (GLMM) with a binomial/ordinal link, as appropriate, will be used.

Model specification. Year is a fixed effect (four levels; the aim is to compare these specific cohorts, not to generalize to a population of years). Age, sex and baseline frailty category are covariates. A random intercept per participant accounts for repeated participation; a random slope for Time is added only if the data and the convergence allow it, because most participants contribute two observations per cycle. A covariate for the number of previous cycles completed is explored to separate the effect of the current cycle from accumulated participation. Estimates (marginal means, differences with 95% confidence intervals and Satterthwaite degrees of freedom) are obtained with lme4 and lmerTest (Bates et al., 2015). Retention over the summer break is estimated, descriptively given the small expected number of returning participants, as the difference between T0 of year n + 1 and T1 of year n in those participants.

12.10 Subgroup Analyses

Exploratory subgroup analyses (not powered for formal hypothesis testing) will examine whether intervention effects differ by: baseline fall-risk status (high vs. low), sex, age tertiles, and adherence category (e.g., ≤50%, 51–70%, >70%). These will be implemented as interaction terms within the mixed-effects models described in Section 12.9, or as stratified paired analyses within each annual cohort, and interpreted descriptively given the exploratory nature of these comparisons.

12.11 Sensitivity Analyses

- Per-protocol / high-adherence sensitivity analysis restricted to participants with > 70% session attendance, replicating the prior published approach, conducted within each annual cohort.
- Comparison of complete-case vs. any specified imputation approach for the primary outcome.
- Exclusion of returning participants who no longer meet the entry criteria and are allowed to continue under the eligibility rule (Section 5.2), for example those who start the cycle at the SPPB ceiling (score of 12), to assess robustness of findings in the core target population.
- First cycle only. Analysis keeping only the first cycle of each participant, so that every observation comes from a different person.

- Pilot cohorts. The primary and confirmatory analyses use the four formal annual cohorts only (2026/2027 to 2029/2030). The 2024–2025 and 2025–2026 pilot cohorts (Fernández-Ortega et al., 2026; Fernández-Revelles et al., 2026, both under review) are added in a sensitivity analysis, with a pilot indicator and its interaction with Time in the model of Section 12.9, to check that the conclusions do not depend on their inclusion.
- Cognitive status. Exclusion of participants with a baseline MMSE of 17 or less, the usual range of moderate-to-severe impairment, since no MMSE threshold is applied for entry.
- Attendance. Because some participants interrupt their training for a period but still complete the assessments, attendance is analyzed as a continuous covariate (dose-response) and participants with interruptions of four consecutive weeks or more are compared with the rest.

12.12 Multiplicity

Multiplicity. The SPPB is the single primary outcome and is tested at $\alpha = 0.05$ without adjustment. The secondary performance outcomes (TUG, handgrip strength, 30-second chair stand and 6MWT) are adjusted with the Holm procedure (Holm, 1979) within that family. The remaining secondary and exploratory outcomes (anthropometry, IFIS, MMSE, fall-risk and frailty categories, subgroup analyses) are reported with unadjusted p-values and labeled exploratory. Both raw and adjusted p-values are reported.

12.13 Sample Size / Power Justification for the SAP

See Section 5.5. In summary, with the variability observed in the 2025–2026 pilot cohort (Fernández-Ortega et al., 2026, under review; SD of the individual SPPB change 1.38), 17 to 18 evaluable participants are enough for 80% power at the one-point minimal clinically important difference. A cohort of about 54 completers per year (10% loss) detects 0.54 points with 80% power and the pooled sample of about 216 completers detects 0.26 points; with 25% loss the figures are 45 completers and 0.59 points, and 180 completers and 0.29 points. Because only 1 to 5 participants per year are expected to be returning participants (Section 5.5), the adjustment for clustering is expected to be small; the power of the mixed model is recalculated with the observed number of returning participants, from the linkage of participant codes across years.

12.14 Software and Reporting Standards

All analyses will be conducted in R (see Section 12.1). Results will be reported following the TREND statement for nonrandomized evaluations of behavioral and public-health interventions (Des Jarlais et al., 2004), with exact p-values reported to three decimal places (or "< .001"), effect sizes and 95% confidence intervals reported alongside all point estimates, and figures/tables prepared in a format suitable for direct inclusion in manuscripts and annual program reports.

12.15 Interim Analyses

Interim and annual analyses. The annual analysis of each cohort is descriptive and part of the service evaluation. It is reported to the Ayuntamiento and may be published, but it does not trigger decisions to stop or modify the program and its p-values are labeled non-confirmatory. The confirmatory analysis is the pooled analysis after year 4 (Section 12.9).

13. Dissemination Plan

Dissemination plan: (a) an annual summary report to the Ayuntamiento de Santoña after each end-of-program assessment; (b) peer-reviewed publications, at least one for the pooled 4-year analysis and possibly one per cohort, as was done in 2022 and 2023 (Fernández-Revelles et al., 2022, 2023) and is under review for 2024–2025 and 2025–2026 (Fernández-Ortega et al., 2026; Fernández-Revelles et al., 2026, both under review); (c) presentations at scientific congresses on physical activity and ageing; and (d) a summary of the results returned to participants in plain language at the end of each cycle.

14. Study Timeline

The 4 annual cycles correspond to the academic years 2026/2027 to 2029/2030. Exact day-level dates within each October–June cycle are set each year in the municipal sports calendar (see Section 4.3); the table below gives the planned structure.

Phase	2026/2027	2027/2028	2028/2029	2029/2030
Recruitment / referral	Sep.–Oct. 2026	Sep.–Oct. 2027	Sep.–Oct. 2028	Sep.–Oct. 2029
Baseline assessment (T0)	≈ Oct. 2026	≈ Oct. 2027	≈ Oct. 2028	≈ Oct. 2029
Intervention period (~36 weeks)	Oct.–Jun.	Oct.–Jun.	Oct.–Jun.	Oct.–Jun.
End-of-program assessment (T1)	≈ Jun. 2027	≈ Jun. 2028	≈ Jun. 2029	≈ Jun. 2030
Annual data analysis / reporting	Jul.–Sep. 2027	Jul.–Sep. 2028	Jul.–Sep. 2029	Jul.–Sep. 2030
Pooled 4-year analysis	—	—	—	Jul.–Dec. 2030

15. Limitations

- The quasi-experimental, single-arm pre–post design (with no concurrent, non-exercising control group) precludes definitive causal inference; observed changes may partly reflect factors external to the intervention, such as regression to the mean or concurrent lifestyle changes.
- A pronounced ceiling effect on the SPPB has been observed in prior cohorts, as a high proportion of participants enter the program already close to the maximum score (in 2025–2026, 20 of 92 participants, 21.7%, started at the maximum of 12; Fernández-Ortega et al., 2026, under review), which may reduce sensitivity to detect further improvement in higher-functioning participants; the SPPB's concordance with alternative performance measures such as a self-paced long-distance walk is itself only moderate (Sayers et al., 2006), and self-reported fitness (IFIS) shows only moderate concordance with performance-based testing (Mayhew et al., 2020), so triangulating across instruments is advisable.
- Assessors are not blinded to participants' intervention status, although standardized testing protocols are followed at both timepoints.
- The 2024–2025 and 2025–2026 pilot cohorts (Fernández-Ortega et al., 2026; Fernández-Revelles et al., 2026, both under review) were run before this protocol and under routine program-evaluation conditions, so they are used only as preliminary evidence and in sensitivity analyses, and not as part of the confirmatory sample.

- Repeated enrollment of the same participants across multiple annual cycles introduces non-independence in the pooled 4-year dataset, addressed statistically via mixed-effects models (Section 12.9) but still limiting the interpretation of "Year" effects as fully independent replications.
- As a real-world, service-embedded program, adverse-event ascertainment and non-completion reasons have not historically been systematically recorded in full detail (see Section 8), which may limit safety-related conclusions unless the reporting procedures in Section 8.3 are formalized; self-managed and community-based falls-prevention exercise programs more broadly face known feasibility and outcome-ascertainment challenges (Mansson et al., 2022; Company-Sancho et al., 2021).

16. References

Aartolahti, E., Tolppanen, A. M., Lönnroos, E., Hartikainen, S., & Häkkinen, A. (2015). Health condition and physical function as predictors of adherence in long-term strength and balance training among community-dwelling older adults. *Archives of Gerontology and Geriatrics*, *61*(3), 452–457. https://doi.org/10.1016/j.archger.2015.06.016

Ambrose, A. F., Paul, G., & Hausdorff, J. M. (2013). Risk factors for falls among older adults: A review of the literature. *Maturitas*, *75*(1), 51–61. https://doi.org/10.1016/j.maturitas.2013.02.009

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Beard, J. R., Officer, A., De Carvalho, I. A., Sadana, R., Pot, A. M., Michel, J. P., Lloyd-Sherlock, P., Epping-Jordan, J. E., Peeters, G. M. E. E., Mahanani, W. R., Thiyagarajan, J. A., & Chatterji, S. (2016). The World report on ageing and health: A policy framework for healthy ageing. *The Lancet*, *387*(10033), 2145–2154. https://doi.org/10.1016/S0140-6736(15)00516-4

Bergland, A., Jørgensen, L., Emaus, N., & Strand, B. H. (2017). Mobility as a predictor of all-cause mortality in older men and women: 11.8 year follow-up in the Tromsø study. *BMC Health Services Research*, *17*(1), Article 22. https://doi.org/10.1186/s12913-016-1950-0

Bergland, A., & Strand, B. H. (2019). Norwegian reference values for the Short Physical Performance Battery (SPPB): The Tromsø Study. *BMC Geriatrics*, *19*(1), Article 216. https://doi.org/10.1186/s12877-019-1234-8

Bohrer, R. C. D., Pereira, G., Beck, J. K., Lodovico, A., & Rodacki, A. L. F. (2019). Multicomponent training program with high-speed movement execution of ankle muscles reduces risk of falls in older adults. *Rejuvenation Research*, *22*(1), 43–50. https://doi.org/10.1089/rej.2018.2063

Borg, G. A. V. (1982). Psychophysical bases of perceived exertion. *Medicine & Science in Sports & Exercise*, *14*(5), 377–381.

Brach, J. S., Studenski, S. A., Perera, S., VanSwearingen, J. M., & Newman, A. B. (2007). Gait variability and the risk of incident mobility disability in community-dwelling older adults. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *62*(9), 983–988. https://doi.org/10.1093/gerona/62.9.983

Cabrero-García, J., Muñoz-Mendoza, C. L., Cabañero-Martínez, M. J., González-Llopís, L., Ramos-Pichardo, J. D., & Reig-Ferrer, A. (2012). Valores de referencia de la Short Physical Performance Battery para pacientes de 70 y más años en atención primaria de salud. *Atención Primaria*, *44*(9), 540–548. https://doi.org/10.1016/j.aprim.2012.02.007

Casas Herrero, A., & Izquierdo, M. (2012). Ejercicio físico como intervención eficaz en el anciano frágil. *Anales del Sistema Sanitario de Navarra*, *35*(1), 69–85. https://doi.org/10.4321/S1137-66272012000100007

Cavanaugh, E. J., Richardson, J., McCallum, C. A., & Wilhelm, M. (2018). The predictive validity of physical performance measures in determining markers of preclinical disability in community-dwelling middle-aged and older adults: A systematic review. *Physical Therapy*, *98*(12), 1010–1021. https://doi.org/10.1093/ptj/pzy109

Cesari, M., Kritchevsky, S. B., Penninx, B. W. H. J., Nicklas, B. J., Simonsick, E. M., Newman, A. B., Tylavsky, F. A., Brach, J. S., Satterfield, S., Bauer, D. C., Visser, M., Rubin, S. M., Harris, T. B., & Pahor, M. (2005). Prognostic value of usual gait speed in well-functioning older people: Results from the Health,

Aging and Body Composition Study. *Journal of the American Geriatrics Society*, *53*(10), 1675–1680. https://doi.org/10.1111/j.1532-5415.2005.53501.x

Chu, W., Chang, S., & Ho, H. (2021). Adverse health effects of frailty: Systematic review and meta-analysis of middle-aged and older adults with implications for evidence-based practice. *Worldviews on Evidence-Based Nursing*, *18*(4), 282–289. https://doi.org/10.1111/wvn.12508

Company-Sancho, M. C., Alonso-Poncelas, E., Rich-Ruiz, M., Cidoncha-Moreno, M. Á., Gonzalez-Pisano, A., Abad-Corpa, E., & the Otago Programme Working Group. (2021). The relation between functional performance, falls and previous falls among participants in the Otago Programme: A secondary data analysis. *International Journal of Environmental Research and Public Health*, *18*(12), Article 6501. https://doi.org/10.3390/ijerph18126501

Cruz-Jentoft, A. J., Bahat, G., Bauer, J., Boirie, Y., Bruyère, O., Cederholm, T., Cooper, C., Landi, F., Rolland, Y., Sayer, A. A., Schneider, S. M., Sieber, C. C., Topinkova, E., Vandewoude, M., Visser, M., & Zamboni, M. (2019). Sarcopenia: Revised European consensus on definition and diagnosis. *Age and Ageing*, *48*(1), 16–31. https://doi.org/10.1093/ageing/afy169

Davis, J. C., Rhodes, R. E., Khan, K. M., Mansournia, M. A., Khosravi, A., Chan, P., Zhao, M., Jehu, D. A., & Liu-Ambrose, T. (2021). Cognitive function and functional mobility predict exercise adherence in older adults who fall. *Gerontology*, *67*(3). https://doi.org/10.1159/000513452

Davis, J. C., Robertson, M. C., Ashe, M. C., Liu-Ambrose, T., Khan, K. M., & Marra, C. A. (2010). International comparison of cost of falls in older adults living in the community: A systematic review. *Osteoporosis International*, *21*(8), 1295–1306. https://doi.org/10.1007/s00198-009-1162-0

De Fátima Ribeiro Silva, C., Ohara, D. G., Matos, A. P., Pinto, A. C. P. N., & Pegorari, M. S. (2021). Short Physical Performance Battery as a measure of physical performance and mortality predictor in older adults: A comprehensive literature review. *International Journal of Environmental Research and Public Health*, *18*(20), Article 10612. https://doi.org/10.3390/ijerph182010612

Des Jarlais, D. C., Lyles, C., Crepaz, N., & the TREND Group. (2004). Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: The TREND statement. *American Journal of Public Health*, *94*(3), 361–366. https://doi.org/10.2105/AJPH.94.3.361

Farrance, C., Tsofliou, F., & Clark, C. (2016). Adherence to community based group exercise interventions for older people: A mixed-methods systematic review. *Preventive Medicine*, *87*, 155–166. https://doi.org/10.1016/j.ypmed.2016.02.037

Fernández-Ortega, J., Loza-Herbella, J., Rangel-García, J. A., & Fernández-Revelles, A. B. (2026). *Restoring physical resilience: Consistent functional and sarcopenia gains across two Vivifrail cohorts.* Manuscript under review at Maturitas.

Fernández-Revelles, A. B., Rangel-García, J. A., Fernández-Ortega, J., & Loza-Herbella, J. (2026). *Long-term multicomponent exercise on physical function and sarcopenia in older adults: A quasi-experimental analysis.* Manuscript under review at Apunts Educación Física y Deportes.

Fernández-Revelles, A. B., Rangel-García, J. A., & Loza-Herbella, J. (2022). Effect of 32 weeks of multicomponent physical exercise for the prevention of fragility in people over 65. *Journal of Sport and Health Research*, *14*, 45–56.

Fernández-Revelles, A. B., Rangel García, J. A., & Loza Herbella, J. (2023). Effect of 36 weeks of multicomponent physical exercise for the prevention of fragility in people over 65. *Journal of Sport and Health Research*, *15*(Supl. 1). https://doi.org/10.58727/jshr.102762

Fielding, R. A., Katula, J., Miller, M. E., Abbott-Pillola, K., Jordan, A., Glynn, N. W., Goodpaster, B., Walkup, M. P., King, A. C., & Rejeski, W. J. (2007). Activity adherence and physical function in older adults with functional limitations. *Medicine & Science in Sports & Exercise*, *39*(11), 1997–2004. https://doi.org/10.1249/mss.0b013e318145348d

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Fried, L. P., Tangen, C. M., Walston, J., Newman, A. B., Hirsch, C., Gottdiener, J., Seeman, T., Tracy, R., Kop, W. J., Burke, G., & McBurnie, M. A. (2001). Frailty in older adults: Evidence for a phenotype. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *56*(3), M146–M157. https://doi.org/10.1093/gerona/56.3.M146

Ganz, D. A., & Latham, N. K. (2020). Prevention of falls in community-dwelling older adults. *New England Journal of Medicine*, *382*(8), 734–743. https://doi.org/10.1056/NEJMcp1903252

Gawel, J., Vengrow, D., Collins, J., Brown, S., Buchanan, A., & Cook, C. (2012). The short physical performance battery as a predictor for long term disability or institutionalization in the community dwelling population aged 65 years old or older. *Physical Therapy Reviews*, *17*(1), 37–44. https://doi.org/10.1179/1743288X11Y.0000000050

Guralnik, J. M., Simonsick, E. M., Ferrucci, L., Glynn, R. J., Berkman, L. F., Blazer, D. G., Scherr, P. A., & Wallace, R. B. (1994). A short physical performance battery assessing lower extremity function: Association with self-reported disability and prediction of mortality and nursing home admission. *Journal of Gerontology*, *49*(2), M85–M94. https://doi.org/10.1093/geronj/49.2.m85

Heinrich, S., Rapp, K., Rissmann, U., Becker, C., & König, H. H. (2010). Cost of falls in old age: A systematic review. *Osteoporosis International*, *21*(6), 891–902. https://doi.org/10.1007/s00198-009-1100-1

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Hortobágyi, T., Lesinski, M., Gäbler, M., VanSwearingen, J. M., Malatesta, D., & Granacher, U. (2015). Effects of three types of exercise interventions on healthy old adults' gait speed: A systematic review and meta-analysis. *Sports Medicine*, *45*(12), 1627–1643. https://doi.org/10.1007/s40279-015-0371-2

Hurvitz, E. A., Richardson, J. K., Werner, R. A., Ruhl, A. M., & Dixon, M. R. (2000). Unipedal stance testing as an indicator of fall risk among older outpatients. *Archives of Physical Medicine and Rehabilitation*, *81*(5), 587–591.

Izquierdo, M. (2019). Prescripción de ejercicio físico. El programa Vivifrail como modelo. *Nutrición Hospitalaria*. https://doi.org/10.20960/nh.02680

Izquierdo, M., Merchant, R. A., Morley, J. E., Anker, S. D., Aprahamian, I., Arai, H., … Singh, M. F. (2021). International exercise recommendations in older adults (ICFSR): Expert consensus guidelines. *The Journal of Nutrition, Health & Aging*, *25*(7), 824–853. https://doi.org/10.1007/s12603-021-1665-8

Kameniar, K., Mackintosh, S., Van Kessel, G., & Kumar, S. (2024). The psychometric properties of the Short Physical Performance Battery to assess physical performance in older adults: A systematic review. *Journal of Geriatric Physical Therapy*, *47*(1), 43–54. https://doi.org/10.1519/JPT.0000000000000337

Kaushal, N., Langlois, F., Desjardins-Crépeau, L., Hagger, M. S., & Bherer, L. (2019). Investigating dose–response effects of multimodal exercise programs on health-related quality of life in older adults. *Clinical Interventions in Aging*, *14*, 209–217. https://doi.org/10.2147/CIA.S187534

Kwon, S., Perera, S., Pahor, M., Katula, J., King, A., Groessl, E., & Studenski, S. (2009). What is a meaningful change in physical performance? Findings from a clinical trial in older adults (the LIFE-P study). *The Journal of Nutrition, Health & Aging*, *13*(6), 538–544. https://doi.org/10.1007/s12603-009-0104-z

Laffon De Mazières, C., Morley, J. E., Levy, C., Agenes, F., Barbagallo, M., Cesari, M., Chavant, J., Chicoulaa, B., De Souto Barreto, P., Fitten, L. J., Chang Fu, L., Iolascon, G., Izquierdo, M., Jaenicke, T., Kanis, J. A., Kraenzlin, M., Radermecker, R., Sinclair, A., Souza, M. Á., … Vellas, B. (2017). Prevention of functional decline by reframing the role of nursing homes? *Journal of the American Medical Directors Association*, *18*(2), 105–110. https://doi.org/10.1016/j.jamda.2016.11.019

Lee, J. M. L., Sung, P., Tan, K. Y., Duncan, P., Sim, R., Koh, K., & Chan, A. (2024). Trajectories of adherence to a falls prevention exercise intervention in Singapore: A mixed methods study. *Aging and Health Research*, *4*(1), Article 100181. https://doi.org/10.1016/j.ahr.2024.100181

Lipkin, D. P., Scriven, A. J., Crake, T., & Poole-Wilson, P. A. (1986). Six minute walking test for assessing exercise capacity in chronic heart failure. *British Medical Journal (Clinical Research Edition)*, *292*(6521), 653–655. https://doi.org/10.1136/bmj.292.6521.653

Liu, M., & Miyawaki, C. E. (2020). What types of physical function predict program adherence in older adults? *Rehabilitation Nursing*, *45*(5), 279–286. https://doi.org/10.1097/rnj.0000000000000209

Mahoney, F. I., & Barthel, D. W. (1965). Functional evaluation: The Barthel Index. *Maryland State Medical Journal*, *14*, 61–65.

Mansson, L., Pettersson, B., Rosendahl, E., Skelton, D. A., Lundin-Olsson, L., & Sandlund, M. (2022). Feasibility of performance-based and self-reported outcomes in self-managed falls prevention exercise interventions for independent older adults living in the community. *BMC Geriatrics*, *22*(1), Article 147. https://doi.org/10.1186/s12877-022-02851-9

Máximo, R. D. O., Luiz, M. M., Lima, S. S., Steptoe, A., Oliveira, C. D., & Alexandre, T. D. S. (2025). Short Physical Performance Battery or chair stand: Which better predicts disability among high-functioning older adults? *Journal of the American Medical Directors Association*, *26*(8), Article 105720. https://doi.org/10.1016/j.jamda.2025.105720

Mayhew, A. J., Griffith, L. E., Gilsing, A., Beauchamp, M. K., Kuspinar, A., & Raina, P. (2020). The association between self-reported and performance-based physical function with activities of daily living disability in the Canadian Longitudinal Study on Aging. *The Journals of Gerontology: Series A*, *75*(1), 147–154. https://doi.org/10.1093/gerona/glz122

McNemar, Q. (1952). Statistical theory and research design. *Annual Review of Psychology*, *3*. https://doi.org/10.1146/annurev.ps.03.020152.002205

McPhate, L., Simek, E. M., & Haines, T. P. (2013). Program-related factors are associated with adherence to group exercise interventions for the prevention of falls: A systematic review. *Journal of Physiotherapy*, *59*(2), 81–92. https://doi.org/10.1016/S1836-9553(13)70160-7

Merellano-Navarro, E., Collado-Mateo, D., García-Rubio, J., Gusi, N., & Olivares, P. R. (2017). Validity of the International Fitness Scale “IFIS” in older adults. *Experimental Gerontology*, *95*, 77–81. https://doi.org/10.1016/j.exger.2017.05.001

Millan-Domingo, F., Tarazona-Santabalbina, F. J., Carretero, A., Olaso-Gonzalez, G., Viña, J., & Gomez-Cabrera, M. C. (2022). Real-life outcomes of a multicomponent exercise intervention in community-

dwelling frail older adults and its association with nutritional-related factors. *Nutrients*, *14*(23), Article 5147. https://doi.org/10.3390/nu14235147

Perera, S., Mody, S. H., Woodman, R. C., & Studenski, S. A. (2006). Meaningful change and responsiveness in common physical performance measures in older adults. *Journal of the American Geriatrics Society*, *54*(5), 743–749. https://doi.org/10.1111/j.1532-5415.2006.00701.x

Podsiadlo, D., & Richardson, S. (1991). The Timed "Up & Go": A test of basic functional mobility for frail elderly persons. *Journal of the American Geriatrics Society*, *39*(2), 142–148. https://doi.org/10.1111/j.1532-5415.1991.tb01616.x

Rockwood, K., & Mitnitski, A. (2007). Frailty in relation to the accumulation of deficits. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *62*(7), 722–727. https://doi.org/10.1093/gerona/62.7.722

Room, J., Hannink, E., Dawes, H., & Barker, K. (2017). What interventions are used to improve exercise adherence in older people and what behavioural techniques are they based on? A systematic review. *BMJ Open*, *7*(12), Article e019221. https://doi.org/10.1136/bmjopen-2017-019221

Sáez De Asteasu, M. L., Martínez-Velilla, N., Zambom-Ferraresi, F., Casas-Herrero, Á., Lucía, A., Galbete, A., & Izquierdo, M. (2019). Physical exercise improves function in acutely hospitalized older patients: Secondary analysis of a randomized clinical trial. *Journal of the American Medical Directors Association*, *20*(7), 866–873. https://doi.org/10.1016/j.jamda.2019.04.001

Sayers, S. P., Guralnik, J. M., Newman, A. B., Brach, J. S., & Fielding, R. A. (2006). Concordance and discordance between two measures of lower extremity function: 400 meter self-paced walk and SPPB. *Aging Clinical and Experimental Research*, *18*(2), 110–115. https://doi.org/10.1007/BF03327424

Serra-Prat, M., & Palomera, E. (2019). Muscle strength, sarcopenia and frailty associations with balance and gait parameters: A cross-sectional study. *European Journal of Geriatrics and Gerontology*, *1*(2), 61–66. https://doi.org/10.4274/ejgg.galenos.2019.99

Silva, N. L., Oliveira, R. B., Fleck, S. J., Leon, A. C., & Farinatti, P. (2014). Influence of strength training variables on strength gains in adults over 55 years-old: A meta-analysis of dose–response relationships. *Journal of Science and Medicine in Sport*, *17*(3), 337–344. https://doi.org/10.1016/j.jsams.2013.05.009

Son, K. Y., Shin, D. W., Lee, J. E., Kim, S. H., Yun, J. M., & Cho, B. (2022). Association between timed up and go test and future incidence of disability: A nationwide representative longitudinal study in Korea. *PLOS ONE*, *17*(7), Article e0270808. https://doi.org/10.1371/journal.pone.0270808

Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Wood, A. M., & Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. *BMJ*, *338*, Article b2393. https://doi.org/10.1136/bmj.b2393

Studenski, S. (2011). Gait speed and survival in older adults. *JAMA*, *305*(1), 50–58. https://doi.org/10.1001/jama.2010.1923

Syddall, H., Cooper, C., Martin, F., Briggs, R., & Aihie Sayer, A. (2003). Is grip strength a useful single marker of frailty? *Age and Ageing*, *32*(6), 650–656. https://doi.org/10.1093/ageing/afg111

Tou, N. X., Wee, S. L., Seah, W. T., Ng, D. H. M., Pang, B. W. J., Lau, L. K., & Ng, T. P. (2021). Effectiveness of community-delivered functional power training program for frail and pre-frail community-dwelling older adults: A randomized controlled study. *Prevention Science*, *22*(8), 1048–1059. https://doi.org/10.1007/s11121-021-01221-y

Vasunilashorn, S., Coppin, A. K., Patel, K. V., Lauretani, F., Ferrucci, L., Bandinelli, S., & Guralnik, J. M. (2009). Use of the Short Physical Performance Battery score to predict loss of ability to walk 400 meters: Analysis from the InCHIANTI Study. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *64A*(2), 223–229. https://doi.org/10.1093/gerona/gln022

Woo, J. (2019). 69 Falls: A manifestation of frailty and sarcopenia. *Age and Ageing*, *48*(Supplement_4), iv18–iv27. https://doi.org/10.1093/ageing/afz164.69

World Health Organization. (2011). *Waist circumference and waist–hip ratio: Report of a WHO expert consultation, Geneva, 8–11 December 2008.* World Health Organization. https://www.who.int/publications/i/item/9789241501491

Wu, S., Nan, J., Chang, J., Jiang, D., Cao, Z., Zhou, S., Feng, H., & Xiao, L. D. (2025). Adherence to exercise intervention for community-dwelling older adults with sarcopenia: A systematic review and meta-analysis. *Age and Ageing*, *54*(4), Article afaf094. https://doi.org/10.1093/ageing/afaf094

Yang, X., Li, S., Xu, L., Liu, H., Li, Y., Song, X., Bao, J., Liao, S., Xi, Y., & Guo, G. (2024). Effects of multicomponent exercise on frailty status and physical function in frail older adults: A meta-analysis and systematic review. *Experimental Gerontology*, *197*, Article 112604. https://doi.org/10.1016/j.exger.2024.112604

Supplementary Material

Appendix S1: Sample Training Session

The following is a fully worked example of a single training session, adapted and translated from the program's session-record template, illustrating the level of detail with which each of the Vivifrail functional-level sessions actually delivered in the program should be described (per Section 6.2, one such fully detailed sample session should be documented for each training modality/functional level used in the program).

Field	Detail
Facility	Mini pabellón de las Instalaciones Deportivas Municipales (IDM), Santoña
Equipment	Chairs; anti-stress rubber balls; elastic resistance bands; weighted ankle cuffs
Session goals	Improve upper- and lower-body strength, balance, joint mobility, and muscular flexibility

Warm-up (≈10 minutes, 8–10 repetitions per exercise)

Performed gently to progressively raise body temperature, activate the musculature, and prepare the joints:

- Neck (flexion-extension, gentle rotations, tilts).
- Shoulders (arm swings backward and forward, lateral swings, circles).
- Elbows (flexion-extension, arm circles).
- Wrists (flexion-extension, rotations).
- Spine (flexion-extension, lateral tilts, trunk rotations).
- Hips (flexion-extension, circles).
- Knees (flexion-extension).
- Ankles (dorsiflexion, plantarflexion, circles).

Main Part: Balance and Gait (2 sets, with ≈1-minute rest between sets)

Exercise	Description	Volume
Single-leg stance	Next to the chair, flex one leg and hold for 5 s; alternate legs.	2 × 10 reps/leg (5-s hold)
Heel-toe standing	Balance a few seconds on the toes, then on the heels.	2 × 10 reps, few seconds each
Weight transfers	Shift body weight straight toward one leg while raising the opposite heel, then to the other side.	2 × 10 reps/side
Straight-line heel-to-toe walking	Near a wall/handrail, walk placing the heel of one foot against the toes of the other.	2 × 10 steps/leg

Main Part: Upper-Body Strength (2 sets × 10 reps)

- Loaded arm flexion-extension: with an elastic band under both feet, grip both ends at knee height and flex the arms toward the shoulders, keeping wrists firm and elbows close to the torso.
- Hand grip: squeeze a rubber ball progressively as hard as possible, relax, and repeat with the other hand. Progression: towel-wringing exercise once the ball exercise is mastered.

- Horizontal arm opening with resistance: holding the band at chest height with both ends, open the arms until the elbows are fully extended.

Main Part: Lower-Body Strength (2 sets × 10 reps)

- Leg flexion without chair: standing behind a table, squat down (flexing hips and knees as if sitting) and return to standing. Variant: same movement starting from seated, using the arms on the thighs or a chair only if strictly necessary.
- Hip abduction: supported on a chair with a straight back, move one leg away from the other without bending the knee, feet facing forward; repeat with the other leg. Progress using weighted ankle cuffs.
- Hip extension / hamstring: supported with arms on the chair and a straight back, flex the knee bringing the foot backward, then return; repeat with the other leg. Progress using weighted ankle cuffs.
- Calf raises: supported on a chair, feet shoulder-width apart, rise onto the toes as high as possible and hold for 3 s, then lower slowly until the heels touch the floor.

Cool-Down / Stretching (≈10 minutes; 2 sets × 3 reps, 10-s hold each)

- Neck: turn the head to the right until slight tension is felt in the neck muscles, hold 10 s, then repeat to the left.
- Arms: raise the arms overhead with hands clasped as if reaching for the ceiling; hold 10 s, relax for 5 s, repeat.
- Shoulders: place one hand on the opposite shoulder with the elbow against the chest; with the other hand, push the elbow toward the opposite hand until tension is felt in the posterior shoulder; repeat on the other side.
- Posterior thigh: seated, extend one leg with the heel on the floor, gently pull the foot upward, place both hands on the opposite knee, and lean the trunk forward until tension is felt in the lower back/posterior thigh; repeat with the other leg.
- Anterior thigh: supported on a chair, bend one leg with the hand assisting the flexion until tension is felt in the front of the thigh; repeat with the other leg.
- Ankles: flex the foot upward, bringing the toes toward the body, hold 10 s, pause, then point the foot in the opposite direction; repeat with the other foot.

Safety Observations (from the source session record)

Breathing should be paced during strength exercises, inhaling before the effort and exhaling during it. Loads should be determined in advance. Stretches should never be performed with violence or abrupt movements. In every exercise requiring a chair for sitting or support, its use is mandatory to prevent falls.

Appendix S2: Borg RPE Scale (0 to 10), adapted version used by the program

The program uses a 0 to 10 scale with verbal anchors, colour bands and facial icons (Annex 1 of the assessment protocol, reproduced below), which is easier to use with older adults than the original Borg scale. The target zone for strength work is 5 to 7 (Section 6.2).

Rating	Descriptor
0	Descanso (Rest)
1	Extremadamente fácil (Extremely easy)
2	Muy fácil (Very easy)
3	Fácil (Easy)
4	Algo fácil (Somewhat easy)
5	Duro (Hard)
6	Algo difícil (Somewhat hard)
7	Muy duro (Very hard)
8	Realmente muy difícil (Really very hard)
9	Extremadamente duro (Extremely hard)
10	Máximo (Maximal)



Figure S2. Adapted Borg scale used in the program (Annex 1 of the assessment protocol).

Appendix S3: Adverse Event Severity Classification

(Reproduced from Section 8.2 for standalone reference in data-collection materials.)

Category	Description	Practical examples
Mild	Minor injuries or discomfort that do not require urgent medical attention and do not prolong-ly interrupt daily activities.	Severe muscle soreness persisting > 48 h; transient post-exertional discomfort; mild dizziness/orthostatic hypotension on standing; minor blisters/chafing.
Severe	Situations that compromise participant health, require medical attention/hospitalization, or cause temporary or permanent disability.	Falls resulting in fracture or head trauma; syncope; hypertensive crisis or cardiovascular events; severe musculoskeletal injury.

Appendix S4: International Fitness Scale (IFIS), Spanish version used by the program

The questionnaire is read aloud by the assessor, who makes sure that the participant understands that the answers refer to people of the same age (Section 7.3). The five items are scored from 1 (very poor) to 5 (very good).

Por favor, piensa sobre tu nivel de condición física (comparado con personas de tu edad) y elige la opción más adecuada.

1. Mi condición física general es:

- ☐ Muy mala (1)
- ☐ Mala (2)
- ☐ Aceptable (3)
- ☐ Buena (4)
- ☐ Muy buena (5)

2. Mi condición física cardio-respiratoria (capacidad para hacer ejercicio, por ejemplo, correr durante mucho tiempo) **es:**

- ☐ Muy mala (1)
- ☐ Mala (2)
- ☐ Aceptable (3)
- ☐ Buena (4)
- ☐ Muy buena (5)

3. Mi fuerza muscular es:

- ☐ Muy mala (1)
- ☐ Mala (2)
- ☐ Aceptable (3)
- ☐ Buena (4)
- ☐ Muy buena (5)

4. Mi velocidad / agilidad es:

- ☐ Muy mala (1)
- ☐ Mala (2)
- ☐ Aceptable (3)
- ☐ Buena (4)
- ☐ Muy buena (5)

5. Mi flexibilidad es:

- ☐ Muy mala (1)
- ☐ Mala (2)
- ☐ Aceptable (3)
- ☐ Buena (4)
- ☐ Muy buena (5)

Figure S4. IFIS questionnaire as administered (Figure 1 of the assessment protocol).