Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Effects of sprint-embedded 3 vs 3 small-sided games on neuromuscular performance, physiological responses, external load and technical actions in female football players: a randomised controlled trial

Bilal Gök[1]

bigok@gelisim.edu.tr

İstanbul Gelişim Üniversitesi

Fehmi Tuncel[1]

İstanbul Gelişim Üniversitesi

Research Article

Keywords: Women's football, Small-sided games, External load, GNSS, Countermovement jump, Rating of perceived exertion

Posted Date: October 27th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11192198/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Coaches often add running bouts to small-sided games (SSGs), but evidence in female players over several weeks is limited. This randomised controlled trial compared 3 vs 3 SSGs with and without an embedded 10-s sprint on jump performance, physiological and perceptual responses, external load and technical actions.

Methods

Twelve licensed female players aged 18–25 years were randomly allocated to sprint-embedded SSGs (EXP, n = 6), in which a 10-s sprint separated two 2-min bouts in each set, or to standard SSGs with 10 s of passive rest (CON, n = 6). Both groups played three sets once a week for 6 weeks on an 18 × 30 m pitch. The primary outcome was countermovement jump (CMJ) height at baseline, week 3 and week 6. Secondary outcomes were GNSS-derived external load, heart rate, rating of perceived exertion (RPE), video-coded technical actions and VO_2max estimated from the 30 − 15 Intermittent Fitness Test (baseline and week 6). Group × time interactions were tested with two-way mixed ANOVA.

Results

CMJ height increased in both groups, with no evidence of a group × time interaction ($F(2,20) = 2.61$, $p = 0.100$, $\eta p^2 = 0.21$; baseline-adjusted difference 1.5 cm, 95% CI − 0.3 to 3.4). Total distance, accelerations, PlayerLoad, RPE and distance in speed zones 2–4 increased over time in both groups ($p \leq 0.001$). Of 19 secondary interaction tests, one was nominally significant: distance covered at 12.1–15.4 $km \cdot h^{-1}$ increased more in EXP ($F(2,20) = 4.99$, $p = 0.041$, $\eta p^2 = 0.33$; difference in change 44.8 m, 95% CI 2.7 to 86.9). This result did not survive adjustment for multiple comparisons or a non-parametric check ($p = 0.093$).

Conclusions

In this small trial, adding a 10-s sprint to 3 vs 3 SSGs once a week for 6 weeks was associated with changes that were largely similar to those of standard SSGs. The possible additional increase in moderate- to high-speed running is exploratory. Because the sample was small and players in each group shared one game, larger cluster-randomised trials are needed.

Trial registration:

ClinicalTrials.gov, NCT07845318. Registered 7 May 2026 – Retrospectively registered.

Background

Small-sided games (SSGs) are a central part of football training because they combine technical, tactical and conditioning demands in a single drill [1, 2]. Coaches control the intensity of SSGs by changing the number of players, pitch dimensions, rules and work-to-rest structure [1, 3]. The relative pitch area per player, for example, alters heart rate responses and running demands [4].

Small pitches rarely allow players to reach high running speeds, so high-speed running and sprinting tend to be under-represented in SSGs compared with match play [4, 5]. For this reason, SSGs have been compared with repeated-sprint training [6] and combined with short running drills [7]. Köklü et al. reported that adding linear running bouts to 3 vs 3 and 4 vs 4 games increased internal and external loads in young players [7]. Most of this evidence, however, comes from male players and from single sessions.

Evidence from female players is more limited, although pitch size and playing surface have been shown to affect their physiological and perceptual responses to SSGs [8]. Methodological guidance has also called for more research that includes women as participants [9]. Whether a short sprint embedded in SSGs changes the responses of female players over several weeks has not been examined.

The aim of this study was therefore to compare 3 vs 3 SSGs with and without an embedded 10-s sprint, performed once a week for 6 weeks, on countermovement jump (CMJ) height, physiological and perceptual responses, external load and technical actions in female football players. We hypothesised that the sprint-embedded format would produce greater increases in high-speed running and physiological load than the standard format.

Methods

Study design

This was a two-arm, parallel-group randomised controlled trial with measurements at pre-, mid- and post-test. It is reported in line with the CONSORT 2010 statement [10]. The study was carried out at the Performance Measurement and Evaluation Laboratory of Istanbul Gelişim University (Istanbul, Türkiye) between 27 February and 10 May 2024. The trial was registered retrospectively at ClinicalTrials.gov (NCT07845318) on 7 May 2026. The study was conducted as part of a doctoral thesis and was approved by the institutional ethics committee before recruitment; registration was completed retrospectively, before submission for publication. The trial protocol is available in the ClinicalTrials.gov record.

Participants

Twelve female football players aged 18–25 years volunteered for the study. All had played in licensed Turkish Football Federation women's leagues for at least three years. Players were included if they were

aged 18–25 years, had at least three years of licensed playing experience, had no health problem that would prevent training during the programme, were able to attend one session per week and took part voluntarily. Players were excluded if they had less than three years of licensed experience, had an injury or health problem, or missed more than two training sessions.

Before the study, players were informed about its aims, procedures, possible risks and their right to withdraw, and all gave written informed consent. The study was approved by the Ethics Committee of Istanbul Gelişim University (İstanbul Gelişim Üniversitesi Rektörlüğü Etik Kurulu; meeting no. 2024-01, 12 January 2024; decision no. 2024-01-107) and was conducted in accordance with the Declaration of Helsinki [11]. Participant characteristics are shown in Table 1.

Table 1
Participant characteristics at baseline

Variable	EXP (n = 6)	CON (n = 6)
Age (years)	21.8 ± 1.5	20.0 ± 2.1
Height (cm)	162.7 ± 4.5	166.0 ± 10.0
Body mass (kg)	52.9 ± 1.4	58.5 ± 14.0
Body fat (%)	21.7 ± 5.3	19.5 ± 5.4
Fat-free mass (kg)	41.5 ± 3.9	46.9 ± 10.9
Skeletal muscle mass (kg)	22.8 ± 2.1	25.6 ± 6.3
V_IFT ($km \cdot h^{-1}$)	19.8 ± 2.1	18.5 ± 1.9

Values are mean ± SD. Baseline differences were not tested for significance. EXP: experimental group; CON: control group; V_IFT: final velocity of the 30 − 15 Intermittent Fitness Test.

Sample size

The sample size was limited by the number of eligible players available. A sensitivity analysis for the group × time interaction was performed following the G*Power 3.1 procedure [12] (F tests; repeated measures, within–between interaction; two groups, three measurements, $\alpha = 0.05$, power = 0.80, correlation among repeated measures = 0.5, nonsphericity correction $\varepsilon = 1$). With a total sample of 12 players (numerator df = 2, denominator df = 20, critical F = 3.49), the smallest detectable interaction effect was $f = 0.40$ (noncentrality parameter $\lambda = 11.24$), equivalent to $\eta p^2 \approx 0.14$. Power to detect a moderate interaction effect ($f = 0.25$) was 0.40.

Randomisation, allocation and blinding

Players were allocated in a 1:1 ratio by drawing lots, with the restriction that each group contained six players. Folded paper slips, an equal number for each group, were placed in an opaque container and drawn once for each player. BG prepared the slips, enrolled the participants and assigned them to groups according to the draw. No independent allocator was involved, so allocation was not concealed from the investigator who enrolled participants. Players and coaches could not be blinded to the intervention, and both investigators (BG and FT), who collected and analysed the data, were aware of group allocation.

Intervention

Both groups trained once a week for 6 weeks using 3 vs 3 possession games without goalkeepers on an 18 × 30 m pitch (540 m²; 90 m² per player). The aim of the game was to keep possession of the ball as a team. Each set consisted of 2 min of play, a 10-s interval and another 2 min of play (4 min 10 s per set). At the whistle ending the first 2-min bout, EXP players sprinted for 10 s within the playing area and then resumed play immediately, whereas CON players rested passively for 10 s. Three sets were completed per session with 4 min 10 s of passive rest between sets (Table 2). Spare balls were placed around the pitch so that play restarted immediately when the ball went out. Players in each group played only with members of their own group, forming two teams of three.

Table 2
Training protocol

Group	Weeks	Pitch	Format	Set structure	Sets	Rest between sets	Total session time
EXP	6 (1 session/week)	18 × 30 m	3 vs 3	2 min play + 10 s sprint + 2 min play	3	4 min 10 s (passive)	25 min
CON	6 (1 session/week)	18 × 30 m	3 vs 3	2 min play + 10 s passive rest + 2 min play	3	4 min 10 s (passive)	25 min

Before each session, players completed a 10-min general warm-up consisting of 5 min of submaximal running and dynamic stretching, followed by three submaximal repetitions of the test movements.

Outcomes

The primary outcome was jump height. In the trial registry it is listed as “change in neuromuscular fatigue (active vertical jump height)”; the test performed was a countermovement jump on force plates, and because it was measured under rested conditions it is interpreted here as an indicator of neuromuscular performance rather than fatigue. Registered secondary outcomes were RPE, maximum heart rate, estimated VO_2max, accelerations, sprint distance (distance covered above 18.4 km·h^{-1}), distance covered in five speed zones, PlayerLoad, total distance and technical actions. Mean heart rate

and peak speed were not registered and are reported as additional outcomes. VO_2max was registered with a week-3 time point, but the 30 − 15 IFT was performed only at baseline and week 6.

Measurements

Anthropometric measurements were taken at baseline. Pre-, mid- and post-test measurements were taken in week 1 (baseline), week 3 and week 6. External load, heart rate, RPE and technical actions were recorded during the SSG session of each test week. CMJ height was measured before the SSG session at each time point. The 30 − 15 Intermittent Fitness Test (30 − 15 IFT) was administered at pre- and post-test only. All tests were performed in the afternoon at the same time of day.

Anthropometry and body composition. Height was measured with a stadiometer (SECA, Hamburg, Germany) to the nearest 0.1 cm. Body mass, body fat percentage and skeletal muscle mass were measured by bioelectrical impedance (InBody 270, InBody Co., Seoul, Korea). Fat-free mass was calculated as body mass × (1 − body fat percentage / 100).

30 − 15 Intermittent Fitness Test. The test was performed as described by Buchheit [13], using a 28-m shuttle course. Running speed started at 8 km·h⁻¹ and increased by 0.5 km·h⁻¹ at each 45-s stage (30 s of running and 15 s of passive recovery). The test ended when a player could no longer keep the required pace or failed to reach the 3-m tolerance zone at the beep three consecutive times. The speed of the last completed stage was recorded as V_IFT. VO_2max was estimated as 28.3 − 2.15 × G − 0.741 × A − 0.0357 × BM + 0.0586 × A × V_IFT + 1.03 × V_IFT, where G is sex (2 for female), A is age in years and BM is body mass in kg [13]. In female football players, V_IFT has shown high test–retest reliability (ICC = 0.91; CV = 1.8%) [14]. For the post-test estimate, body mass measured at the final session was used.

External load. Players wore a Catapult Vector S7 unit (Catapult Sports, Melbourne, Australia) with 10-Hz global navigation satellite system (GNSS) and inertial sensors, placed between the scapulae in a manufacturer-supplied vest. Each player wore the same unit throughout the study. Data were downloaded and processed in OpenField (Catapult Sports; 2024 version) with the manufacturer's default settings. The variables were total distance; distance covered in five speed zones (zone 1: 0–6 km·h⁻¹; zone 2: 6.1–12 km·h⁻¹; zone 3: 12.1–15.4 km·h⁻¹; zone 4: 15.5–18.3 km·h⁻¹; zone 5: >18.4 km·h⁻¹); peak speed; number of accelerations (default OpenField threshold); and PlayerLoad (arbitrary units), an accelerometer-derived index of overall mechanical load [15]. Vector S7 units show good inter-device reliability for distance and peak speed, whereas acceleration counts are less stable across sessions [16]. The 10-s sprint and rest intervals were excluded from the GNSS analysis, so the analysed period comprised the two 2-min bouts of each set.

Heart rate and RPE. Heart rate was recorded continuously with the Polar Team Pro system (Polar Electro Oy, Kempele, Finland), and mean and maximum heart rate were calculated for each session. RPE was collected 30 min after each session with the CR-10 scale by asking "How was your workout?" [17].

Countermovement jump. CMJ was performed on dual force plates (ForceDecks, VALD Performance, Brisbane, Australia). Players kept their hands on their hips, performed a rapid countermovement to approximately 90° knee flexion and jumped as high as possible. Three trials were performed with 2 min of rest between them, and the highest jump height was used for analysis. Trials with technical errors were repeated. CMJ height measured with ForceDecks has shown high test–retest reliability (ICC = 0.97) [18], and CMJ is commonly used to monitor neuromuscular status [19].

Technical actions. Sessions were filmed with a GoPro Hero 12 Black camera (GoPro, San Mateo, CA, USA) fixed about 5 m above the ground at one corner of the pitch so that the whole playing area was visible. One researcher coded all videos twice [20]. Accurate passes, inaccurate passes (passes intercepted by an opponent or played out of the pitch), balls out of play, one-on-one duels (shielding the ball from an opponent with part or all of the body), feints (body or ball movements to deceive an opponent) and dribbles were counted. Values were expressed per player per set.

Statistical analysis

Data are presented as mean ± standard deviation. Baseline characteristics are reported descriptively without significance testing [10]. The primary analysis was a two-way mixed ANOVA with group (EXP, CON) as the between-subject factor and time (pre, mid, post) as the within-subject factor. The group × time interaction was the test of the intervention effect, and CMJ height was the primary outcome. For VO_2max, time had two levels (pre, post) because the 30 − 15 IFT was not performed at mid-test. When the main effect of time was significant, pairwise comparisons with Bonferroni correction were made with both groups pooled.

Normality was examined with the Shapiro–Wilk test for each group and time point, and sphericity with Mauchly's test. Because these tests have little power with six participants per group, the Greenhouse–Geisser correction was applied to all within-subject effects regardless of the Mauchly result; sphericity was violated ($p < 0.05$) for 10 of the 20 variables with three time points. Partial eta squared (ηp^2) was reported as the effect size and interpreted as small (0.01), moderate (0.06) or large (0.14) [21]; for the primary outcome and for any nominally significant interaction, a 90% CI for ηp^2 was calculated from the noncentral F distribution [22].

Three post hoc sensitivity analyses, which were not specified in the trial registry, were performed. First, post-test values were compared with ANCOVA using pre-test values as the covariate, which estimates the between-group difference after accounting for baseline imbalance. Second, the between-group difference in pre-to-post change was calculated with its 95% CI and Hedges' g. Third, change scores were compared with the Mann–Whitney U test, because several variables departed from normality and RPE is ordinal. These analyses were used to judge the robustness of the primary analysis, not to replace it.

Nineteen secondary group × time interactions were tested. Interaction p values are reported without adjustment, and secondary outcomes are interpreted as exploratory. To indicate the risk of false-positive

findings, Holm-adjusted p values across the 19 secondary interactions [23] are reported for any nominally significant result.

Players in each group trained and were tested within the same 3 vs 3 game, so their GNSS, heart rate and technical data are unlikely to be fully independent. Because each arm consisted of a single game, a game-level (cluster) analysis was not possible, and players were analysed as independent units. This may underestimate standard errors and overstate precision [24], and the results should be interpreted with this in mind.

The doctoral thesis from which this article is derived analysed these data with group entered as a within-subject factor. The present analysis treats group as a between-subject factor, in line with the parallel-group design, and supersedes the thesis results. Significance was set at $p < 0.05$. Analyses were performed in Python 3 with the pingouin package (version 0.7.0) [25], SciPy and statsmodels.

Results

All 12 players completed the intervention and all test sessions, and no adverse events were reported. Baseline characteristics are shown in Table 1; body mass varied more in CON, and CMJ height was higher in EXP at baseline (Table 3). The participant flow is shown in Fig. 1.

Primary outcome: jump height

CMJ height increased over time ($F(2,20) = 47.43$, $p < 0.001$, $\eta p^2 = 0.83$; all pairwise comparisons $p \leq 0.036$). There was no evidence of a group × time interaction ($F(2,20) = 2.61$, $p = 0.100$, $\eta p^2 = 0.21$, 90% CI 0.00 to 0.39; Table 3; Fig. 2A). From pre- to post-test, CMJ height increased by 3.6 cm in EXP and 2.6 cm in CON (difference in change 1.1 cm, 95% CI − 0.4 to 2.6; Hedges' g = 0.85, 95% CI − 0.50 to 2.19). The baseline-adjusted post-test difference was 1.5 cm (95% CI − 0.3 to 3.4; ANCOVA $F(1,9) = 3.42$, $p = 0.098$). All interval estimates include zero, so the data are compatible both with no difference and with a moderately larger increase in EXP. The group main effect ($F(1,10) = 10.80$, $p = 0.008$) reflects the difference already present at baseline (24.7 ± 3.3 vs. 21.3 ± 1.5 cm) and not an effect of the intervention.

Table 3
Jump height (primary outcome) and external load at pre-, mid- and post-test

Group	Pre-test	Mid-test	Post-test	Group	Time	Group × time	Δ EXP − CON (95% CI)	Hedges' g (95% CI)
CMJ height (cm) – primary outcome								
EXP	24.7 ± 3.3	26.5 ± 2.4	28.4 ± 2.9	F = 10.80	F = 47.43	F = 2.61	1.1	0.85
				p = 0.008	p < 0.001	p = 0.100	(− 0.4 to 2.6)	(− 0.50 to 2.19)
CON	21.2 ± 1.5	21.6 ± 1.4	23.8 ± 2.0	ηp² = 0.52	ηp² = 0.83	ηp² = 0.21		
Total distance (m)								
EXP	1025.8 ± 109.0	1104.2 ± 114.4	1250.5 ± 122.9	F = 2.29	F = 82.90	F = 1.22	49.5	0.63
				p = 0.161	p < 0.001	p = 0.302	(− 43.9 to 142.9)	(− 0.69 to 1.95)
CON	976.5 ± 38.3	1027.8 ± 47.7	1151.7 ± 79.1	ηp² = 0.19	ηp² = 0.89	ηp² = 0.11		
Peak speed (km·h⁻¹) †								
EXP	17.2 ± 1.5	17.8 ± 1.0	17.7 ± 0.8	F = 0.04	F = 0.51	F = 1.08	0.5	0.35
				p = 0.850	p = 0.541	p = 0.341	(− 1.2 to 2.2)	(− 0.95 to 1.65)
CON	17.5 ± 1.6	17.3 ± 1.0	17.5 ± 0.5	ηp² = 0.00	ηp² = 0.05	ηp² = 0.10		
Accelerations (n)								
EXP	487.7 ± 28.1	537.5 ± 45.7	587.3 ± 68.2	F = 0.73	F = 39.74	F = 1.91	35.5	0.79
				p = 0.413	p < 0.001	p = 0.192	(− 17.6 to 88.6)	(− 0.54 to 2.13)
CON	484.8 ± 22.0	521.5 ± 40.6	549.0 ± 36.4	ηp² = 0.07	ηp² = 0.80	ηp² = 0.16		
Zone 1, 0–6 km·h⁻¹ (m)								
EXP	502.0 ± 25.7	522.5 ± 34.3	511.0 ± 47.0	F = 0.33	F = 3.19	F = 0.72	−11.2	−0.35
				p = 0.579	p = 0.098	p = 0.428	(− 49.0 to 26.7)	(− 1.65 to 0.95)

Group	Pre-test	Mid-test	Post-test	Group	Time	Group × time	Δ EXP – CON (95% CI)	Hedges' g (95% CI)
CMJ height (cm) – primary outcome								
CON	513.2 ± 42.9	527.3 ± 40.7	533.3 ± 51.9	ηp² = 0.03	ηp² = 0.24	ηp² = 0.07		
Zone 2, 6.1–12 km·h⁻¹ (m)								
EXP	170.3 ± 29.3	185.7 ± 26.2	224.8 ± 28.0	F = 2.15 p = 0.173 ηp² = 0.18	F = 49.23 p < 0.001 ηp² = 0.83	F = 0.07 p = 0.837 ηp² = 0.01	2.3 (− 29.2 to 33.9)	0.09 (− 1.20 to 1.37)
CON	155.0 ± 19.7	166.2 ± 15.7	207.2 ± 18.7					
Zone 3, 12.1–15.4 km·h⁻¹ (m)								
EXP	294.5 ± 78.7	318.2 ± 86.0	406.7 ± 76.6	F = 2.74 p = 0.129 ηp² = 0.21	F = 73.57 p < 0.001 ηp² = 0.88	F = 4.99 p = 0.041 ηp² = 0.33	44.8 (2.7 to 86.9)	1.26 (− 0.14 to 2.67)
CON	251.8 ± 35.3	270.0 ± 42.2	319.2 ± 47.1					
Zone 4, 15.5–18.3 km·h⁻¹ (m)								
EXP	55.5 ± 21.7	71.5 ± 23.7	99.3 ± 36.4	F = 0.50 p = 0.497 ηp² = 0.05	F = 29.37 p < 0.001 ηp² = 0.75	F = 0.50 p = 0.526 ηp² = 0.05	10.2 (− 20.2 to 40.6)	0.40 (− 0.90 to 1.70)
CON	50.8 ± 24.6	60.0 ± 26.1	84.5 ± 29.5					
Sprint distance, zone 5 > 18.4 km·h⁻¹ (m)								
EXP	3.8 ± 4.5	3.7 ± 3.4	7.0 ± 5.1	F = 0.14 p = 0.714 ηp² = 0.01	F = 4.96 p = 0.032 ηp² = 0.33	F = 0.23 p = 0.726 ηp² = 0.02	1.3 (− 4.7 to 7.4)	0.26 (− 1.03 to 1.55)
CON	5.7 ± 7.1	4.3 ± 4.9	7.5 ± 4.7					

Values are mean ± SD (n = 6 per group). Statistics are F; p; ηp². Group effect F(1,10); time and group × time effects F(2,20) unless stated otherwise; p values for within-subject effects are Greenhouse–Geisser

corrected. Δ: between-group difference in pre-to-post change. Shaded row: significant group × time interaction.

Interaction p values are not adjusted for multiplicity. Zone 3 distance: Holm-adjusted interaction p = 0.78 (19 secondary interactions); ANCOVA F(1,9) = 5.50, p = 0.044; Mann–Whitney U on change scores p = 0.093. CMJ: countermovement jump. † Not a registered outcome (additional).

Physiological and perceptual responses

Mean heart rate ($F(2,20) = 18.14$, $p < 0.001$, $\eta p^2 = 0.64$), PlayerLoad ($F(2,20) = 11.56$, $p = 0.001$, $\eta p^2 = 0.54$) and RPE ($F(2,20) = 13.00$, $p < 0.001$, $\eta p^2 = 0.57$) increased over time, with no evidence of a group × time interaction ($p \geq 0.29$; Table 4; Fig. 2D). PlayerLoad was higher at mid-test ($p = 0.022$) and post-test ($p = 0.004$) than at pre-test, RPE was higher at post-test than at pre-test ($p = 0.002$), and mean heart rate increased between each pair of time points ($p \leq 0.024$). Mean heart rate was higher in CON than in EXP at all time points, including baseline (group effect: $F(1,10) = 9.52$, $p = 0.012$). There was no evidence of time, group or interaction effects for maximum heart rate or estimated VO_2max ($p \geq 0.10$).

Table 4
Physiological and perceptual responses at pre-, mid- and post-test

Group	Pre-test	Mid-test	Post-test	Group	Time	Group × time	Δ EXP − CON (95% CI)	Hedges' g (95% CI)
Maximum heart rate (bpm)								
EXP	180.8 ± 10.2	183.0 ± 8.4	183.8 ± 4.4	F = 0.44 p = 0.521 ηp² = 0.04	F = 2.34 p = 0.156 ηp² = 0.19	F = 0.15 p = 0.711 ηp² = 0.01	1.2 (− 5.7 to 8.1)	0.20 (− 1.09 to 1.49)
CON	178.7 ± 8.6	179.8 ± 8.0	180.5 ± 6.1					
Mean heart rate (bpm) †								
EXP	135.8 ± 7.6	141.0 ± 5.7	143.0 ± 3.7	F = 9.52 p = 0.012 ηp² = 0.49	F = 18.14 p < 0.001 ηp² = 0.64	F = 1.28 p = 0.294 ηp² = 0.11	−1.2 (− 8.9 to 6.6)	−0.18 (− 1.47 to 1.11)
CON	148.2 ± 9.2	150.5 ± 8.0	156.5 ± 6.9					
PlayerLoad (AU)								
EXP	176.5 ± 18.7	187.5 ± 8.9	194.0 ± 9.5	F = 2.92 p = 0.118 ηp² = 0.23	F = 11.56 p < 0.001 ηp² = 0.54	F = 0.31 p = 0.715 ηp² = 0.03	−6.7 (− 28.8 to 15.4)	−0.36 (− 1.65 to 0.94)
CON	162.0 ± 18.3	175.0 ± 13.0	186.2 ± 16.2					
VO2max (ml·kg⁻¹·min⁻¹)								
EXP	51.6 ± 4.6	–	52.0 ± 4.3	F = 3.04 p = 0.112 ηp² = 0.23	F = 3.18 p = 0.105 ηp² = 0.24	F = 2.27 p = 0.162 ηp² = 0.19	0.4 (− 0.2 to 0.9)	0.80 (− 0.53 to 2.14)
CON	47.7 ± 3.6	–	47.8 ± 3.7					
RPE (CR-10)								
EXP	5.3 ± 0.5	5.7 ± 0.8	6.0 ± 0.6	F = 0.14 p = 0.713	F = 13.00 p < 0.001	F = 1.00 p = 0.380	−0.5 (− 1.3 to 0.3)	−0.72 (− 2.04 to 0.61)

Group	Pre-test	Mid-test	Post-test	Group	Time	Group × time	Δ EXP − CON (95% CI)	Hedges' g (95% CI)
Maximum heart rate (bpm)								
CON	5.2 ± 0.4	5.8 ± 0.8	6.3 ± 0.5	ηp^2 = 0.01	ηp^2 = 0.57	ηp^2 = 0.09		

Values are mean ± SD (n = 6 per group). Statistics are F; p; ηp^2. Group effect F(1,10); time and group × time effects F(2,20) unless stated otherwise; p values for within-subject effects are Greenhouse–Geisser corrected. Δ: between-group difference in pre-to-post change. Shaded row: significant group × time interaction.

VO_2max was estimated from the 30 − 15 IFT at pre- and post-test only; time and group × time effects are F(1,10). RPE: rating of perceived exertion (CR-10); AU: arbitrary units. † Not a registered outcome (additional).

External load

Total distance (F(2,20) = 82.90, ηp^2 = 0.89), accelerations (F(2,20) = 39.74, ηp^2 = 0.80) and distance in zones 2 (F(2,20) = 49.23, ηp^2 = 0.83), 3 (F(2,20) = 73.57, ηp^2 = 0.88) and 4 (F(2,20) = 29.37, ηp^2 = 0.75) increased over time (all $p < 0.001$; Table 3; Fig. 2C). For each of these variables, all pairwise differences between time points were significant ($p \leq 0.007$). Sprint distance (zone 5) showed a time effect (F(2,20) = 4.96, p = 0.032, ηp^2 = 0.33) that was due to an increase from mid- to post-test (p = 0.025); the pre-to-post difference was not significant (p = 0.246). There was no evidence of change in peak speed or zone 1 distance ($p \geq 0.10$).

Of the 19 secondary interaction tests, only zone 3 distance showed a nominally significant group × time interaction (F(2,20) = 4.99, p = 0.041, ηp^2 = 0.33, 90% CI 0.04 to 0.50; Fig. 2B). From pre- to post-test, zone 3 distance increased by 112.2 m in EXP and by 67.3 m in CON (difference in change 44.8 m, 95% CI 2.7 to 86.9; Hedges' g = 1.26, 95% CI − 0.14 to 2.67). The baseline-adjusted difference was 49.1 m (95% CI 1.8 to 96.5; ANCOVA p = 0.044). However, the Mann–Whitney U test on change scores was not significant (p = 0.093), and the interaction did not remain significant after Holm adjustment for the 19 secondary interactions (adjusted p = 0.78). This result is therefore regarded as exploratory.

Technical actions

There was no evidence of a group × time interaction for any technical variable ($p \geq 0.23$; Table 5). The number of accurate passes changed over time (F(2,20) = 4.18, p = 0.039, ηp^2 = 0.29), increasing from pre- to mid-test (p = 0.017), and was higher in EXP than in CON at all time points, including pre-test (group effect: F(1,10) = 6.36, p = 0.030), which indicates a pre-existing difference. Dribbles and one-on-one duels increased numerically in EXP but not in CON. The baseline-adjusted ANCOVA was nominally significant for dribbles (p = 0.020) and duels (p = 0.047), but the interaction tests (p = 0.336 and p = 0.231) and the

Mann–Whitney U tests ($p = 0.093$ and $p = 0.180$) were not, and these post hoc results are not interpreted as evidence of an effect.

Table 5
Technical actions at pre-, mid- and post-test

Group	Pre-test	Mid-test	Post-test	Group	Time	Group × time	Δ EXP – CON (95% CI)	Hedges' g (95% CI)
Accurate passes (n/set)								
EXP	8.55 ± 1.17	9.77 ± 1.44	9.38 ± 1.41	F = 6.36 p = 0.030 ηp² = 0.39	F = 4.18 p = 0.039 ηp² = 0.29	F = 0.26 p = 0.740 ηp² = 0.03	0.28 (− 1.57 to 2.12)	0.18 (− 1.11 to 1.47)
CON	7.44 ± 0.46	8.16 ± 0.81	8.00 ± 1.38					
Inaccurate passes (n/set)								
EXP	2.05 ± 0.83	2.39 ± 0.54	2.11 ± 0.62	F = 0.10 p = 0.756 ηp² = 0.01	F = 1.32 p = 0.288 ηp² = 0.12	F = 1.16 p = 0.326 ηp² = 0.10	0.61 (− 0.47 to 1.69)	0.67 (− 0.65 to 1.99)
CON	2.33 ± 0.79	2.16 ± 0.62	1.77 ± 0.54					
Ball out of play (n/set)								
EXP	1.83 ± 0.72	1.50 ± 0.59	1.66 ± 0.60	F = 0.01 p = 0.919 ηp² = 0.00	F = 0.69 p = 0.507 ηp² = 0.06	F = 0.18 p = 0.831 ηp² = 0.02	0.11 (− 0.96 to 1.18)	0.12 (− 1.16 to 1.41)
CON	1.83 ± 0.35	1.66 ± 0.60	1.55 ± 0.34					
One-on-one duels (n/set)								
EXP	0.33 ± 0.30	0.61 ± 0.39	0.89 ± 0.40	F = 2.88 p = 0.120 ηp² = 0.22	F = 1.57 p = 0.238 ηp² = 0.14	F = 1.61 p = 0.231 ηp² = 0.14	0.56 (− 0.20 to 1.32)	0.88 (− 0.47 to 2.22)
CON	0.44 ± 0.50	0.44 ± 0.27	0.44 ± 0.17					
Feints (n/set)								
EXP	0.50 ± 0.59	0.44 ± 0.34	0.39 ± 0.39	F = 3.91 p = 0.076	F = 0.48 p = 0.606	F = 0.12 p = 0.865	0.00 (− 0.67 to 0.68)	0.00 (− 1.28 to 1.29)

Group	Pre-test	Mid-test	Post-test	Group	Time	Group × time	Δ EXP − CON (95% CI)	Hedges' g (95% CI)
Accurate passes (n/set)								
CON	0.83 ± 0.41	0.89 ± 0.35	0.72 ± 0.39	ηp² = 0.28	ηp² = 0.05	ηp² = 0.01		
Dribbles (n/set)								
EXP	0.72 ± 0.65	0.94 ± 0.77	1.22 ± 0.55	F = 4.53 p = 0.059 ηp² = 0.31	F = 1.10 p = 0.344 ηp² = 0.10	F = 1.13 p = 0.336 ηp² = 0.10	0.50 (0.04 to 0.97)	1.28 (− 0.13 to 2.70)
CON	0.44 ± 0.40	0.44 ± 0.27	0.44 ± 0.46					

Values are mean ± SD (n = 6 per group). Statistics are F; p; ηp². Group effect F(1,10); time and group × time effects F(2,20) unless stated otherwise; p values for within-subject effects are Greenhouse–Geisser corrected. Δ: between-group difference in pre-to-post change. Shaded row: significant group × time interaction.

Technical actions are expressed per player per set.

Discussion

Main findings

In this small randomised trial, adding a 10-s sprint to each set of 3 vs 3 SSGs once a week for 6 weeks did not produce clear differences from standard SSGs. Both groups showed similar increases in CMJ height and in most measures of external and internal load. The only between-format difference, a larger increase in distance covered at 12.1–15.4 km·h⁻¹ in EXP, was nominally significant but did not withstand adjustment for multiple testing. The hypothesis that the sprint-embedded format would increase high-speed running and physiological load was therefore not supported.

Primary outcome

CMJ height increased by 2.6–3.6 cm in both groups, with no evidence of a group × time interaction. The confidence intervals for the between-group difference ranged from a small disadvantage to a moderate advantage for EXP. With a power of about 0.40 to detect a moderate interaction effect, the trial cannot exclude a moderate effect of the sprint-embedded format, and the absence of a significant interaction should not be read as evidence that the formats are equivalent [26]. The higher values in EXP were present from baseline and should not be attributed to the intervention. CMJ is commonly used to

monitor neuromuscular status [19], but acute neuromuscular fatigue requires measurements taken shortly before and after a session [27]. Because CMJ was measured under rested conditions, the outcome registered as “neuromuscular fatigue” is interpreted here as neuromuscular performance, and the data do not show whether either format increased or reduced fatigue. The similar increase in both groups may reflect familiarisation with the test, training outside the study or seasonal changes; without a group that did not play SSGs, it cannot be attributed to SSG training.

Secondary outcomes

Total distance, accelerations, PlayerLoad, mean heart rate and RPE increased in parallel in both groups. Because RPE and mean heart rate rose together with running output, these changes are more consistent with players working harder in later sessions than with more economical movement. Familiarisation with the rules and pitch and better collective organisation may have contributed. The larger increase in zone 3 distance in EXP is compatible with the sprint bout altering running during the games that followed it. Because the 10-s sprints were excluded from the GNSS analysis, this difference describes running during play, not the load of the sprints themselves. The finding is exploratory: it was one of 19 secondary interaction tests, it did not survive Holm adjustment, and the non-parametric test did not support it. There was no evidence of a difference in peak speed or in distance above 18.4 km·h^{-1}. Estimated VO_2max, maximum heart rate and technical actions also showed no evidence of a group × time interaction; the between-group differences in mean heart rate and accurate passes were present from baseline.

Comparison with previous studies

Köklü et al. reported higher blood lactate and RPE, and greater total and moderate- to high-speed running distance, when two 15-s running drills were added to 3 vs 3 and 4 vs 4 SSGs [7]. In the present study, the added running amounted to three 10-s bouts per session, which is small compared with 12 min of play, and it was not included in the analysed GNSS data. This may explain why heart rate, PlayerLoad and RPE did not differ between formats. Pitch size and surface affect the physiological and perceptual responses of female players to SSGs [8], and the relative area per player constrains high-speed running [4]. Replicating match-derived high-intensity demands requires careful choice of format and pitch size [5]. Technical demands in SSGs also depend on format and bout duration [20, 28].

Possible mechanisms

A short maximal effort might raise running tempo when play resumes, which would be consistent with the zone 3 result, but this mechanism was not measured and remains speculative. The relative area of 90 m² per player probably limited the speeds that could be reached during play in both formats, which may explain the lack of differences in peak speed and zone 5 distance.

Practical applications

For coaches, the results suggest that inserting a brief sprint into 3 vs 3 SSGs on an 18 × 30 m pitch does not by itself change heart rate, perceived exertion, jump performance or estimated aerobic fitness

compared with the standard format. It may slightly increase moderate- to high-speed running during play, but this effect was uncertain, and there was no evidence of more very high-speed running or higher peak speeds. Coaches who want players to reach near-maximal speeds may need larger pitches or separate sprint drills. Because the trial was small and each format was delivered as a single game, these suggestions need confirmation in larger samples and other game formats.

Strengths

The study used a randomised parallel-group design in female players, who are under-represented in SSG research. Jump height and external load were measured with force plates and 10-Hz GNSS units whose reliability has been reported [16, 18]. Follow-up was complete, with no missing data. Results are reported with effect sizes, confidence intervals, sensitivity analyses and individual data (Fig. 2).

Limitations

The limitations are listed in order of their likely influence on the conclusions. First, the sample was small (n = 6 per group). The trial could detect only large interaction effects ($f \geq 0.40$), and its power to detect a moderate effect was about 0.40, so moderate but practically relevant differences may have been missed. Second, each group played as a single 3 vs 3 game, so players' outputs were probably not independent. The effective sample size may therefore be smaller than 12, and the p values and confidence intervals may be too optimistic; this cannot be resolved retrospectively. Third, there was no group that did not play SSGs, so changes over time cannot be attributed to the training. Fourth, the trial was registered retrospectively, after data collection, and the registry descriptions of some outcomes were refined for this report. The sensitivity analyses were not specified in the registry.

Fifth, 19 secondary interactions were tested without formal adjustment, which increases the risk of false-positive findings; the zone 3 result should be viewed in this light. Sixth, technical actions were coded twice by a single researcher, but no formal estimate of intra-rater reliability was calculated. Seventh, acceleration counts from the GNSS units used here are less reliable across sessions than distance and speed [16]. Eighth, menstrual cycle phase and hormonal contraceptive use were not recorded [9]. Ninth, VO_2max was estimated from the 30 − 15 IFT rather than measured directly. Tenth, players and coaches could not be blinded. Finally, one investigator enrolled participants, prepared the allocation and delivered the intervention, and both investigators who assessed outcomes and analysed the data were aware of group allocation, which may have introduced allocation or assessment bias. The findings apply to young adult licensed players and should not be generalised to elite or professional women's football.

Future research

Future trials should randomise several teams or games to each format and analyse them as clusters, include a group that does not play SSGs, and recruit larger samples of female players. Measuring CMJ immediately before and after sessions would allow acute neuromuscular fatigue to be assessed, and

including the sprint period in the GNSS analysis would quantify the total load of the sprint-embedded format. Prospective registration of the protocol and analysis plan, recording of menstrual cycle status and formal reliability testing of technical coding are also recommended.

Conclusions

In a small sample of female football players, 3 vs 3 SSGs with and without an embedded 10-s sprint, performed once a week for 6 weeks, were associated with similar changes in CMJ height and in most physiological, perceptual, external-load and technical variables. There was weak, exploratory evidence of a larger increase in distance covered at 12.1–15.4 $km \cdot h^{-1}$ with the sprint-embedded format. These findings should be interpreted cautiously and need confirmation in larger, cluster-randomised trials.

Abbreviations

CI: confidence interval; CMJ: countermovement jump; CON: control group; EXP: experimental group; GNSS: global navigation satellite system; ηp^2: partial eta squared; RPE: rating of perceived exertion; SSG: small-sided game; V_IFT: final velocity of the 30-15 Intermittent Fitness Test; VO_2max: maximal oxygen uptake; 30-15 IFT: 30-15 Intermittent Fitness Test.

Declarations

Ethics approval: The study protocol was reviewed and approved by the Ethics Committee of Istanbul Gelişim University (İstanbul Gelişim Üniversitesi Rektörlüğü Etik Kurulu; meeting no. 2024-01, 12 January 2024; decision no. 2024-01-107). All procedures were performed in accordance with the Declaration of Helsinki and its later amendments.

Consent to participate: Written informed consent to participate was obtained from all participants before enrolment. All participants were adults (18–25 years) and were informed that they could withdraw from the study at any time without giving a reason.

Consent for publication: Not applicable. This manuscript does not contain any individual person's identifiable data, images or videos; individual data in Figure 2 are presented anonymously.

Availability of data and materials: The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

Competing interests: The authors declare that they have no competing interests.

Funding: This study was supported by the Istanbul Gelişim University Scientific Research Projects Coordination Office (project no. TDK-2024-135). The funder had no role in the design of the study, data collection, analysis, interpretation or writing of the manuscript.

Authors' contributions: BG conceived and designed the study, recruited the participants, prepared the randomisation, supervised the training sessions, collected and analysed the data, prepared the tables and figures and drafted the manuscript. FT supervised the study, contributed to the study design, collected and analysed the data, contributed to the interpretation of the results and critically revised the manuscript for important intellectual content. Both authors read and approved the final manuscript and agree to be accountable for all aspects of the work.

Acknowledgements: This article is derived from the first author's doctoral thesis, completed in 2025 at the Institute of Graduate Studies, Istanbul Gelişim University, under the supervision of FT. The authors thank the players who took part in the study.

References

1. Hill-Haas SV, Dawson B, Impellizzeri FM, Coutts AJ. Physiology of small-sided games training in football: a systematic review. Sports Med. 2011;41(3):199–220. https://doi.org/10.2165/11539740-000000000-00000.
2. Davids K, Araújo D, Correia V, Vilar L. How small-sided and conditioned games enhance acquisition of movement and decision-making skills. Exerc Sport Sci Rev. 2013;41(3):154–61.
3. Rampinini E, Impellizzeri FM, Castagna C, Abt G, Chamari K, Sassi A, et al. Factors influencing physiological responses to small-sided soccer games. J Sports Sci. 2007;25(6):659–66.
4. Castellano J, Puente A, Echeazarra I, Casamichana D. Influence of the number of players and the relative pitch area per player on heart rate and physical demands in youth soccer. J Strength Cond Res. 2015;29(6):1683–91. https://doi.org/10.1519/JSC.0000000000000788.
5. Riboli A, Esposito F, Coratella G. Small-sided games in elite football: practical solutions to replicate the 4-min match-derived maximal intensities. J Strength Cond Res. 2023;37(2):366–74. https://doi.org/10.1519/JSC.0000000000004249.
6. Eniseler N, Şahan Ç, Özcan I, Dinler K. High-intensity small-sided games versus repeated sprint training in junior soccer players. J Hum Kinet. 2017;60:101–11. https://doi.org/10.1515/hukin-2017-0104.
7. Köklü Y, Cihan H, Alemdaroğlu U, Dellal A, Wong DP. Acute effects of small-sided games combined with running drills on internal and external loads in young soccer players. Biol Sport. 2020;37(4):375–81. https://doi.org/10.5114/biolsport.2020.96943.
8. López-Fernández J, Sánchez-Sánchez J, Rodríguez-Cañamero S, Ubago-Guisado E, Colino E, Gallardo L. Physiological responses, fatigue and perception of female soccer players in small-sided games with different pitch size and sport surfaces. Biol Sport. 2018;35(3):291–9. https://doi.org/10.5114/biolsport.2018.77829.
9. Elliott-Sale KJ, Minahan CL, de Jonge XAKJ, Ackerman KE, Sipilä S, Constantini NW, et al. Methodological considerations for studies in sport and exercise science with women as

participants: a working guide for standards of practice for research on women. Sports Med. 2021;51(5):843–61.

10. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c332.
11. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human participants. JAMA. 2025;333(1):71–4.
12. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods. 2007;39(2):175–91.
13. Buchheit M. The 30 – 15 intermittent fitness test: accuracy for individualizing interval training of young intermittent sport players. J Strength Cond Res. 2008;22(2):365–74.
14. Čović N, Jelešković E, Alić H, Rađo I, Kafedžić E, Sporiš G, et al. Reliability, validity and usefulness of 30–15 intermittent fitness test in female soccer players. Front Physiol. 2016;7:510. https://doi.org/10.3389/fphys.2016.00510.
15. Boyd LJ, Ball K, Aughey RJ. The reliability of MinimaxX accelerometers for measuring physical activity in Australian football. Int J Sports Physiol Perform. 2011;6(3):311–21.
16. Crang ZL, Duthie G, Cole MH, Weakley J, Hewitt A, Johnston RD. The inter-device reliability of global navigation satellite systems during team sport movement across multiple days. J Sci Med Sport. 2022;25(4):340–4. https://doi.org/10.1016/j.jsams.2021.11.044.
17. Foster C, Florhaug JA, Franklin J, Gottschall L, Hrovatin LA, Parker S, et al. A new approach to monitoring exercise training. J Strength Cond Res. 2001;15(1):109–15.
18. Collings TJ, Lima YL, Dutaillis B, Bourne MN. Concurrent validity and test–retest reliability of VALD ForceDecks' strength, balance, and movement assessment tests. J Sci Med Sport. 2024;27(8):572–80. https://doi.org/10.1016/j.jsams.2024.04.014.
19. Claudino JG, Cronin J, Mezêncio B, McMaster DT, McGuigan M, Tricoli V, et al. The countermovement jump to monitor neuromuscular status: a meta-analysis. J Sci Med Sport. 2017;20(4):397–402.
20. Fanchini M, Azzalin A, Castagna C, Schena F, McCall A, Impellizzeri FM. Effect of bout duration on exercise intensity and technical performance of small-sided games in soccer. J Strength Cond Res. 2011;25(2):453–8.
21. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates; 1988.
22. Steiger JH. Beyond the F test: effect size confidence intervals and tests of close fit in the analysis of variance and contrast analysis. Psychol Methods. 2004;9(2):164–82.
23. Holm S. A simple sequentially rejective multiple test procedure. Scand J Stat. 1979;6(2):65–70.
24. Hurlbert SH. Pseudoreplication and the design of ecological field experiments. Ecol Monogr. 1984;54(2):187–211. https://doi.org/10.2307/1942661.

25. Vallat R. Pingouin: statistics in Python. J Open Source Softw. 2018;3(31):1026. https://doi.org/10.21105/joss.01026.
26. Altman DG, Bland JM. Absence of evidence is not evidence of absence. BMJ. 1995;311(7003):485.
27. Gathercole R, Sporer B, Stellingwerff T, Sleivert G. Alternative countermovement-jump analysis to quantify acute neuromuscular fatigue. Int J Sports Physiol Perform. 2015;10(1):84–92.
28. Owen AL, Wong DP, McKenna M, Dellal A. Heart rate responses and technical comparison between small- vs. large-sided games in elite professional soccer. J Strength Cond Res. 2011;25(8):2104–10. https://doi.org/10.1519/JSC.0b013e3181f0a8a3.

Figures

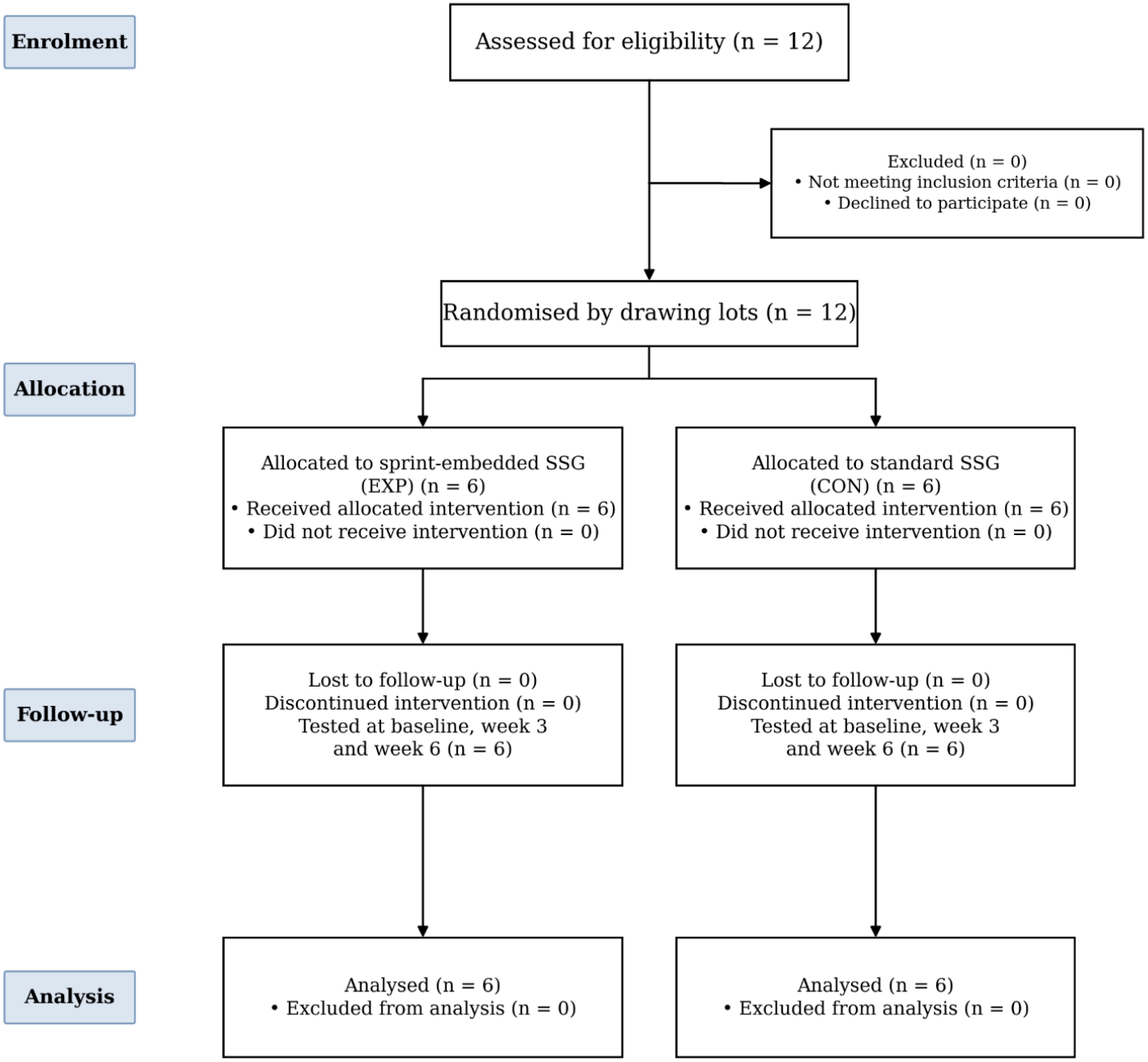


Figure 1

CONSORT flow diagram of participants through the trial. EXP: sprint-embedded small-sided game group; CON: standard small-sided game group.

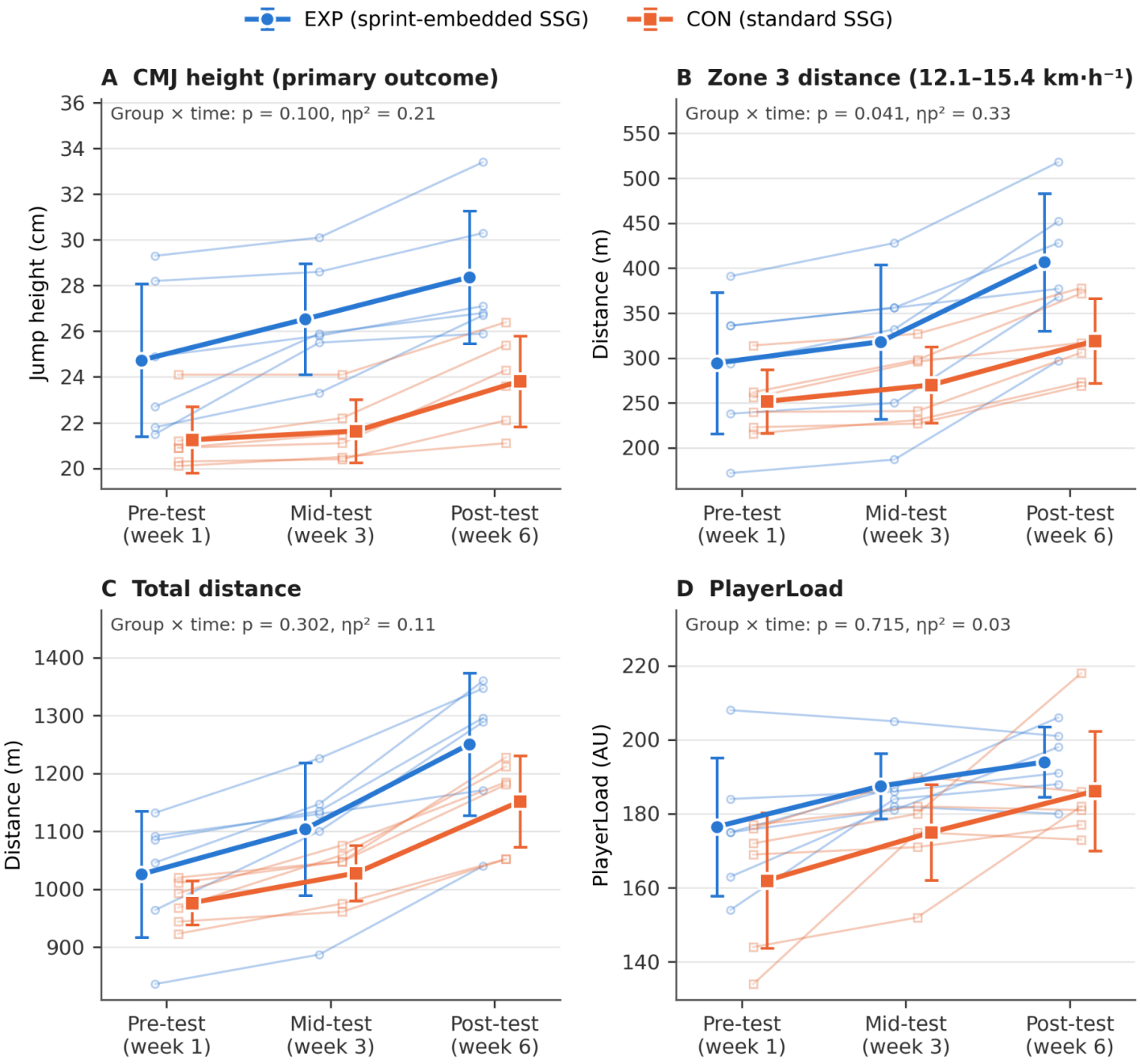


Figure 2

Individual and group responses at pre-, mid- and post-test. Thin lines show individual players (circles, EXP; squares, CON); thick lines show group means ± SD. (A) Countermovement jump height (primary outcome); (B) distance covered at 12.1–15.4 km·h⁻¹ (zone 3); (C) total distance; (D) PlayerLoad. Group × time p values are Greenhouse–Geisser corrected. EXP: sprint-embedded small-sided game group (n = 6); CON: standard small-sided game group (n = 6); AU: arbitrary units.